

Università della Calabria  
Dipartimento di linguistica

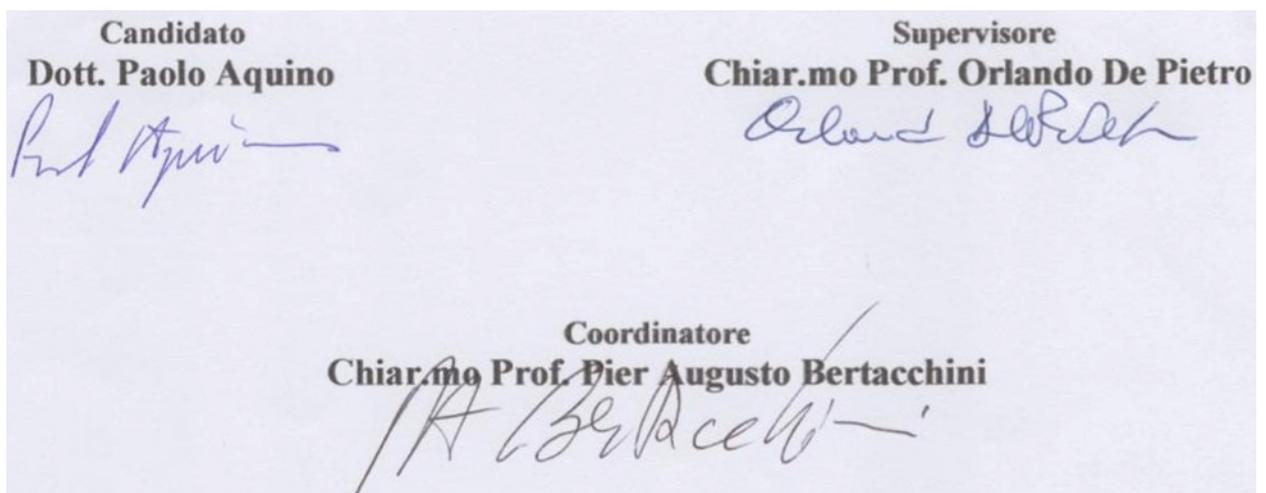
Dottorato di ricerca in  
“Psicologia della Programmazione e Intelligenza Artificiale”

**XXI ciclo**

**Tesi di dottorato**

*Processi decisionali, anomalie cognitive, e intelligenza artificiale  
analisi teoriche e applicazioni empiriche*

(Settore scientifico disciplinare prevalente: M-PSI/01 Psicologia generale)





# Indice

## Introduzione

pag. 6

## Capitolo 1 Analisi economiche e psicologiche delle decisioni

1.1 Introduzione	11
1.2 Gli economisti e la teoria dell'utilità attesa	11
1.3 Gli psicologi e la razionalità limitata e le anomalie cognitive	13
1.3.1 La teoria della razionalità limitata	15
1.3.2 Le anomalie cognitive nei processi decisionali	16
1.4 Fattori che impediscono il processo decisionale	19
1.4.1 Acquistare seguendo dei processi decisionali semplificati	21
1.5 I Conflitti decisionali e le influenze in una decisione	25
1.5.1 Le influenze nel processo decisionale derivanti dal contesto	26
1.6 Strategie decisionali per superare i conflitti decisionali	28
1.6.1 Il modello decisionale della <i>eliminazione per aspetti</i> e l' <i>avversione per le perdite</i>	29
1.6.2 La teoria del <i>campo decisionale</i>	30
1.6.3 Confronto dei diversi criteri decisionali	31
1.7 I conflitti decisionali nelle strategie aziendali e nelle strategie politiche	32
1.7.1 Distorsioni cognitive nelle strategie aziendali	33
1.7.2 Anomalie cognitive nelle strategie politiche: il <i>paternalismo progressista</i>	34
1.8 Conclusioni	35

## Capitolo 2 L'intelligenza artificiale e gli agenti intelligenti

2.1 Introduzione	37
2.2 Nozioni di <i>intelligenza artificiale</i>	37
2.2.1 Principali campi d'indagine dell'intelligenza artificiale	43
2.3 Le principali tecniche di intelligenza artificiale	44
2.4 Gli agenti intelligenti	48
2.4.1 Agenti intelligenti: tecniche per fronteggiare l'incertezza.	50
2.4.2 La Teoria dei Giochi per fronteggiare l'incertezza	53
2.4.3 La ricerca delle azioni migliori	56
2.4.4 Agenti per il riconoscimento del linguaggio naturale	57
2.4.5 Agenti a supporto delle decisioni di acquisto online	59
2.5 L'apprendimento automatico degli agenti intelligenti	60
2.5.1 L'apprendimento automatico con <i>alberi decisionali</i>	61

2.5.2 L'apprendimento automatico <i>per conoscenze</i>	62
2.5.3 L'apprendimento automatico <i>statistico</i>	63
2.5.4 L'apprendimento automatico mediante <i>reti neurali</i>	64
2.5.5 L'apprendimento automatico <i>per rinforzo</i>	65
2.5.6 Il comportamento ottimale di un agente intelligente	66
2.6 Conclusioni	66

### **Capitolo 3 Sistemi di supporto alle decisioni di acquisto**

3.1 Introduzione	68
3.2 Il filtraggio automatico delle informazioni	68
3.3 Le principali metodologie di filtraggio delle informazioni	72
3.4 I Sistemi di estrazione delle informazioni	77
3.5 I sistemi di supporto alle decisioni di acquisto	77
3.5.1 Sistemi compensatori e non compensatori	80
3.6 Decisioni di acquisto di prodotti con molte caratteristiche	81
3.6.1 Le caratteristiche di un telefono cellulare	82
3.7 Valutare le relazioni tra le caratteristiche attraverso le tecniche di regressione lineare	85
3.8 Recuperare informazioni rilevanti tra tante informazione superflue, attraverso le tecniche di <i>data mining</i>	87
3.9 Conclusioni	89

### **Capitolo 4 Commercio elettronico e reputazione online**

4.1 Introduzione	91
4.2 Le difficoltà di un sito e-commerce	91
4.3 I siti "comparatori" di prodotti online	94
4.3.1 Cosa può fare un sito e-commerce per agevolare i motori di ricerca dedicati allo shopping	95
4.3.2 Phonescoop.com: un portale che seleziona i cellulari con un sistema <i>content based</i> .	96
4.3.3 Retrevo.com: con un sistema <i>content based</i> , mette in relazione prezzo e performance delle caratteristiche per diverse tipologie di prodotto.	99
4.3.4 Twenga.it: uno dei più completi <i>price comparison</i>	101
4.4 La reputazione online	104
4.4.1 L'influenza dei blog nelle decisioni di acquisto dei consumatori	107
4.4.2 Le attività di correzione della reputazione online	108
4.4.3 Monitorare le preferenze dei consumatori	109
4.5 Indagini sugli acquisti online nei principali paesi europei	109
4.5.1 Indagini sugli acquisti online italiani	111
4.6 Sviluppi nelle tecnologie dell'informazione	115
4.7 Conclusioni	116

## **Capitolo 5 Analisi empiriche e indagini esplorative sulle decisioni di acquisto nel mercato della telefonia mobile**

5.1 Introduzione	117
5.2 Analisi sui cellulari venduti online sul sito <i>kelkoo.it</i>	117
5.2.1 Statistiche descrittive e matrice di correlazione	118
5.2.2 Analisi di regressione multivariata	121
5.2.2.1 Analisi di regressione con eliminazione progressiva delle variabili	127
5.2.2.2 Analisi di <i>stepwise regression</i>	129
5.2.3 Analisi di <i>data mining</i>	130
5.2.3.1 Analisi di <i>clustering</i>	130
5.2.3.2 Analisi di <i>decision trees</i>	134
5.2.3.3 La rete delle dipendenze	138
5.3 Analisi sui cellulari venduti in Italia	139
5.3.1 Analisi di <i>data mining</i>	139
5.3.1.1 Analisi di <i>clustering</i>	139
5.3.1.2 Analisi di <i>decision trees</i>	143
5.3.1.3 La rete delle dipendenze	143
5.3.2 La penetrazione sul mercato di diverse caratteristiche dei telefoni cellulari	144
5.4 Analisi delle opinioni degli utenti del sito <i>Kelkoo.it</i>	146
5.4.1 Sintesi dei risultati	147
5.4.2 Un'analisi di <i>Data Mining</i>	152
5.4.2.1 Analisi di <i>decision trees</i>	152
5.5 Conclusioni	154

## **Capitolo 6 Progettazione e sviluppo di un nuovo sistema di supporto alle decisioni**

6.1 Introduzione	157
6.2 Una diversa metodologia decisionale	157
6.3 La progettazione dell'applicazione	159
6.4 Valutazione del sistema di aiuto alle decisioni "Acquisto Facile"	165
6.4.1 Metodologia e risultati dell'indagine	166
6.4.2 Il Questionario	168
6.5 Alcuni approfondimenti sulla metodologia di supporto alle decisioni	172
6.6 Conclusioni	177

<b>Conclusioni</b>	<b>179</b>
--------------------	------------

<b>Riferimenti bibliografici</b>	<b>185</b>
----------------------------------	------------

## Introduzione

Il principale obiettivo delle attività di ricerca svolte durante i tre anni di dottorato è stato quello di analizzare il comportamento dei consumatori (*consumer behaviour*) durante la fase di scelta dei prodotti da acquistare. A questo scopo sono stati analizzati i processi decisionali in condizioni di incertezza, tipicamente attuati dagli acquirenti nel momento in cui devono decidere il prodotto da acquistare fra i tanti che il mercato offre. In particolare sono stati analizzati i diversi sistemi di supporto alle decisioni, tecnologici e non, oggi a disposizione delle aziende, e quindi di chi deve acquistare; sono stati studiati anche alcuni sviluppi dell'intelligenza artificiale e analizzate alcune metodologie del cosiddetto "filtraggio" delle informazioni. In considerazione di ciò, la tesi si pone come obiettivo finale quello di evidenziare la validità di una diversa metodologia di filtraggio, che riprende la metodologia adottata da diversi agenti *content based filtering*, ma che attenua i problemi di compensazione tra attributi positivi e negativi dei prodotti. Questo sistema di supporto alle decisioni di acquisto è basato sul principio della minimizzazione degli scarti negativi fra i livelli effettivi di performance delle diverse caratteristiche di un prodotto e i livelli minimi di performance richiesti; esso è stato sperimentato sui processi di acquisto di un prodotto "molto complesso", perché caratterizzato da molteplici attributi/funzioni, come il telefono cellulare. Il sistema è stato sottoposto a verifica, sia online, sia intervistando i clienti di un punto di vendita tradizionale, ed è disponibile sul sito creato ad hoc "*SceltaGiusta.it*" all'indirizzo <http://www.sceltagiusta.it>.

Al fine di esemplificare il percorso di studio e ricerca seguito durante il dottorato e nello specifico nell'elaborazione della tesi finale, vengono qui di seguito richiamati i focus degli studi effettuati con i principali argomenti trattati e analizzati; per una descrizione più analitica degli argomenti trattati si rimanda alla sezione "Conclusioni" alla fine di questa tesi.

Le analisi economiche dei processi decisionali sono impostate in genere in termini astratti di razionalità delle scelte; le analisi psicologiche invece cercano di capire le caratteristiche concrete dei processi cognitivi, e le motivazioni di scelte che appaiono non conformi ad astratti criteri di razionalità. Le prime analisi economiche dei processi decisionali in condizioni d'incertezza mettevano in evidenza il principio della massimizzazione del valore atteso, definito come la somma dei valori delle alternative possibili moltiplicati per le rispettive probabilità. Alcune debolezze delle analisi economiche dei processi decisionali furono messe in evidenza sin dal 1947 da Herbert Simon, psicologo, informatico ed economista americano, che fu tra i pionieri dell'intelligenza artificiale, e ottenne il premio Nobel per l'economia nel 1978. Simon introdusse la nozione di "razionalità limitata", e indicò tre fattori che impediscono processi decisionali astrattamente razionali: l'incompletezza delle conoscenze, la difficoltà di avere un quadro completo di tutte le possibili decisioni, la difficoltà di prevedere le conseguenze delle diverse decisioni.

Daniel Kahneman e Amos Tversky, entrambi di origine israeliana e docenti di psicologia in prestigiose università americane, pubblicarono, a partire dai primi anni settanta del secolo scorso, alcuni saggi che gettarono le fondamenta della moderna teoria economica e psicologica delle decisioni, soprattutto in condizioni d'incertezza. Grazie a queste ricerche Kahneman ottenne il premio Nobel per l'economia nel 2002; probabilmente anche Tversky lo avrebbe ottenuto, se non fosse morto prematuramente nel 1996. Kahneman e Tversky mostrarono che i giudizi delle persone sono il prodotto dell'azione di meccanismi cognitivi che, a volte in modo inconsapevole, influenzano molte decisioni. Secondo l'euristica della *disponibilità*, le persone, quando devono assegnare una probabilità a un evento futuro fanno riferimento principalmente alle esperienze passate, e in particolare a quelle rimaste più impresse nella memoria, perché associate a esperienze emotive più significative. Ciò significa che la probabilità di eventi che

impressionano maggiormente, per esempio gli incidenti aerei, è sovrastimata da gran parte delle persone.

Spesso le opzioni fra cui scegliere non sono tutte disponibili per il confronto simultaneamente, ma si presentano sequenzialmente, oppure debbono essere cercate appositamente; in queste situazioni spesso si osserva che l'aumento delle informazioni disponibili può complicare invece che agevolare il processo decisionale. Per esempio, se si presenta l'opportunità di acquistare un certo prodotto a condizioni particolarmente vantaggiose, molte persone decidono abbastanza rapidamente di acquistarlo, se invece la stessa opportunità è accompagnata da un'altra altrettanto vantaggiosa, molte persone, nella difficoltà di scegliere fra le due opzioni, spesso rinviando la decisione, magari cercando altre opportunità; la decisione può invece essere facilitata dalla presenza di una opzione per qualche aspetto analoga ma chiaramente meno vantaggiosa (per esempio l'offerta nello stesso negozio di due modelli di televisori della stessa marca, con prestazioni analoghe, in vendita a prezzi significativamente diversi, può stimolare artificialmente la vendita del televisore che costa meno. Un'altra anomalia è che in presenza di conflitto decisionale fra due opzioni, per esempio un televisore di migliore qualità e prezzo più alto e uno di minore qualità e prezzo più basso, l'aggiunta di una opzione con caratteristiche più estreme, orienta la decisione verso l'opzione intermedia ("di compromesso"). Un'altra anomalia decisionale riguarda il cosiddetto "effetto disgiunzione", per cui l'incertezza riguardo un evento che di per sé non influenza la sostanza della decisione, ma può influenzare la natura della motivazione della decisione, spesso ritarda la decisione stessa; dotare una opzione di un attributo positivo può rendere meno probabile la scelta di quella opzione da parte di persone per le quali quell'attributo non ha valore, anche nel caso in cui le persone non debbano pagare niente per esso.

Daniel McFadden, premio Nobel per l'economia nel 2000 insieme a James Heckman, per i suoi contributi allo sviluppo delle teorie e dei metodi per l'analisi delle scelte discrete, ricorda un proverbio olandese secondo cui "*chi ha una scelta ha un problema*", per affermare che i consumatori spesso si trovano a disagio di fronte alla necessità di scegliere, anche se ciò contrasta con il fatto che sia la logica sia l'esperienza mostrano che la possibilità di scegliere rappresenta un vantaggio. La contraddizione si può spiegare con il fatto che operare sui mercati reali è spesso complicato, per la variabilità delle caratteristiche dei prodotti, l'imperfezione delle informazioni, la volatilità dei prezzi, la scarsa affidabilità delle controparti, il rischio di sbagliare nelle decisioni, con conseguente rammarico dopo l'acquisto. Molte ricerche di psicologia e di economia cognitiva e sperimentale hanno mostrato che le persone spesso prendono delle decisioni fra loro contraddittorie, non riescono a imparare dall'esperienza, mostrano riluttanza ad effettuare scambi, valutano il loro livello di felicità raffrontando la loro situazione con quella di altri. Diverse ricerche nell'ambito della psicologia cognitiva hanno inoltre mostrato che la memoria è imperfetta e che le percezioni sono spesso caratterizzate da errori sistematici o distorsioni (*bias cognitivi*); si consideri, per esempio, la relazione fra memoria reale e memoria emozionale, i ricordi orientano la percezione delle alternative, le preferenze, e quindi le decisioni; distorsioni sistematiche nel ricordare fatti e sensazioni possono quindi provocare errori di prospettiva e decisioni contraddittorie e successivo rammarico (*regret*).

Ciò che le persone conservano e poi recuperano dalla memoria è influenzato dagli umori e dalle emozioni, per esempio, gli spot pubblicitari rimangono più forti in memoria, e quindi influenzano maggiormente le scelte quando il loro contenuto emozionale corrisponde meglio agli umori del consumatore. Le sensazioni emotive di gioia e dolore rimangono in memoria più delle altre, e possono maggiormente condizionare le decisioni future (una volta che si è rimasti "scottati" da un tipo di decisioni si tende a essere più cauti), ma la loro intensità tende a diminuire col tempo.

Diversi esperimenti di psicologi ed economisti hanno mostrato che spesso le decisioni dei consumatori sono influenzate dal contesto decisionale, dalle modalità di presentazione delle informazioni e dai riferimenti (*àncore*) offerti per la decisione, soprattutto nel caso di beni

complessi da valutare. Per esempio, un dato di riferimento che può influenzare la decisione di acquistare un'automobile, un televisore, un telefono cellulare, o altri prodotti simili in un supermercato può essere il prezzo di vendita al dettaglio suggerito dal produttore. Quando un prodotto è presentato in modo da far apparire come un affare il suo acquisto, spesso molti consumatori si affrettano a comprarlo. Un secondo tipo di anomalie decisionali sistematiche derivano dal cosiddetto “*effetto dotazione*”, per cui la disponibilità a pagare per avere una quantità aggiuntiva di un prodotto che già si possiede è in genere significativamente minore del prezzo richiesto per rinunciare a una parte del prodotto posseduto. Ciò evidenzia probabilmente un fenomeno psicologico per cui le persone spesso si abituano alla situazione in cui si trovano e considerano con sospetto i cambiamenti, oppure una certa avversione agli scambi in assenza di evidenti significativi vantaggi, per il timore di trovarsi peggio di prima.

In sintesi, le principali cause di distorsioni decisionali sistematiche in rapporto al processo di valutazione delle diverse opzioni decisionali sono: l'influenza che l'ambiente e le modalità di presentazione delle informazioni hanno sul processo decisionale, e in particolare l'abitudine di ancorare le decisioni a dei riferimenti che possono essere inseriti ad arte nella formulazione del contesto decisionale per indirizzare la decisione; il fatto che di solito l'attenzione dei consumatori è attratta da aspetti caratteristici particolari, invece di essere rivolta alla globalità delle informazioni rilevanti; le asimmetrie di valutazione fra guadagni e perdite, nel senso che si assegna maggiore importanza all'esigenza di evitare eventuali perdite che alle opportunità di guadagno; la tendenza a pensare che l'associazione casuale fra due eventi (coincidenze) sia in realtà di natura causale.

I risultati di diverse ricerche hanno mostrato che un numero eccessivamente elevato di attributi, indipendentemente dal numero di opzioni, può determinare un “sovraccarico informativo” (*information overload*) che complica il processo decisionale.

I problemi dell'eccesso di informazioni, in termini sia di attributi sia di opzioni, hanno cominciato ad apparire sempre più rilevanti a partire da primi anni duemila soprattutto in connessione con lo sviluppo degli acquisti su Internet; mentre nei negozi tradizionali esistono dei vincoli fisici che limitano il numero delle varietà di ciascun bene che possono essere mostrate ai potenziali acquirenti, nei siti di vendite online non esistono praticamente limiti, e possono essere rese disponibili migliaia di varietà per ciascun bene. Ma un'ampia possibilità di scelta appare come un'arma a doppio taglio; da un lato le persone amano avere ampie possibilità di scelta, d'altro lato, però, si sentono in imbarazzo di fronte alla scelta fra un numero molto elevato di opzioni disponibili.

Ancora più difficile risulta il processo decisionale in presenza di correlazione negativa fra gli attributi; vale a dire quando le opzioni che appaiono buone per alcuni attributi, appaiono non buone dal punto di vista di altri attributi; per esempio, nei telefoni cellulari c'è in genere una correlazione negativa fra la leggerezza e l'autonomia (a una maggiore autonomia corrisponde in genere una minore leggerezza). In questi casi si ha un ambiente decisionale “ostile” (*unfriendly*), poiché l'acquirente non riesce a individuare una scelta che domina le altre per tutti gli aspetti, ma è costretto a sacrificare alcune esigenze a favore di altre. Le correlazioni negative comportano difficili problemi decisionali, e una diminuzione della soddisfazione e della sicurezza di aver compiuto la scelta giusta. Un'altra importante situazione ricorrente di correlazioni negative fra attributi, che richiede il ricorso a operazioni di compensazione (*trade-off*), è quella in cui si deve scegliere tra un'elevata qualità associata ad una scarsa economicità o viceversa.

Un'altra caratteristica soggettiva rilevante per i processi decisionali è l'importanza assegnata dal consumatore ai diversi attributi dei prodotti; spesso l'importanza dei diversi attributi ha un andamento a gradini, con alcuni attributi considerati molto importanti e gli altri sempre meno importanti.

Secondo i risultati di alcune ricerche, un consumatore può effettuare delle buone scelte anche considerando pochi attributi, perché in alcune circostanze con delle semplici euristiche di scelta basate su poche informazioni si possono ottenere benefici significativi a costi

relativamente bassi. Le euristiche decisionali (“*scorciatoie cognitive*”), come, per esempio, accontentarsi di scelte soddisfacenti invece che perseguire quelle ottimali, oppure la eliminazione progressiva delle opzioni per aspetti, potrebbero consentire al consumatore di effettuare rapidamente delle scelte sulla base di alcuni attributi importanti, ma allo stesso tempo potrebbero comportare violazioni delle regole delle decisioni razionali e comportare delle scelte che non riflettono pienamente le preferenze del consumatore. In particolare ciò è probabile che si verifichi in presenza di attributi inversamente correlati.

Nell’era di Internet e dell’accesso facile a quantità elevatissime di informazioni, il problema del sovraccarico di informazioni per i consumatori sta diventando sempre più rilevante. Una possibile reazione dei consumatori al sovraccarico informativo può essere il differimento della decisione di acquisto, per il timore di non riuscire ad assumere la decisione migliore.

Attualmente la maggior parte delle ricerche nel campo delle decisioni privilegiano gli ambienti online e le grandi opportunità offerte da Internet. Si pensi, per esempio, allo sviluppo di nuove metodologie di filtraggio delle informazioni, in grado di selezionare automaticamente le opzioni per ciascuno più interessanti, su cui si basano gli algoritmi di nuovi sistemi di raccomandazione o agenti intelligenti. Le principali metodologie di riferimento sono due: metodologie “sociali” o “collaborative”( che si basano sulle opinioni di consumatori o esperti del settore) e metodologie cognitive (*content based*, perchè si basano sulle preferenze dei consumatori e sulle performance degli attributi); negli ultimi anni sono state avanzate anche delle proposte metodologiche di sintesi di questi due approcci.

Secondo Häubl e Trifts, le matrici di comparazione e gli agenti *content-based filtering* consentono ai consumatori di prendere decisioni migliori con un minore sforzo cognitivo, allentando il *trade-off* tra grado di accuratezza delle decisioni e sforzo sostenuto[Häubl, Trifts 2000]. Alcuni agenti *content-based*, in caso di opzioni con attributi sia positivi che negativi per il consumatore, a volte selezionano opzioni non soddisfacenti, poiché tendono a compensare le performance negative per alcuni aspetti con performance positive da altri punti di vista.

Principale obiettivo di questa tesi è quello di evidenziare la validità di una diversa metodologia di filtraggio, che riprende la metodologia adottata da diversi agenti *content based filtering*, ma attenua i problemi di compensazione tra attributi positivi e negativi. Il prodotto “complesso” preso in esame per analizzare i processi decisionali è il telefono cellulare, caratterizzato da molteplici attributi/funzioni. Durante la ricerca sono state svolte delle analisi approfondite del mercato italiano dei telefoni cellulari, in particolar modo di quello online.

Rimandando, come detto all’inizio, alla lettura delle Conclusioni per la descrizione dettagliata degli argomenti trattati, si fa presente in sintesi che nel primo capitolo vengono esaminate le principali differenze fra analisi economiche e analisi psicologiche delle decisioni. La rassegna critica delle analisi economiche e psicologiche delle decisioni ha messo in evidenza in particolare le critiche degli psicologici alle analisi economiche dei processi decisionali. La rilevanza di queste critiche è stata evidenziata dall’attribuzione del premio Nobel per l’economia ad alcuni studiosi di formazione psicologica, come Herbert Simon e Daniel Kahneman. In particolare queste critiche hanno messo in evidenza diverse anomalie cognitive che possono influenzare profondamente i processi decisionali, non soltanto dei consumatori, ma anche dei manager; in questo capitolo sono anche evidenziate delle proposte di utilizzazione delle scoperte psicologiche sulle anomalie cognitive nella definizione delle strategie politiche. Il secondo capitolo della tesi è dedicato a una rassegna critica di alcuni sviluppi dell’intelligenza artificiale particolarmente rilevanti per le ricerche presentate negli articoli successivi. In particolare, l’attenzione si è concentrata su alcune metodologie di apprendimento automatico degli agenti intelligenti. Nel terzo capitolo vengono analizzate le metodologie di filtraggio delle informazioni adottate dagli agenti di raccomandazione presenti online. Particolare attenzione è dedicata ai sistemi di supporto decisionale compensatori e non compensatori, e ai sistemi di comparazione online. Il quarto capitolo è dedicato a delle analisi sul commercio elettronico e la reputazione online. Il quinto capitolo presenta i risultati di analisi empiriche e indagini esplorative sulle

decisioni di acquisto di telefoni cellulari; in queste indagini sono state applicate alcune metodologie di apprendimento automatico sviluppate nell'ambito delle applicazioni di intelligenza artificiale. In particolare, sono state utilizzate metodologie di regressione multivariata, di *stepwise regression*, di *decision tree analysis* e di *cluster analysis*, per cercare di comprendere la logica delle decisioni di acquisto dei prodotti multi-attributo, e in particolare dei telefoni cellulari. Nel sesto ed ultimo capitolo viene presentato un sistema originale di supporto alle decisioni di acquisto, basato sul principio della minimizzazione degli scarti negativi fra i livelli effettivi di performance delle diverse caratteristiche di un prodotto e i livelli minimi di performance richiesti. Il sistema è stato sottoposto a verifica sia online, sia intervistando i clienti di un punto di vendita tradizionale.

# Capitolo 1

## Analisi economiche e psicologiche delle decisioni

### 1.1 Introduzione

La decisione è la scelta di intraprendere un'azione, tra più alternative considerate (*opzioni*), da parte di un individuo o di un gruppo (*decisore*). Perché si possa parlare propriamente di decisione è necessario che il decisore abbia di fronte a sé una pluralità di opzioni che condizionino la “scelta obbligata”. Il processo decisionale è oggetto di studio di una pluralità di discipline: filosofia, logica, matematica, statistica, psicologia, sociologia, economia, ecc. Alcune di queste adottano un approccio *descrittivo* ed alcune un approccio *normativo*. La psicologia adotta ad esempio un approccio descrittivo, cercando di scoprire come effettivamente vengono prese le decisioni nei diversi contesti; invece, l'economia adotta un approccio normativo cercando di individuare il modo con cui le decisioni dovrebbero essere prese facendo riferimento ad ideali decisori razionali.

D'altra parte, per poter decidere in modo razionale il decisore deve conoscere le opzioni disponibili e le conseguenze che possono scaturire da ciascuna. Spesso, però, il decisore non dispone di informazioni complete, nel senso che ignora talune opzioni o non è in grado di prevedere tutte le conseguenze ad esse associate.

Questo capitolo espone le principali teorie economiche e psicologiche inerenti i processi decisionali, sottolineando come un legame tra i due diversi approcci consenta, per quanto possibile, di evitare al decisore, di compiere scelte sbagliate.

### 1.2 Gli economisti e la teoria dell'utilità attesa

Il fatto che ci siano persone razionali che decidono di giocare alla roulette o acquistano il biglietto di una lotteria dimostra che ciascuno è disposto a rischiare più di quanto consigli una valutazione razionale, basata sul calcolo delle probabilità. Un'azione o decisione in condizioni di incertezza può generare diversi esiti o eventi, ognuno con una certa probabilità di verificarsi. Se si investono 10.000 euro acquistando azioni di una certa società, dopo un anno il valore potrebbe diventare 12.000 con una probabilità del 50% oppure 9.000 con probabilità 50%.

Nel caso in cui gli individui non si preoccupano affatto del rischio connesso alla scelta di un prospetto incerto, le loro decisioni sono guidate unicamente dal valore atteso, che rappresenta la media ponderata dei possibili valori monetari di un prospetto, i cui pesi sono dati dalle rispettive probabilità. Il prospetto selezionato sarà quello con il più alto valore atteso. Uno degli esempi più famosi di analisi di questo tipo nella storia del pensiero risale alla celeberrima “scommessa” di Blaise Pascal [Pascal 1976]<sup>1</sup>.

La maggior parte degli individui, invece, non è indifferente al rischio, quindi per spiegare le scelte in condizioni di incertezza gli economisti fanno uso della *teoria dell'utilità attesa*, secondo la quale tra diversi prospetti alternativi gli individui scelgono quello con *l'utilità attesa più alta*, piuttosto che quello con il *maggior valore atteso*. Il concetto di utilità attesa, fu introdotto da Nicolas Bernoulli nel 1738 per proporre una soluzione per il paradosso di San

---

<sup>1</sup> Secondo Pascal (1670, pensiero 233), non si può fare a meno di scegliere tra vivere come se Dio ci fosse o come se Dio non ci fosse; dovendo necessariamente scegliere, conviene “scommettere” sull'esistenza di Dio: “...Valutiamo questi due casi: se vincete, vincete tutto, se perdete non perdetevi nulla. Scommettete, dunque, che Dio esiste, senza esitare...”. I “Pensieri” sono una raccolta di riflessioni che riguardavano probabilmente l'idea di scrivere una apologia del Cristianesimo; Pascal, morto a 39 anni nel 1662, non riuscì a completarla, e i frammenti ritrovati furono pubblicati per la prima volta da suoi amici e parenti nel 1670.

Pietroburgo<sup>2</sup> ed un'analisi rigorosa della teoria dell'utilità attesa quale base teorica dei processi decisionali razionali è stata sviluppata da Von Neumann e Morgenstern [ Neumann e Morgenstern, 1944].

La teoria dell'utilità attesa consente di tener conto dell'atteggiamento degli individui rispetto al rischio e di trattare le scelte in un ambiente incerto in modo analogo alle scelte in condizioni di certezza. Secondo questo approccio, ad ogni possibile risultato  $x$  di un prospetto è assegnato un certo valore sulla base della funzione di utilità dell'individuo  $u(x)$ . L'utilità attesa è ottenuta come la media ponderata *delle utilità* associate ad ogni possibile esito, con il peso determinato dalle rispettive probabilità.

Prendiamo in considerazione un prospetto  $X$  che consente di ottenere un pagamento di 100 euro con una probabilità pari al 60 per cento e un pagamento di 20 euro con una probabilità del 40 per cento. Il valore atteso,  $E(X)$  o  $\bar{x}$ , di questo prospetto è pari a:

$$E(X) = \bar{x} = 0,6 \cdot 100 + 0,4 \cdot 20 = 68$$

Per calcolare l'utilità attesa è necessario usare la funzione di utilità dell'individuo, che supponiamo assuma la seguente forma:  $u(x) = \sqrt{x}$  e trasformare i valori monetari relativi ai vari esiti negli indici di utilità. L'utilità attesa della lotteria  $X$ ,  $UA(X)$ , è così uguale a:

$$UA(X) = 0,6 \cdot \sqrt{100} + 0,4 \cdot \sqrt{20} = 7,79$$

Secondo la funzione di utilità  $u(x)$ , l'utilità che l'individuo percepirebbe se ottenesse, invece del prospetto incerto, un pagamento certo pari al valore atteso  $\bar{x}$  sarebbe invece data da:

$$u(\bar{x}) = \sqrt{68} = 8,25$$

Un aspetto fondamentale delle scelte in condizioni di incertezza è l'atteggiamento degli individui rispetto al rischio, che può essere di avversione, neutralità o propensione.

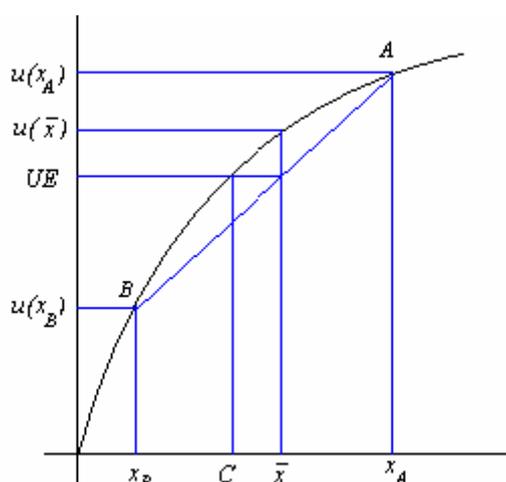
Un individuo avverso al rischio preferisce ricevere un reddito certo  $x$  piuttosto che disporre di un prospetto incerto il cui valore atteso corrisponde a  $x$ . Nell'esempio esaminato, si è visto che l'utilità del valore certo pari a  $x$  è maggiore dell'utilità attesa dal prospetto. Allo stesso modo, tra un pagamento certo di 1.000 euro e una lotteria che dà 2.000 euro con probabilità 0,5 oppure 0 con probabilità 0,5 (che ha quindi un valore atteso di 1.000), un individuo avverso al rischio preferisce strettamente il reddito certo.

I soggetti avversi al rischio presentano una funzione di utilità concava, cioè una funzione che aumenta a tassi decrescenti all'aumentare dei payoff (come la funzione  $u(x) = \sqrt{x}$  usata nell'esempio precedente), rappresentata nella Figura 1.1. Nella Figura  $x_A$  e  $x_B$  sono i due possibili risultati; l'utilità attesa del prospetto è data dal punto medio sulla corda  $BA$ , che corrisponde sull'asse delle ascisse al valore atteso  $x$ . Dal grafico si evince anche che l'utilità del valore  $x$  è maggiore dell'utilità attesa del prospetto.

---

<sup>2</sup> Il paradosso di San Pietroburgo, proposto nel 1713 da Nicolas Bernoulli, cugino di Daniel Bernoulli, riguarda il valore atteso di un ipotetico gioco d'azzardo presso il casinò di San Pietroburgo consistente nel lanciare ripetutamente una moneta un infinito numero di volte, raddoppiando a ogni lancio la posta in gioco, e vincendo un valore pari alla posta quando esce croce. Il paradosso consiste nel fatto che, benché il valore atteso di questo gioco sia infinito, nessuna persona ragionevole è disposta a pagare una somma molto grande per potervi partecipare. La soluzione proposta da Daniel Bernoulli si basa sull'idea che il criterio decisionale razionale consiste nel massimizzare non il valore atteso della ricchezza ma la sua utilità attesa, e che l'utilità marginale della ricchezza è decrescente; in particolare, Bernoulli ipotizzò una funzione di utilità totale uguale al logaritmo della ricchezza, ottenendo una utilità attesa del gioco pari a 2. Nel 1728 una soluzione analoga era stata proposta da Gabriel Cramer (matematico svizzero inventore della regola di Cramer per la soluzione dei sistemi di equazioni lineari) che, ipotizzando una funzione di utilità totale uguale alla radice quadrata della ricchezza, ottenne un valore pari a circa 3 per l'utilità attesa del gioco.

Figura 1.1 - La funzione di utilità per un individuo avverso al rischio



### 1.3 Gli psicologi e la razionalità limitata e le anomalie cognitive

Alcune debolezze dell'analisi economica (razionale) dei processi decisionali furono messe in evidenza magistralmente da Herbert Simon nel quinto capitolo del suo libro sul comportamento amministrativo dedicato alla psicologia delle decisioni amministrative [Simon H. A., 1947].

In particolare Simon mise in evidenza tre fattori che impediscono processi decisionali astrattamente razionali: l'incompletezza delle conoscenze, la difficoltà di avere un quadro completo di tutte le possibili decisioni e la difficoltà di prevedere le conseguenze delle diverse decisioni. Secondo Simon, mentre nei modelli di scelta razionale si suppone di poter valutare tutte le opzioni prima di assumere una decisione, in realtà spesso le diverse opzioni si presentano sequenzialmente, e si decide quando si ritiene che una delle opzioni disponibili sia soddisfacente, interrompendo il processo di acquisizione di nuove informazioni; un caso concreto menzionato da Simon è quello della vendita di una casa, in cui diversi possibili prezzi di vendita vengono conosciuti in sequenza, e l'attesa di eventuali opportunità migliori può comportare la perdita di alcune opportunità. Simon manifestò quindi seri dubbi sull'utilità concreta delle teorie economiche dei comportamenti razionali, esprimendo invece l'opinione che per capire i processi decisionali effettivi siano più utili le teorie psicologiche della percezione e della cognizione.

Critiche più radicali di quelle di Simon alla teoria economica delle decisioni basata sul principio della massimizzazione dell'utilità attesa, sono state avanzate a partire dagli anni settanta del secolo scorso da altri due studiosi di psicologia, Daniel Kahneman e Amos Tversky, entrambi di origine israeliana e docenti di psicologia in prestigiose università americane. Kahneman ottenne il premio Nobel per l'economia nel 2002 insieme a Vernon Smith, il fondatore dell'economia sperimentale (anche Tversky l'avrebbe probabilmente ricevuto se non fosse morto prematuramente nel 1996)<sup>3</sup>.

Shafir, Simonson e Tversky, dopo aver sottolineato la rilevanza concreta dei conflitti decisionali, proposero la distinzione fra criteri decisionali formalizzati che cercano di attribuire dei valori alle diverse opzioni alternative, e criteri decisionali fondati su comparazioni ragionate informali, concentrando poi la loro attenzione sulle anomalie decisionali osservate in concreto quando si utilizzano dei criteri decisionali informali [Shafir, Simonson, Tversky, 1993].

Un'anomalia che spesso si verifica è l'inversione della decisione a seconda che il problema sia posto in termini positivi (quale opzione scegliere) o negativi (quale opzione

<sup>3</sup> Per delle approfondite analisi dei collegamenti fra l'Economia cognitiva di Kahneman e Tversky e l'Economia sperimentale di Vernon Smith si veda, per esempio, Motterlini e Guala (2005), e Motterlini e Palmarini (2005).

rigettare); ciò accade perché nel primo caso l'attenzione si concentra principalmente sugli aspetti positivi delle opzioni, e nel secondo caso su quelli negativi; se una delle due opzioni presenta alcuni aspetti fortemente positivi e altri marcatamente negativi, mentre nella seconda le diverse caratteristiche si presentano con valori medi, è probabile che se il problema è posto in termini positivi sia scelta la prima opzione, e che questa sia però rigettata nel caso si chieda di individuare l'opzione da non scegliere.

Solitamente le opzioni fra cui scegliere non sono tutte disponibili per il confronto simultaneo, ma si presentano sequenzialmente, oppure debbono essere cercate appositamente; in queste situazioni spesso si osserva che l'aumento delle informazioni disponibili può complicare invece che agevolare il processo decisionale. Per esempio, se si presenta l'opportunità di acquistare un certo prodotto a condizioni particolarmente vantaggiose, molte persone decidono abbastanza rapidamente di acquistarlo, se invece la stessa opportunità è accompagnata da un'altra, altrettanto vantaggiosa, molte persone, nella difficoltà di scegliere fra le due occasioni, spesso rinviando la decisione, magari cercando altre opportunità; la decisione può invece essere facilitata dalla presenza di una opzione per qualche aspetto analoga ma chiaramente meno vantaggiosa (per esempio l'offerta nello stesso negozio di due modelli di televisori della stessa marca, con prestazioni analoghe, in vendita a prezzi significativamente diversi può stimolare artificialmente la vendita del televisore che costa meno [Simonson e Tversky, 1992]). Un'altra anomalia individuata da Simonson è che in presenza di conflitto decisionale fra due opzioni, per esempio un televisore di migliore qualità e prezzo più alto e uno di minore qualità e prezzo più basso, l'aggiunta di una opzione con caratteristiche più estreme orienta la decisione verso l'opzione intermedia ("di compromesso"). Ultima anomalia decisionale messa in evidenza da Tversky e Shafir riguarda il cosiddetto "*effetto disgiunzione*", per cui l'incertezza riguardo un evento che di per sé non influenza la sostanza della decisione, ma può influenzare la natura della motivazione della decisione, spesso ritarda la decisione stessa [Tversky, Shafir, 1992]. Secondo Simonson, Carmon e O'Curry dotare una opzione di un attributo positivo, può rendere meno probabile la scelta di quella opzione da parte di persone per le quali quella opzione non ha valore, anche nel caso in cui le persone non debbano pagare niente per quell'attributo [Simonson, Carmon, O'Curry, 1994].

Kahneman e Tversky pubblicarono, a partire dai primi anni settanta del secolo scorso, alcuni saggi che gettarono le fondamenta della moderna teoria economica e psicologica delle decisioni, soprattutto in condizioni d'incertezza. Essi mostrarono che i giudizi delle persone sono il prodotto dell'azione di meccanismi cognitivi che, a volte in modo inconsapevole, influenzano molte decisioni. Secondo l'euristica della *disponibilità*, le persone, quando devono assegnare una probabilità a un evento futuro fanno riferimento principalmente all'esperienze passate, e in particolare a quelle rimaste più impresse nella memoria, perché associate a esperienze emotive più significative. Ciò significa che la probabilità di eventi che impressionano maggiormente, per esempio gli incidenti aerei, è sovrastimata da gran parte delle persone.

Le ricerche di Kahneman e Tversky misero in evidenza diversi esempi di apparenti anomalie e contraddizioni nelle decisioni di molte persone in condizioni d'incertezza, che li portarono a mettere in discussione la rilevanza empirica del modello dell'utilità attesa di von Neumann-Morgenstern [Neumann, Morgenstern, 1944]. Kahneman e Tversky misero in evidenza che spesso le persone decidono in modo diverso, a seconda di come vengono presentate le opzioni. Per esempio, molte persone sono disposte ad attraversare tutta la città per risparmiare 5 euro su un prodotto che costa 15 euro, ma non per un prodotto che costa 125 euro. Nelle scelte finanziarie, inoltre, spesso le persone mostrano amore per il rischio quando c'è la possibilità di evitare una perdita sicura, e avversione per il rischio nella generalità dei casi. Quindi, in situazioni di incertezza le decisioni delle persone sembrano spesso essere determinate da valutazioni soggettive delle probabilità, che possono essere anche molto diverse dalle probabilità oggettive.

Uno degli articoli più importanti di Kahneman e Tversky fu pubblicato nel 1979 su *Econometrica*; in quell'articolo fu presentata la “teoria del prospetto” (*prospect theory*) che rappresenta un dei punti di riferimento più importanti della moderna teoria delle decisioni in condizioni d'incertezza [Kahneman e Tversky, 1979].

### 1.3.1 La teoria della razionalità limitata

L'analisi dei limiti che in concreto le persone incontrano nei loro processi decisionali fu in gran parte approfondita da Simon. Egli cercò di individuare delle procedure decisionali soddisfacenti mediante dei meccanismi dinamici in grado di adattare i livelli di aspirazione alla realtà, sulla base di informazioni sull'ambiente decisionale. Simon chiarì che gli aggettivi “soddisfacente” e “ottimizzante” sono delle etichette per due approcci generali all'analisi del comportamento razionale in situazioni in cui la complessità e l'incertezza rendono impossibile la razionalità globale.

Gli ottimizzanti o “massimizzatori” decidono soltanto dopo avere esaminato tutte le alternative disponibili, i soddisfacentisti, invece, valutano un numero limitato di alternative, e ne scelgono una “sufficientemente buona”. I criteri di ottimizzazione sono stati sostituiti da criteri di prestazione soddisfacente, utilizzando metodi “euristici”, o per tentativi ed errori, per aiutare la ricerca di alternative plausibili, anche grazie alla disponibilità di strumenti computazionali sempre più potenti [Simon, 1972].

La profonda analisi critica della razionalità astratta propria delle teorie decisionali della scienza economica ortodossa valse ad Herbert Simon il premio Nobel per l'Economia nel 1978, nonostante la sua attività didattica universitaria fosse prevalentemente concentrata nel campo della psicologia e dell'informatica. Nella conferenza tenuta l'8 dicembre 1978 in occasione dell'assegnazione del premio Nobel, pubblicata nel 1979 sull'*American Economic Review*, Simon introdusse la nozione di “razionalità limitata” per sintetizzare le procedure utilizzate per rendere concretamente trattabili i problemi decisionali complessi: perseguire scelte soddisfacenti invece che ottimali, sostituire obiettivi globali e astratti con obiettivi tangibili di cui si possa osservare e misurare la realizzazione, suddividere i compiti decisionali complessi fra più specialisti, coordinandone il lavoro per mezzo di una struttura di comunicazioni e di rapporti di autorità. Nella *Nobel lecture* del 1978 Simon chiarì che la razionalità dei processi decisionali è limitata quando manca la conoscenza di tutte le alternative, c'è incertezza su eventi esogeni rilevanti, non si è in grado di prevedere tutte le conseguenze delle possibili decisioni. Aspetti caratteristici dei processi decisionali con razionalità limitata sono la ricerca di informazioni, la cui analisi fu approfondita da George Stigler [Stigler, 1961], premio Nobel per l'economia nel 1982, e la nozione di decisioni soddisfacenti (*satisficing*) con riferimento a un qualche livello di “aspirazione”, piuttosto che astrattamente ottimali<sup>4</sup>. L'analisi dei processi decisionali soddisfacenti risale alle ricerche psicologiche sui livelli di aspirazione di Lewin e altri, che mostrarono che i livelli di aspirazione non sono statici, ma tendono a variare con l'esperienza; in condizioni ambientali favorevoli essi aumentano, mentre diminuiscono in situazioni ambientali negative. Le analisi dei processi decisionali che tengono conto delle attività di ricerca delle informazioni e di obiettivi decisionali soddisfacenti con riferimento a livelli di aspirazione che si adattano dinamicamente alle condizioni ambientali mostrarono la possibilità di arrivare a decisioni ragionevoli sulla base di informazioni e di elaborazioni incomplete, evitando il rischio di una paralisi decisionale nel miraggio di decisioni astrattamente ottimali. Simon mise in evidenza che già intorno alla metà degli anni cinquanta era stata proposta una teoria della razionalità limitata, in alternativa alla classica razionalità onnisciente, ed erano state effettuate delle ricerche empiriche che mostravano la corrispondenza dei processi decisionali concreti alle assunzioni della razionalità limitata piuttosto che a quelle della razionalità perfetta [Simon, 1985].

---

<sup>4</sup> Per discussioni sulla nozione generale di “razionalità limitata” si veda anche Simon (1983, traduzione italiana, La ragione nelle vicende umane, 1984), e March (1994, edizione italiana, Prendere decisioni, 1998, capitoli 1).

La relazione fra processi decisionali soddisfacenti e ottimizzanti fu ulteriormente chiarita da Simon nel 1981: “*Se una persona non può fare una cosa non la farà, indipendentemente da quanto la desideri. L’economia normativa ha mostrato che soluzioni esatte ai problemi più complessi di ottimizzazione non sono né raggiungibili né individuabili. .... Così la microeconomia normativa, evidenziando come sia impossibile ottimizzare le situazioni reali, dimostra che l’uomo economico è in realtà una persona che «soddisfa», che cioè accetta alternative «abbastanza buone», non perché preferisca il meno al più, ma perché non ha scelta. .... L’economia è una teoria della razionalità umana che ha a che fare sia con la razionalità procedurale – e cioè con i modi in cui si prendono le decisioni – sia con la razionalità sostanziale – e cioè con i contenuti di queste decisioni. Questo interesse per le procedure conduce, a livello individuale, all’esame delle tecniche per «soddisfare» e queste, a loro volta, conducono alla psicologia dei processi cognitivi. .... La ricerca operativa e l’intelligenza artificiale hanno introdotto nuove tecniche che aumentano la razionalità procedurale degli attori economici e li aiutano a prendere decisioni migliori*” [Simon, 1981].

### **1.3.2 Le anomalie cognitive nei processi decisionali**

Daniel McFadden premio Nobel per l’economia nel 2000 insieme a James Heckman, per i suoi contributi allo sviluppo delle teorie e dei metodi per l’analisi delle scelte discrete, ha individuato ben 25 “anomalie” cognitive, e quindi decisionali, messe in evidenza principalmente dalle ricerche di Kahneman e Tversky [McFadden, 2005]. McFadden classifica queste anomalie decisionali in quattro classi concernenti la formazione delle percezioni e la elaborazione delle informazioni (contesto, punto di riferimento, disponibilità e superstizione), e due classi riguardanti il processo di strutturazione del compito cognitivo (processo e proiezione). Gli effetti di contesto riguardano il fatto che il processo di elaborazione delle informazioni, e quindi il processo decisionale, sembra essere spesso influenzato dalle modalità in cui esse sono presentate; in particolare, Kahneman e Tversky hanno mostrato sperimentalmente che molte persone decidono diversamente fra due alternative identiche a seconda che esse siano presentate in termini di guadagni o di perdite rispetto alla situazione esistente [Kahneman, Tversky, 1984].

Innanzitutto McFadden, ricorda un proverbio olandese secondo cui “*chi ha una scelta ha un problema*”, per affermare che i consumatori spesso si trovano a disagio di fronte alla necessità di scegliere, anche se ciò contrasta con il fatto che sia il calcolo razionale sia l’esperienza mostrano che la possibilità di scegliere rappresenta un vantaggio. La contraddizione si può spiegare con il fatto che operare sui mercati reali é spesso complicato, per la variabilità delle caratteristiche dei prodotti, l’imperfezione delle informazioni, la volatilità dei prezzi, la scarsa affidabilità delle controparti, il rischio di sbagliare nelle decisioni, con conseguente rammarico dopo l’acquisto.

McFadden afferma che molte ricerche di psicologia e di economia cognitiva e sperimentale hanno mostrato che le persone spesso prendono delle decisioni fra loro contraddittorie, non riescono a imparare dall’esperienza, mostrano riluttanza ad effettuare scambi, valutano il loro livello di felicità raffrontando la loro situazione a quella di altri. Diverse ricerche nell’ambito della psicologia cognitiva hanno inoltre mostrato che la memoria è imperfetta e che le percezioni sono spesso caratterizzate da errori sistematici o distorsioni (*bias cognitivi*); si consideri, per esempio, la relazione fra memoria reale e memoria emozionale, i ricordi orientano la percezione delle alternative, le preferenze, e quindi le decisioni; distorsioni sistematiche nel ricordare fatti e sensazioni possono quindi provocare errori di prospettiva e decisioni contraddittorie e successivo rammarico (*regret*).

Ciò che le persone conservano e poi recuperano dalla memoria è influenzato dagli umori e dalle emozioni, per esempio, gli spot pubblicitari rimangono più forti in memoria, e quindi influenzano maggiormente le scelte, quando il loro contenuto emozionale corrisponde meglio agli umori del consumatore. Le sensazioni emotive di gioia e dolore rimangono in memoria più delle altre, e possono maggiormente condizionare le decisioni future (una volta che si è rimasti

“scottati” da un tipo di decisioni si tende a essere più cauti), ma la loro intensità tende a diminuire col tempo.

Il recupero di informazioni dalla memoria è un fenomeno cognitivo alquanto complesso; ciò che si ricorda è influenzato dal contesto e dagli umori del momento, e questi dipendono da come sono presentate le alternative decisionali. La memoria selettiva porta a ricordare le cose che attraggono di più l'attenzione; le coincidenze, in particolare, rimangono maggiormente impresse nella memoria, e ciò influenza le valutazioni probabilistiche (probabilità soggettive; in conseguenza di ciò le persone tendono a immaginare relazioni causali inesistenti tra fenomeni associati soltanto per effetto di coincidenze casuali. Altri problemi derivano dalla difficoltà di avere una percezione oggettiva del tempo (durante le esperienze piacevoli il tempo ‘vola’, e il contrario accade per le esperienze dolorose; gli eventi passati appaiono spesso meno lontani di quanto in realtà siano, per effetto di una visione ‘telescopica’ del passato). Kahneman e Krueger, per esempio, hanno mostrato che episodi di gioia o di sofferenza sono ricordati in modo selettivo a seconda della loro intensità e distanza nel tempo; ciò può condurre a decisioni sulla base di ciò che si ricorda soggettivamente, invece che sulla base delle esperienze oggettive [Kahneman e Krueger, 2006]. Laibson ha cercato di spiegare il fenomeno *hyperbolic discounting*, per cui il tasso di sconto diminuisce gradualmente man mano che si considerano alternative sempre più lontane nel tempo (per esempio molte persone preferiscono 1000 euro oggi a 1005 euro fra un anno, ma 1005 euro fra 51 anni a 1000 euro fra 50 anni [Laibson et al., 2005]).

In sintesi, le principali “anomalie” cognitive, e quindi decisionali, legate al funzionamento della memoria umana sono, secondo McFadden, le seguenti: l'intensità delle emozioni diminuisce nella memoria con il passar del tempo, le informazioni più recenti e significative sono quelle più presenti in memoria in ciascun istante, la ricostruzione in memoria di informazioni passate è influenzata dal contesto e dagli stimoli attuali, le coincidenze sono memorizzate in modo più forte e persistente, la nozione soggettiva del tempo tende a comprimere nella memoria i periodi lontani nel tempo [McFadden, 2005].

Diversi esperimenti di psicologi e di economisti hanno mostrato che spesso le decisioni dei consumatori sono influenzate dal contesto decisionale e dalle modalità di presentazione delle informazioni e dai riferimenti (*àncore*) offerti per la decisione, soprattutto nel caso di beni complessi da valutare. Per esempio, un dato di riferimento che può influenzare la decisione di acquistare un'automobile, un televisore, un telefono cellulare, o altri prodotti simili, in un supermercato può essere il prezzo di vendita al dettaglio suggerito dal produttore. Kahneman e Tversky hanno scoperto che quando un prodotto è presentato in modo da far apparire come un affare il suo acquisto, spesso molti consumatori si affrettano a comprarlo<sup>5</sup> [Kahneman, Tversky, 1992]. Un secondo tipo di anomalie decisionali sistematiche derivano dal cosiddetto “*effetto dotazione*”, per cui la disponibilità a pagare per avere una quantità aggiuntiva di un prodotto che già si possiede è in genere significativamente minore del prezzo richiesto per rinunciare a una parte del prodotto posseduto, in una misura che difficilmente potrebbe essere spiegata dal principio dell'utilità marginale decrescente derivante dalla sostituibilità imperfetta fra i diversi beni. Ciò evidenzia probabilmente un fenomeno psicologico per cui le persone spesso si abitano alla situazione in cui si trovano e considerano con sospetto i cambiamenti, oppure una certa avversione agli scambi in assenza di evidenti significativi vantaggi, per il timore di trovarsi peggio di prima.

L'*Economist*, il settimanale economico più prestigioso del mondo, ha dedicato alcuni anni fa la sezione “*Economic focus*” (30 agosto 2003, p. 56) all'analisi dell'effetto dotazione (*endowment effect*). In termini semplici questa distorsione fa sì che le persone in genere attribuiscono un valore aggiuntivo ai beni che già posseggono, per il semplice fatto di possederli.

---

<sup>5</sup> McFadden (2006, p. ) scrive che quando suggerì a un suo amico commerciante di utilizzare questa scoperta per incrementare vendite e profitti, l'amico gli rispose che questa tecnica gli era stata già insegnata quando frequentava la scuola professionale.

Questa distorsione è stata confermata ormai da moltissimi esperimenti, soprattutto nelle università americane. Per esempio, nella *Cornel University* furono distribuiti a caso fra gli studenti due regali aventi uguale valore di mercato: una tazza da caffè oppure una barra di cioccolato; successivamente fu data agli studenti la possibilità di scambiarsi reciprocamente i regali. In assenza di bias, i ricercatori, che avevano precedentemente verificato che metà degli studenti preferiva la tazza da caffè e l'altra metà la barra di cioccolato, si aspettavano che circa il 50 per cento degli studenti si sarebbero scambiati vicendevolmente i regali (la probabilità di ricevere casualmente un regalo diverso da quello preferito era appunto del 50 per cento); in realtà appena il 10 per cento degli studenti decisero di scambiarsi il regalo. L'interpretazione più plausibile di questo risultato è che il semplice possesso di qualcosa modifica in breve tempo le preferenze a favore del bene posseduto, anche se esso è stato ottenuto non per scelta ma per caso.

In sintesi, le principali cause di distorsioni decisionali sistematiche evidenziate da McFadden in rapporto al processo di valutazione delle diverse opzioni decisionali sono l'influenza che l'ambiente decisionale e le modalità di presentazione delle informazioni hanno sul processo decisionale, e in particolare l'abitudine di ancorare le decisioni a dei riferimenti che possono essere inseriti ad arte nella formulazione del contesto decisionale per indirizzare la decisione, il fatto che di solito l'attenzione dei consumatori è attratta da aspetti caratteristici particolari, invece di essere rivolta alla globalità delle informazioni rilevanti, le asimmetrie di valutazione fra guadagni e perdite, nel senso che si assegna maggiore importanza alla esigenza di evitare eventuali perdite, piuttosto che alle opportunità di guadagno, la tendenza a pensare che l'associazione casuale fra due eventi (coincidenze) sia in realtà di natura causale [McFadden, 2005].

La teoria economica delle decisioni fa quindi riferimento a consumatori ideali che hanno una grande capacità di elaborazione per poter valutare opzioni alternative anche numerose e complesse, ed elevata capacità di prevedere e analizzare le conseguenze delle diverse decisioni sul proprio livello di felicità. Già Simon aveva tuttavia evidenziato che in pratica le capacità computazionali e logiche di gran parte dei consumatori sono in realtà relativamente piccole in rapporto a quelle necessarie per poter individuare le decisioni ottimali e molte volte i consumatori non sono neppure consapevoli di molte potenziali opportunità decidendo inconsapevolmente in modo automatico e abitudinario [Simon, 1947].

Spesso le decisioni dei consumatori sono influenzate dall'ambiente sociale, sia per quel che riguarda il processo di acquisizione di informazioni e consigli<sup>6</sup>, sia per la tendenza delle persone a cercare l'approvazione degli altri alle proprie decisioni; ciò limita in qualche misura l'influenza sui processi decisionali delle motivazioni esclusivamente egoistiche. Le implicazioni delle modalità di trasmissione delle informazioni sono state sempre considerate con grande attenzione nelle attività di marketing, e in particolare nelle fasi di lancio e penetrazione dei nuovi prodotti, per il cui successo è essenziale riuscire a stimolare un flusso interpersonale di opinioni favorevoli. Oltre a fornire informazioni, le relazioni sociali influenzano le decisioni dei consumatori attraverso le norme generali di comportamento largamente condivise.

Gli effetti di reciprocità rappresentano una forma di interazione sociale largamente diffusa; essa è semplice nei casi di interazioni simultanee, come negli scambi bilaterali, mentre è più sofisticata nel caso di interazioni distanti nel tempo, per le quali assumono importanza la reputazione e la fiducia. Un celebre esempio di comportamento che sembra contrastare con la razionalità economica ma che può essere spiegato in termini di ipotesi di interazioni differite e di norme sociali è quello del gioco dell'ultimatum in una sola mossa [Fehr e Schmidt, 1999]. Si suppone che vi siano due giocatori anonimi A e B, A propone una divisione a sua scelta di un premio di 1000 euro, se B accetta, la suddivisione dei 1000 euro viene effettuata, altrimenti

---

<sup>6</sup> I supporti alle decisioni di acquisto di tipo collaborativo o sociale fanno appunto riferimento alle valutazioni espresse da persone con preferenze simili, nel selezionare le opzioni da proporre a ciascuna persona.

nessuno dei due ottiene niente. Secondo un ragionamento strettamente razionale, B dovrebbe accettare qualsiasi somma positiva, sia pur piccola, decisa da B; sapendo questo A dovrebbe proporre una suddivisione che comporta un importo minimo per B. Diversi esperimenti hanno però mostrato che in giochi di questo tipo le offerte sono in genere comprese fra il 42 e il 50 per cento del premio, e che offerte inferiori al 20 per cento sono rifiutate metà delle volte. Questi risultati, inspiegabili in termini di comportamenti strettamente razionali, possono essere spiegati in termini di norme sociali di correttezza, che possono comportare che una persona accetti anche di perdere volontariamente del denaro per punire comportamenti eccessivamente egoistici.

#### **1.4 Fattori che impediscono il processo decisionale**

Misuraca, Fasolo e Cardaci, affermano che la difficoltà nelle decisioni di acquisto possono dipendere da fattori interni e da fattori esterni rispetto ai soggetti che devono prendere la decisione [Misuraca, Fasolo, Cardaci, 2007]. Secondo loro, i fattori esterni, oltre al numero di opzioni, sono il numero di attributi che caratterizzano le diverse opzioni, il loro livello di dettaglio, le interrelazioni fra gli attributi. I fattori interni sono costituiti invece dallo stile decisionale di ciascuna persona, e dagli indicatori dell'importanza che per ciascuna persona hanno i diversi attributi.

Per quel che riguarda i fattori esterni del processo decisionale, un numero eccessivamente elevato di opzioni può provocare difficoltà in quanto chi deve assumere la decisione può sentirsi in difficoltà per la necessità di dover considerare troppe possibilità, e potrebbe, paradossalmente, preferire un numero più limitato di possibilità di scelta [Iyengar, Lepper, 2000]. Già nei primi anni settanta del secolo scorso, Jacoby, Speller e Berning avevano messo in evidenza che a un numero minore di opzioni è a volte associata una migliore qualità delle scelte [Jacoby, Speller, Berning, 1974]; Schwarts ha messo in evidenza che quando il numero di opzioni è minore le persone appaiono in genere più soddisfatte e meno dubbiose delle scelte effettuate [Schwartz, 2004].

Un altro rilevante fattore esterno di difficoltà del processo decisionale può essere costituito da un numero eccessivamente elevato di attributi o caratteristiche delle diverse alternative di scelta. I risultati di diverse ricerche hanno mostrato che un numero eccessivamente elevato di attributi o caratteristiche, indipendentemente dal numero di opzioni, può determinare un "sovraccarico informativo" (*information overload*) che complica il processo decisionale<sup>7</sup>. I risultati di un esperimento di Jacoby, Speller e Kohn hanno evidenziato che all'aumentare del numero degli attributi, diminuisce l'abilità di scegliere il prodotto migliore, ma sembra aumentare, tuttavia, il livello di soddisfazione e il grado di certezza di aver compiuto la scelta giusta, nonostante un aumento della confusione percepita; a un maggior numero di attributi corrisponde, quindi, da un lato l'effetto positivo di far sentire meglio gli acquirenti, ma anche una maggiore confusione e decisioni di acquisto spesso peggiori [Jacoby, Speller, Kohn, 1974]. I problemi dell'eccesso di informazioni, in termini sia di attributi sia di opzioni sono particolarmente rilevanti per gli acquisti su internet<sup>8</sup>.

Un fattore rilevante di "sovraccarico informativo" è costituito spesso dal livello di specificazione di ciascuna caratteristica. Soltanto in pochi casi la variabile descrittiva di una caratteristica è di natura "binaria" (0-1 = assente-presente); in gran parte dei casi le diverse caratteristiche sono rappresentate da variabili che possono assumere molti valori (variabili discrete), o addirittura infiniti valori (variabili continue). Alcuni esperimenti di Lurie (2004) hanno mostrato che, nel caso dell'acquisto di una calcolatrice, i processi decisionali sono più semplici ed efficaci quando le diverse caratteristiche sono rappresentate da variabili binarie

<sup>7</sup> Si veda, per esempio, Uacoby, Speller e Kohn (1974) e Malhotra (1982).

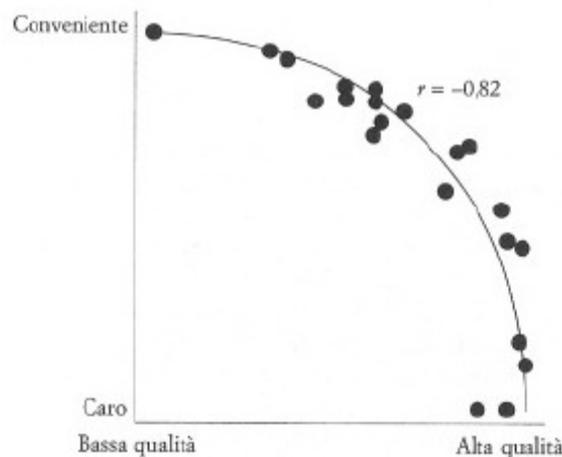
<sup>8</sup> Fasolo, McClelland, Todd e Misuraca (2007, p. 53) riportano che un sito web inglese di supporto agli acquisti di telefoni cellulari contiene informazioni su oltre 80 diverse caratteristiche per ciascun modello. Al contrario. Sia la rivista americana "Consumer reports", sia la rivista italiana "Altro consumo", concentrano in genere l'attenzione su poche caratteristiche.

invece che da variabili discrete con 4 possibili valori.

Un'altra rilevante difficoltà "esterna" dei processi decisionali può essere costituita dalla correlazione negativa fra diversi attributi; per esempio, nei telefoni cellulari c'è in genere una correlazione negativa fra la leggerezza e l'autonomia (a una maggiore autonomia corrisponde in genere una minore leggerezza). In questi casi si ha secondo McClelland un ambiente decisionale "ostile" (*unfriendly*), poiché l'acquirente non riesce a individuare una scelta che domina le altre per tutti gli aspetti, ma è costretto a sacrificare alcune esigenze a favore di altre [McClelland, 1978].

Le correlazioni negative comportano difficili problemi decisionali, e una diminuzione della soddisfazione e della sicurezza di aver compiuto la scelta giusta. Una situazione ricorrente di correlazioni negative fra attributi, che richiede il ricorso a operazioni di compensazione (*trade-off*), è quella in cui si deve scegliere tra un'elevata qualità associata ad una scarsa economicità o viceversa. L'abituale correlazione negativa qualità/convenienza è richiamata dal proverbio «*compri quello che paghi*» (*you get what you pay for*). Un esempio concreto che illustra l'alta correlazione negativa presente in scelte comuni è rappresentato nella Figura 1.1 dov'è riportata la correlazione tra qualità/convenienza di 25 *mountain bike* passate in rassegna recentemente dalla rivista americana «*Consumer Reports*». In questo caso, emerge una forte correlazione negativa (-0,82) fra convenienza e qualità delle *mountain bike*. Ciò costringe i consumatori ad operare una compensazione fra queste due dimensioni e a gestire il conflitto, un'operazione che essi preferiscono evitare il più possibile [Hogarth, 1987].

Figura 1.1 - Correlazione negativa fra convenienza e qualità di 25 *mountain bike*.



Fonte: Misuraca, Fasolo, Cardaci (2007), pag. 56.

Le correlazioni tra attributi possono cambiare da positive a negative nel corso del processo di scelta. Ciò avviene perché i consumatori spesso scelgono tra le diverse opzioni in due fasi distinte: in una prima fase, di «soltimento» (*winning*), i consumatori *eliminano* tutte le opzioni non desiderabili; nella seconda fase (*comparison*), comparano le opzioni rimanenti e operano una scelta tra queste ultime [McClelland e altri 1987]. Questo processo in due fasi, se da un lato è opportuno perché consente di ridurre il sovraccarico informativo, dall'altro può diventare motivo di ulteriori difficoltà, perché, in seguito all'operazione di soltimento, i consumatori devono scegliere fra prodotti i cui attributi sono correlati negativamente. Le correlazioni possono cambiare soprattutto in Internet, se efficaci aiuti decisionali consentono di rimuovere le opzioni non desiderate e passare così da correlazioni positive a negative o viceversa [Edwards e Fasolo 2001; Fasolo, Motta e Misuraca 2005].

Un'altra caratteristica soggettiva rilevante per i processi decisionali è l'importanza

assegnata ai diversi attributi dei prodotti; spesso l'importanza dei diversi attributi ha un andamento a gradini, con alcuni attributi considerati molto importanti e gli altri sempre meno importanti.

Quando il numero di attributi ritenuti rilevanti è molto grande la decisione può essere particolarmente difficile; per semplificare la decisione si può pensare allora di non considerare alcuni degli attributi; Hertwig e Todd hanno evidenziato che ciò non necessariamente peggiora significativamente la qualità della decisione [Hertwig e Todd, 2003]; addirittura Gigerenzer e Goldstein hanno evidenziato che a volte “meno è meglio” (*less is more effect*). La motivazione di questo effetto deriva dai meccanismi cognitivi (euristiche) utilizzati per trarre delle inferenze; quando si utilizzano euristiche semplici che considerano soltanto poche variabili importanti è possibile ridurre il rischio di “*overfitting*”, vale a dire di interpretare come significative delle variabili in realtà irrilevanti (*noise* o “rumore”); un esempio di euristica semplificata è l'euristica “prendi il migliore” (*take the best heuristic*), che prende decisioni tra due opzioni sulla base della variabile in grado di differenziarle meglio [Gigerenzer, Goldstein, 1996].

Fasolo, McClelland e Todd, tuttavia, hanno mostrato che riducendo drasticamente il numero di attributi da considerare non influenza significativamente la qualità della decisione soltanto se gli attributi considerati sono di gran lunga più importanti di quelli trascurati, o se tutti gli attributi sono tra loro positivamente correlati [Fasolo, McClelland, Todd, 2007].

Hogart e Karelaia hanno mostrato che euristiche che utilizzano pochi attributi con il minimo livello di dettaglio (per esempio con due soli valori) possono essere più efficienti delle euristiche che utilizzano attributi a più livelli o continui [Hogart, Karelaia, 2006].

Nel caso di correlazioni negative, se i diversi attributi hanno uguale importanza, secondo la teoria della decisione multiattributo [Keeney e Raiffa, 1976], l'approccio migliore è cercare le informazioni rilevanti per le diverse opzioni, calcolare i “valori additivi pesati” per ciascuna opzione, e scegliere l'opzione con il valore additivo pesato più elevato; il valore additivo pesato si ottiene moltiplicando il peso di ogni attributo per il suo valore, e sommando i diversi prodotti.

Secondo Misuraca, Fasolo e Cardaci, una possibile strategia per gestire il conflitto generato da correlazioni negative potrebbe essere quella di essere soddisfacentisti invece che massimizzatori; in particolare può essere opportuno semplificare l'ambiente decisionale riducendo il sovraccarico di informazioni offrendo meno opzioni; può anche essere utile semplificare la quantità di informazioni presentate simultaneamente ai consumatori, concentrandole su un piccolo numero di attributi, e riducendo il livello di dettaglio, pur consentendo il reperimento di informazioni aggiuntive da parte degli interessati. Un sovraccarico di informazioni è più probabile in presenza di molte opzioni simultaneamente disponibili, come, per esempio, su Internet; questi sono gli ambienti decisionali più difficili, in cui maggiormente potrebbe essere utile uno sfolgimento delle opzioni, e una riduzione del numero di attributi valutati non comporterebbe necessariamente una riduzione della qualità della scelta; fornire meno informazioni, soprattutto quelle di importanza minore, può rendere il processo decisionale meno difficile.

#### **1.4.1 Acquistare seguendo dei processi decisionali semplificati**

Uno degli argomenti più importanti delle ricerche sulle scelte dei consumatori è il problema della complessità decisionale; esso deriva sia dal numero elevato di opzioni, sia dai molti attributi che caratterizzano le diverse opzioni, sia dalla difficoltà di capire come compensare aspetti positivi e negativi delle diverse opzioni (*trade-off*). Nicola J. Bown, nell'introduzione al numero speciale di *Marketing Theory* dedicato all'analisi della rilevanza delle ricerche sulle valutazioni e analisi delle decisioni (*judgment and decision making*), ha messo in evidenza che queste ricerche cercano di capire il modo in cui le persone assumono decisioni complesse come quelle dei consumatori sui mercati [Bown, 2007]. Da un lato Bown ha mostrato che le persone desiderano avere ampie possibilità di scelta; secondo Schwartz tuttavia, possibilità di scelte eccessivamente numerose possono essere in ultima analisi dannose per i

consumatori, e quindi una limitazione del numero di opzioni disponibili può aumentare l'utilità derivabile dalla scelta. Alcune imprese, in effetti, hanno migliorato i loro risultati restringendo la gamma di prodotti; per esempio, Schwarts ha messo in evidenza che Procter & Gamble ottenne un aumento significativo delle vendite in seguito a una riduzione della sua gamma di prodotti. In generale tuttavia i consumatori si trovano spesso di fronte a numeri elevatissimi di possibilità di scelta; per esempio, venditori on line offrono circa cento diversi tipi di caffè, eppure gran parte dei consumatori decidono in genere in pochi secondi il tipo di caffè da acquistare; ciò significa che, a causa del poco tempo a disposizione, le persone in genere scelgono senza valutare tutte le opzioni disponibili.

Un tipo di semplificazione consiste nel passare da criteri decisionali *compensatori* a criteri decisionali *non compensatori*; vale a dire, invece di valutare tutte le opzioni di scelta attuando una compensazione fra performance migliori per alcuni attributi e performance meno buone per altri, le persone spesso utilizzano strategie decisionali in cui questa compensazione non opera. I criteri decisionali non compensatori si basano sull'uso di "euristiche", cioè di regole cognitive semplificate. Una strategia decisionale non compensatoria adottata di frequente è quella dell'eliminazione per aspetti [Tversky, 1972]<sup>9</sup>; secondo questa strategia, ciascun consumatore considera le diverse opzioni attributo per attributo, cominciando da quello più importante, ed eliminando le opzioni caratterizzate da una performance non soddisfacente dal punto di vista dell'attributo considerato. Per esempio, si ipotizzi che per un potenziale acquirente di un telefono cellulare gli attributi più importanti siano in primo luogo il prezzo e poi l'autonomia, e in particolare l'esigenza che il prezzo non superi i 100 euro, e che l'autonomia in *stand by* non sia inferiore a 300 ore; egli in primo luogo scarterà tutti i telefonini che hanno un prezzo maggiore di 100 euro, fra quelli rimasti scarterà quelli che hanno un'autonomia in *stand by* inferiore a 300 ore; a questo punto, se più telefonini soddisfano questi due requisiti, il consumatore potrà prendere in considerazione le performance da altri punti di vista, eventualmente continuando a seguire un ordine di importanza delle altre caratteristiche e scartare successivamente i telefonini che non hanno una performance ritenuta soddisfacente dal punto di vista considerato. Questo processo decisionale è relativamente semplice dal punto di vista cognitivo, poiché il consumatore valuta uno per volta gli attributi dei telefonini e non è costretto a dei trade-off fra performance contrastanti da diversi punti di vista. Il punto debole di una simile strategia decisionale è che essa trascura molte informazioni, e può condurre a scelte evidentemente discutibili, come per esempio quella di preferire un telefonino che costa 100 euro ad uno che costa 101 euro ma che ha delle performance molto migliori da diversi altri punti di vista; una questione importante è capire quale può essere la rilevanza concreta di questo problema.

---

<sup>9</sup> Amos Tversky (1937 -1999) pioniere della psicologia cognitiva, e più in generale della scienza cognitiva, collaborò per molti anni con Daniel Kahneman, premio Nobel per l'economia nel 2002 insieme a Vernon L. Smith, "per avere integrato intuizioni della ricerca psicologica nella teoria economica", nella ricerca di euristiche nello studio degli errori sistematici delle persone (*cognitive bias*), nell'analisi delle decisioni in condizioni d'incertezza. Tversky sviluppò assieme a Kahneman la *Prospect Theory*, per cercare di spiegare l'irrazionalità di alcune decisioni economiche; se non fosse morto prematuramente, quasi sicuramente Tversky avrebbe condiviso con Kahneman il premio Nobel per l'economia. Un altro grande economista e operatore finanziario che ha collaborato intensamente con Kahneman nell'analisi dei comportamenti economici che sembrano contrastare con la razionalità economica è Richard Thaler, fondatore, fra l'altro, di un fondo di investimenti che cerca di sfruttare i bias cognitivi degli investitori, e in particolare l'effetto dotazione, l'avversione per le perdite, e la preferenza per lo "status quo". Vernon L. Smith è considerato il fondatore dell'economia sperimentale, i cui risultati sono spesso convergenti con quelli di Kahneman, Tversky, e Thaler.

Fasolo, McClelland e Todd hanno analizzato la relazione fra il numero di attributi considerati e la bontà della scelta. Essi hanno misurato la bontà delle scelte semplificate confrontandole con quella che sarebbe effettuata attraverso una strategia compensatoria che tenesse conto di tutte le informazioni disponibili, e hanno trovato che, se tutti gli attributi sono positivamente correlati, vale a dire se a una buona performance per un attributo corrispondono buone performance anche per gli altri attributi, e se alcuni attributi sono più importanti di altri, una strategia decisionale semplificata che considera soltanto gli attributi più importanti non peggiora significativamente la scelta [Fasolo, McClelland, Todd, 2007].

Strategie decisionali semplificate, in cui le opzioni sono considerate attributo per attributo possono andar bene quando le scelte sono simultanee, come in un supermercato o su internet; una strategia diversa potrebbe essere preferibile quando le opzioni sono presentate sequenzialmente o in modi imprevedibili, e l'insieme delle scelte si riduce col tempo, per esempio quando si cerca una casa. Nella ricerca di una casa si possono individuare dei livelli soddisfacenti per i principali attributi, e scegliere la prima che soddisfa queste aspirazioni, senza perder tempo a cercare di acquisire e valutare tutte le informazioni teoricamente disponibili [Simon, 1947].

Mediante alcune simulazioni Fasolo, McClelland e Todd hanno mostrato che i consumatori possono effettuare delle buone scelte, anche trascurando molte delle informazioni teoricamente disponibili; ciò in particolare è vero se non c'è conflitto fra gli attributi dei prodotti o se alcuni attributi sono molto più importanti di altri. In questi casi è sufficiente considerare un solo attributo per effettuare una buona scelta, intendendo per buona scelta una il cui il valore non si discosta di più del 10 per cento da quello della scelta migliore possibile.

Secondo Fasolo, McClelland e Todd un'ampia possibilità di scelta appare come un'arma a doppio taglio; da un lato le persone amano avere ampie possibilità di scelta, d'altro lato, però, si sentono in imbarazzo di fronte alla scelta fra un numero molto elevato di opzioni disponibili. Da un lato, Bown e altri hanno messo in evidenza che una opzione ha maggiori probabilità di essere accettata quando è offerta insieme a un'altra opzione che quando è offerta isolatamente; analogamente, diverse ricerche hanno documentato la preferenza dei consumatori per la varietà. Per esempio, Iyengar e Lepper hanno mostrato che un numero maggiore di consumatori si fermarono davanti a uno stand con 24 varietà di marmellata che di fronte a uno in cui ve ne erano soltanto sei varietà [Iyengar, Lepper, 2000]. D'altro lato una maggiore possibilità di scelta a volte comporta dei problemi, poiché il dover scegliere fra un numero di opzioni eccessivamente elevato può far sì che i consumatori avvertano la difficoltà della scelta e temono di non essere in grado di effettuare una scelta soddisfacente e saranno scoraggiati nell'acquisto.

Il problema di una eccessiva varietà di scelte ha cominciato ad apparire sempre più rilevante a partire da primi anni duemila, soprattutto in connessione con lo sviluppo degli acquisti su Internet; mentre nei negozi tradizionali esistono dei vincoli fisici che limitano il numero delle varietà di ciascun bene che possono essere mostrate ai potenziali acquirenti, nei siti di vendite online non esistono praticamente limiti, e possono essere rese disponibili quindi migliaia di varietà per ciascun bene.

Negli ultimi anni, in effetti, diverse ricerche hanno mostrato che possono essere ottenute inferenze accurate, o addirittura più accurate, utilizzando soltanto una parte delle informazioni disponibili [Gigerenzer, Goldstein, 2002]. In una delle loro simulazioni, Gigerenzer e Goldstein misurarono l'accuratezza delle risposte alla domanda su quale di due città tedesche hanno più abitanti, sulla base di nove elementi informativi; essi trovarono che delle risposte euristiche "veloci e frugali" che si basavano su poche informazioni possono essere quasi altrettanto accurate delle risposte che si basavano sulla considerazione di tutte le informazioni disponibili. Inoltre, quando le diverse euristiche furono applicate a un altro gruppo di città, quelle "veloci e frugali" ebbero risultati addirittura migliori di quelle che utilizzavano tutte le informazioni disponibili. La spiegazione di questi risultati apparentemente paradossali è che mentre le euristiche più semplici si basavano sulle informazioni più rilevanti, le regressioni multiple

utilizzavano anche informazioni irrilevanti, suscettibili anche, eventualmente, di indurre in errore. Nel cercare di dare un senso a tutte le informazioni disponibili, i criteri decisionali basati su regressioni multiple possono condurre a decisioni sbagliate. Le persone inoltre, utilizzano spesso dei criteri decisionali euristici basati su un solo elemento decisionale, quello di solito più efficace, oppure su euristiche ricognitive, che si basano su potenti elementi di conoscenza ricognitiva [Gigerenzer, Goldstein, 2002].

Due conclusioni importanti emergono da queste ricerche; in primo luogo, il fatto di considerare soltanto una parte delle informazioni disponibili non comporta necessariamente decisioni peggiori, e può persino migliorare i risultati quando le procedure decisionali sono applicate a nuove scelte; in secondo luogo, il beneficio di utilizzare soltanto alcune delle informazioni disponibili dipende dalla bontà della relazione fra l'euristica utilizzata e l'ambiente decisionale. Per esempio, l'euristica "prendi il migliore" conduce a decisioni corrette soltanto se utilizzata in ambienti decisionali in cui esistono alcune informazioni altamente rilevanti. Per quel che riguarda le scelte dei consumatori, il problema è cercare di capire in quali situazioni un consumatore può effettuare delle buone scelte, anche ignorando gran parte delle informazioni disponibili.

Un tipico paradigma delle ricerche sul comportamento del consumatore riguarda il caso di opzioni con due attributi importanti; la qualità e il prezzo [Simonson, Tversky, 1992; Tversky, Shafir, 1992]. In concreto però la qualità di una opzione dipende spesso da valutazioni su un elevato numero di attributi delle diverse opzioni. In diverse ricerche che hanno analizzato le metodologie di elaborazione e sintesi delle informazioni riguardanti i diversi attributi delle differenti opzioni, l'attenzione è stata concentrata sul trade-off fra accuratezza della scelta e costi di ricerca e di elaborazione delle informazioni. Nei casi in cui gli attributi delle diverse opzioni sono correlati positivamente, e/o esistono pochi attributi molto più importanti di tutti gli altri, si può effettuare una buona scelta anche sulla base di informazioni relative a pochi attributi [Bettman e altri, 1993].

Gli attributi delle diverse opzioni sono correlati positivamente, se una opzione che appare buona con riferimento a un attributo, appare buona anche con riferimento all'altro attributo ed in questo caso, l'ambiente decisionale viene definito amichevole (*friendly*). Invece, gli attributi sono negativamente correlati, quando le opzioni che appaiono buone per alcuni attributi, appaiono non buone dal punto di vista di altri attributi ed in questo caso, l'ambiente decisionale viene definito ostile (*unfriendly*) [Shanteau, Thomas, 2000]; nei casi di correlazioni negative fra attributi le decisioni sono più difficili, poiché si presenta un trade-off fra vantaggi e svantaggi delle diverse opzioni; sono queste situazioni in cui si manifestano dei conflitti decisionali che comportano una diminuzione dell'utilità della scelta, e dubbi sulla bontà delle decisioni. La correlazione negativa più frequente è quella fra qualità ed economicità delle diverse opzioni, intendendo per economicità il reciproco, o l'opposto, del valore del prezzo<sup>10</sup>.

Le valutazioni sull'importanza relativa dei diversi attributi, essendo eminentemente soggettive, variano ovviamente fra un consumatore e l'altro; inoltre, il peso attribuito ai diversi attributi aumenta all'aumentare della variabilità degli indicatori di performance relativi a ciascun attributo<sup>11</sup>. Alcune ricerche hanno messo in evidenza che spesso le valutazioni d'importanza dei diversi attributi evidenziano un andamento "a gradini", con pochi attributi molto importanti, e tanti altri attributi molto meno rilevanti. Comunque, in genere può essere individuato per ciascun bene un nucleo di attributi importanti su cui le persone in genere concentrano l'attenzione [Edwards e Barron, 1994].

I processi decisionali semplificati non sempre sono affidabili. In passato alcune ricerche sulle scelte effettuate utilizzando poche informazioni avevano ottenuto risultati non

---

<sup>10</sup> Tuttavia Curry e Riesz (1988), confrontando gli indicatori di qualità e i prezzi riportati sulla rivista americana *Consumer Reports*, trovarono per alcune categorie di prodotti una correlazione positiva fra qualità ed economicità.

<sup>11</sup> Per dei risultati diversi si veda Fischer (1995).

particolarmente positivi. Barron e Barret avevano simulato delle scelte di una città in cui vivere sulla base di nove attributi (clima, abitazioni, cure sanitarie, criminalità, trasporti, istruzione, attività ricreative, arti, economia) [Barron, Barret, 1999]. Il database utilizzato comprendeva 15 città degli Stati Uniti che presentavano un contesto decisionale “ostile”, nel senso che le correlazioni fra gli attributi erano prevalentemente negative (quindi ciascuna città presentava in genere rispetto alle altre sia vantaggi, sia svantaggi); nelle ricerche si ipotizzò che gli attributi avessero tutti la stessa importanza. Nelle simulazioni furono esaminati gli effetti di scegliere sulla base di un sottoinsieme dei nove attributi, in termini di proporzione del valore perso, ottenuta confrontando il valore dell’opzione scelta in questo modo, con quello dell’opzione che appariva migliore sulla base di tutti gli attributi. Il valore delle diverse opzioni fu stimato moltiplicando per ciascuna opzione i valori degli attributi per i rispettivi pesi, e sommando i risultati dei prodotti per ciascuna opzione, così come indicato dalla teoria dell’utilità in presenza di più attributi [Keeney, Raiffa, 1993].

I risultati ottenuti mostrarono che in presenza di attributi negativamente correlati e ugualmente importanti, le decisioni assunte sulla base di una parte soltanto delle informazioni disponibili sugli attributi possono comportare significative perdite di valore; ciò è conforme alla regola generale condivisa da economisti ed esperti di teoria delle decisioni, secondo cui trascurare alcune delle informazioni viola uno dei principi fondamentali della teoria delle decisioni. Barron e Barret trovarono invece che la perdita di valore è minore, ma ancora significativa, quando gli attributi non sono ugualmente importanti, e la scelta è effettuata sulla base delle informazioni relative agli attributi più importanti [Barron, Barret, 1999].

### **1.5 I Conflitti decisionali e le influenze in una decisione**

Una possibile reazione dei consumatori al sovraccarico informativo può essere il differimento della decisione di acquisto, per il timore di non riuscire ad assumere la decisione migliore; ciò significa che i consumatori dal canto loro trovano difficoltà a rinunciare a considerare una parte rilevante delle informazioni disponibili. Coloro che commercializzano i prodotti dovrebbero quindi cercare di mettere in evidenza le informazioni più importanti relative a un piccolo numero di attributi, e dare la possibilità di ottenere ulteriori informazioni soltanto alle persone disposte a cercarle attivamente; la probabilità che i consumatori differiscano l’acquisto è presumibile che diminuisca se ci sono meno informazioni sui prodotti da valutare.

Anche se il fatto di considerare soltanto una parte delle informazioni disponibili comporta il rischio di trascurare qualcosa d’importante, le euristiche decisionali “*scorciatoie cognitive*”, come accontentarsi di scelte soddisfacenti invece che perseguire quelle ottimali, e la eliminazione progressiva delle opzioni per aspetti, consentono al consumatore di effettuare rapidamente delle scelte sulla base di alcuni aspetti importanti e quindi in genere riflettono le sue effettive preferenze. Altre euristiche, tuttavia, possono comportare violazioni delle regole delle decisioni razionali e comportare delle scelte che non riflettono le vere preferenze del consumatore. Secondo Shafir il comportamento dei compratori può spesso essere influenzato dalla maggiore o minore facilità di giustificare le scelte fatte; ciò può determinare scelte inconsistenti, poiché le giustificazioni possibili dipendono dal contesto in cui le scelte sono effettuate [Tversky, Shafir, 1992].

Boothe, Schwartz e Chapman hanno messo in evidenza che informazioni relativamente poco importanti possono provocare delle inversioni di preferenze razionalmente non giustificabili. Nel loro esperimento essi offrirono delle coppie di beni di consumo differenti in valore oggettivo e desiderabilità. Le coppie di prodotti furono scelte in modo tale che un articolo avesse un prezzo di mercato elevato ma fosse meno desiderabile, mentre l’altro articolo avesse un basso prezzo di mercato ma fosse più desiderabile. Gran parte delle persone scelsero il prodotto più desiderabile, ma al prodotto indesiderabile, che aveva un prezzo di mercato più elevato, fu assegnato un prezzo minimo di vendita più elevato. I risultati suggeriscono che quando le persone non sono sicure su come tradurre le loro preferenze in un valore monetario

esse fanno riferimento al prezzo di mercato piuttosto che alle loro preferenze. L'implicazione di questa "euristica - prezzo di mercato" é che se i commercianti stabiliscono il prezzo al dettaglio di un prodotto a un certo livello, i consumatori utilizzeranno quel prezzo come base delle loro valutazioni di prezzo. In altre parole, le scelte tendono a riflettere le preferenze delle persone, mentre i prezzi minimi a cui si è disposti a vendere riflettono i prezzi di mercato; questa euristica del prezzo di mercato può contribuire a spiegare l'attrazione delle svendite, e la correlazione fra prezzo e valore [Boothe, Schwartz e Chapman, 2007].

Le ricerche sul comportamento dei consumatori hanno mostrato ripetutamente che le scelte dei consumatori dipendono dagli specifici insiemi di alternative in cui ciascuna opzione è considerata, in quanto lo stesso prodotto può essere più desiderabile in un contesto piuttosto che in un altro [Simonson e Tversky, 1992].

### **1.5.1 Le influenze nel processo decisionale derivanti dal contesto**

Le implicazioni pratiche degli effetti di contesto per le strategie di marketing sono numerose, in particolar modo per le strategie di posizionamento, di prezzo e di promozione dei prodotti. Busemeyer, Barkan, Mehta e Chaturvedi hanno presentato una rassegna delle principali spiegazioni teoriche proposte per le "influenze di contesto" delle decisioni. Gli effetti di contesto considerati sono: *l'effetto similarità*, *l'effetto attrazione*, *l'effetto compromesso*, e *l'effetto del punto di riferimento* [Busemeyer, Barkan, Mehta, Chaturvedi, 2007] .

Gli effetti di similarità, di attrazione, e di compromesso sono stati illustrati da questi studiosi, ipotizzando un contesto iniziale caratterizzato dalla possibilità di scelta fra due tipi di automobile, una Volkswagen Jetta, valutata 7 dal sito Edmunds.com, e messa in vendita al prezzo di 17.675 euro (opzione A) e un'Audi A4, valutata 8,3 e messa in vendita al prezzo di 25.760 euro (opzione B); successivamente viene aggiunta un'altra opzione, rappresentata da una Suzuki Aero, valutata 6,7, e messa in vendita al prezzo di 14.019 euro (opzione S). Fra gli attributi che caratterizzano queste tre opzioni c'è una evidente correlazione negativa, poiché all'aumentare dell'indice di qualità l'indice di economicità (rappresentato, per esempio, dal reciproco o dall'opposto del prezzo) diminuisce; l'esempio rappresenta quindi un contesto difficile di scelta (*unfriendly*), poiché l'acquirente deve accettare una peggiore performance in termini di uno dei due attributi (minore economicità) per avere una migliore performance in termini dell'altro attributo (migliore qualità). In ordine decrescente di qualità, e quindi crescente in termini di economicità le tre opzioni possono essere elencate nel modo seguente: S, A, B. In termini sia di economicità sia di qualità, la differenza fra A e B è significativamente maggiore della distanza fra A ed S; ciò significa che A ed S appaiono abbastanza "simili" rispetto a B e A.

L'effetto "similarità" consiste nel fatto che se, considerando le opzioni a due a due le probabilità di scelta sono uguali sia fra le opzioni A e B, sia fra le opzioni A ed S, il solo fatto di aggiungere alle due opzioni iniziali A e B una terza opzione abbastanza simile ad A, fa sì che la probabilità di scelta di B rimanga invariata, mentre la probabilità di scelta sia di A sia di S diminuisca; in altri termini, le due opzioni di scelta relativamente simili si danneggiano vicendevolmente. L'effetto similarità viola il principio razionale di indipendenza della scelta di una opzione dalla presenza o meno di alternative irrilevanti; in base a questo principio, se, in assenza di altre opzioni, la probabilità di scelta dell'opzione A è uguale alla probabilità di scelta dell'opzione B, questa eguaglianza non dovrebbe essere influenzata dall'aggiunta di una terza opzione irrilevante dal punto di vista del confronto fra A e B. In base al principio di razionalità, se le probabilità di scelta delle tre opzioni sono eguali nel caso di confronti diretti, esse dovrebbero essere uguali anche nel caso in cui ciascuna delle tre opzioni è confrontata con le altre due. In realtà invece si verifica che passando dai confronti diretti fra due sole opzioni a un contesto di scelta con tre opzioni, di cui due più simili fra di loro rispetto all'altra, le probabilità di scelta di ciascuna delle due opzioni relativamente simili risultano minori della probabilità di scelta dell'opzione "diversa".

L'effetto attrazione può essere illustrato ipotizzando di aggiungere alle due opzioni iniziali dell'esempio precedente una Toyota Prius, con un indicatore di performance pari a 7 (come l'opzione A), ma messa in vendita a un prezzo superiore (20.480 invece di 17.675 euro). L'aggiunta di una "opzione esca" D simile ad A ma da essa dominata, in concreto spesso fa aumentare la probabilità di scelta di A, che adesso può apparire come una buona occasione in confronto a D. Ciò viola il principio razionale della regolarità, secondo cui l'aggiunta di una terza opzione dovrebbe o lasciare invariate le probabilità di scelta di tutte le opzioni precedente (se D non è mai scelta) o farle diminuire tutte (se qualche volta D è scelta); in realtà invece spesso si osserva che l'aggiunta di D fa aumentare la probabilità che sia scelta l'opzione più vicina a D, ma oggettivamente più conveniente di D.

L'effetto compromesso può essere illustrato ipotizzando di aggiungere alle due opzioni iniziali A e B una terza opzione D caratterizzata da una maggiore economicità ma da un minore indicatore di qualità rispetto sia ad A che a B. Simonson e Tversky hanno evidenziato che in genere l'aggiunta dell'opzione di basso prezzo e bassa qualità fa aumentare la probabilità di scelta della opzione A rispetto all'opzione B, poiché l'aggiunta di D fa apparire A come una scelta intermedia o "di compromesso" fra B e D, in termini sia di qualità sia di prezzo. Anche questi comportamenti sembrano in contrasto con il principio razionale di indipendenza delle scelte da alternative irrilevanti. Più in generale, secondo Simonson e Tversky se le probabilità di scelta delle due opzioni iniziali erano uguali, dopo l'aggiunta di una terza opzione che fa apparire intermedia una delle opzioni iniziali, la frequenza con cui è scelta l'opzione intermedia è maggiore delle frequenze di scelta delle due opzioni "estreme" [Simonson e Tversky, 1992].

L'effetto del "punto di riferimento" è stato illustrato da Busemeyer, Barkan, Mehta e Chaturvedi ipotizzando che un consumatore abbia ordinato un'automobile HE dal concessionario, con un indice di performance pari 7,4 e un prezzo di 13.440 euro. A causa di un ritardo nella disponibilità dell'auto ordinata, il concessionario propone o di mantenere invariato l'ordine o di scegliere fra due nuove opzioni, una HC di qualità 8,1 al prezzo di 13.470 euro, e una SI di qualità 5,6, al prezzo di 12.000 euro. Da questo punto di vista, HC sembra essere la scelta preferibile; potrebbe la scelta essere diversa se inizialmente fosse stata ordinata un'automobile diversa?

Tversky e Kahneman fecero due esperimenti con due diversi punti di riferimento. In un primo esperimento ai partecipanti fu chiesto di immaginare di possedere E, un bene di alta qualità e bassa economicità, dando la scelta fra mantenere E o scambiarlo con A o con B; rispetto ad E B ha un piccolo vantaggio in termini di qualità e nessuno svantaggio in termini di economicità, mentre A ha un significativo vantaggio in termini di economicità, ma un rilevante svantaggio di qualità; in queste condizioni E fu raramente scelto, e B fu scelto molto più di A. In un altro esperimento fu chiesto ai partecipanti di immaginare di possedere F, un bene di bassa qualità e alta economicità, dando loro la scelta fra mantenerlo, oppure scambiarlo con A o con B. Rispetto ad F adesso A ha un piccolo vantaggio in termini di economicità e analoga qualità, mentre B è di qualità molto più alta e di economicità molto più bassa; in questo caso pochissimi decisero di mantenere F, ed A risultò essere, sia pur leggermente, preferito a B. Questo esperimento mostra come le preferenze della maggioranza delle persone fra A e B possano invertirsi, a seconda che il punto di partenza sia leggermente peggiore di B o leggermente peggiore di A. Di fronte alla difficoltà di valutare la convenienza di un mutamento radicale della combinazione qualità-prezzo rispetto a quella di partenza, la maggioranza delle persone sembra accontentarsi di piccoli miglioramenti rispetto alla situazione di partenza, qualunque essa sia (alta economicità e bassa qualità, oppure bassa economicità e alta qualità).

In un'altra coppia di esperimenti Tversky e Kahneman chiesero ai partecipanti di immaginare di essere stati addestrati per un certo tipo di lavoro T, di elevata qualità e di bassa economicità,<sup>12</sup> non più disponibile, e di avere la possibilità di scelta fra due diverse occupazioni

---

<sup>12</sup> Per esempio un lavoro molto interessante e poco faticoso, ma pagato relativamente poco.

A e B, B caratterizzato da piccoli vantaggi e svantaggi rispetto a T, e A caratterizzato da rilevanti vantaggi e svantaggi rispetto a T; gran parte delle persone dichiararono di scegliere B rispetto ad A. Alle persone fu poi chiesto di immaginare di essere state addestrate per un tipo di lavoro S, di bassa qualità ma di elevata economicità, non più disponibile, e di avere la possibilità di scelta sempre fra A e B; in questo caso A è analogo a S in termini sia di economicità sia di qualità, mentre B è di qualità molto più elevata ma di economicità molto più bassa; in questo caso gran parte delle persone dichiararono di preferire A rispetto a B. A seconda del punto di riferimento, l'esperimento mostrò quindi una inversione delle preferenze fra A e B [Tversky, Kahneman, 1981].

Alla Conferenza di Roma sulla simulazione dei comportamenti adattivi (Roma, CNR, settembre 2006), Zeeshan-ul-hassan Usmani, uno scienziato informatico dell'Università di Princeton e Ronaldo Menezes del *Florida Institute of Technology* hanno presentato un lavoro in cui si mostra la possibilità di aumentare le vendite sfruttando l'istinto delle persone a fare quello che altri hanno già fatto (*herd instinct*). Per informare i clienti sugli acquisti precedenti, Usmani e Menezes hanno proposto di inserire in ogni prodotto messo in vendita in un supermercato una targhetta identificatrice con un codice a barra che viene letto ogni volta che il prodotto viene acquistato; le informazioni sugli acquisti sono immediatamente trasmesse a un computer centrale, che le elabora e le invia agli schermi collocati sugli scaffali, così che i clienti possono conoscere in tempo reale il numero di persone che hanno appena acquistato il prodotto che stanno guardando. Il fatto di sapere che tante altre persone hanno appena acquistato quel prodotto è rassicurante per i clienti più incerti, e spesso li convince ad acquistare. Questo sistema è stato recentemente adottato in via sperimentale dalle catene di supermercati Wall Mart in America e Tesco in Inghilterra.

Matthew Salganik, del Dipartimento di Sociologia dell'Università di Princeton, nell'ambito delle sue ricerche di Ph.D presso la Columbia University di New York, sperimentò il funzionamento di un mercato musicale in cui 14.000 persone scaricavano delle canzoni mai sentite prima. I ricercatori scoprirono che quando le persone potevano vedere le canzoni ordinate in base al numero di volte in cui erano state già scaricate, esse tendevano a "seguire la folla", vale a dire ciascuno tendeva a scaricare di più le canzoni che erano state scaricate da un numero più elevato di persone.

Questo effetto psicologico, "*herd effect*" o "effetto gregge o effetto folla" si manifesta anche negli acquisti online; per sfruttarlo Amazon e altri venditori online forniscono informazioni dettagliate ai potenziali acquirenti riguardo la popolarità dei diversi prodotti. (*The Economist*, 11 novembre 2006, p. 90).

"L'effetto-gregge" si manifesta spesso soprattutto nei mercati azionari; una regola spesso seguita dagli investitori, ricordata in un articolo su L'Espresso del 23 agosto 2007 (p. 126) è la seguente: "quando vedi che tutti vendono, vendi anche tu"; da questo effetto derivano gli episodi di "*panic selling*", in cui si manifesta una diffusa tendenza a vendere azioni derivante principalmente dall'osservazione che molti altri le stanno vendendo. Analogamente, quando si sparge la voce, più o meno fondata, che una banca si trova in difficoltà, spesso si forma una coda sempre più lunga di correntisti che chiedono di riavere indietro i soldi depositati, principalmente per il fatto di osservare che tanti altri lo hanno fatto.

### **1.6 Strategie decisionali per superare i conflitti decisionali**

Una idea interessante della letteratura sui processi decisionali è che le persone hanno un insieme di strategie decisionali, e che esse possono cambiare strategia a seconda della dimensione dell'insieme di scelta e del contesto di scelta [Gigerenzer e Selten, 2001; Payne e altri, 1993]. Esaminare tutti i possibili modelli di cambiamento delle strategie è molto difficile; se ne possono tuttavia considerare alcuni. Si ipotizzi, per esempio, che una persona possa scegliere fra una strategia decisionale compensatoria e una strategia non compensatoria; la probabilità di scelta di una strategia compensatoria aumenta all'aumentare della dimensione

dell'insieme di scelta. Per la strategia decisionale compensatoria si può ipotizzare che a ciascuna opzione venga assegnato un certo valore di un indicatore di utilità, e che sia scelta l'opzione a cui corrisponde il massimo valore di utilità; se due opzioni hanno la stessa utilità, la scelta può essere casuale. Per la strategia non compensatoria si può ipotizzare una regola lessicografica; in questo caso una persona prima considera l'aspetto più importante, e sceglie l'opzione migliore per questo aspetto; se più opzioni sono analoghe per quel che riguarda l'aspetto più importante, si considera il secondo attributo in ordine di importanza, e viene selezionata l'opzione migliore per il secondo aspetto, e così via, fino a quando non rimane una sola opzione. Due opzioni possono essere considerate equivalenti per un aspetto, se la differenza di performance per quell'aspetto è minore di un certo valore-soglia; l'ordine degli attributi in termini d'importanza può anche essere diverso fra una scelta e l'altra. Per tener conto di effetti di contesto nel caso di scelte binarie, è necessario ipotizzare che strategie lessicografiche non compensatorie siano spesso utilizzate per scelte binarie. Se si utilizzano strategie decisionali lessicografiche, il modello dei cambiamenti di strategia può spiegare anche l'effetto similarità, se si ipotizza che la differenza fra una opzione e quella ad essa simile sia inferiore del valore-soglia, così che le due opzioni possano essere considerate equivalenti. Supponendo che le opzioni simili siano quelle di elevata economicità e bassa qualità, se l'aspetto qualità è considerato per primo, viene scelta l'altra opzione; se invece si considera prima l'aspetto economicità, sarebbero preferite le opzioni simili, e la scelta fra di esse sarebbe casuale. Supponendo che, prima di aggiungere la terza opzione, le probabilità di scelta fra le due opzioni originarie siano uguali (0,5 ciascuna); l'aggiunta della terza opzione fa dimezzare la probabilità di scelta originaria simile alla terza, e lascia inalterata la probabilità di scelta dell'altra opzione.

Il modello dei cambiamenti di strategia di scelta non può fornire spiegazioni soddisfacenti per l'effetto compromesso; in presenza di tre opzioni, la strategia lessicografica non porterebbe mai alla scelta dell'opzione intermedia, poiché essa non è la migliore da nessun punto di vista; ipotizzando una strategia decisionale compensatoria, l'opzione intermedia sarebbe scelta se essa avesse il più alto valore dell'indicatore di *performance*, ma ciò contrasta con l'ipotesi che le probabilità di scelta delle due opzioni originarie siano uguali. Busemeyer e altri hanno dimostrato inoltre che questo modello non può spiegare neppure l'effetto attrazione né l'effetto punto di riferimento [Busemeyer et al., 2007].

### **1.6.1 Il modello decisionale della “eliminazione per aspetti” e l'avversione per le perdite**

Il modello decisionale della “eliminazione per aspetti” fu proposto in origine da Tversky [Tversky, 1972] e da Tversky e Sattah [Tversky, Sattah, 1979]; in questo modello ogni opzione è caratterizzata da un certo numero di aspetti e a ciascun aspetto è assegnato un valore che ne misura l'importanza; gli aspetti comuni sono presenti in diverse opzioni, mentre gli aspetti unici sono presenti in una sola opzione. Il meccanismo di scelta comporta due stadi: a un primo stadio viene scelto uno degli aspetti con un criterio probabilistico, così che gli aspetti più importanti hanno una maggiore probabilità di essere selezionati; al secondo stadio, le opzioni che non contengono l'aspetto selezionato sono escluse. Il processo di selezionare un aspetto ed eliminare tutte le opzioni che non lo contengono prosegue fino a quando rimane una sola opzione, che viene quindi scelta. Questo modello fu in origine pensato per spiegare l'effetto similarità; Tversky dimostrò che il modello della eliminazione per aspetti soddisfa il principio di regolarità, e quindi non può tener conto dell'effetto attrazione; questo modello decisionale non può neppure spiegare né l'effetto punto di riferimento, né l'effetto compromesso. Per poter tener conto di questi effetti Tversky e Kahnemann svilupparono un modello di scelta basato sul concetto di “avversione alle perdite”: il “*componential context model*” [Tversky e Kahnemann, 1991]. Secondo questo modello vi è una differenza fondamentale fra gli insiemi di scelta che contengono soltanto due opzioni, e quelli che ne includono più di due; quando l'insieme di scelta

contiene soltanto due opzioni, il valore assegnato a ciascuna di esse non dipende dal contesto<sup>13</sup>; il contesto diventa invece rilevante quando tre o più opzioni sono incluse nell'insieme di scelta. In questo caso, il valore di ciascuna opzione ha due componenti, una indipendente dal contesto e una dipendente da esso: quest'ultima componente si basa sul concetto di vantaggi e svantaggi di una opzione rispetto all'altra; per esempio l'opzione A, rispetto all'opzione B, può avere un grande vantaggio in termini di qualità e un forte svantaggio in termini di economicità. Un'assunzione cruciale di questo modello è che gli svantaggi hanno un impatto più forte dei vantaggi (avversione alle perdite). Negli esempi utilizzati per illustrare l'effetto punto di riferimento, l'opzione più vicina a quella di riferimento risulta essere in genere favorita rispetto a quella più lontana; l'ipotesi di avversione alle perdite può spiegare questo risultato, se si considera che l'opzione più vicina al punto di riferimento ha rispetto a questa piccoli vantaggi e piccoli svantaggi, mentre l'opzione più lontana presenta grandi vantaggi e grandi svantaggi. Se gli svantaggi hanno un impatto più forte dei vantaggi, a parità di altre condizioni, gli svantaggi delle opzioni più lontane dal punto di riferimento hanno forti effetti negativi che non possono essere compensati dai loro vantaggi.

Utilizzando il concetto di "avversione per le perdite", con il modello del contesto componenziale si cercò di spiegare l'effetto attrazione, l'effetto compromesso, e l'effetto punto di riferimento; tuttavia questo modello non è in grado di fornire una spiegazione per l'effetto similarità; ciò perché l'ipotesi di avversione per le perdite implica che una opzione sarebbe favorita dall'aggiunta di un'altra opzione ad essa simile, e ciò è in contrasto con molte evidenze empiriche.

### **1.6.2 La teoria del "campo decisionale"**

La teoria del campo decisionale (*decision field theory*), cerca di spiegare la dinamica temporale delle preferenze; secondo questa teoria, ciascuna persona ha in ciascun momento un diverso sistema di preferenze per le opzioni disponibili. La struttura delle preferenze muta nel corso del tempo perché l'attenzione si sposta verso aspetti diversi delle opzioni disponibili; si consideri, per esempio, la scelta fra diversi tipi di automobili; mentre in un primo momento l'attenzione potrebbe essere concentrata sulla qualità, successivamente essa potrebbe spostarsi sulla economicità, e così via. Il processo di aggiornamento della valutazione delle diverse opzioni prosegue, fino a quando la preferenza per una di esse supera una certa soglia; questa opzione è quella che viene effettivamente scelta [Busemeyer, Townsend, 1993] [Roe et al, 2001].

Gli effetti preferenziali di contesto in questo modello si basano sul grado di similarità fra diverse opzioni; il contrasto fra diverse opzioni aumenta man mano che esse diventano più simili, questo contrasto fa aumentare l'attrattiva dell'opzione migliore e diminuire quella dell'opzione meno buona. Dhar e Blazer hanno mostrato che la valutazione del grado di similarità fra due opzioni può cambiare quando si introduce una terza opzione; ciò è in accordo con gli effetti di contesto [Dhar, Blazer, 1996]. La teoria del campo decisionale rappresenta questo risultato mediante reti neurali competitive regressive, spesso utilizzate in modelli connessionisti per produrre effetti di esaltazione dei contrasti nella percezione [Haykin, 1994]; questo processo di "inibizione ricorrente" si evolve col tempo, diventando più forte man mano che si procede nel processo decisionale. È stato dimostrato che la teoria del campo decisionale, basandosi sul grado di similarità fra le alternative, fornisce delle spiegazioni plausibili per gli effetti similarità, attrazione, e compromesso, sulla base di analoghi valori dei parametri decisionali. Per esempio, se l'opzione D è simile, ma inferiore, all'opzione A, mentre l'opzione B è molto diversa da entrambe, il contrasto fra D e A fa aumentare la probabilità di scelta di A e diminuire quella di D, mentre la probabilità di scegliere B rimane invariata; di conseguenza la probabilità di scelta di A aumenta rispetto a quella di B [Roe et al, 2001].

---

<sup>13</sup> Ciò comporta che questo modello non è in grado di spiegare gli effetti di contesto che a volte vengono osservati anche nel caso di scelte binarie (Mellers e Biagini, 1994).

Secondo la teoria del campo decisionale, sia la dimensione sia la direzione degli effetti di contesto possono dipendere dal tempo a disposizione per la decisione; in particolare gli effetti di attrazione e di compromesso diventano più forti all'aumentare della durata del processo decisionale [Simonson, 1989; Dahr e altri, 2000]. La teoria del campo decisionale è un esempio di un modello decisionale connessionista; altri due modelli decisionali connessionisti sono stati proposti da Guo e Holyoak, per cercare di spiegare gli effetti di attrazione e di similarità [Guo, Holyoak, 2002], e un modello di rete neurale artificiale che utilizza alcune ipotesi della teoria del campo decisionale che potrebbe spiegare tutti e quattro gli effetti di contesto qui considerati [McLelland, 2001].

La principale contraddizione di questa analisi teorica riguarda gli effetti dell'aggiunta di una terza opzione sulla opzione più simile (vicina) alla nuova opzione; secondo gli effetti attrazione, compromesso e punto di riferimento, l'aggiunta della terza opzione fa aumentare le probabilità di scelta della opzione simile a quella che si aggiunge; secondo l'effetto similarità, invece, la fa diminuire. In effetti, gran parte dei modelli teorici che spiegano l'effetto similarità non spiegano gli altri effetti, e viceversa; in particolare, i modelli dell'utilità probabilistica, dell'eliminazione per aspetti, del mutamento delle strategie spiegano l'effetto similarità, ma non gli altri effetti, mentre il modello del contesto componenziale non spiega l'effetto similarità ma spiega tutti gli altri effetti. Anche alla luce di questo risultato appare sorprendente la capacità della teoria del campo decisionale di fornire una spiegazione per tutti e quattro gli aspetti qui considerati, superando la contraddizione concernente l'effetto similarità.

### **1.6.3 Confronto di diversi criteri decisionali**

A volte i consumatori conoscono bene il valore per essi di ciascuna opzione, per esempio il menu di un ristorante ben conosciuto o tipi di automobili sperimentate personalmente o da amici e parenti. Il modello della semplice scalabilità e quello dei cambiamenti di strategie decisionali ipotizzano in effetti che non vi siano incertezze riguardo l'utilità per i consumatori delle diverse opzioni; tuttavia i processi decisionali sono diversi. Per il modello della semplice scalabilità la scelta è probabilistica, e quindi il consumatore può anche scegliere una opzione diversa da quella che gli dà la massima utilità; invece il modello dei cambiamenti di strategia prevede un'analisi sistematica delle diverse opzioni, e la scelta di quella che massimizza l'utilità del consumatore. Il criterio di scelta probabilistico sembra più ragionevole per scelte frequentemente ripetute di beni di modesto valore come per il menu di un ristorante; nel loro complesso queste caratteristiche consentono al consumatore di sperimentare una varietà di opzioni senza rischiare molto. Il modello dei cambiamenti di strategia, che comporta un'analisi più sistematica delle opzioni disponibili, è più utilizzato per le scelte di beni importanti, che durano per un periodo di tempo abbastanza lungo e che costano molto, come le automobili e altri beni durevoli; in questi casi il consumatore avverte anche l'esigenza di poter giustificare la razionalità della scelta.

A volte i consumatori iniziano un processo di ricerca avendo soltanto idee vaghe su ciò che cercano; per esempio una persona potrebbe cercare qualcosa di "carino" da indossare per una festa, senza sapere bene cosa; oppure una persona che pensa di uscire, senza sapere se andare al cinema o al ristorante. Quando gli obiettivi non sono ben definiti, i criteri di valutazione delle opzioni sono necessariamente ambigui, poiché le utilità delle diverse opzioni possono cambiare rapidamente; in situazioni ambigue e incerte è presumibile che aumenti la probabilità di scelta di opzioni intermedie, o di compromesso, mentre è meno probabile che emerga un contesto caratterizzato da valutazioni chiare di dominanza necessarie per l'operare dell'effetto attrazione. A volte invece le persone sanno bene cosa cercano, ma emergono delle situazioni di conflitto decisionale per il fatto che ciascuna delle opzioni disponibili corrisponde soltanto parzialmente agli obiettivi perseguiti (trade-off fra qualità, prezzo, tempo di attesa, ecc.). Per queste situazioni potrebbero essere utili modelli dell'eliminazione per aspetti e del contesto componenziale. Il modello dell'eliminazione per aspetti spiega l'effetto similarità,

mentre quello del contesto componenziale spiega gli effetti attrazione, compromesso e punti di riferimento. Nel modello dell'eliminazione per aspetti ciascuna opzione è valutata per i singoli aspetti presi singolarmente uno per volta; nel modello del contesto componenziale, invece, ciascuna opzione è considerata nella sua globalità come un pacchetto di vantaggi e svantaggi, che si cerca di valutare adottando delle tecniche di compensazione, tenendo anche conto del principio di avversione alle perdite.

Le scelte dei consumatori sono spesso influenzate dalla scarsa disponibilità di tempo [Svenson e Maule, 1993]; in questi casi potrebbe essere particolarmente utile un modello dinamico come la teoria del campo decisionale. Una persona, per esempio, sia incerta fra una casa ottima per l'aspetto ritenuto più importante (per esempio l'ubicazione), ma insoddisfacente per diversi aspetti ritenuti meno importanti (intonaco, dimensione della cucina, esposizione, ecc.) e una casa molto insoddisfacente per l'aspetto ritenuto più importante ma ottima per diversi aspetti meno importanti. La teoria del campo decisionale ipotizza che l'attenzione sia rivolta prima agli aspetti più importanti e poi a quelli meno importanti; ciò significa che se il tempo disponibile per la decisione è molto breve è più probabile che venga scelta la casa migliore per l'aspetto più importante, se il tempo disponibile per la scelta è sufficientemente lungo è più probabile che venga scelta l'altra casa [Busemeyer e Jonson, 2004; Diederich, 2003].

La fretta di scegliere sembra ridurre la rilevanza degli effetti di contesto, mentre la possibilità di riflettere sembra potenziarli [Dhar et al., 2000]; ciò perché la riflessione, paradossalmente, di solito fa aumentare l'incertezza sulle preferenze. D'altronde bisogna considerare che la riflessione di solito è più frequente quando inizialmente è maggiore l'incertezza, per cui la persona può rivolgersi per consigli ad amici, parenti o esperti. Il processo di riflessione comporta di solito anche più accurate verifiche dei prezzi, e confronti dei diversi aspetti delle opzioni disponibili; è possibile quindi che in effetti l'incertezza aumenti invece di diminuire. La teoria del campo decisionale, per la sua attenzione agli aspetti temporali dei processi decisionali, sembra essere quella meglio in grado di descrivere accuratamente il processo di decisione meditata e la sua influenza sugli effetti di contesto.

Gli acquirenti, sia di beni di consumo, sia di beni strumentali per l'attività delle imprese, si trovano in questi anni spesso confusi dalle ampie possibilità di scelta disponibili, in termini sia di varietà dei beni sia di prezzi e altre condizioni di vendita. I venditori, da parte loro, da un lato hanno l'esigenza di vendere i prodotti effettivamente disponibili, sia di far sì che i loro clienti rimangano soddisfatti, in quanto soltanto in questo modo potranno contare in futuro sulla fedeltà di questi clienti.

L'aumento della concorrenza, anche tramite Internet, ha reso sempre più difficile per le imprese attrarre nuovi clienti e mantenere la fedeltà dei vecchi clienti. Internet, non soltanto ha introdotto nuove forme di concorrenza in mercati prima relativamente protetti, ma ha introdotto anche nuove modalità concorrenziali. Nel mercato dei computer, per esempio, Dell ha conquistato una posizione di leadership con un modello di produzione dei computer in base alle specificazioni richieste da ciascun cliente; i produttori di automobili, elettrodomestici e altri beni complessi hanno cercato di imitare in parte il modello di produzione e vendita sperimentato con successo da Dell, consistente essenzialmente in una sorta di "personalizzazione di massa" dei prodotti.

### **1.7 I conflitti decisionali nelle strategie aziendali e nelle strategie politiche**

Secondo la teoria del prospetto di Kahneman e Tversky le persone mostrano *avversione al rischio* quando devono scegliere tra una vincita sicura ed una che è solo probabile, anche quando quest'ultima ha un'utilità attesa maggiore di quella della vincita sicura. Per esempio le persone preferiscono vincere con certezza la somma di 900 euro piuttosto che accettare una scommessa che offre il 70% di probabilità di vincere 1500 euro e il 30% di probabilità di non vincere nulla.

Se le persone sono avverse al rischio nel caso di possibili vincite, esse sono invece propense all'assunzione di rischi nel caso di possibili perdite. Ciò vuol dire che davanti ad una scelta tra una perdita sicura di 900 euro ed una scommessa che offre la possibilità di perdere 1500 euro con il 70% di probabilità e di non perdere nulla con il 30% di probabilità, in media le persone accettano la scommessa. Le persone preferiscono assumersi dei rischi maggiori se devono recuperare una situazione negativa poiché rischiando si può cercare di ridurre o di evitare le perdite [Kahneman e Tversky, 1979].

Gli investitori a volte mostrano un'eccessiva fiducia nelle proprie abilità (*optimistic bias*); in molte attività le persone si giudicano più brave della media anche quando ciò non è vero; per esempio, la maggioranza degli automobilisti ritengono di essere guidatori migliori della media. Secondo una indagine, nei mercati finanziari l'eccessiva fiducia nelle proprie capacità porta ad ottenere risultati più scarsi; chi ha un più elevato livello di fiducia in sé stesso ottiene profitti inferiori. Inoltre gli investitori più sicuri di sé tendono a modificare di frequente il loro portafoglio, forse spinti dalla loro impressione di essere più capaci degli altri [Fenton O'Creevy, Nicholson, Soane, William, 1998].

Olsen ritiene che l'eccessiva sicurezza nelle proprie capacità sia una strategia di ragionamento che si è sviluppata per ragioni evolutive e di adattamento alla vita quotidiana. In primo luogo secondo questo autore ci sarebbe un'asimmetria tra emozioni positive e negative per cui la sicurezza in sé associata alla soppressione di emozioni negative permetterebbe di affrontare anche situazioni ambientali molto critiche mantenendo un buon livello di motivazione [Olsen, 2000].

Le persone inoltre, a volte agiscono cercando di confermare le loro ipotesi sull'andamento futuro delle borse. In realtà sarebbe molto meglio andare alla ricerca di informazioni che possono falsificare le ipotesi iniziali. Infatti solo in assenza di informazioni contrarie rispetto alle proprie opinioni si può ragionevolmente sostenere di avere ragione. Un esperimento condotto da Forsythe, Nelson, Neumann e Wright che ha utilizzato dei mercati sperimentali ha suggerito che gli investitori che utilizzano una strategia non confermativa ottengono i profitti migliori. Inoltre questi investitori hanno una minore tendenza al falso consenso, cioè a ritenere che la maggioranza delle persone abbia le loro stesse opinioni. Le strategie di tipo confermativo potrebbero essere meno utilizzate dagli investitori professionali poiché la loro esperienza potrebbe averli condotti a comprendere l'inadeguatezza di questo tipo di comportamento [Forsythe, Nelson, Neumann, Wright, 1992].

### **1.7.1 Distorsioni cognitive nelle strategie aziendali**

Roxburgh, responsabile della sede di Londra della McKinsey, ha cercato di analizzare in che modo le scoperte dell'economia comportamentale (*behavioral economics*) possono aiutare a capire perché spesso dei bravi manager possono adottare delle cattive strategie. Innanzitutto Roxburgh ricorda che dopo circa 40 anni la teoria sulle strategie delle imprese risulta essere ben sviluppata e largamente conosciuta e diffusa; il lavoro pionieristico di Porter, Mintzberg e altri ha originato un'ampia letteratura teorica su come le imprese possono adottare delle buone strategie. Gran parte dei manager delle imprese hanno ricevuto un'adeguata formazione strategica, e le grandi imprese hanno dei dipartimenti di strategia ben attrezzati. Ciononostante, l'esperienza concreta offre tanti esempi di errori strategici. Roxburgh si chiede come mai i manager adottano a volte delle cattive strategie, nonostante le vaste conoscenze strategiche disponibili. Alcune spiegazioni possibili possono essere analisi non corrette, eccessiva ambizione, avidità, e altri difetti dei manager; Roxburgh si sofferma tuttavia su una causa particolare: il funzionamento della mente umana.

La mente umana non è quella macchina razionale che spesso si immagina che sia; nel corso della sua evoluzione il cervello ha sviluppato "scorciatoie cognitive", semplificazioni, distorsioni, e cattive abitudini; il cervello può quindi a volte essere una guida ingannevole dal punto di vista dei processi decisionali razionali. La razionalità dei comportamenti delle persone,

postulato fondamentale dell'analisi economica, non sempre trova conferma nei comportamenti concreti. Le implicazioni di alcune inadeguatezze della mente umana sono state studiate in modo rigoroso da diversi ricercatori nell'ambito delle scienze cognitive, e in particolare nell'ambito dell'economia cognitiva o comportamentale; le loro ricerche hanno messo in evidenza diversi casi in cui il postulato di razionalità dell'analisi economica tradizionale appare in contrasto con le decisioni concrete delle persone [Roxburgh, 2003].

Alcune scoperte ottenute nell'ambito delle ricerche di economia cognitiva o comportamentale sono state utilizzate per cercare di spiegare gli errori spesso presenti nei processi decisionali delle imprese, e in particolare gli errori nelle decisioni di investimento. Alcune imprese, soprattutto quelle finanziarie, sono riuscite, grazie ai risultati delle ricerche di economia cognitiva o comportamentale, a rimodellare i loro processi decisionali per cercare di contrastare alcune distorsioni cognitive e decisionali evidenziate in queste ricerche.

Roxburgh ha messo in evidenza otto distorsioni cognitive e decisionali evidenziate dalle ricerche di economia comportamentale o cognitiva che possono spiegare alcuni errori strategici osservati nell'ambito del settore dei servizi finanziari in Europa, in particolare fra gli ultimi anni del secolo scorso e i primi anni duemila (*dot-com era*). Roxburgh sottolinea che gli errori strategici osservati in quel periodo non rappresentano delle eccezioni irripetibili, ma degli errori sistematici tipici del funzionamento della mente umana, che probabilmente si verificheranno ancora in futuro. Le otto distorsioni cognitive considerate da Roxburgh sono le seguenti:

- 1) eccesso di fiducia in se stessi (*overconfidence*);
- 2) calcolo mentale (*mental accounting*);
- 3) il bias dello "status quo" (*status quo bias*);
- 4) la distorsione dell'ancoraggio (*anchoring bias*);
- 5) considerare rilevanti costi sostenuti non più recuperabili (*sunk-cost bias*);
- 6) l'istinto di seguire la folla (*herding instinct*);
- 7) distorsioni nella stima di dolori e piaceri futuri (*misestimating future hedonic states*);
- 8) la distorsione del falso consenso (*false consensus*).

Per una più approfondita analisi delle distorsioni elencate precedentemente, si consiglia la lettura del testo di Roxburgh [Roxburgh, 2003].

In sintesi, la consapevolezza riguardo i punti di debolezza della mente umana può aiutare i responsabili strategici a cercare di evitarne le conseguenze negative più rilevanti; tutti i responsabili strategici dovrebbero conoscere le intuizioni fondamentali dell'economia cognitiva o comportamentale altrettanto bene di quelle degli altri campi dell'economia. Queste conoscenze non sono sufficienti tuttavia a evitare gli errori strategici; avidità, arroganza, e negligenza continueranno a causarne ancora; una buona comprensione delle più frequenti distorsioni cognitive può tuttavia aiutare a diminuire la probabilità che dei bravi manager sostengano delle cattive strategie.

### **1.7.2 Anomalie cognitive nelle strategie politiche: il "paternalismo progressista"**

Richard Thaler, economista presso l'università di Chicago pioniere dell'economia comportamentale e in particolare della finanza comportamentale, candidato, fino al 2008 senza successo, al premio Nobel per l'Economia, e Cass R. Sunstein<sup>14</sup>, celebre giurista, hanno presentato nel libro "*Nudge. Improving Decisions About Health, Wealth, and Happiness*"<sup>15</sup>, una

<sup>14</sup> Cass Sunstein é probabilmente il più citato docente di diritto del mondo; dall'autunno del 2008 è passato a insegnare presso la Harvard Law School, dove dirige il progetto sulla regolamentazione dei rischi.

<sup>15</sup> Richard H. Thaler and Cass Sunstein *Nudge: Improving Decisions About Health, Wealth, and Happiness*, Yale University Press, 2008. In sintesi l'approccio era stato già presentato nel 2003 in un intervento alla riunione annuale dell'American Economic Association pubblicato sull'American Economic Review: R. H. Thaler and C. R. Sunstein. 2003. Libertarian Paternalism, American Economic Review (AEA Papers and Proceedings, 2003, 94 (2), pp. 175-79.

sorta di manifesto del “paternalismo progressista”. Thaler e Sunstein si sono autodefiniti “consiglieri informali” della campagna di Obama per la presidenza degli Stati Uniti. I principali obiettivi sono stati illustrati da Richard Thaler in una intervista su L’Espresso del 4 settembre 2008 (pp. 124-125). L’idea di fondo è cercare di utilizzare le scoperte sulle distorsioni cognitive per individuare dei meccanismi in grado di stimolare le persone a prendere, magari inconsapevolmente, delle decisioni in grado di migliorare la loro salute, il loro benessere economico, la loro felicità.

Il tentativo si basa sul convincimento che non è vero che le persone sono in grado di pensare come degli esperti di logica, capaci di immagazzinare ed elaborare con saggezza quantità enormi di informazioni. Secondo Thaler “.. l’individuo medio .... È un essere pigro, impulsivo, troppo occupato per prestare attenzione ai particolari, e tende ad esprimere preconcetti che sono difficilmente sradicabili.” Thaler quindi propone “Un paternalismo progressista, che non impone alla gente di comportarsi in una certa maniera, ma la incoraggia. Bisogna catturare la fantasia delle persone e rendere il comportamento automatico. In America il dibattito economico si divide ancora tra quelli che sostengono il liberismo e quelli che invece propongono l’intervento governativo. Dovremmo superare questo tipo di dibattito, dovremo capire che si può scegliere la via del mercato e, allo stesso tempo, si possono spingere le persone a prendere decisioni positive”. (L’Espresso, 4 settembre 2008, p. 125)

Gli esempi proposti da Richard Thaler nella sua intervista sono i seguenti:

- l’idea di alcune organizzazioni *non profit* di stimolare riduzioni nel consumo di energia elettrica mediante l’installazione di “lampadine incandescenti”, vale a dire di lampadine che diventano sempre più incandescenti man mano che aumenta l’energia elettrica consumata;
- il tentativo di alcune scuole di stimolare gli alunni consumare più frutta esponendola nei distributori automatici sui ripiani più vicini ai loro occhi;
- la trovata, sperimentata con successo nell’aeroporto olandese di *Schiphol*, di aumentare la pulizia dei bagni disegnando l’immagine di una mosca al centro degli orinai per i maschi;
- la legge dello Stato di New York che impone ai ristoranti di pubblicare il contenuto calorico delle pietanze, nel ambito della battaglia contro l’obesità. Secondo Thaler questa legge è utile sia perché informa in modo più esplicito i consumatori sui possibili effetti negativi di alcune pietanze, sia perché stimola le imprese a diminuire il contenuto calorico di alcuni prodotti;
- un sito web dell’Illinois che cerca di incoraggiare le persone a donare gli organi, chiedendo ai donatori di invitare anche i loro amici a donare gli organi; in questo modo si cerca di sfruttare a fin di bene la tendenza al conformismo delle persone.

## 1.8 Conclusioni

Il problema di una eccessiva varietà di scelte ha cominciato ad apparire sempre più rilevante a partire da primi anni duemila, soprattutto in connessione con lo sviluppo degli acquisti su Internet; mentre nei negozi tradizionali esistono dei vincoli fisici che limitano il numero delle varietà di ciascun bene che possono essere mostrate ai potenziali acquirenti, nei siti di vendite online non esistono praticamente limiti, e possono essere rese disponibili quindi migliaia di varietà per ciascun bene. Una possibile reazione dei consumatori al sovraccarico informativo può essere il differimento della decisione di acquisto, per il timore di non riuscire ad assumere la decisione migliore; ciò significa che i consumatori dal canto loro trovano difficoltà a rinunciare a considerare una parte rilevante delle informazioni disponibili. Le principali “anomalie” cognitive, e quindi decisionali, legate al funzionamento della memoria umana sono, secondo McFadden le seguenti: l’intensità delle emozioni diminuisce nella memoria con il passar del tempo, le informazioni più recenti e significative sono quelle più presenti in memoria in ciascun istante, la ricostruzione in memoria di informazioni passate è influenzata dal contesto e dagli stimoli attuali, le coincidenze sono memorizzate in modo più forte e persistente, la nozione soggettiva del tempo tende a comprimere nella memoria i periodi lontani nel tempo. Secondo

Misuraca, Fasolo e Cardaci, inoltre, le difficoltà nelle decisioni di acquisto possono dipendere da fattori interni e da fattori esterni, rispetto ai soggetti che devono prendere la decisione. I fattori esterni, oltre al numero di opzioni, sono il numero di attributi che caratterizzano le diverse opzioni, il loro livello di dettaglio, le interrelazioni fra gli attributi. I fattori interni sono costituiti invece dallo stile decisionale di ciascuna persona, e dagli indicatori dell'importanza che per ciascuna persona hanno i diversi attributi.

Alcune aziende, tra le quali Dell, hanno cercato di ovviare all'*information overload*, conquistando una posizione di leadership con un modello di produzione dei computer in base alle specificazioni richieste da ciascun cliente; i produttori di automobili, elettrodomestici e altri beni complessi stanno cercando di imitare in parte il modello di produzione e vendita sperimentato con successo da Dell, consistente essenzialmente in una sorta di "personalizzazione di massa" dei prodotti.

Ma la vera rivoluzione, è rappresentata dalla diffusione di agenti intelligenti, in grado di apprendere da modifiche dell'ambiente circostante. Nel prossimo capitolo tratteremo infatti di intelligenza artificiale, un settore in pieno sviluppo, soprattutto in ambienti online dove tutte le informazioni sono a portata di mouse. In particolar modo nelle vendite online, gli agenti intelligenti sono utilissimi per aiutare il consumatore ad effettuare scelte di acquisto in grado di soddisfare le sue esigenze al prezzo più conveniente.

## Capitolo 2

### L'Intelligenza artificiale

#### 2.1 Introduzione

L'intelligenza artificiale rappresenta uno dei campi di ricerca più affascinanti e innovativi, dal punto di vista sia scientifico sia applicativo. L'obiettivo dell'intelligenza artificiale è quello di sistematizzare e rendere automatiche alcune attività intellettive; essa può quindi essere applicata a tutti gli ambiti del pensiero umano. Il principale tema unificante delle ricerche di intelligenza artificiale è l'idea di agente intelligente, vale a dire di agenti che ricevono percezioni dall'ambiente e compiono delle azioni. L'interazione fra ricerca teorica e applicazioni concrete è concretamente evidenziata dal fatto che il libro di testo sull'intelligenza artificiale attualmente più diffuso nel mondo è stato scritto da un docente universitario di informatica dell'Università di Berkeley, Stuart Russell, e dal responsabile della *search quality* di Google, l'impresa di maggior successo del mondo negli ultimi dieci anni, Peter Norvig. Il grande fascino intellettuale dell'intelligenza artificiale è evidenziato dal fatto che, secondo Russell e Norvig, l'intelligenza artificiale è citata, insieme alla biologia molecolare, come “*il campo in cui vorrei maggiormente lavorare*”, da parte di gran parte degli scienziati di altre discipline [Russell, Norvig, 2005, vol. 1]. I processi decisionali analizzati nel precedente capitolo, sono strettamente legati agli sviluppi compiuti nel settore dell'intelligenza artificiale; gli agenti intelligenti saranno indispensabili nei prossimi anni per aiutare le persone a gestire meglio il sovraccarico informativo, in particolar modo negli ambienti online; uno dei settori web più dinamici è quello degli acquisti online.

#### 2.2 Nozioni di “intelligenza artificiale”.

Russell e Norvig presentano otto diverse definizioni di “intelligenza artificiale” fornite in otto libri di testo, classificandole in quattro coppie, ottenute incrociando la distinzione fra sistemi razionali e sistemi umani, con quella fra sistemi che pensano e sistemi che agiscono (Figura 2.1).

L'approccio all'intelligenza artificiale che cerca di sviluppare agenti artificiali in grado di *pensare come degli esseri umani* è stato perseguito principalmente nell'ambito delle scienze cognitive, associando modelli per computer sviluppati dall'intelligenza artificiale e tecniche di sperimentazione psicologica, nel tentativo di costruire teorie verificabili sul funzionamento della mente umana. Un esempio importante di lavoro nell'ambito di questo approccio è il General Problem Solver (GPS) sviluppato da Allen Newell e Herbert Simon nel 1961; essi infatti non si limitarono a scrivere un programma in grado di risolvere correttamente i problemi, ma confrontarono le procedure risolutive del loro programma con quelle della mente umana. L'interazione fra scienze cognitive e intelligenza artificiale ha consentito progressi importanti soprattutto nelle aree della visione e del linguaggio naturale.

Un geniale ricercatore che fornì contributi importantissimi allo sviluppo dell'intelligenza artificiale fu Alan Turing, un grande logico inglese, la cui principale occupazione durante la seconda guerra mondiale riguardò le tecniche per individuare dei sistemi di decriptazione dei messaggi in codice dei tedeschi<sup>16</sup>. L'esempio più famoso dell'approccio all'intelligenza artificiale che cerca di sviluppare agenti artificiali in grado di *agire come degli esseri umani* è il test proposto da Alan Turing nel 1950: un agente artificiale, per esempio un computer, può essere considerato intelligente se un esaminatore umano che rivolge delle domande in forma

---

<sup>16</sup> Turing inventò un congegno elettromeccanico denominato *bombe*, in grado di decifrare i messaggi tedeschi cifrati mediante il sistema *enigma*.

scritta all'agente non è in grado di distinguere se le risposte provengono da una persona o da un agente artificiale. Per poter superare questo test d'intelligenza l'agente artificiale dovrebbe essere in grado di interpretare il linguaggio umano, rappresentare la conoscenza, ragionare in modo automatico, apprendere.

Figura 2.1 – Le otto diverse definizioni di “intelligenza artificiale” presentate da Russel e Norvig

Nozioni di intelligenza artificiale		
Criteri	Umanità	Razionalità
Pensare	Sistemi che <u>pensano</u> come esseri umani: - "L'eccitante, nuovo tentativo di far sì che i computer arrivino a pensare ... <i>macchine dotate di mente</i> , nel pieno senso della parola. (Haugeland, 1985). - "[L'automazione delle] attività che associamo al pensiero umano, come il processo decisionale, la risoluzione di problemi, l'apprendimento.."(Bellman, 1978).	Sistemi che <u>pensano</u> razionalmente: "Lo studio delle facoltà mentali, attraverso l'uso di modelli computazionali." (Charniak e McDermont,1985). - "Lo studio dei processi di calcolo che rendono possibile percepire, ragionare e agire." (Winston, 1992)
Agire	Sistemi che <u>agiscono</u> come esseri umani: - "L'arte di creare macchine che eseguono attività che richiedono intelligenza quando vengono svolte da persone" (Kurzweil, 1990). - "Lo studio di come fare eseguire ai computer le attività in cui, al momento, le persone sono più brave." (Rich e Knight, 1991).	Sistemi che <u>agiscono</u> razionalmente: - "L'intelligenza computazionale è lo studio della progettazione di agenti intelligenti." (Poole e altri, 1998) - "L'intelligenza artificiale .. riguarda il comportamento intelligente negli artefatti." (Nilsson, 1998)
Fonte: Russel e Norvig, 2005, volume 1, pag. 2.		

L'approccio che cerca di sviluppare agenti artificiali in grado di *pensare razionalmente* si ispira agli studi di Aristotele intorno al 400 a. C. sui processi di ragionamento razionalmente corretti (*irrefutabili*). L'esempio più famoso è il sillogismo, che fornisce un schema di ragionamento che consente di ottenere delle deduzioni corrette da alcune premesse. Per esempio: premesse: a) Socrate è un uomo, b) tutti gli uomini sono mortali; deduzione: Socrate è mortale. Se le premesse sono vere, anche la deduzione è vera.

La disciplina che studia le “*leggi del pensiero razionale*”, vale a dire i meccanismi di ragionamento razionalmente corretti è la logica. Sviluppi importanti nell'analisi delle “leggi del pensiero razionale” si sono avuti nel corso del diciannovesimo secolo, con lo sviluppo di sistemi di notazioni logiche molto più potenti dell'aritmetica, che consentono di formulare proposizioni riguardanti tutti gli aspetti del mondo e le loro interrelazioni. In particolare, George Boole formulò nel 1847 i principi della logica proposizionale o booleana, e Gottlob Frege estese nel 1879 la logica booleana in modo da includere oggetti e relazioni, creando la logica del primo ordine, utilizzata attualmente come il miglior sistema di rappresentazione della conoscenza. I principali limiti di questo approccio sono due: 1) non sempre è possibile esprimere le conoscenze concrete nei termini precisi delle notazioni logiche formali; 2) la soluzione di molti problemi concreti applicando l'approccio astrattamente razionale richiederebbe risorse computazionali enormi, molto più grandi di quelle attualmente disponibili, nonostante l'eccezionale sviluppo negli ultimi decenni delle potenzialità dei computer.

L'approccio degli agenti artificiali che *agiscono razionalmente* può essere considerato uno sviluppo dell'approccio degli agenti che *pensano razionalmente*. Il termine "agente" da l'idea di qualcosa che compie delle azioni; un "agente computerizzato" dovrebbe quindi essere qualcosa di più di un semplice programma, in quanto in grado, per esempio, di percepire alcune caratteristiche ambientali, adattarsi ai cambiamenti, scambiarsi degli obiettivi con altri agenti. Un agente razionale dovrebbe comportarsi in modo da ottenere il migliore risultato possibile (il miglior risultato atteso in condizioni di incertezza), in rapporto, ovviamente, agli obiettivi perseguiti. Nell'approccio all'intelligenza artificiale basato sulle leggi del pensiero razionale l'attenzione è concentrata sulla correttezza delle inferenze; naturalmente per agire razionalmente è necessario essere in grado anche di pensare razionalmente, così da capire quali azioni possono consentire di raggiungere meglio gli obiettivi perseguiti. Tuttavia, Russel e Norvig avvertono che la razionalità perfetta dei comportamenti spesso non è concretamente raggiungibile, anche per l'insufficienza delle risorse computazionali; è quindi opportuno individuare dei meccanismi comportamentali di razionalità limitata, utilizzabili quando non ci sono le condizioni per attuare dei comportamenti perfettamente razionali [Russel, Norvig, 2005, vol. 1].

Russel e Norvig mettono inoltre in evidenza i contributi che diverse discipline hanno fornito alla nascita dell'intelligenza artificiale. In primo luogo la scienza più antica, la filosofia, che con Aristotele formulò intorno al 400 a. C. le prime leggi del pensiero razionale, in grado di consentire di ottenere meccanicamente deduzioni dalle premesse; successivamente intorno al 1300 Raimondo Lullo ideò un meccanismo in grado di eseguire dei ragionamenti corretti (le "ruote della logica"); intorno al 1550 Leonardo da Vinci progettò, ma non costruì, un calcolatore meccanico. Le prime calcolatrici furono effettivamente costruite nella prima metà del '600, prima da uno scienziato tedesco e poi, nel 1642, dal grande matematico e filosofo francese Blaise Pascal (la "pascalina"); Pascal scrisse che la "macchina aritmetica produce effetti che sembrano più vicini al pensiero di tutte le azioni degli animali". Come anticipato precedentemente in Aristotele, si trova quindi il fondamento logico dell'approccio degli agenti che operano razionalmente, vale a dire il collegamento fra conoscenza e azione; Aristotele evidenziò infatti che le azioni sono giustificate da un collegamento logico fra gli obiettivi dell'agente e la conoscenza dei risultati delle diverse azioni. Da questa intuizione che risale ad Aristotele partirono Newell e Simon nell'impostare negli anni cinquanta il General Problem Solver (GPS).

La matematica ha contribuito allo sviluppo dell'intelligenza artificiale fornendo strumenti di formalizzazione dei concetti elaborati dai filosofi in tre aree fondamentali: la logica, la computazione, e la probabilità. Nel campo della logica i principali contributi furono forniti da George Boole che nel 1847 formulò i principi della logica proposizionale, e da Gottlob Frege che nel 1879 estese la logica proposizionale in modo da includere oggetti e relazioni, creando la logica del primo ordine, considerata attualmente come il più efficace sistema di rappresentazione della realtà.

Per quel che riguarda la computazione, intorno alla metà del diciannovesimo secolo Boole e altri contribuirono allo sviluppo di algoritmi utilizzabili per ottenere delle deduzioni logiche, nel 1900 Hilbert formulò il "problema della decisione", vale a dire il problema di sapere se esistono dei limiti invalicabili alla potenza delle procedure di dimostrazione, Nel 1930 Gödel dimostrò il "teorema dell'incompletezza", secondo cui all'interno di qualsiasi linguaggio abbastanza potente da descrivere le proprietà dei numeri naturali esistono proposizioni vere che sono "in decidibili", nel senso che non possono essere dimostrate da nessun algoritmo; ciò significa che esistono funzioni sui numeri interi che non possono essere rappresentate da un algoritmo, e quindi non sono computabili. Nel 1936 Turing ideò un meccanismo in grado di calcolare tutte le funzioni computabili ("macchina di Turing"), e mostrò che esistono delle funzioni che nessuna macchina può calcolare, per esempio, dire se in generale un certo programma, ricevuto un input, fornirà un risultato, oppure continuerà l'esecuzione per sempre. Un altro problema computazionale concretamente rilevante è quello della "intrattabilità", evidenziato a metà degli anni sessanta in rapporto alla distinzione fra crescita polinomiale e

crescita esponenziale della complessità al crescere della dimensione. I problemi caratterizzati da una crescita esponenziale della complessità al crescere della dimensione possono risultare intrattabili nel senso che esemplificazioni anche moderatamente complesse non possono essere risolte in tempi ragionevoli; in questi casi è necessario cercare di suddividere il problema generale in sottoproblemi trattabili. Questi risultati mettono in evidenza l'esigenza di fare un uso molto accurato delle risorse computazionali disponibili, nonostante l'aumento straordinario negli ultimi decenni della potenza dei computer.

Nel campo della probabilità, uno dei contributi più significativi e concretamente importanti per le applicazioni dell'intelligenza artificiale è il teorema di Bayes, formulato intorno alla metà del '700 dal reverendo inglese Thomas Bayes, e riscoperto in modo indipendente alcuni anni dopo dal grande matematico francese Blaise Pascal; il teorema di Bayes consente di aggiornare gradualmente la stima della probabilità di un evento man mano che si ottengono informazioni aggiuntive.

Il principale contributo della scienza economica allo sviluppo dell'intelligenza artificiale riguarda probabilmente l'analisi delle decisioni razionali nel senso di massimizzare l'utilità o la felicità delle persone, questa analisi fu impostata da Walras nella seconda metà del diciannovesimo secolo con riferimento a sistemi economici con tantissimi agenti, e da von Neumann e Morgenstern negli anni quaranta del ventesimo secolo per sistemi economici con pochi agenti in cui ciascuno deve tener conto delle presumibili reazioni degli altri (teoria dei giochi); Von Neumann e Morgenstern impostarono anche l'analisi delle decisioni razionalmente ottimali in condizioni d'incertezza. Paradossalmente, Herbert Simon, uno dei pionieri dell'intelligenza artificiale, ottenne il premio Nobel per l'economia nel 1978, allontanandosi dall'impostazione astrattamente massimizzante della teoria economica ortodossa per analizzare i processi decisionali concreti volti in genere a individuare decisioni "soddisfacenti" (razionalità limitata).

Le neuroscienze, invece, analizzando i processi di elaborazione delle informazioni all'interno del cervello, cercano di fornire una risposta a uno dei più grandi e affascinanti misteri della scienza sull'origine del pensiero: com'è che un semplice insieme di cellule può condurre al pensiero, alla consapevolezza, all'azione? Nella seconda metà del diciannovesimo secolo si scoprì che il cervello è costituito da cellule nervose interconnesse mediante punti di congiunzione chiamati *sinapsi* (neuroni); gli sviluppi della risonanza magnetica funzionale negli anni novanta del secolo scorso hanno consentito di monitorare le attività cerebrali durante i processi cognitivi. Lo sviluppo dei computer (*cervelli elettronici*) ha stimolato confronti fra il loro funzionamento e quello del cervello umano. Attualmente il numero di neuroni in un tipico cervello umano supera ancora di molto il numero di porte logiche dei computer più potenti, ma secondo la legge di Moore i due valori si eguaglieranno intorno al 2020. Inoltre, la velocità di elaborazione dei chip dei computer è circa un milione di volte più grande dei neuroni; in complesso però il cervello umano risulta ancora circa 50 mila volte più veloce di un tipico computer nell'esecuzione delle attività, perché mentre i computer possono utilizzare un piccolo numero di processori (di solito soltanto uno), tutti i neuroni del cervello sono attivi contemporaneamente (mentre il computer elabora le informazioni in sequenza, il cervello le elabora in parallelo).

Particolarmente intense sono le interrelazioni fra intelligenza artificiale e psicologia, la scienza della psiche, dell'anima, della mente, dell'intelligenza, del pensiero, delle emozioni. Secondo Russel e Norvig la visione del cervello come un dispositivo per l'elaborazione di informazioni può farsi risalire al lavoro di William James, professore di psicologia nell'Università di Harvard, autore nel 1890 dei "*Principi di psicologia*", uno dei testi più rilevanti nella storia della psicologia; fondatore presso l'Università di Harvard di uno dei primi laboratori di psicologia sperimentale degli Stati Uniti [Russel, Norvig, 2005, vol. 1, pag. 20]. Negli Stati Uniti gli orientamenti cognitivisti furono per molto tempo oscurati dal dominante indirizzo comportamentista, ma in Europa, e in particolare nell'unità di ricerca di psicologia

applicata dell'Università di Cambridge, le intuizioni cognitive continuarono ad essere sviluppate nel corso della prima metà del novecento, per merito di Fredric Bartlett, Kennet Craig, e Donal Broadbent. In particolare, Craik individuò i tre requisiti fondamentali di un agente basato sulla conoscenza: 1) lo stimolo deve essere tradotto in una rappresentazione interna; 2) la rappresentazione deve essere manipolata da processi cognitivi per ottenere nuove rappresentazioni interne; 3) le nuove rappresentazioni interne devono essere trasformate in azione. Se un agente ha in testa un modello della realtà esterna e delle proprie possibili azioni sarà in grado di provare diverse alternative, decidere quali di esse sia la migliore, reagire a situazioni future prima che si manifestino, utilizzando la conoscenza di eventi passati per gestire gli eventi presenti e futuri [Craig, 1943]. Il lavoro di Craig fu proseguito da Donald Broadbent, nominato direttore dell'unità di ricerca di psicologia applicata dell'Università di Cambridge nel 1958, che nel libro *"Percezione e comunicazione"* del 1958, considerato un classico della psicologia cognitiva, illustrò alcuni dei primi modelli di fenomeni psicologici fondati sulla elaborazione di informazioni. Negli Stati Uniti lo sviluppo di modelli cognitivi basati su computer stimolò la nascita della scienza cognitiva. La nascita di quest'area di ricerca interdisciplinare può farsi risalire a un workshop svoltosi nel settembre del 1956 al Massachusetts Institute of technology (MIT), due mesi dopo del workshop che aveva segnato a Dartmouth la nascita dell'intelligenza artificiale. Al workshop del MIT George Miller presentò "Il magico numero sette"<sup>17</sup>, Noam Chomsky "Tre modelli per la decisione del linguaggio", Allen Newel ed Herbert Simon "La macchina Logic Theory". Quei lavori mostrarono magistralmente come dei modelli basati sui computer potevano essere usati per discutere la psicologia della memoria, la psicologia del linguaggio, e la psicologia del pensiero logico. Negli anni successivi l'idea che una teoria cognitiva dovesse essere come un programma per computer, vale a dire descrivere dettagliatamente un meccanismo di elaborazione dell'informazione che potrebbe implementare una qualche funzione cognitiva, divenne largamente condivisa.

Altro aspetto caratteristico della psicologia cognitiva è che si interessa di processi cognitivi (percezione, attenzione, memoria, linguaggio, pensiero, creatività), trascurati dai comportamentisti; inoltre, la mente è concepita come un elaboratore di informazione. Nei primi modelli cognitivistici, l'elaborazione dell'informazione era concepita come un processo che avviene per stadi consecutivi, terminate le operazioni proprie di uno stadio si passa al successivo, e così via; negli anni '70 furono presentati nuovi modelli che mettevano in evidenza sia la possibilità di retroazione di uno stadio di elaborazione su quelli precedenti, sia la possibilità che si attivassero le operazioni di uno stadio successivo senza che quelli precedenti avessero già elaborato del tutto le informazioni. Un altro aspetto importante fu l'accentuazione del carattere finalizzato dei processi mentali; il comportamento viene concepito come una serie di atti guidati dai processi cognitivi al fine della soluzione di un problema, con continui aggiustamenti per garantire la migliore soluzione. La nozione di *"feedback"* (retroazione), sviluppata dalla cibernetica divenne centrale in questa concezione del comportamento orientato verso una meta; lo psicologo sperimentale del linguaggio George Miller con le sue opere rappresentò un'autentica svolta nella rappresentazione del comportamento: il comportamento era visto come il prodotto di una elaborazione dell'informazione, com'è compiuta da un calcolatore, per lo svolgimento di un piano utile alla soluzione del problema. Il comportamento è visto come il risultato di un processo di continua verifica retroattiva del piano di comportamento secondo l'unità TOTE (*test, operate, test, exit*): l'atto finale (*exit*) non consegue direttamente ad un input sensoriale o a un comando

---

<sup>17</sup> In quel lavoro, poi pubblicato nell'articolo *"The magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information"* (Psychological Review, n. 63, pp. 81-97) lo psicologo cognitivista George A. Miller evidenziò nel 1956 che la capacità di memoria delle persone giovani è pari a circa sette elementi (cifre, lettere, parole, o altre unità); ricerche successive hanno sostanzialmente confermato l'intuizione di Miller, precisando però che la capacità di memoria è circa sette per le cifre, ma circa sei per le lettere e cinque per le parole, e che essa è minore per le parole lunghe che per quelle brevi.

motorio, ma è il risultato di precedenti operazioni di verifica (*test*) delle condizioni ambientali, di esecuzione (*operate*) intermedie e di nuove verifiche (*test*). Nel 1967 uscì il libro dello psicologo statunitense Ulric Neisser, “psicologia cognitiva”, nel quale venivano sintetizzate le ricerche condotte nei 10 anni precedenti secondo la prospettiva che fu definitivamente chiamata cognitivista. La letteratura sperimentale sui processi cognitivi aumentò rapidamente, sostituendo le prospettive passate con la nuova prospettiva che si diffuse anche in campo della psicologia sociale e della psicopatologia; nei primi anni '70 si parlava addirittura di una rivoluzione cognitivista nella ricerca psicologica.

Una disciplina che per diversi aspetti è simile all'intelligenza artificiale è la cibernetica, scienza interdisciplinare che studia il funzionamento di qualsiasi sistema dinamico. Nel 1947 Norbert Wiener, un matematico americano che aveva lavorato anche con il premio Nobel Bertrand Russell prima di sviluppare interesse per i sistemi di controllo e le loro relazioni con i processi cognitivi, definì la cibernetica come lo studio del controllo e della comunicazione nell'animale e nella macchina. Ciò riflette la convinzione che molti fenomeni di cui si occupano varie scienze, dalla biologia all'economia, si fondano sugli stessi principi di funzionamento che governano tutti i sistemi, al di là della dicotomia fra sistemi naturali e sistemi artificiali. Questi principi riguardano essenzialmente il controllo, mediante l'elaborazione e la trasmissione di informazioni, di una parte del sistema su un'altra, al fine di guidarne l'andamento verso gli obiettivi stabiliti. Le macchine di cui si occupa la cibernetica sono in particolare quelle in cui è presente una forma di controllo e correzione delle proprie operazioni durante la loro stessa esecuzione (retroazione o *feedback*).

La psicologia entra in relazione con la cibernetica nel caso dei sistemi psicofisici di ricezione, analisi, ed elaborazione delle informazioni. Si tratta in particolare dei processi cognitivi (percezione, linguaggio, apprendimento, formazione di concetti, ecc.), che comportano continui scambi di informazioni fra gli organismi e l'ambiente. Le applicazioni cibernetiche in psicologia si fondano principalmente sulle teorie dell'informazione, della regolazione e degli automatismi. La teoria dell'informazione è stata applicata in psicologia sperimentale nelle indagini sui processi di riconoscimento dei segnali, sulle risposte selettive agli stimoli, sulla trasmissione e ricezione del linguaggio. Per la teoria della regolazione, all'interno dei sistemi complessi è presente un sistema di controllo e regolazione del comportamento, per cui, quando intendiamo effettuare un movimento secondo un determinato modello, la differenza fra questo modello e l'effettivo svolgersi del movimento viene adoperata come un nuovo segnale che determina una regolazione del movimento stesso tale da mantenerlo quanto più possibile vicino a quello dato dal modello. La teoria degli automatismi si basa sul fatto che i sistemi di elaborazione delle informazioni possono venire rappresentati mediante circuiti trasferibili in apparati meccanici detti *automi*, che effettuano delle operazioni proprie degli esseri viventi. L'applicazione di questa teoria alla psicologia riguarda principalmente la simulazione nei computer dei processi cognitivi della percezione, del pensiero, della soluzione dei problemi, al fine di analizzare le strategie di elaborazione delle informazioni e dell'apprendimento, considerato che anche l'apprendimento si basa in parte su automatismi. La teoria degli automatismi è utilizzata nelle applicazioni di intelligenza artificiale, in base all'ipotesi di Turing secondo cui i circuiti utilizzati nelle macchine elettroniche possono possedere proprietà analoghe a quelle del sistema nervoso, e quindi possono essere in grado di trasmettere l'informazione e di memorizzarla, rivelandosi così capaci di simulare molte attività del pensiero umano, comprese quelle che richiedono degli adattamenti alla complessità dell'ambiente.

La prima applicazione concreta di “intelligenza artificiale”, intesa come svolgimento automatico di funzioni tipiche della mente umana può farsi risalire, come anticipato precedentemente, alla “*pascalina*”, la calcolatrice ideata e costruita nel 1642 dal grande filosofo e matematico francese Blaise Pascal, per fornire un aiuto al padre nel suo lavoro di esattore delle tasse. Un contributo importantissimo allo sviluppo dell'intelligenza artificiale fu fornito quasi

due secoli dopo dall'ingegnere inglese Charles Babbage e da Ada Lovelace Byron<sup>18</sup>; fra il 1822 e il 1870 Babbage ideò delle macchine in grado di calcolare automaticamente delle funzioni polinomiali attraverso il metodo delle differenze finite (*difference engine*), e poi un vero e proprio computer meccanico funzionante con schede perforate (*analytical engine*). Ada Lovelace scrisse il primo programma per computer con riferimento al progetto di *analytical engine*, ma Babbage non riuscì a completare la costruzione del calcolatore che aveva ideato; la *difference engine* fu costruita in Inghilterra nel 1991 dal *London science museum*, in occasione del bicentenario della nascita di Babbage, evidenziando una precisione di calcolo fino a oltre 30 cifre decimali<sup>19</sup>.

Secondo Nilsson, l'Intelligenza Artificiale si occupa del comportamento intelligente in sistemi artificiali. Quindi, tale comportamento, a sua volta, coinvolge capacità quali ragionamento, apprendimento, comunicazione e azione in ambienti complessi. L'intelligenza artificiale ha fra i suoi obiettivi a lungo termine lo sviluppo di macchine che possano fare queste cose come, o addirittura meglio degli esseri umani, nonché la comprensione di questo tipo di comportamento sia che esso si presenti in macchine sia in esseri umani o in altri animali.” [Nilsson, 2002]

Riportiamo una citazione particolarmente ispirata di James Albus sulla comprensione dell'intelligenza: “*Credo che comprendere l'intelligenza significhi comprendere l'abilità di acquisire, rappresentare e memorizzare la conoscenza, di generare e apprendere il comportamento intelligente, di sviluppare e usare stimoli, emozioni e priorità, di trasformare i segnali sensoriali in simboli, di manipolare i simboli per agire secondo logica, di ragionare sul passato ed elaborare piani per il futuro; e che significhi anche comprendere come i meccanismi dell'intelligenza producano i fenomeni dell'illusione, del credere, della speranza, della paura e dei sogni e perfino della bontà e dell'amore*”. E quindi una citazione di Russel e Wefald che sottolineano la capacità di apprendere, come uno dei principali segni distintivi di un sistema intelligente: “*L'apprendimento è una parte importante dell'autonomia. Un sistema si può considerare autonomo fintantoché il suo comportamento è determinato dai suoi ingressi immediati e dalle esperienze passate piuttosto che dal suo progettista. .... Le strutture di conoscenza interna al sistema dovrebbero quindi essere costruibili, in principio, partendo dalla sua esperienza del mondo.*” [Russel, Wefald, 1991].

### 2.2.1 Principali campi d'indagine dell'intelligenza artificiale

L'interazione con sistemi automatici di elaborazione delle informazioni coinvolge oggi tramite internet una parte significativa della popolazione in tanti paesi; in questo ambito agenti intelligenti artificiali possono muoversi e interagire, anche con esseri umani<sup>20</sup>. Molte ricerche, svolte sia nelle università che in imprese, pubbliche e private, nel settore della *new economy* e del commercio elettronico, hanno evidenziato la grande rilevanza di questo fenomeno. La ricerca teorica ha cercato di dare all'intelligenza artificiale un apparato formale di riferimento che possa permettere la progettazione di sistemi intelligenti artificiali sempre più sofisticati. I principali campi di indagine sono lo studio di logiche specifiche per formalizzare e analizzare il

---

<sup>18</sup> Figlia del grande poeta inglese George Byron, che abbandonò la madre dopo la nascita di Ada, deluso, forse, per non avere avuto un figlio maschio.

<sup>19</sup> La persistente attualità dei progetti di Babbage fu messa in evidenza nel 1999 in un articolo su *the Economist*, secondo cui i progressi nel campo delle nanotecnologie potrebbero consentire di costruire oggi dei calcolatori meccanici del tipo di quelli ideati da Babbage che sarebbero più affidabili dei moderni calcolatori elettronici in condizioni di alte temperature o di forti radiazioni. In riconoscimento del contributo di Ada Lovelace, nel 1979 il Dipartimento della difesa degli Stati Uniti diede il nome “Ada” a un linguaggio di programmazione per computer, per alcuni aspetti simile ai linguaggi C e C++.

<sup>20</sup> Secondo Tapscott e Williams (2007, p. 41), “.. la capacità di riunire la conoscenza di milioni (per non dire miliardi) di utenti autonomamente organizzati dimostra che la collaborazione di massa sta trasformando la nuova rete in un cervello globale.”

ragionamento automatico, i modelli ispirati a sistemi biologici, come le reti neurali, l'apprendimento automatico, le tecniche decisionali in situazioni in cui le informazioni sono incomplete o incerte, la percezione e la rappresentazione del mondo esterno. Un interessante campo di indagine riguarda l'integrazione di tecniche diverse in architetture di agenti intelligenti, in grado di svolgere compiti specifici sia nel mondo fisico (robot) che nel mondo virtuale di internet (softbot); con riferimento a internet, un tema di grande attualità è quello dei sistemi multiagenti, con le problematiche connesse alla loro coordinazione e alla condivisione di conoscenza.

Nel testo sull'intelligenza artificiale di Nilsson del 1998, la cui edizione italiana è stata curata dal professor Gaglio dell'Università di Palermo nel 2002, è stato proposto il paradigma degli agenti intelligenti per una "*nuova sintesi*", che potrebbe consentire di inquadrare in una prospettiva unitaria una molteplicità di teorie e metodologie apparentemente molto distanti [Nilsson, 2002]. Un tentativo analogo era stato effettuato da Nilsson con la versione precedente del libro pubblicata nei primi anni ottanta, proponendo come paradigma unificante quello dei sistemi a produzioni, che costituivano l'ossatura dei "*sistemi esperti*" di cui parleremo più avanti. Il paradigma degli agenti intelligenti sposta l'enfasi sull'interazione dei sistemi di intelligenza artificiale con il mondo esterno, sia fisico che virtuale; ciò significa attribuire maggiore centralità alle funzionalità e alle prestazioni di questi sistemi; il paradigma degli agenti intelligenti ha permesso di superare in parte le controversie del passato, concentrando l'attenzione sulla possibilità di ottenere diverse funzionalità utilizzando approcci differenziati<sup>21</sup>.

Secondo il professor Gaglio la storia dell'intelligenza artificiale è stata sempre fortemente intrecciata con quella dell'informatica. Realizzare e programmare i primi computer creava delle forti aspettative sulle loro capacità di risolvere problemi sempre più complessi; l'intelligenza artificiale si è sempre occupata degli aspetti di frontiera della ricerca nel settore informatico, in particolare di ciò che gli uomini sanno fare e i computer ancora no; diverse tecniche prima considerate di pertinenza dell'intelligenza artificiale adesso fanno parte dell'informatica. Un esempio importante è la programmazione a oggetti, che costituisce la metodologia fondamentale per la costruzione di sistemi software complessi; essa nacque proprio negli ambienti *LISP*, il linguaggio più usato per la realizzazione di gran parte dei programmi di intelligenza artificiale, e ha incorporato molte delle metodologie già sviluppate per i sistemi di rappresentazione della conoscenza. Secondo Gaglio, una delle più promettenti applicazioni pratiche dell'intelligenza artificiale riguarda la selezione delle informazioni più rilevanti dall'enorme quantità accessibile tramite la rete; l'intelligenza artificiale ha contribuito ad estendere l'applicazione delle tecniche di *data mining*, al fine di scoprire eventuali correlazioni fra i dati contenuti in diversi archivi a cui si può accedere in rete, o nella gran mole di ipertesti che formano il *world wide web*. Sempre nell'ambito delle applicazioni per internet, la realizzazione di interfacce più intelligenti ha lo scopo di permettere una più agevole interazione con la rete. Su internet operano tanti agenti intelligenti (*softbot*) che interagiscono e fanno tante cose utili; diverse imprese di informatica applicano e utilizzano tecniche di intelligenza artificiale per aspetti connessi alla sicurezza e alla certificazione, alla ricerca di informazioni in rete, al commercio elettronico, alla formazione attraverso la rete, alla gestione delle conoscenze.

### **2.3 Le principali tecniche di intelligenza artificiale**

Una delle applicazioni di maggior successo delle tecniche di intelligenza artificiale è costituita dai "sistemi esperti" o "sistemi basati sulla conoscenza", che racchiudono le conoscenze su settori specializzati come la medicina, l'ingegneria, l'economia, ecc. Le parti principali di un sistema esperto sono la base di conoscenza e il motore inferenziale; la base di conoscenza consiste in genere in fatti e regole del calcolo dei predicati riguardanti l'argomento

---

<sup>21</sup> Bilotta (1999, p. 11) ricorda che "dal 1980 in poi un nuovo indirizzo di ricerca compare all'interno dell'intelligenza artificiale, quello degli agenti autonomi o intelligenza artificiale basata sul comportamento."

da trattare; il motore inferenziale consiste in tutti i processi che manipolano la base di conoscenza per dedurre l'informazione richiesta all'utente. Nella costruzione di un "sistema esperto" un "ingegnere della conoscenza" lavora con esperti del campo di applicazione al fine di rappresentare le conoscenze specifiche in una forma che possa essere introdotta nella base di conoscenza. Il processo di costruzione del sistema viene iterato di solito attraverso molti cicli; a ogni iterazione un prototipo viene esaminato per vedere se esso effettua inferenze analoghe a quelle dell'esperto[Nilsson, 2002].

Sistemi di questo tipo possono essere adottati, per esempio, da una banca per valutare le richieste di prestiti. Una delle principali difficoltà di applicazione delle tecniche di intelligenza artificiale è rappresentata dalla rappresentazione della "conoscenza comune"; è quindi molto difficile costruire "macchine dotate di senso comune", che potrebbero trovare in concreto moltissime applicazioni nella vita di tutti i giorni. La conoscenza comune potrebbe anche consentire di aumentare la performance dei sistemi esperti ed è essenziale per costruire macchine in grado di comprendere il linguaggio naturale. I ricercatori di intelligenza artificiale hanno affrontato il problema di rappresentare la conoscenza comune per diversi aspetti: descrizione delle proprietà dei materiali, dimensione spaziale, proprietà fisiche degli oggetti, processi ed eventi fisici, rappresentazione del tempo. Le entità che costituiscono i domini, relativi sia al senso comune che alla conoscenza esperta, possono essere organizzate in strutture gerarchiche che consentono di guidare e semplificare il ragionamento. Le conoscenze possono essere codificate in reti semantiche o in strutture di dati chiamati "frame"; le reti semantiche sono strutture di grafi che codificano la conoscenza tassonomica relativa agli oggetti e alle loro proprietà; un *frame* ha un nome e un insieme di coppie attributo-valore.

Gli agenti hanno spesso soltanto informazioni incerte sul loro compito e sul loro ambiente; risulta molto utile quindi nella realizzazione fare riferimento anche alla teoria della probabilità ed all'inferenza probabilistica, in particolar modo alle reti Bayesiane: grafi i cui nodi sono etichettati con variabili casuali. Gli schemi d'inferenza tipici delle reti bayesiane sono l'inferenza casuale (*top-down*), l'inferenza diagnostica (*bottom-up*), e l'inferenza volta a fornire spiegazioni alternative. L'inferenza bayesiana si basa sulla regola di Bayes, secondo cui la probabilità di validità di un modello teorico sulla base degli ultimi risultati disponibili è uguale al rapporto fra la probabilità di ottenere questi risultati nell'ipotesi che il modello sia corretto e la probabilità di ottenere gli stessi risultati indipendentemente dalla validità del modello, moltiplicato per la probabilità di validità del modello valutata prima di ottenere gli ultimi risultati (probabilità a priori).

L'approccio bayesiano fornisce anche un contesto per la progettazione di sistemi di apprendimento artificiali. Secondo la regola di Bayes,  $P(M/D) = P(D/M)*P(M)/P(D)$ , vale a dire: la probabilità di validità di un modello teorico sulla base degli ultimi risultati disponibili è uguale al rapporto fra la probabilità di ottenere questi risultati nell'ipotesi che il modello sia corretto e la probabilità di ottenere gli stessi risultati indipendentemente dalla validità del modello, moltiplicato per la probabilità di validità del modello valutata prima di ottenere gli ultimi risultati. Secondo una interpretazione logica della nozione bayesiana di probabilità, sviluppata dal fisico e statistico americano Richard Cox (*Teorema di Cox*), la regola di Bayes consente di variare il grado di fiducia in un certo modello o ipotesi, utilizzando le probabilità come misure del grado di fiducia. La probabilità della validità di un certo modello prima degli ultimi risultati, rappresenta quindi una misura del grado di fiducia nel modello prima delle ultime informazioni, mentre la probabilità a posteriori rappresenta una misura del grado di fiducia nello stesso modello dopo le ultime informazioni. L'interpretazione logica proposta da Cox, della probabilità Bayesiana è considerata da Jaynes l'estensione naturale della logica aristotelica (in cui le affermazioni possono essere soltanto vere o false) ai ragionamenti in presenza di incertezza; un'altra interpretazione della probabilità bayesiana è quella soggettivista [Jaynes, 2003].

Nel contesto bayesiano il processo di apprendimento automatico inizia con l'individuazione di tutte le possibili spiegazioni plausibili dei dati, assegnando a ciascuna di esse una probabilità "a priori". Quindi, dopo avere osservato dei dati, si valuta la probabilità di ottenerli alla luce di ognuna delle possibili spiegazioni prima individuate. Moltiplicando questa probabilità per la probabilità "a priori", e normalizzando i risultati dividendo per la probabilità di ottenere comunque gli stessi dati, si ottiene una stima delle probabilità "a posteriori" che ciascuna delle possibili spiegazioni sia corretta sulla base dei dati osservati. Per comparare la plausibilità di due possibili spiegazioni ( $M1$  ed  $M2$ ) è sufficiente comparare  $P(M1)*P(D/M1)$  e  $P(M2)*P(D/M2)$ , considerato che il fattore di normalizzazione, vale a dire la probabilità di ottenere comunque i dati osservati, è identico per tutte le possibili spiegazioni. Questo risultato ha implicazioni importanti per i processi di apprendimento automatico; se l'obiettivo è di progettare un agente razionale ideale, esso deve essere in grado di apprendere utilizzando le regole del calcolo probabilistico. In pratica, per molti problemi concreti l'applicazione della regola di Bayes è di solito difficilmente praticabile poiché comporta delle procedure di sommatoria o di integrazione con riferimento a un numero molto elevato di possibili spiegazioni. Questi difficili problemi computazionali possono essere evitati utilizzando delle metodologie bayesiane approssimate, come, per esempio, l'approssimazione di Laplace, la catena di Markov, o il metodo Monte Carlo.

La teoria delle decisioni bayesiana si occupa del problema di come prendere decisioni ottimali, vale a dire decisioni o azioni che minimizzano la perdita attesa. L'inferenza bayesiana è un tipo di inferenza statistica in cui le osservazioni sono utilizzate per aggiornare la probabilità che una certa ipotesi sia vera, o anche per formulare nuove ipotesi. L'inferenza Bayesiana comporta la raccolta di informazioni che possono corroborare o smentire una data ipotesi esplicativa; man mano che l'evidenza si accumula, il grado di fiducia in ciascuna spiegazione cambia, fino a diventare di solito molto alta o molto bassa. I proponenti dell'inferenza bayesiana sostengono che essa può essere utilizzata per discriminare fra ipotesi contrastanti: le ipotesi con un grado di credibilità molto elevato dovrebbero essere accettate come vere, e quelle con un grado di credibilità molto basso dovrebbero essere rigettate perché false. I critici obiettano tuttavia che questo metodo inferenziale può essere distorto dalle opinioni iniziali da cui parte il processo di progressiva verifica. L'inferenza bayesiana di solito deriva da probabilità soggettive e in generale non pretende di offrire una metodologia induttiva oggettiva; tuttavia alcuni statistici bayesiani ritengono che le probabilità possano avere una base oggettiva, e che quindi i processi induttivi di apprendimento possano avere carattere di oggettività. L'inferenza bayesiana è utilizzata per calcolare le probabilità per i processi decisionali in condizioni d'incertezza; oltre alle probabilità può essere comunque utile una funzione di perdita per stimare le conseguenze di eventuali errori decisionali; quindi le probabilità rappresentano la possibilità di sbagliare, mentre una funzione di perdita misura le conseguenze degli errori. Le applicazioni concretamente più rilevanti dell'inferenza bayesiana hanno da sempre riguardato l'affidabilità dei risultati delle analisi mediche, la valutazione complessiva delle prove nelle controversie giudiziarie, la ricerca di sottomarini sperduti, ecc.

Particolarmente interessanti appaiono le applicazioni dell'inferenza bayesiana nel campo dell'intelligenza artificiale e dei sistemi esperti. Le tecniche di inferenza bayesiana hanno avuto un ruolo importante nelle tecniche di riconoscimento computerizzato sin dalla fine degli anni cinquanta. Le metodiche d'inferenza bayesiane sono state anche utilizzate per sviluppare degli algoritmi in grado di identificare i messaggi di posta elettronica indesiderati inviati in grandi quantità (*spam*). *DSPAM*, *Bogofilter*, *SpamAssassin*, *InBoxer* e *Monzilla* sono filtri anti-spam che utilizzano metodiche d'inferenza bayesiana.

Il teorema di Bayes è comunemente attribuito al reverendo Thomas Bayes (1701-1761), che nel suo testamento lasciò cento sterline a Richard Price. Price scoprì fra le carte di Bayes due saggi non pubblicati e li inviò alla Royal Society, che li pubblicò. Questi saggi non ricevettero tuttavia molta attenzione, fino a quando lo stesso teorema non fu scoperto in modo indipendente

dal grande matematico francese Laplace. I matematici inglesi allora rapidamente riscoprirono i lavori di Bayes. Secondo Stephen Stigler, un grande storico della statistica, il teorema di Bayes sarebbe stato scoperto in realtà da Nicholas Saunderson, un matematico cieco sin dal primo anno di vita, che fu professore di matematica nell'Università di Cambridge, su presentazione di Isaac Newton.

Gli eccezionali progressi dei computer, e le limitazioni dei metodi statistici tradizionali, sono alcune delle ragioni della ripresa di popolarità dell'approccio Bayesiano; secondo i sostenitori di questo approccio, esso dà risposte che sono più facili da comprendere e stimola i ricercatori a esplicitare le distorsioni nascoste nelle analisi degli approcci "*frequentisti*" ancora dominanti. I critici, dal canto loro, temono che le analisi bayesiane, partendo da valutazioni aprioristiche, possano condurre a stime meno obiettive dei risultati sperimentali.

Il teorema di Bayes indica come modificare le precedenti valutazioni sulla base di nuova evidenza. Il dibattito tra bayesiani e "*frequentisti*" riflette due diversi approcci verso la ricerca scientifica. L'approccio bayesiano è più adatto ai ricercatori che cercano di utilizzare tutte le informazioni disponibili per ottenere rapidamente dei progressi; essi tendono ad essere aggressivi e ottimisti nelle loro ipotesi; gli statistici frequentisti sono più cauti; essi mirano a conclusioni universalmente accettabili, in grado di far fronte alle obiezioni dei critici. I bayesiani considerano non soltanto ciò che dicono i dati, ma anche le conoscenze acquisite; la visione bayesiana di probabilità è collegata al grado di fiducia; essa è una misura della plausibilità di un evento in una situazione di conoscenza incompleta.

Le reti bayesiane sono utilizzate in diverse applicazioni dei sistemi esperti come ad esempio PATHFINDER, un sistema esperto che assiste i patologi per le diagnosi di malattie ai linfonodi. Nilsson illustra il problema dell'apprendimento di reti bayesiane, che consiste nel trovare una rete che corrisponde al meglio (secondo un dato criterio di valutazione) a un insieme di addestramento di dati, e l'utilizzazione delle inferenze probabilistiche mediante reti bayesiane per decidere le azioni di un agente, ottimali rispetto a un dato criterio di valutazione. Quest'analisi rappresenta una generalizzazione del ciclo percezione/pianificazione/azione in condizioni di certezza. La fase di pianificazione calcola la migliore azione successiva, la fase di azione esegue l'azione, la fase di percezione cerca di capire la situazione ambientale risultante per il ciclo successivo. L'obiettivo può essere configurato in termini di "ricompense" e "penalità" in corrispondenza dell'utilità di diversi stati ambientali. Nell'ipotesi di certezza, l'agente è in grado di determinare accuratamente il suo stato attuale attraverso la percezione, e può prevedere esattamente gli effetti di tutte le sue azioni.

Questo approccio può essere giustificato ritenendo che anche se le azioni non hanno sempre gli effetti previsti, e anche se le percezioni non sono a volte corrette, i sensori avrebbero tenuto informato l'agente sul suo progredire, e la pianificazione ripetuta avrebbe riorientato l'agente verso l'obiettivo. Utilizzando le reti bayesiane, l'agente conosce le probabilità di trovarsi nei diversi stadi, e le probabilità associate a diverse possibili conseguenze delle sue azioni. In genere i calcoli necessari per trattare gran parte dei problemi concreti mediante reti bayesiane richiedono elevatissime potenzialità computazionali, e quindi spesso si è costretti a introdurre delle approssimazioni.

Secondo Nilsson, tecniche diverse di intelligenza artificiale possono essere integrate in architetture di agenti intelligenti. Uno dei primi sistemi integrati per un agente intelligente è stata una raccolta di programmi per computer e hardware specifico nota come "*Shakey il robot*", l'hardware consisteva in una piattaforma mobile, delle dimensioni di un piccolo frigorifero, con sensori sensibili al contatto, una telecamera e un telemetro ottico, tutti controllati da un computer con un collegamento radio/video a due vie. La piattaforma mobile era in grado di girare intorno in un ambiente fatto di grandi scatole in stanze separate da pareti e collegate da porte; Shakey era in grado di spingere le scatole da un punto all'altro. Shakey fu il primo robot mobile in grado di ragionare sulle sue azioni, frutto di ricerche nel campo della robotica, della visione computerizzata, e dell'elaborazione del linguaggio naturale. Fu sviluppato fra il 1966 e il 1972

presso il laboratorio di Intelligenza Artificiale di SRI international (*Stanford Research Institute* fino al 1970), un'organizzazione "non profit", che svolge ricerche per conto di enti pubblici e imprese private. Il gruppo di ricerca era diretto da Charles Rosen, e ne faceva parte fra gli altri, anche Nils Nilsson; il linguaggio di programmazione utilizzato era il LISP<sup>22</sup>

La conoscenza comune, che è la principale carenza che limita le applicazioni di Intelligenza Artificiale dovrebbe essere codificata negli agenti intelligenti per consentire ad essi di svolgere tanti compiti di routine che le persone trovano facili ma che sono in genere impossibili per gli agenti artificiali. Quest'obiettivo è stato perseguito da "Cyc"<sup>23</sup>, un progetto di intelligenza artificiale che cerca di assemblare in un database la conoscenza comune per consentire agli agenti di ragionare il più possibile come persone. Il progetto iniziò nel 1984 presso la *Microelectronics and Computer Technology Corporation* (MCC)<sup>24</sup>. La speranza è che, raggiunto un certo stadio critico, il sistema sarà in grado di apprendere maggiori conoscenze, leggendo testi, conversando con le persone, e venendo istruito.

Nilsson infine, ritiene che le persone apprendono molte delle conoscenze che utilizzano con un processo di "bootstrapping" dalla conoscenza ottenuta in precedenza, prima come bambini, da tendenze innate, in seguito, man mano che attraversano i diversi stadi piagetiani, da abilità e concetti acquisiti in precedenza, e poi da adulti, attraverso la pratica, la lettura e la comunicazione. Ma sembra, per adesso, che la quantità di conoscenza che sarà richiesta dagli agenti di intelligenza artificiale per esibire una intelligenza di livello umano sia così grande, che sarà necessario un simile processo di bootstrapping ancora da scoprire [Nilsson, 2002].

## 2.4 Gli agenti intelligenti

Secondo Russel e Norvig "un agente è qualsiasi cosa possa essere vista come un sistema che percepisce il suo ambiente attraverso sensori e agisce su di esso mediante attuatori". Un agente robotico potrebbe avere telecamere e dispositivi a infrarossi per sensori, e diversi motori per attuatori; un agente software riceve in input come sensori le battute dei tasti, il contenuto dei file e dei pacchetti di rete, e può intervenire sull'ambiente cambiando la visualizzazione di uno schermo, scrivendo file, e inviando pacchetti di dati; per esempio, una calcolatrice è un agente che visualizza 4 sul display quando gli viene fornita la sequenza 2 + 2. La scelta dell'azione di un agente in un certo istante può dipendere dall'intera sequenza percettiva osservata fino a quel momento; in termini matematici, il comportamento di un agente è descritto dalla "funzione agente", che descrive la corrispondenza tra una qualsiasi sequenza percettiva e una specifica azione; la funzione di un agente artificiale è implementata da un "programma agente" [Russel, Norvig, 2005, vol. 1, pag. 46].

Un agente razionale è un agente che fa la cosa giusta, vale a dire la cosa che consente all'agente di ottenere il massimo successo. Ovviamente è necessario disporre di qualche tecnica che consenta di misurare il successo dell'agente; questa tecnica, insieme alle descrizioni dell'ambiente, dei sensori e degli attuatori dell'agente fornirà la specifica completa dell'attività che l'agente è chiamato a svolgere. Il concetto di razionalità ebbe un ruolo marginale nelle

---

<sup>22</sup> Il LISP (*List Processor*) è un linguaggio di programmazione funzionale ideato nel 1958 da John McCarthy; si basa sul concetto di programma come funzione, ed è ancora oggi molto utilizzato nelle applicazioni di intelligenza artificiale.

<sup>23</sup> Il termine Cyc deriva da "encyclopedia" e si pronuncia come *psych*.

<sup>24</sup> *Microelectronics and Computer Technology Corporation* (MCC) fu il primo consorzio fra imprese informatiche americane nel campo della ricerca e sviluppo, costituito nel 1982, con sede in Texas, come la risposta americana alla sfida giapponese concretizzata con il "Fifth Generation Project", che mirava a produrre un nuovo tipo di computer per il 1991, suscitando il timore che il Giappone volesse acquisire il predominio mondiale nel segmento alto dell'industria dei computer. Prima del 1984 i consorzi fra imprese erano illegali negli Stati Uniti in base alla normativa antitrust, e per costituire MCC fu quindi necessaria l'approvazione da parte del Congresso americano del "National Cooperative Research Act". I principali argomenti di ricerca di MCC sono stati i database, le architetture avanzate per computer, le interfacce uomo-computer, i sistemi di elaborazione in parallelo, l'ingegneria del software, l'intelligenza artificiale.

ricerche di intelligenza artificiale fino agli anni ottanta; fu John Doyle uno dei primi a predire che la progettazione di agenti razionali sarebbe stata considerata la missione principale dell'intelligenza artificiale [Doyle, 1982]. Un agente razionale dovrebbe raccogliere le informazioni utili per poter agire razionalmente, ed essere in grado di imparare il più possibile sulla base delle proprie percezioni. Gli agenti di successo suddividono il calcolo della funzione agente in tre periodi: al momento della progettazione una parte dei suoi calcoli viene svolta dai programmatori; quando l'agente sta decidendo la sua prossima azione, svolge altri calcoli; imparando dall'esperienza compie ulteriori elaborazioni per decidere come modificare il proprio comportamento. Quando un agente si appoggia alla conoscenza pregressa inserita dal programmatore invece che sulle sue percezioni, si dice che l'agente manca di autonomia. Un agente razionale dovrebbe essere autonomo, e imparare il più possibile per compensare la presenza di conoscenze parziali o erranee. In pratica in genere non si pretende che un agente sia autonomo sin dall'inizio: finché la sua esperienza è limitata o nulla un simile agente sarebbe costretto ad agire casualmente, a meno di ricevere un aiuto dal progettista; è ragionevole fornire a un agente artificiale, oltre all'abilità di imparare, anche un pò di conoscenza iniziale. Dopo avere accumulato una sufficiente esperienza dell'ambiente, il comportamento dell'agente razionale può diventare a tutti gli effetti indipendente dalla conoscenza pregressa. Incorporare l'apprendimento nel suo progetto, permette di sviluppare un agente razionale capace di muoversi efficacemente in una grande varietà di ambienti differenti.

Gli ambienti sono i problemi di cui gli agenti razionali rappresentano le soluzioni. Il criterio di misura delle prestazioni, l'ambiente esterno, gli attuatori e i sensori dell'agente costituiscono l'ambiente operativo degli agenti intelligenti. Gli agenti software (o *software robot* o *sofbot*), per esempio, operano in domini complessi e senza confini. Nel caso di un *sofbot* progettato per agire all'interno del simulatore di volo di un grande aereo commerciale, il simulatore rappresenta un ambiente complesso e molto dettagliato che include altri velivoli e operazioni a terra, e l'agente software dovrà scegliere in tempo reale tra una vasta gamma di possibili azioni. Un *sofbot* progettato per consultare diversi siti internet e mostrare ai suoi utenti le notizie a cui sono più interessati, per lavorare bene dovrà avere una certa capacità di comprensione del linguaggio naturale, imparare gli argomenti che interessano maggiormente ogni utente ed essere capace di cambiare dinamicamente i propri piani, per esempio quando un sito non è più raggiungibile o quando vengono aperti nuovi siti; Internet è un ambiente altrettanto complesso del mondo reale, con molti agenti artificiali.

Dal punto di vista delle caratteristiche più rilevanti per le applicazioni di intelligenza artificiale gli ambienti possono essere completamente o parzialmente osservabili, deterministici o stocastici, episodici o sequenziali, statici o dinamici, discreti o continui, con agente singolo o multiagente; gli ambienti più complessi sono ovviamente quelli parzialmente osservabili, stocastici, sequenziali, dinamici, continui e multiagente. In particolar modo, negli ambienti multiagente, ogni agente deve considerare le azioni degli altri, e in particolare gli effetti di queste azioni sul proprio benessere

Il compito dell'intelligenza artificiale è progettare il programma agente che implementa la funzione agente, mettendo in relazione percezioni e azioni; ovviamente il programma sarà eseguito da un computer dotato di sensori fisici e attuatori che costituiscono l'architettura dell'agente; un agente è quindi definito dalla sua architettura e dal suo programma. L'architettura può essere costituita da un semplice computer, o da un veicolo robot dotato di vari computer, telecamere e altri sensori; l'architettura rende le percezioni disponibili al programma, esegue il programma, e passa le azioni scelte dal programma agli attuatori man mano che vengono generate. I programmi agente prendono come input la percezione corrente dei sensori e restituiscono un'azione agli attuatori; mentre il programma agente prende come input soltanto la percezione corrente, l'input della funzione agente è costituito dall'intera storia delle percezioni.

I principali tipi di agente intelligente sono gli agenti reattivi semplici, gli agenti reattivi basati su modello<sup>25</sup>, gli agenti basati su obiettivi<sup>26</sup>, gli agenti basati sull'utilità.

Gli agenti reattivi semplici scelgono le azioni sulla base delle percezioni correnti; gli agenti reattivi basati su modello hanno una conoscenza sul funzionamento del mondo, implementata mediante semplici circuiti logici o sviluppata in una teoria completa; gli agenti basati su obiettivi utilizzano le informazioni sulle situazioni desiderabili congiuntamente a quelle relative agli effetti delle sue azioni per scegliere le azioni che portano al soddisfacimento degli obiettivi. Scegliere un'azione in base a un obiettivo è semplice quando l'obiettivo può essere raggiunto in un solo passo, è più difficile quando l'agente deve considerare lunghe sequenze di azioni alternative. La ricerca e la pianificazione sono dei sottocampi dell'intelligenza artificiale dedicati a identificare le sequenze di azioni che permettono a un agente di raggiungere i propri obiettivi. Una funzione di utilità assegna a uno stato (o una sequenza di stati) un numero reale che quantifica il grado di utilità ad esso associato. Una funzione di utilità permette di prendere decisioni razionali quando ci sono più obiettivi in conflitto (per esempio velocità e sicurezza), e quando più obiettivi sono raggiungibili ma nessuno può essere ottenuto con certezza; nel primo caso la funzione di utilità può consentire di conoscere quale obiettivo privilegiare, nel secondo può consentire di confrontare le probabilità di successo e l'importanza degli obiettivi; ogni agente razionale deve comportarsi come se possedesse una funzione di utilità di cui cerca di massimizzare il valore atteso<sup>27</sup>.

Già Turing nel 1950 aveva intuito che il modo migliore di ottenere macchine intelligenti è costruire macchine capaci di apprendere e poi addestrarle; un agente capace di apprendere comprende quattro componenti astratti: l'elemento di apprendimento, responsabile del miglioramento interno, l'elemento esecutivo, che si occupa della selezione delle azioni esterne, prende come input le percezioni e decide le azioni. L'elemento di apprendimento utilizza informazioni provenienti dall'elemento critico riguardo le prestazioni correnti dell'agente e determina se e come modificare l'elemento esecutivo. Qualunque progetto di agente può essere migliorato dall'aggiunta di meccanismi di apprendimento; l'elemento critico dice a quello di apprendimento come si sta comportando l'agente rispetto a degli standard di prestazione prefissati. Il quarto componente di un agente capace di apprendere è il generatore di problemi, il cui scopo è suggerire azioni che portino a esperienze nuove e significative; se l'agente è disposto a esplorare nuove possibilità, magari con il rischio di azioni non ottimali nel breve termine, potrebbe scoprire l'esistenza di azioni migliori nel lungo termine; l'obiettivo del generatore di problemi è di suggerire delle azioni esplorative, nello stesso modo in cui gli scienziati compiono degli esperimenti.

#### **2.4.1 Agenti intelligenti: tecniche per fronteggiare l'incertezza.**

Quando un agente ha sufficienti informazioni, l'approccio logico gli consente di derivare piani razionali; in concreto però gli agenti non hanno quasi mai tutte le informazioni necessarie. La pianificazione condizionale può mitigare in parte la presenza di incertezza soltanto se le azioni di percezione dell'agente gli consentono di ottenere le informazioni necessarie. In una situazione di incertezza, la cosa giusta da fare, vale a dire la decisione razionale, dipende sia

---

<sup>25</sup> Gli agenti reattivi sono stati il riferimento principale degli psicologi di ispirazione behaviorista o comportamentale che concentravano l'attenzione sul meccanismo stimolo-risposta.

<sup>26</sup> Gli agenti artificiali basati su obiettivi fanno principalmente riferimento alla psicologia cognitiva; un lavoro importante per il loro sviluppo fu "*Human Problem Solving di Newell e Simon*" del 1972, che ebbe grande influenza, anche per effetto dei successivi lavori di Newell che ne svilupparono le implicazioni. La distinzione degli obiettivi in "desideri" (generali) e "intenzioni" (perseguite in un certo momento), proposta da Bratman (1987), ha avuto molta importanza nei campi della comprensione del linguaggio naturale e dei sistemi multiagente.

<sup>27</sup> Horwitz e altri (1988) suggerirono esplicitamente di usare come base dell'intelligenza artificiale una razionalità intesa come massimizzazione dell'utilità attesa.

dall'importanza relativa dei diversi obiettivi, sia dalla probabilità e dalla misura del loro raggiungimento.

Secondo Russel e Norvig l'incertezza deriva principalmente dalla pigrizia e dall'ignoranza, e la probabilità fornisce un modo per riassumere l'incertezza. Per esempio, si potrebbe dire che un paziente che ha mal di denti, nell'80 per cento dei casi avrà una carie; questa stima di probabilità può derivare da indagini statistiche sulle frequenze (l'80 per cento dei pazienti con mal di denti considerati in una indagine statistica aveva una carie), o da qualche altra fonte. Il restante 20 per cento di probabilità riassume tutte le altre possibili cause di mal di denti che la pigrizia e/o l'ignoranza non consentono di conoscere. Assegnare una probabilità pari a zero a una certa credenza significa esser sicuri che è falsa, probabilità pari a 1 significa esser sicuri che è vera, probabilità comprese fra 0 e 1 corrispondono a gradi intermedi di fiducia nella verità di una credenza. Una formula del tipo "la probabilità che il paziente abbia una carie è dell'80 per cento" riguarda le credenze dell'agente e non direttamente il mondo. Ciascun singolo paziente la carie può soltanto averla o non averla, non può "averla all'80 per cento". Le probabilità possono ovviamente modificarsi man mano che si acquisiscono nuove informazioni (vale a dire man mano che diminuisce l'ignoranza) [Russel, Norvig, 2005, vol. 2].

Per prendere decisioni razionali in condizioni di incertezza, è necessario conoscere sia le probabilità dei diversi effetti delle decisioni, sia le preferenze dell'agente per questi diversi possibili effetti. Per rappresentare e utilizzare le preferenze si può applicare la teoria dell'utilità, secondo cui ogni possibile stato ha per l'agente un certo grado di utilità, e l'agente preferisce sempre stati con maggiore utilità. Non c'è alcun modo per valutare obiettivamente i gusti o le preferenze di diversi agenti; si può dire che alcune preferenze sono strane, non si può dire che sono irrazionali. Una funzione di utilità può anche incorporare un comportamento altruistico, se include il benessere di altri come fattore che contribuisce in modo diretto all'utilità dell'agente considerato. Le preferenze quindi, nella teoria delle decisioni vengono combinate con le probabilità. Come abbiamo spiegato precedentemente, l'idea fondamentale della teoria delle decisioni è che un agente è razionale se sceglie l'azione che porta alla più alta utilità attesa, calcolata sulla media di tutti i possibili esiti dell'azione; questo è il principio della massimizzazione dell'utilità attesa (MEU = *maximization of expected utility*) sviluppata da Von Neumann e Morgenstern.

L'agente basato sulla teoria delle decisioni ha una conoscenza incerta del mondo; il suo "*stato credenza*" è una rappresentazione delle probabilità di tutti i possibili stati del mondo; con il passar del tempo l'agente accumula informazioni che possono modificare le sue credenze. Dato lo *stato credenza*, l'agente può formulare predizioni probabilistiche sugli esiti delle azioni e selezionare l'azione con la massima utilità attesa.

Le reti bayesiane sono state molto utili nel calcolo delle probabilità nel campo dell'intelligenza artificiale. L'intelligenza artificiale comunque ha cominciato in ritardo rispetto ad altre scienze come la fisica, la genetica, l'economia, ecc., a utilizzare la teoria delle probabilità per modellare l'incertezza; i primi sistemi esperti degli anni settanta, ignoravano l'incertezza; soltanto la seconda generazione di sistemi esperti utilizzavano tecniche probabilistiche, in particolare in ambito medico. I primi risultati furono incoraggianti, ma le difficoltà derivanti dal numero crescente in modo esponenziale delle stime di probabilità richieste dalla distribuzione congiunta completa, fece sì che fra il 1975 e il 1988 gli approcci probabilistici fossero praticamente abbandonati, e furono sperimentati altri approcci al ragionamento incerto. La teoria della probabilità è essenzialmente numerica, mentre i giudizi umani sono spesso di natura qualitativa; anche se le persone di solito non effettuano coscientemente calcoli sui gradi di credenza, esse svolgono comunque ragionamenti logici. Uno degli schemi qualitativi più utilizzati in concreto è il "ragionamento di *default*", che considera le conclusioni non credute in una certa misura, bensì credute finché non si trova una ragione per credere in qualcos'altro.

Alcune applicazioni importanti delle reti bayesiane sono derivate dalle ricerche svolte presso Microsoft sulla deduzione degli obiettivi di un utente di personal computer in base alle sue azioni, e sul filtraggio della posta elettronica indesiderata, presso l'*electric power research institute* sul monitoraggio dei generatori, e presso il controllo di missione di Houston della NASA sulla visualizzazione tempestiva di informazioni critiche.

In caso di incertezza, quando un agente intelligente prende delle decisioni, facendo riferimento alla teoria delle decisioni, si comporta diversamente, a seconda se ci troviamo in un caso di decisioni semplici, oppure di decisioni complesse (sequenziali e con agenti multipli). Nell'analisi delle decisioni semplici si cerca di combinare la teoria dell'utilità con la teoria della probabilità per ottenere agenti basati sulla teoria delle decisioni, in grado di scegliere razionalmente l'azione da intraprendere in base a ciò che l'agente crede e desidera; questi agenti dovrebbero essere in grado di operare in contesti in cui incertezza e conflitti tra obiettivi contrastanti non consentono a un agente puramente logico di arrivare ad alcuna decisione. Un agente basato su obiettivi può soltanto dividere gli stati in buoni e cattivi, mentre un agente basato sulla teoria delle decisioni è in grado di valutare la qualità degli stati in modo continuo. Questo accade perché, per massimizzare l'utilità attesa che è il principio base della teoria delle decisioni, le preferenze di un agente per gli stati del mondo vengono rappresentate da una funzione di utilità, che esprime con un singolo valore la desiderabilità di uno stato e le utilità vengono combinate con le probabilità degli esiti delle azioni per ottenere l'utilità attesa di ogni azione. Quindi, tutto ciò che deve fare un agente intelligente è calcolare le varie quantità e massimizzare l'utilità delle sue azioni, ma questo non significa che il problema dell'intelligenza artificiale sia dunque del tutto risolto.

Benché sia vero che il principio della massima utilità attesa definisce l'azione giusta in ogni contesto decisionale, i calcoli potrebbero essere proibitivi, e spesso è difficile formulare completamente il problema. Conoscere lo stato iniziale del mondo richiede percezione, apprendimento, rappresentazione della conoscenza, inferenza; per calcolare la probabilità di un certo risultato è necessario possedere un modello completo del mondo, ed effettuare inferenze spesso complicatissime su reti bayesiane; per calcolare l'utilità di ogni stato spesso si deve ricorrere alla ricerca o alla pianificazione, perché un agente non sa quanto è buono uno stato finché non appura dove si può arrivare partendo da esso. La teoria delle decisioni comunque fornisce un'infrastruttura in cui collocare ordinatamente tutti i componenti di un sistema intelligente. Se un agente massimizza una funzione di utilità che riflette correttamente la misura di prestazione utilizzata per valutare il suo comportamento, allora otterrà la migliore prestazione possibile, calcolata sulla media degli ambienti in cui potrà operare.

Facendo riferimento all'analisi dei processi decisionali degli individui, descritti nel capitolo precedente, anche gli agenti intelligenti spesso si trovano a dover scegliere fra opzioni caratterizzate da diversi attributi. Quindi, per sfoltire le opzioni a volte può essere utile applicare la nozione di dominanza; una opzione è strettamente dominata se è peggiore di altre opzioni dal punto di vista di tutti gli attributi; ovviamente le opzioni strettamente dominate possono essere eliminate, semplificando il problema di scelta; in caso di incertezza, può essere utile applicare un concetto di dominanza meno forte: la dominanza stocastica. L'opzione A è dominata stocasticamente dall'opzione B dal punto di vista di un certo indicatore di performance, se, per esempio, l'indicatore di performance può variare stocasticamente in modo uniforme fra un valore di 20 e un valore di 60 per l'opzione A, e fra un valore di 25 e un valore di 61 per l'opzione B<sup>28</sup>.

---

<sup>28</sup> Keeney a Raiffa (1976) contiene un'introduzione approfondita alla teoria dell'utilità multiattributo; tra le altre cose contiene la descrizione delle prime implementazioni su computer dei metodi per estrarre i parametri necessari per una funzione di utilità multiattributo, e include ampi resoconti di applicazioni reali della teoria. Per quel che riguarda l'intelligenza artificiale, Welman (1985) introdusse il sistema URP (*utility reasoning package*), che può utilizzare una collezione di asserzioni di indipendenza delle preferenze e di indipendenza condizionale per analizzare la struttura dei problemi decisionali. La teoria delle decisioni è stata uno strumento standard dell'economia, della finanza e del management fin dagli anni cinquanta; fino agli anni ottanta lo strumento

Per la scelta fra le opzioni che rimangono dopo avere eliminato quelle dominate può essere utile l'ipotesi di indipendenza mutua per la preferenza (MPI = *mutual preference independence*). Questa ipotesi è soddisfatta se, indipendentemente dall'importanza di ciascun attributo, le sue variazioni non influenzano le scelte che coinvolgono gli altri attributi dell'insieme. Per esempio, se si preferisce il cellulare A che ha 200 ore di autonomia in conversazione e 380 ore di autonomia in standby al cellulare B che ha 180 ore di autonomia in conversazione e 410 ore di autonomia in standby quando entrambi pesano 130 grammi, il cellulare A deve continuare a essere preferito al cellulare B anche se entrambi pesano 170 grammi. In base a un teorema dimostrato da Debreu (1960), se gli attributi  $X_1, X_2, X_3, \dots, X_n$  sono mutuamente indipendenti per la preferenza, un agente è razionale se massimizza la somma dei valori dei diversi attributi singolarmente considerati (funzione di valore additiva). Anche nei casi in cui l'ipotesi di indipendenza mutua per la preferenza non vale strettamente (potrebbe per esempio essere violata per valori estremi di qualche variabile), alcune funzioni di valore additive possono fornire di solito delle buone rappresentazioni delle preferenze degli agenti.

Una delle parti più importanti dei processi decisionali è sapere quali informazioni cercare; a questa domanda cerca di rispondere la teoria del valore dell'informazione. In sintesi, il valore dell'informazione è tanto più elevato quanto maggiore è la probabilità che essa provochi un cambiamento della decisione, e quando più è alta la differenza fra il valore della decisione presa in presenza dell'informazione rispetto a quella che sarebbe stata presa in sua assenza. Il valore dell'informazione non può mai essere negativo, considerato che nel caso peggiore essa può essere ignorata e quindi non influenzare la decisione; ovviamente il processo di ricerca dell'informazione può avere effetti negativi se il costo da sostenere per ottenerla supera il suo valore. Un agente intelligente che raccoglie informazioni dovrebbe tenere in considerazione il valore di ogni informazione in rapporto al suo costo; un semplice algoritmo potrebbe operare selezionando ripetutamente l'osservazione con il più alto valore dell'informazione, fino a quando il costo dell'osservazione successiva supera il beneficio atteso.

#### **2.4.2 La Teoria dei Giochi per fronteggiare l'incertezza**

L'incertezza derivante da decisioni di altri agenti, e in particolare dall'interazione competitiva fra diversi agenti, può essere analizzata mediante la *Teoria dei Giochi*. I giochi più semplici sono quelli sequenziali con informazioni perfette; per questi giochi è possibile utilizzare la ricerca *minimax* per trovare le mosse ottime; altri aspetti della teoria dei giochi riguardano i giochi a mosse simultanee. La teoria dei giochi è una teoria concretamente utilizzata per decisioni molto importanti in diversi campi: la finanza, la vendita delle frequenze radio, la difesa nazionale, e anche nel commercio per lo sviluppo dei prodotti e le decisioni di prezzo dei prodotti. La teoria dei giochi può essere utilizzata per la progettazione di agenti (analisi delle decisioni dell'agente e calcolo dell'utilità attesa per ciascuna decisione nell'ipotesi che anche gli altri giocatori stiano agendo in modo ottimo secondo la stessa teoria) e la progettazione di meccanismi (quando un ambiente è popolato da più agenti può essere possibile definirne le regole (ovvero il gioco a cui gli agenti devono partecipare) in modo tale da massimizzare il bene comune quando ogni agente adotta i comportamenti che, in base alla teoria dei giochi, massimizzano la sua utilità.

---

principale per rappresentare problemi semplici è stato l'albero delle decisioni. Le reti di decisioni, o diagrammi d'influenza, sono state introdotte da Howard e Matheson (1984), basandosi sul lavoro svolto da un gruppo di ricerca allo *Stanford Research Institute* (SRI). Nilson e Lauritzen (2000) hanno collegato gli algoritmi per le reti di decisioni agli sviluppi degli algoritmi di *clustering* per reti bayesiane. Articoli sulle reti di decisione e sulla modellazione dell'utilità compaiono regolarmente nella rivista *Management Science*. Dopo il ritorno di interesse nei metodi probabilistici in intelligenza artificiale negli anni ottanta, i sistemi esperti basati sulla teoria delle decisioni hanno avuto un'ampia diffusione.

L'esempio più famoso di teoria dei giochi è il “*dilemma del prigioniero*”<sup>29</sup>: due presunti ladri, A e B, sono arrestati dalla polizia e interrogati separatamente. Ognuno dei due sa che se entrambi confessano verranno condannati a 5 anni di prigione ciascuno, mentre se nessuno dei due confessa potranno essere condannati a un anno di prigione ciascuno per dei reati minori. La polizia offre a ognuno dei due l'opportunità di confessare, testimoniando contro il complice, promettendo che, se uno soltanto dei due confessa, a chi confessa sarà condonata completamente la pena, mentre al complice che non confessa verrà inflitta la pena di 10 anni di reclusione. Supponendo che l'utilità di ciascuno dei due complici non dipende in alcun modo di ciò che accadrà al complice, quali sono le decisioni razionali dei due agenti? La soluzione proposta per questo dilemma dalla teoria dei giochi è che per ciascuno dei due agenti la strategia dominante è confessare; per strategia dominante per A si intende una strategia che, indipendentemente dalla decisione dell'altro agente, comporta una utilità maggiore per A, analogamente è definita una strategia dominante per B. Quando ogni agente ha una strategia dominante, la combinazione delle strategie dominanti prende il nome di “equilibrio con strategie dominanti”; in questo esempio l'equilibrio di strategie dominanti corrisponde alla decisione di confessare sia per A che per B, e quindi a una pena di 5 anni di reclusione sia per A che per B. In generale, un profilo di strategie forma un equilibrio se a nessun agente conviene una strategia diversa se gli altri agenti mantengono invariata la loro strategia. Il “dilemma” sta nel fatto che l'equilibrio con strategie dominanti è peggiore per entrambi gli agenti di quello che si potrebbe ottenere se entrambi decidessero di non confessare e testimoniare contro il complice. Certamente è *possibile* che entrambi si rifiutino di testimoniare, ma è molto *improbabile*; entrambi i giocatori, considerando l'opzione di *rifiutarsi*, si rendono conto che farebbero meglio a *testimoniare*; è il potere di attrazione del punto di equilibrio.”

La soluzione (*testimoniare, testimoniare*) può essere evitata se si suppone che gli agenti sappiano che si incontreranno ancora un numero indefinito di volte, o che abbiano una certa predisposizione alla cooperazione, nel senso che l'utilità di ogni agente dipende positivamente anche dall'utilità dell'altro agente, o limitando la potenza computazionale degli agenti in modo che non possano ragionare in modo perfettamente razionale, oppure se uno degli agenti ritiene che l'altro abbia una “razionalità limitata”.

Un campo di ricerca di grande interesse è costituito dall'impiego della teoria dei giochi per progettare ambienti o “meccanismi” idonei a ottenere la massimizzazione di una funzione di utilità globale, attraverso l'interazione fra agenti razionali<sup>30</sup>. La progettazione di meccanismi, o teoria dei giochi inversa, rappresenta un argomento di grande importanza e attualità per l'economia e la politica, oltre che per l'intelligenza artificiale. Dato un insieme di agenti, e utilizzando le tecniche della teoria dei giochi, si cerca di costruire sistemi intelligenti partendo da sistemi più limitati, anche non cooperativi, in maniera simile a quella per cui squadre di esseri umani possono raggiungere obiettivi che vanno molto al di là delle capacità dei singoli individui che le compongono. Esempi di progettazione di meccanismi includono la messa all'asta di biglietti aerei economici, l'instradamento di pacchetti di informazioni tra computer, la suddivisione di un gruppo di medici in più ospedali, la cooperazione all'interno di una squadra di calciatori robot. La ricerca accademica sulla progettazione dei meccanismi è stata stimolata a

---

<sup>29</sup> Il “dilemma del prigioniero” fu inventato come esercizio per gli studenti da Albert W. Tucker nel 1950.

<sup>30</sup> Leonid Hurwicz (1973) presentò i fondamenti matematici per la progettazione di meccanismi; nel 2007 Hurwicz ha ottenuto, a 90 anni, il premio Nobel per l'economia, insieme a Roger B. Myerson ed Eric S. Maskin, di oltre 30 anni più giovani, che contribuirono a sviluppare il lavoro di Hurwicz. Leonid Hurwicz aveva iniziato le sue ricerche negli anni cinquanta, quando era molto vivo il dibattito sugli effetti delle riforme di ispirazione socialista sull'efficienza economica. Nella motivazione per l'assegnazione del premio Nobel si afferma che Hurwicz fu un pioniere nell'applicare a questo problema una più rigorosa, matematica, analisi, e riporta una frase di Myerson, secondo cui “per diversi decenni Leo Hurwicz ha lavorato per mostrare in che modo modelli economici matematici possono fornire una impostazione generale per analizzare diverse istituzioni economiche, come quelle del capitalismo e del socialismo, come meccanismi per coordinare gli individui in una società”.

partire dagli anni novanta anche dal fatto che alcuni governi ritennero di aver perso centinaia di milioni di dollari per non essere riusciti a progettare meccanismi adeguati per la vendita all'asta delle licenze di trasmissione in diverse bande di frequenza. Formalmente un meccanismo è costituito da un linguaggio per la descrizione dell'insieme, potenzialmente infinito, delle possibili strategie adottabili dagli agenti, e da una regola per i risultati che determina le vincite degli agenti dato un profilo di strategie adottabili. Il meccanismo di gran lunga più rilevante riguarda l'organizzazione complessiva di un sistema economico; secondo l'intuizione di Adam Smith della "mano invisibile", un'organizzazione economica in cui ogni agente persegue la massimizzazione della propria utilità realizza indirettamente il massimo benessere per tutti. L'analisi economica e l'esperienza storica hanno mostrato che questa intuizione ha una grande capacità esplicativa, ma che vi sono delle cause che possono far fallire" questo meccanismo, in particolare le esternalità, i beni pubblici, le asimmetrie informative, le forme di mercato non concorrenziali (fallimento dei mercati) [Russel, Norvig, 2005, vol. 2, pag 300].

Attenzione particolare è stata dedicata alle aste, un meccanismo per la compravendita di beni<sup>31</sup>; le strategie corrispondono alle offerte dei diversi partecipanti all'asta, il risultato determina chi ottiene i beni e quanto deve pagarli. Un possibile esempio di applicazione di un'asta in intelligenza artificiale si ha quando una collezione di agenti deve decidere se cooperare su un piano congiunto. Hunsberger e Grosz hanno dimostrato che questo risultato si può ottenere in modo efficiente con un'asta in cui gli agenti puntano sui rispettivi ruoli all'interno del piano congiunto [Hunsberger, Grosz, 2000]. Il tipo più comune di asta è l'asta inglese standard al primo prezzo, in cui il banditore aumenta progressivamente il prezzo del bene, fino a quando rimane un solo agente disposto a pagare il prezzo proposto dal banditore. Con questo meccanismo il vincitore dell'asta paga un prezzo pari alla puntata più alta effettuata da almeno un altro agente, e l'incremento di prezzo fra una puntata e l'altra. La strategia dominante dei partecipanti all'asta inglese è semplice: continuare a puntare fino a quando il prezzo proposto dal banditore è più basso della propria valutazione del bene, senza tener conto delle strategie degli altri agenti; l'asta inglese è un meccanismo "a prova di strategia", nel senso che stimola i partecipanti a rivelare le loro valutazioni del bene. Un limite dell'asta inglese è che i partecipanti devono essere riuniti tutti contemporaneamente, eventualmente anche stando in posti diversi, ma collegati con linee di comunicazione ad alta velocità. Un meccanismo alternativo è l'asta segreta (*sealed bid auction* o asta a busta chiusa); in questo caso ogni partecipante comunica segretamente al venditore la propria puntata, e quella più alta vince. Con questo meccanismo la strategia di dichiarare il valore reale del bene per l'agente non è più dominante; se l'agente *i* valuta il bene 100.000 euro, ma è convinto che la puntata massima degli altri agenti sarà 60.000 euro, non è razionale per *i* offrire un prezzo significativamente maggiore di 60.000 euro. Se l'agente con la valutazione più elevata sottovaluta le valutazioni degli altri agenti, è possibile che il bene venga venduto a un agente diverso da quello che lo valuta di più. Un meccanismo che consente di evitare questo rischio è l'asta segreta al secondo prezzo o "asta Vickrey", dal nome dello studioso che la propose, vincitore del premio Nobel per l'economia nel 1996. In queste aste vince l'agente che offre il prezzo più elevato, ma paga non il prezzo offerto, ma il prezzo corrispondente alla seconda puntata più alta. Con questo meccanismo la strategia dominante di ciascun agente è offrire un prezzo uguale al valore effettivo del bene per sé. Per la sua semplicità, e per i requisiti computazionali minimi richiesti sia al venditore che ai partecipanti, l'asta Vickrey è usata largamente nella costruzione di sistemi intelligenti artificiali distribuiti.

Nel caso dell'instradamento di pacchetti su internet, gli agenti corrispondono agli archi nel grafo delle connessioni di rete; ogni agente conosce il costo di inviare il messaggio sul proprio arco, il costo di non avere alcun messaggio da inviare è zero. L'obiettivo è trovare il cammino più economico per far viaggiare ciascun messaggio dall'origine alla destinazione; il

---

<sup>31</sup> Paul Milgrom (1997) ha presentato un meccanismo per gestire aste per valori di miliardi di dollari; egli ha successivamente fornito importanti contributi, sia teorici che professionali, all'analisi dei meccanismi d'asta.

costo totale del cammino è uguale alla somma dei costi dei diversi archi. Il cammino più breve può essere calcolato se si conoscono i costi di passo, è quindi essenziale che ogni agente comunichi il suo costo reale; il problema è che ciascun agente ha interesse a segnalare costi elevati per far sì che il traffico sia instradato altrove. Un meccanismo “a prova di strategia” potrebbe esser ottenuto facendo pagare a ogni agente  $i$  una somma pari alla lunghezza del cammino più breve che non contiene l’arco  $i$ , meno la lunghezza del cammino più breve in cui il costo dell’arco  $i$  è uguale a zero. Con questo meccanismo la strategia dominante di ciascun agente è di dichiarare il valore reale del costo di inviare un messaggio sul proprio arco, in modo da consentire di individuare il cammino più economico. In pratica questo meccanismo non è tuttavia utilizzato, a causa dei suoi alti costi di comunicazione e di calcolo; il progettista del meccanismo dovrebbe comunicare con tutti gli altri giocatori, e poi risolvere il problema di ottimizzazione; potrebbe valerne la pena se questi costi potessero essere ammortizzati su più messaggi, ma in realtà i costi fluttuano in continuazione per via delle congestioni nel traffico, e perché alcune macchine si scollegano mentre altre andranno online.

### 2.4.3 La ricerca delle azioni migliori

Molti problemi in intelligenza artificiale possono essere risolti cercando fra molte possibili soluzioni. Uno dei più famosi problemi trattati dalla ricerca operativa fin dall’inizio del diciannovesimo è il “problema del commesso viaggiatore” o “*travelling salesman problem*” (TSP) al quale sono state riconosciute numerose soluzioni. Il problema è il seguente: dato un certo numero di città e i costi di spostamento fra una città e l’altra, qual è l’itinerario che consente di visitare una volta tutte le città e tornare in quella di partenza al minimo costo? Si tratta di un problema che comporta elevate difficoltà computazionali, ma per il quale sono stati sviluppati metodi euristici che consentono di ottenere con elevata probabilità soluzioni distanti non più del 3 per cento da quella ottimale anche per milioni di città. Una soluzione euristica molto semplice è quella dell’algoritmo *greedy* o “*algoritmo avido*”: partendo da una città a caso, scegliere ogni volta la città più vicina non ancora visitata; in alcuni casi, tuttavia, quest’algoritmo può generare itinerari molto diversi da quelli ottimali, e, al limite, anche l’itinerario più costoso di tutti. Le strategie di ricerca non informate sono in genere fortemente inefficienti, una strategia di ricerca informata, che utilizza conoscenze specifiche del problema può trovare le soluzioni in modo molto più efficiente. La ricerca “prima il migliore” o “*best-first*”, consiste nella scelta ogni volta della soluzione che allora sembra la migliore. Per esempio, la ricerca *Best-first greedy* o golosa o avida consiste nello scegliere ogni volta il nodo più vicino all’obiettivo finale. Nel caso della scelta di un itinerario da una città di partenza verso una città di arrivo passando per diverse altre città, questa strategia di ricerca potrebbe essere applicata dirigendosi ogni volta verso la città che in linea d’aria è più vicina a quella di destinazione finale. L’algoritmo si chiama “goloso” perché cerca a ogni passo di arrivare il più possibile vicino all’obiettivo; in alcuni casi questo algoritmo può comportare la scelta di itinerari molto più lunghi, e se non ci sono rilevatori di stati ripetuti può esserci il rischio di andare avanti e indietro fra due città all’infinito. L’algoritmo di “ricerca a raggio locale” o ricerca “*local beam*”, comincia invece, con  $k$  stati generati casualmente ed a ogni passo sono generati i successori di tutti i  $k$  stati. Se uno qualsiasi di essi è un obiettivo l’algoritmo termina, altrimenti sceglie i  $k$  successori migliori della lista e ricomincia e nel caso in cui uno stato genera successori migliori degli altri stati l’algoritmo abbandona le ricerche infruttuose e sposta le risorse dove si stanno ottenendo i risultati migliori. Per ridurre il rischio di una troppa rapida concentrazione delle ricerche in uno spazio limitato, il *local beam* invece di scegliere i  $k$  successori migliori, sceglie  $k$  successori casualmente, assegnando ai migliori una maggiore probabilità di essere scelti. Quest’approccio ricorda il processo di selezione naturale: i discendenti di un organismo popolano la generazione successiva in base alle loro capacità di adattamento.

Più o meno simile è ciò che accade con gli algoritmi presenti negli agenti intelligenti online. Nella ricerca online un agente opera alternando computazione e azione: prima esegue un’azione,

poi osserva l'ambiente e determina l'azione successiva. Dopo ogni azione un agente online riceve una percezione che gli comunica lo stato raggiunto; in base a questa percezione può arricchire la mappa dell'ambiente. La versione corrente della mappa viene utilizzata per decidere l'azione successiva: quest'alternanza tra pianificazione e azione fa sì che gli algoritmi di ricerca online siano molto diversi da quelli offline. Un algoritmo online può espandere soltanto il nodo che sta occupando fisicamente; la ricerca in profondità ha questa proprietà, poiché il nodo successivo è sempre figlio di quello appena espanso.

#### 2.4.4 Agenti per il riconoscimento del linguaggio naturale

La comprensione del linguaggio naturale è uno dei campi di ricerca più importanti dell'intelligenza artificiale; esso coinvolge filosofia, linguistica, rappresentazione della conoscenza, ragionamento logico e probabilistico; la comprensione del linguaggio naturale richiede anche analisi empiriche dei comportamenti umani. Gli agenti possono inviarsi segnali (atti linguistici) per informare, avvertire, chiedere aiuto, condividere conoscenze, ecc.; gli esseri umani sembrano essere gli unici animali a usare una grammatica per produrre una varietà illimitata di linguaggi strutturati. L'ambiguità è un problema di grande rilievo per la comprensione del linguaggio naturale; moltissime frasi hanno molte possibili interpretazioni, di cui in genere soltanto una è appropriata; la disambiguazione dipende dalla conoscenza del mondo, dalla situazione corrente, e dall'uso del linguaggio. Gran parte dell'uso del linguaggio avviene nel contesto di frasi multiple; lo studio del discorso si dedica a testi collegati, al fine di risolvere riferimenti pronominali da una frase all'altra, e unire più frasi in segmenti coerenti.

Quando Turing ha proposto il test sull'intelligenza di un agente artificiale lo ha basato sul linguaggio, perché il linguaggio è intimamente legato al pensiero. La comprensione del linguaggio naturale è uno dei campi più importanti dell'intelligenza artificiale; tra le sue influenze si possono includere la filosofia e la linguistica, oltre alla rappresentazione della conoscenza e al ragionamento logico e probabilistico; a differenza di altre aree dell'intelligenza artificiale, la comprensione del linguaggio naturale richiede una investigazione empirica del comportamento umano reale. Gli agenti si possono inviare segnali per diversi scopi: informare, avvertire, richiedere aiuto, condividere conoscenza, promettere qualcosa. Inviare un segnale origina un atto linguistico; tutti gli atti linguistici rappresentano un tentativo di far sì che un altro agente creda o faccia qualcosa<sup>32</sup>. I primi algoritmi al computer per l'analisi sintattica furono dimostrati da Yngve [Yngve, 1955]; algoritmi efficienti sono stati sviluppati alla fine degli anni sessanta e successivamente perfezionati sempre più; Church e Patil ad esempio, hanno affrontato il problema delle ambiguità sintattiche [Church, Patil, 1982].

Uno dei primi studi sulla "disambiguazione" fu la teoria di Wilks sulla "semantica della preferenza", che cerca di trovare le interpretazioni che minimizzano il numero di anomalie semantiche [Wilks, 1975].

Le frasi in linguaggio naturale sono ben diverse dalle formule logiche; in logica, se si affermano le formule A, B, e C a una base di conoscenza, si avrà comunque una congiunzione fra A, B, e C; in linguaggio naturale, invece, l'ordine delle frasi è molto importante.

Il riconoscimento sintattico di un linguaggio non contestuale può essere effettuato utilizzando un "*chart parser*"; è utile aumentare una grammatica per gestire problemi come la concordanza tra oggetto e verbo e il caso dei pronomi. Il formalismo delle grammatiche a clausole definite (DGC = *definite clauses grammar*) permette di esprimere gli aumenti; con una grammatica a clausole definite il riconoscimento sintattico e l'interpretazione semantica (e persino la generazione del linguaggio) possono essere svolti utilizzando l'inferenza logica. Anche l'interpretazione semantica può essere gestita ricorrendo a una grammatica aumentata; una forma quasi-logica può servire da utile passo intermedio tra analisi sintattica e semantica.

---

<sup>32</sup> L'idea del linguaggio come azione nasce con la filosofia del linguaggio del ventesimo secolo (Wittgenstein, 1953), e in particolare con il libro "Speech acts" di Searle (1969)

L'ambiguità è un problema molto grave nell'interpretazione del linguaggio naturale; moltissime frasi possono avere diversi significati, ma di solito soltanto uno è quello appropriato. La soluzione delle ambiguità dipende dalla conoscenza del mondo, dalla situazione corrente, e dall'uso del linguaggio. Gran parte dell'uso del linguaggio avviene nel contesto di frasi multiple; lo studio del discorso si dedica quindi ai testi collegati; è importante sapere come risolvere i riferimenti pronominali fra una frase e l'altra, e come unire più frasi in segmenti coerenti. Mediante l'induzione di grammatiche l'agente potrebbe apprendere una grammatica dagli esempi, ma i diversi test effettuati hanno avuto esiti negativi, essendoci limitazioni quanto si va a generalizzare la grammatica appresa.

Il riconoscimento del parlato è considerato anche un problema di inferenza probabilistica per il fatto che le comunicazioni vocali sono caratterizzate da variazioni nel modo in cui le parole sono pronunciate, rumori di fondo, parole diverse con pronuncia analoga, ecc. L'agente quindi identifica una sequenza di parole pronunciate da un oratore in base al corrispondente segnale acustico. I modelli probabilistici "acustici" di riconoscimento del parlato mirano a massimizzare la probabilità condizionata che una certa parola corrisponda al segnale acustico osservato; i modelli probabilistici "linguistici" di riconoscimento del parlato si basano sulla probabilità a priori di frasi diverse. Il modello "a bigrammi" si basa sulla probabilità che una parola compaia subito dopo un'altra parola. I modelli "acustici" sono quelli più complessi; essi si basano su un'importante coperta nel campo della fonologia (lo studio del linguaggio), secondo cui tutti i linguaggi umani usano un repertorio limitato di circa 40-50 suoni chiamati "foni". Approssimativamente un fono corrisponde a una singola vocale o consonante, alcune combinazioni di lettere producono a volte tuttavia foni distinti, e alcune lettere producono foni diversi in contesti diversi. Il modello inferenziale tipicamente utilizzato per il riconoscimento del parlato è il modello di Markov nascosto. Un modello di Markov nascosto è un modello statistico in cui viene modellizzato un processo markoviano con parametri sconosciuti; esso può essere considerato il più semplice modello di rete bayesiana dinamica [Markov, 1971; Russel, Norvig, 2005]

I sistemi di riconoscimento vocale utilizzano insiemi di dati giganteschi, ed enormi risorse computazionali. Nel riconoscimento di parole isolate in buone condizioni acustiche, con un vocabolario di poche migliaia di parole e un singolo osservatore, l'accuratezza può superare il 99 per cento. Nel caso di parlato continuo e senza restrizioni e una varietà di oratori, l'accuratezza di solito varia fra il 60 e l'80 per cento, anche in buone condizioni acustiche; l'accuratezza è ancora minore nel caso di trasmissioni telefoniche con rumori di fondo. Alcuni ricercatori hanno esplorato la possibilità di utilizzare reti bayesiane dinamiche invece del modello bayesiano nascosto, al fine di sfruttare il maggior potere espressivo delle reti dinamiche bayesiane, e catturare una parte maggiore del complesso stato nascosto dall'apparato che produce il parlato [Zweig e Russel, 1998; Richardson et al, 2000]. Questi tentativi non sembrano tuttavia avere avuto fino ad ora successo, così che il modello nascosto di Markov rimane ancora il modello di gran lunga più utilizzato per il riconoscimento automatico del parlato.

Una delle aree più rilevanti per quel che riguarda le utilizzazioni commerciali dei sistemi artificiali di riconoscimento vocale è la sanità, in particolare negli Stati Uniti. Altre aree di applicazione sono l'aeronautica, in particolare per gli aerei da combattimento, e per l'addestramento dei controllori di volo, la telefonia (i programmi di trasferimento voce-testo richiedono tuttavia troppa memoria per essere implementati efficacemente sui cellulari, anche se Symbian e Windows hanno ideato dei programmi di riconoscimento vocale per smartphones), le persone diversamente abili, le traduzioni automatiche, il controllo delle automobili, la telematica, la documentazione dei processi nei tribunali, l'automazione domestica, la posta elettronica, l'interazione multimediale, la robotica, ecc.

### 2.4.5 Agenti a supporto delle decisioni di acquisto online

Tra gli agenti di ricerca realizzati per applicazioni in ambienti online, risultano di particolare importanza gli agenti di supporto alle decisioni di acquisto. L'obiettivo del nostro lavoro è infatti, quello di semplificare il compito arduo di ogni consumatore al momento di acquistare un prodotto o servizio. Quindi, ciò che riguarda la codificazione della conoscenza relativa allo shopping su internet, mediante la costruzione di un agente di ricerca in grado di aiutare un acquirente a trovare in rete offerte di prodotti. Negli ambienti online, quando l'acquirente fornisce una descrizione del prodotto, l'agente per lo shopping ha il compito di produrre una lista di pagine web che ne offrono la vendita. Se la descrizione del prodotto è precisa, il compito dell'agente sarà di cercare i negozi che offrono i prezzi più convenienti; se la descrizione è parziale, l'agente dovrà confrontare prodotti diversi. Ovviamente, l'ambiente dell'agente per lo shopping in internet è costituito dall'intero World Wide Web, quindi un ambiente complesso e in costante evoluzione utilizzato ogni giorno da milioni di persone. Le percezioni dell'agente sono pagine web; mentre l'utente le vedrebbe presentate in forma grafica, l'agente le percepisce come stringhe di caratteri comprendenti parole comuni mescolate a etichette del linguaggio HTML (Hyper Text Markup Language)<sup>33</sup>. L'agente quindi, deve essere in grado di estrarre informazioni utili da questo tipo di percezioni; non sempre ciò è facile, se si considera che le pagine web possono includere cookie, codici Java, Javascript e Flash, protocolli di esclusione dei softbots, HTML scorretto, file sonori, filmati e testo integrato nelle immagini grafiche. Il primo compito dell'agente è trovare offerte di prodotti rilevanti che siano effettivamente in vendita; in genere conviene cominciare dalla home page di un negozio online e considerare tutte le pagine che possono essere raggiunte seguendo collegamenti rilevanti, che a un certo punto giungeranno fino alle offerte. Una pagina è rilevante se può essere raggiunta da una catena di collegamenti di categoria rilevanti partendo dalla home page di un negozio online e se da lì si può raggiungere un'offerta di prodotto. Per riuscire a seguire i collegamenti rilevanti è necessario conoscere una complessa gerarchia di prodotti e non è facile elencare tutte le possibili categorie dello shopping perché un acquirente potrebbe sempre escogitare un nuovo desiderio, e le aziende cercheranno continuamente di inventare nuovi prodotti. In genere, una ontologia che comprende un migliaio di prodotti potrà soddisfare le necessità di gran parte dei potenziali acquirenti [Russel, Norvig, 2005]. Problemi possono derivare dal fatto che a volte la stessa categoria ha più nomi (sinomia), e lo stesso nome si riferisce a più categorie (omonimia o ambiguità). Un problema ancora più rilevante è rappresentato dall'enorme varietà delle descrizioni che un utente può inserire o dei nomi di categoria utilizzati dai diversi negozi. La sinomia e l'ambiguità possono provocare un rilevante aumento del numero di cammini dell'agente, e a volte può rendere difficile determinare se una pagina è effettivamente rilevante. Difficoltà maggiori possono derivare dalla grande varietà di descrizioni che un utente può inserire e dalla loro possibile discrepanza rispetto alle denominazioni utilizzate dai negozi online.

Si supponga a questo punto che l'agente abbia "prodotto" un certo numero di pagine a cui corrisponde altrettante offerte per il tipo di prodotto cercato. Per confrontarle dovrà innanzitutto estrarvi le informazioni rilevanti per la scelta: in particolare i prezzi e le diverse caratteristiche qualitative dei prodotti; i programmi wrapper in genere consentono ai softbots di estrarre informazioni dalle pagine web. Altri problemi possono derivare dal fatto che di solito il prezzo dipende dal metodo di pagamento o dall'eventuale diritto di alcuni clienti di ottenere degli sconti. Una volta ottenute le diverse offerte l'agente dovrà confrontarle; un semplice metodo di confronto consiste nell'esclusione delle offerte dominate, vale a dire le offerte uguali o migliori delle altre per gran parte degli aspetti e peggiori per alcuni aspetti. In genere alla fine di questo

---

<sup>33</sup> HTML é il "markup language" predominante per le pagine web. Esso fornisce uno strumento di descrizione della struttura delle informazioni testuali in un documento, indicando collegamenti, intestazioni, paragrafi, liste, ecc., e aggiungendo al testo forme interattive, immagini, codici, ecc. HTML fu ideato verso la fine degli anni ottanta da Tim Berners-Lee e Robert Cailliau presso il CERN di Ginevra.

tipo di confronto rimane ancora un rilevante numero di offerte, ciascuna delle quali è migliore di altre per alcuni aspetti e peggiore per altri aspetti.

Per scegliere fra tante offerte residue l'agente dovrebbe conoscere il grado d'importanza che per l'utente hanno i diversi attributi del prodotto considerato.

## **2.5 L'apprendimento automatico degli agenti intelligenti**

L'idea alla base dell'apprendimento è che le percezioni non dovrebbero essere utilizzate soltanto per decidere quale azione intraprendere, ma anche per migliorare la capacità futura dell'agente. L'agente apprende quando osserva le proprie interazioni con il mondo e i suoi processi decisionali, dalla semplice memorizzazione delle esperienze, fino alla formulazione di vere e proprie teorie scientifiche; l'apprendimento induttivo deriva dalle osservazioni.

L'analisi dei processi di apprendimento degli agenti può essere suddivisa in quattro parti: una riguarda l'apprendimento dalle osservazioni, in cui vengono descritti agenti in grado di migliorare il proprio comportamento attraverso l'analisi dell'esperienza; una riguarda la conoscenza con l'apprendimento, in cui viene esaminato il problema di come apprendere quando si sa già qualcosa; una riguarda i metodi di apprendimento statistico, in cui si considera l'apprendimento come una forma di ragionamento incerto sulle osservazioni; un'altra parte dell'analisi considera l'apprendimento per rinforzo, in cui si esamina come un agente può imparare dalle ricompense e dalle punizioni che ottiene.

Un agente in grado di apprendere contiene sia un elemento esecutivo, che decide quali azioni intraprendere, sia un elemento di apprendimento, che modifica quello esecutivo in modo da prendere decisioni sempre migliori. I ricercatori nel campo dell'apprendimento automatico hanno escogitato una grande varietà di elementi di apprendimento; per comprenderli è utile esaminare come la loro progettazione dipende dal contesto in cui gli agenti operano. Il progetto di un elemento di apprendimento è influenzato da tre aspetti principali: quali componenti dell'elemento esecutivo dovranno essere appresi; il tipo di feedback disponibile per l'apprendimento; la tipologia di rappresentazione utilizzata. I principali progetti di agente includono in genere un meccanismo per far corrispondere direttamente le condizioni associate allo stato corrente delle azioni; un metodo per dedurre proprietà importanti del mondo partendo dalla sequenza delle percezioni; informazioni circa le modalità di evoluzione del mondo e i risultati delle possibili azioni; informazioni sull'utilità di diversi stati del mondo; informazioni sul valore delle azioni che indicano quanto ognuna di esse sia desiderabile; obiettivi che descrivono classi di stati il cui raggiungimento massimizza l'utilità dell'agente.

Il tipo di feedback disponibile è il fattore più importante per determinare la natura del problema di apprendimento affrontato dall'agente; da questo punto di vista possono essere distinti tre casi: apprendimento supervisionato, apprendimento non supervisionato; apprendimento per rinforzo. Nel caso dell'apprendimento supervisionato si apprende una funzione partendo da esempi di input e output. L'apprendimento non supervisionato consiste nell'imparare a riconoscere *pattern*, schemi, o regolarità nell'input senza alcuna indicazione specifica dei valori in uscita; un agente in grado di apprendere in modo non supervisionato non può imparare cosa fare, perché non possiede alcuna informazione su cosa costituisce un'azione corretta o uno stato desiderabile. Nel caso dell'apprendimento per rinforzo, il più generale, invece di sentirsi dire cosa deve fare da un insegnante, un agente deve apprendere basandosi sul "rinforzo", vale a dire su premi o punizioni in conseguenza di comportamenti desiderabili o non desiderabili.

Fattore importante per la progettazione di sistemi capaci di apprendere è la disponibilità di conoscenza pregressa; gran parte della ricerca in intelligenza artificiale, in informatica e in psicologia ha considerato casi in cui l'agente ha cominciato a operare senza la minima conoscenza di ciò che sta cercando di imparare, e ha accesso soltanto agli esempi che gli si presentano in conseguenza delle sue esperienze. Benché questo sia un caso speciale molto importante, esso non è generale; gran parte dell'apprendimento umano avviene in contesti in cui

è presente una grande quantità di conoscenza di fondo; alcuni psicologi e linguisti sostengono addirittura che anche i neonati possiedono una certa conoscenza del mondo; in ogni caso la conoscenza pregressa può aiutare molto l'apprendimento.

Un algoritmo per l'apprendimento deterministico supervisionato riceve in ingresso il valore corretto di una funzione sconosciuta in alcuni punti e deve cercare di ricostruirla nel modo più fedele possibile<sup>34</sup>. Il compito fondamentale dell'inferenza induttiva pura (o induzione) è il seguente: data una collezione di esempi di una certa funzione  $f$ , restituire una funzione  $h$  che approssima  $f$ . La funzione  $h$  è chiamata ipotesi, il motivo per cui l'apprendimento è un'attività concettualmente difficile è che non è facile stabilire se una particolare funzione  $h$  è una buona approssimazione di  $f$ . Requisito fondamentale di una buona ipotesi è che essa possa generalizzarsi bene, vale a dire che possa predire correttamente esempi non ancora incontrati; è questo il problema fondamentale dell'induzione. Un esempio è costituito dal problema di individuare una funzione con una sola variabile esplicativa in grado di rappresentare una ipotesi consistente (vale a dire in grado di rappresentare i dati osservati).

### **2.5.1 L'apprendimento automatico con alberi decisionali**

L'induzione di alberi di decisione<sup>35</sup> è una delle forme di apprendimento più semplici e di maggior successo. Un albero di decisione prende in input un oggetto o una situazione descritta da un insieme di attributi, e restituisce una decisione, ovvero il valore predetto di uscita per tale input. L'apprendimento di una funzione a valori discreti prende nome di "classificazione"; nel caso continuo si usa il termine "regressione". L'attenzione è concentrata spesso sulla classificazione booleana, in cui ogni esempio è classificato come vero o falso. Un albero di decisione funziona eseguendo in sequenza una serie di test; ogni nodo interno corrisponde a un test sul valore di una delle proprietà, e le diramazioni uscenti dal nodo sono etichettate con tutti i possibili risultati; ogni foglia specifica il valore da fornire in uscita. Gli alberi di decisione sono completamente espressivi all'interno della classe dei linguaggi proposizionali, nel senso che non esiste alcuna funzione booleana che non possa essere scritta sotto forma di albero di decisione; per molti funzioni, tuttavia, sarebbero necessari alberi di decisione esponenzialmente grandi. Un esempio per un albero di decisione booleano consiste in un vettore di attributi in input e un singolo valore booleano in output; gli esempi positivi sono quelli in cui l'obiettivo è vero, quelli negativi quelli in cui l'obiettivo è falso; l'insieme completo degli esempi prende il nome di insieme di addestramento o di training. Una semplice euristica che può consentire di trovare un albero di decisione relativamente piccolo che si accorda con l'insieme di addestramento è l'algoritmo apprendimento-albero-decisione; l'idea base di questo algoritmo è di verificare per primo il valore dell'attributo più importante, vale a dire quello che ha il maggiore impatto sulla classificazione; in questo modo in genere si arriva alla classificazione corretta con un numero

---

<sup>34</sup> Sono stati proposti diversi approcci algoritmici per l'apprendimento. La tecnica della migliore ipotesi corrente tiene in memoria una singola ipotesi, specializzandola quando appare troppo generica, e generalizzandola quando appare limitata; secondo le prime ricerche di psicologia cognitiva (Bruner e altri, 1957), questa tecnica sembra essere una forma naturale di apprendimento umano dei concetti. In intelligenza artificiale questo concetto è stato sviluppato da Winston (1970), che affrontò il problema dell'apprendimento della descrizione di oggetti complessi. Il metodo dello "spazio delle versioni" (Mitchell, 1977, 1982), invece, tiene in memoria l'intero insieme delle ipotesi consistenti, ed elimina gradualmente quelle che si rivelano inconsistenti con i nuovi esempi. Questo approccio è stato utilizzato nel sistema esperto Meta-DENDRAL per la chimica, nel sistema LEX, capace di risolvere problemi di analisi matematica, nella serie di algoritmi AQ, che apprendono insiemi di regole logiche. L'approccio della "identificazione al limite" (Gold, 1967), ispirato a modelli di scoperta scientifica tratti dalla filosofia della scienza (Popper, 1962), è stato applicato principalmente all'apprendimento di grammatiche partendo da esempi di frasi.

<sup>35</sup> L'EPAM (*Elementary Perceiver And Memorizer*), un modello cognitivo-simulativo dell'apprendimento di concetti da parte delle persone proposto da Feigenbaum (1961), fu uno dei primi sistemi a utilizzare alberi di decisione (o reti di discriminazione). ID3, proposto da Quinlan (1979), introdusse l'idea fondamentale di utilizzare il contenuto informativo come base della funzione euristica. Per quel che riguarda la letteratura statistica sull'apprendimento di alberi di decisione un riferimento importante è *Classification and Regression Trees o CART*, Breiman e altri, (1984).

piccolo di prove. Naturalmente l'albero di decisione che si ottiene per induzione applicando l'algoritmo all'insieme di addestramento spesso sarà diverso dall'albero da cui sono stati generati i dati, poiché l'algoritmo lavora sugli esempi e cerca quindi un albero di decisione che si accorda con gli esempi; utilizzando più esempi di solito si ottiene per induzione un albero di decisione più simile all'albero originale.

Lo schema utilizzato per selezionare gli attributi nell'apprendimento di alberi di decisione è progettato in modo da minimizzare la profondità dell'albero risultante; l'idea è scegliere l'attributo che fornisce la classificazione più esatta possibile degli esempi: un attributo perfetto suddivide gli esempi in insiemi tutti positivi o tutti negativi; un attributo è inutile se dà come risultati degli insiemi che hanno più o meno la stessa proporzione di esempi positivi e negativi dell'insieme originale. Un algoritmo di apprendimento è buono se produce ipotesi che riescono a predire con accuratezza la classificazione di esempi mai incontrati in precedenza; è ovvio che una predizione è buona se si rivela esatta, per cui è possibile valutare la qualità di un'ipotesi confrontando le sue predizioni con le classificazioni corrette; questo viene fatto su un insieme di esempi noto come insieme di test. Per capire quanto deve essere grande il guadagno di informazione per dividere l'albero in base a un determinato attributo, può essere utile un test di significatività statistica. Si comincia ipotizzando che non ci sia alcun pattern nei dati (ipotesi nulla), e si analizzano i dati per vedere in che misura essi si discostano dalla totale assenza di regolarità. Se il grado di deviazione risulta statisticamente improbabile (normalmente a livello-soglia di probabilità del 5 per cento o meno), ciò viene considerato un buon indizio della presenza di un pattern significativo nei dati. Le probabilità sono calcolate in base alle distribuzioni standard delle deviazioni che ci si aspetterebbe di misurare in un campionamento casuale; in questo caso l'ipotesi nulla è che il campione sia irrilevante, e che quindi il guadagno di informazione per un campione infinitamente grande sia pari a zero. In presenza di rumore il rumore può essere tollerato: gli errori di classificazione causano un incremento lineare dell'errore di predizione, mentre quelli nella descrizione degli esempi hanno un effetto asintotico che peggiora man mano che l'albero si suddivide in insiemi più piccoli e quindi più facili da comprendere.

Nel caso di attributi caratterizzati da valori continui, invece di generare un numero infinito di rami, gli algoritmi di apprendimento di alberi di decisione in genere trovano un punto di divisione (*split point*) che dà il maggior guadagno di informazione. Se invece di una classificazione discreta si cerca di predire un valore numerico (attributi di output a valori continui), è opportuno ricorrere a un albero di regressione, che in ogni foglia contiene non già un singolo valore ma una funzione lineare di un sottoinsieme degli attributi numerici. L'algoritmo di apprendimento deve decidere quando terminare le divisioni dell'albero e cominciare ad applicare la regressione lineare utilizzando gli attributi rimanenti. Le variabili a valori continui rappresentano un problema particolarmente importante, perché sia i processi fisici che i modelli finanziari forniscono dati numerici. Sono stati costruiti diversi pacchetti commerciali che soddisfano questi criteri, con i quali sono stati sviluppati moltissimi sistemi. Nel quarto capitolo presenteremo Microsoft Analysis Manager, un sistema esperto che integra algoritmi di clustering e di decision trees che ci sono stati molto utili per effettuare analisi empiriche che riguardavano in particolar modo il mercato della telefonia mobile. In diverse aree industriali e commerciali gli alberi di decisione sono normalmente il primo metodo sperimentato quando si deve estrarre un metodo di classificazione da un insieme di dati. Una proprietà importante degli alberi di decisione è che, a differenza delle reti neurali, un essere umano può comprendere l'output dell'algoritmo di apprendimento (questo è un requisito legale per le decisioni finanziarie).

### **2.5.2 L'apprendimento automatico per conoscenze**

L'approccio moderno è rivolto alla progettazione di agenti che fanno già alcune cose, e cercano di impararne altre. Un agente di questo tipo potrebbe anche partire da zero, iniziando con dei programmi di induzione pura; dopo avere acquisito alcune conoscenze, l'agente dovrà

tuttavia essere in grado di utilizzarle per imparare altre cose in modo più efficiente. Presentiamo alcuni esempi molto semplici di come le conoscenze pregresse possono facilitare l'apprendimento; per esempio, una semplice vignetta che raffigura un uomo primitivo che arrostisce un animale dopo averlo infilato in un bastone appuntito può essere sufficiente a far apprendere una modalità di cottura relativamente efficace; un viaggiatore che per la prima volta arriva in Brasile, sentendo parlare portoghese il primo brasiliano che incontra, può inferire che tutti i brasiliani parlano portoghese; uno studente in medicina, ascoltando un paziente descrivere i sintomi di una malattia e il farmaco successivamente prescritto dal medico, inferisce che quel farmaco è efficace nel trattamento della malattia che presenta quei sintomi. Nel caso della cottura degli animali infilati in un bastone si ha un processo di generalizzazione basato sulle spiegazioni (EBL = *explanation based learning*): il bastone appuntito consente di tenere l'animale sul fuoco, pur mantenendo la mano lontana dal fuoco; l'apprendimento basato sulle spiegazioni è un metodo alquanto efficace per convertire i principi base delle teorie in utile conoscenza specializzata. Nel caso del viaggiatore che apprende la lingua parlata in Brasile, la conoscenza a priori rilevante è che in ogni paese, gran parte delle persone parlano la stessa lingua; la generalizzazione dalla lingua parlata da una persona a quella parlata nel paese in cui quella persona si trova, rappresenta un apprendimento basato sulla rilevanza (RBL = *relevance based learning*). Nell'esempio dello studente, si presume che egli sia in grado, in base a conoscenze pregresse, di riconoscere la malattia di cui il paziente descrive i sintomi, e, sulla base della prescrizione del medico, apprende con quali farmaci essa in generale può essere curata; si parla in questo caso di apprendimento induttivo basato sulla conoscenza (KBIL = *knowledge based inductive learning*) [Russel, Norvig, 2005, vol. 2].

### 2.5.3 L'apprendimento automatico statistico

L'apprendimento statistico è una forma di ragionamento incerto sulle osservazioni; gli agenti possono gestirla ricorrendo a metodi probabilistici e alla teoria delle decisioni, ma prima è necessario imparare dall'esperienza una teoria probabilistica del mondo. I metodi di apprendimento statistico vanno dal semplice calcolo delle medie alla costruzione di modelli complessi come le reti bayesiane e le reti neurali; il loro campo di applicazione comprende l'informatica, l'ingegneria, la neurobiologia, la psicologia, e la fisica. L'apprendimento di "massima verosimiglianza" sceglie l'ipotesi che massimizza la verosimiglianza dei dati; essa è equivalente all'ipotesi massima a posteriori, con distribuzione a priori uniforme. In casi semplici, come la regressione lineare in reti bayesiane completamente osservabili, soluzioni di massima verosimiglianza possono essere trovate facilmente in forma chiusa. Attualmente l'algoritmo *expectation maximization*, che in effetti può essere considerato uno schema, dal momento che prima di essere utilizzabile a una nuova famiglia di distribuzioni richiede molto lavoro matematico, è uno degli algoritmi più utilizzati al mondo. In intelligenza artificiale, il primo sistema che ha applicato con successo l'algoritmo *expectation maximization* per la modellizzazione di miscugli è stato AUTOCLASS, utilizzato per diverse applicazioni scientifiche concrete (scoperta di nuovi tipi di stelle partendo dai dati spettrali, e di nuove classi di proteine nei database di sequenze DNA/proteine (Hunter e States, 1992).

L'apprendimento statistico è un'area di ricerca molto attiva, sono stati fatti grandi passi avanti sia in teoria che in pratica, tanto che oggi è possibile apprendere quasi ogni modello che consente un'inferenza esatta o approssimata. Alla fine degli anni novanta si è verificata una tendenziale convergenza tra l'apprendimento automatico, la statistica e le reti neurali, incentrata sui metodi per la creazione di grandi modelli probabilistici partendo dai dati.

L'apprendimento bayesiano calcola la probabilità di ogni ipotesi condizionandola ai dati osservati, e su tale base formula predizioni; le predizioni sono quindi basate su tutte le ipotesi, pesate secondo la rispettiva probabilità, e non soltanto su quella "migliore"; l'apprendimento si riduce quindi a un problema di inferenza probabilistica. Una prima caratteristica fondamentale dell'apprendimento bayesiano è che per qualsiasi distribuzione a priori che non esclude l'ipotesi

vera, la probabilità a posteriori di ciascuna ipotesi falsa prima o poi andrà a zero, poiché la probabilità di generare dati “non caratteristici” per un tempo indefinito tende a zero; altra importantissima caratteristica della predizione bayesiana è che essa è ottima, indipendentemente dalle dimensioni dell’insieme dei dati; ciò significa che essa sarà corretta più spesso di qualsiasi altra predizione. Tuttavia, nei problemi di apprendimento reali, lo spazio delle ipotesi è generalmente troppo grande, o addirittura infinito; di conseguenza in gran parte dei casi è necessario ricorrere a metodi approssimati.

L’apprendimento di massima verosimiglianza, molto comune in statistica per la diffidenza nei confronti di attribuzioni soggettive di probabilità, appare opportuno quando non ci sono ragioni per preferire a priori una ipotesi invece di un’altra, per esempio quando tutte le ipotesi hanno la stessa complessità. L’apprendimento di massima verosimiglianza rappresenta una buona approssimazione quando l’insieme dei dati è grande, perché in questo caso la distribuzione a priori conta poco rispetto ai dati. L’apprendimento di massima verosimiglianza permette di utilizzare procedure molto semplici, ma evidenzia gravi carenze quando gli insiemi di dati sono piccoli; per esempio, dopo avere osservato un certo tipo di caramella estratta da un sacchetto l’ipotesi di massima verosimiglianza è che tutte le caramelle siano di quel tipo; questa inferenza appare poco ragionevole, a meno che la distribuzione a priori delle ipotesi non preveda un solo tipo di caramelle. L’apprendimento bayesiano invece, non richiede alcun principio aggiuntivo, nel senso che in fondo c’è un solo processo di apprendimento: quello che esegue l’inferenza nelle reti bayesiane.

#### **2.5.4 L’apprendimento automatico mediante reti neurali**

Un neurone è una cellula cerebrale la cui funzione principale è raccogliere, elaborare, e propagare segnali elettrici; la capacità del cervello di elaborare informazioni deriva principalmente da reti di neuroni. Le prime ricerche di intelligenza artificiale avevano l’obiettivo di creare reti neurali artificiali (connessionismo, elaborazione parallela distribuita, computazione neurale, ecc.) Un semplice modello matematico di un neurone fu proposto da McCulloch e Pitts (1943), in cui il neurone scatta quando una combinazione lineare dei suoi input supera una certa soglia. Negli ultimi decenni, nell’ambito delle neuroscienze computazionali, sono stati sviluppati modelli sempre più dettagliati e realistici, sia per singoli neuroni che per sottosistemi complessi del cervello.

Le reti neurali rimangono una delle forme più popolari ed efficaci di sistemi di apprendimento. In sintesi, le reti neurali sono costituite da nodi o unità, unite da collegamenti diretti, che propagano l’attivazione fra i diversi nodi; a ogni collegamento è associato un peso numerico, che determina la forza e il segno della connessione; successivamente ogni unità calcola innanzitutto una somma pesata dei propri input, e poi applica una “funzione di attivazione” alla somma ottenuta per derivare l’output.

Le reti neurali sono capaci di svolgere compiti di apprendimento molto complessi; esistono già nella letteratura scientifica decine di migliaia di pubblicazioni su specifiche applicazioni delle reti neurali. Così come per le reti bayesiane, se si sceglie una rete neurale troppo grande essa potrà memorizzare tutti gli esempi formando una grande tabella, ma non sarà probabilmente in grado di generalizzare bene e predire con accuratezza il valore della funzione per input che non ha già incontrato; ciò significa che, come tutti i modelli statistici, anche le reti neurali sono soggette a problemi di sovradattamento quando nel modello sono presenti troppi parametri; i modelli con molti parametri si accordano bene con i dati, ma potrebbero non generalizzare altrettanto bene dei modelli con pochi parametri.

Il riconoscimento di cifre scritte a mano, ad esempio, è un problema importante per molte applicazioni, tra cui la suddivisione automatica della posta in base al codice postale, la lettura automatica di disegni, l’inserimento di dati nei computer, ecc. In quest’area rapidi progressi sono stati ottenuti grazie allo sviluppo di algoritmi di apprendimento. L’approccio più semplice è il classificatore ai 3 vicini più prossimi, che ha il vantaggio di non richiedere tempo di

addestramento; essendo un algoritmo basato sulla memoria, deve memorizzare moltissime immagini e la sua esecuzione è piuttosto lenta; la sua frequenza di errore è di poco superiore al 2 per cento. Una rete neurale, utilizzando la validazione incrociata, con circa 300 unità nascoste e interconnessioni complete fra gli strati ha consentito di ridurre la frequenza di errore a meno del 2 per cento. Una serie di reti neurali specializzate (“LeNet”) sono state progettate in modo da sfruttare la struttura del problema, e cioè la conoscenza del fatto che l’input consiste in un insieme di pixel disposti in una matrice bidimensionale, e che piccoli cambiamenti nella posizione o nell’inclinazione di un’immagine non sono importanti; ciò ha consentito di ridurre la frequenza degli errori a meno dell’1 per cento. Questi risultati appaiono notevoli, se si considera che le persone in genere si stima che abbiano in media una frequenza di errore compresa fra lo 0,2 e lo 0,3 per cento.

### **2.5.5 L’apprendimento automatico per rinforzo**

Un agente può imparare anche dai suoi successi e dai suoi fallimenti, dalle ricompense e dalle punizioni, questo si chiama apprendimento “*per rinforzo*”. Mentre l’apprendimento dalle osservazioni e l’apprendimento statistico riguarda metodi di apprendimento che consentono di apprendere funzioni e modelli di probabilità partendo dagli esempi, l’apprendimento per rinforzo riguarda le modalità con cui gli agenti possano imparare cosa fare; in altre parole si tratta di capire come un agente può imparare a comportarsi in modo efficace in un ambiente sconosciuto, date le sue percezioni e occasionali ricompense<sup>36</sup>.

Quando l’agente ha il dovere di scegliere azioni già durante l’apprendimento, esso deve gestire il trade-off tra il valore stimato delle azioni e la possibilità di apprendere nuove informazioni utili; non è possibile trovare una soluzione esatta al problema dell’esplorazione, ma esistono semplici euristiche che permettono di ottenere risultati accettabili. Considerato che potenzialmente permette di eliminare la necessità di codificare a mano le strategie di controllo, l’apprendimento per rinforzo continua a essere una delle aree di ricerca più attive nel campo dell’apprendimento automatico; particolarmente preziose promettono di essere le applicazioni alla robotica; in questo campo infatti devono essere gestiti ambienti continui, in cui i comportamenti di successo possono essere composti da milioni di azioni elementari. L’apprendimento per rinforzo si basa sull’idea che un agente può apprendere cosa fare se in seguito alle sue azioni gli si fornisce qualche tipo di feedback (ricompensa o rinforzo) che gli consenta di capire ciò che è buono e ciò che non è buono, così da consentirgli di scegliere opportunamente le sue azioni; è importante che le ricompense non soltanto siano percepite dall’agente, ma che esse siano considerate come conseguenza delle sue azioni. Il compito dell’apprendimento per rinforzo è di utilizzare le ricompense osservate per ricostruire una politica ottima per l’ambiente, anche in situazioni in cui l’agente non ha una conoscenza pregressa né dell’ambiente né della funzione di ricompensa<sup>37</sup>. Per esempio, si può ipotizzare di partecipare a un gioco di cui non si conoscono le regole, in cui a un certo punto ci si rende conto di aver perso o vinto; in base a queste sole informazioni, l’agente deve cercare di apprendere una funzione di valutazione che gli consenta di ottenere stime ragionevoli della sua probabilità di vittoria nelle diverse circostanze.

---

<sup>36</sup> Turing (1948, 1950) fu uno dei primi a proporre applicazioni dell’apprendimento per rinforzo in intelligenza artificiale, sebbene non fosse molto convinto della sua efficacia. Nell’ambito della psicologia, la letteratura dell’apprendimento per rinforzo risale ai primi decenni del ventesimo secolo; sviluppi particolarmente importanti sono dovuti a Skinner (1953) nell’ambito della sua nozione di “apprendimento operante”. Uno dei primi lavori di successo nel campo dell’apprendimento automatico fu quello di Athur Samuel (1959); esso includeva diverse idee moderne sull’apprendimento per rinforzo, incluso il calcolo alle differenze temporali e l’approssimazione della funzione. Sutton e Barto (1998) hanno mostrato come l’apprendimento per rinforzo unisca concetti di apprendimento, pianificazione, e azione.

<sup>37</sup> Il collegamento fra apprendimento per rinforzo e processi decisionali di Markov fu enunciato da Werbos (1977); lo sviluppo dell’apprendimento per rinforzo in intelligenza artificiale scaturisce dal lavoro svolto nei primi anni ottanta presso l’università del Massachusetts da Barto e altri (1981, 1983) e da Sutton (1988, 1990).

### 2.5.6 Il comportamento ottimale di un agente intelligente

Il comportamento di un agente che apprende in modo passivo è predeterminato da una politica fissa; un agente attivo, invece, deve decidere quali azioni intraprendere, al fine di apprendere un modello completo, con le probabilità di tutti gli esiti delle azioni, e non soltanto il modello della politica prefissata. Un agente che a ogni passo segue la politica ottima per il modello appreso (agente goloso o *greedy*), rischia di non apprendere la vera politica ottima; ciò perché il modello appreso può non coincidere con l'ambiente vero, e ciò che appare ottimo nel modello potrebbe non esserlo nell'ambiente reale. L'agente goloso non tiene conto del fatto che le azioni non si limitano a fornire ricompense in base al modello appreso corrente, ma contribuiscono anche ad apprendere il modello vero influenzando le percezioni ricevute; migliorando il modello l'agente riceverà ricompense maggiori in futuro; un agente deve quindi cercare un compromesso fra lo sfruttamento, che massimizza la ricompensa calcolata in base alle stime correnti di utilità, e l'esplorazione, che può consentire di massimizzare l'utilità a lungo termine; concentrarsi sullo sfruttamento delle conoscenze correnti rischia di portare in un vicolo cieco, la continua esplorazione migliora le conoscenze ma serve a poco se queste non sono mai sfruttate. Ad esempio nella realtà ci si trova continuamente di fronte alla scelta fra la continuazione di una vita abbastanza confortevole, o il tentativo di cambiare nella speranza di scoprire opportunità migliori. Uno schema ragionevole di comportamento ottimale in termini di sfruttamento delle conoscenze già acquisite e l'esplorazione per migliorare le conoscenze è lo schema GLIE (*greedy in the limit of infinite exploration*), che tende a diventare goloso quando l'esplorazione tende a infinito. Una specificazione semplice, ma molto lenta di uno schema GLIE prevede che l'agente scelga un'azione casuale in una certa frazione dei passi, e segua la politica golosa nei rimanenti passi; un approccio più efficace prevede di assegnare un peso maggiore alle azioni che sono state provate di meno, ciò può essere sintetizzato in una funzione di esplorazione, che specifica i termini del *trade-off* tra la golosità e la curiosità.

In generale, l'apprendimento per rinforzo con approssimazione generica delle funzioni rimane un'arte alquanto difficile<sup>38</sup>.

La prima applicazione significativa dell'apprendimento per rinforzo fu il anche primo programma di apprendimento in assoluto: il sistema per giocare a dama di Athur Samuel [Samuel, 1959, 1967]; il sistema TD-Gammon di Gerry Tesauro [Tesauro, 1992] ha illustrato le grandi potenzialità delle tecniche di apprendimento per rinforzo: successivi perfezionamenti consentirono di raggiungere standard di gioco paragonabili a quelli dei più forti campioni umani.

## 2.6 Conclusioni

L'interazione con sistemi automatici di elaborazione delle informazioni coinvolge oggi tramite Internet una parte significativa della popolazione in tanti paesi; in questo ambito agenti intelligenti artificiali riescono a muoversi e interagire, anche con esseri umani. Molte ricerche, svolte sia nelle università che in imprese, pubbliche e private, nel settore della new economy e del commercio elettronico, hanno evidenziato la grande rilevanza di questo fenomeno. La ricerca teorica ha cercato di dare all'intelligenza artificiale un apparato formale di riferimento che possa permettere la progettazione di sistemi intelligenti artificiali sempre più sofisticati; i principali campi di indagine sono lo studio di logiche specifiche per formalizzare e analizzare il ragionamento automatico degli agenti intelligenti; riuscire a percepire e rappresentare il mondo esterno, decidere in modo razionale attraverso l'utilizzo di tecniche decisionali in situazioni in cui le informazioni sono incomplete o incerte e quindi apprendere in modo automatico

---

<sup>38</sup> Le reti neurali sono la forma più diffusa di approssimazione di funzioni; un problema evidenziato dai sistemi alle differenze temporali basati su reti neurali è che tendono a dimenticare le esperienze più antiche; ciò può determinare fallimenti catastrofici nel caso queste circostanze dovessero ripresentarsi; l'approssimazione della funzione ottenuta attraverso l'apprendimento basato sulle istanze può evitare questo problema (Forbes, 2000).

servendosi addirittura di sistemi biologici, come le reti neurali. Tutto ciò stimola l'integrazione di tecniche diverse in architetture di agenti intelligenti, in grado di svolgere compiti specifici sia nel mondo fisico (*robot*) che nel mondo virtuale di internet (*softbot*). Con riferimento a internet, ad esempio, un tema di grande attualità è quello dei sistemi multi-agenti, con le problematiche connesse alla loro coordinazione e alla condivisione di conoscenza.

Illustrando un approccio interdisciplinare tipico delle applicazioni di intelligenza artificiale, in questo capitolo è stato evidenziato come può essere utile per un agente intelligente fare riferimento alla teoria delle decisioni; in base al principio della massimizzazione dell'utilità attesa, un agente è razionale se sceglie l'azione che porta alla più alta utilità attesa. Le decisioni, spesso, riguardano la scelta fra opzioni caratterizzate da molti attributi; in questo capitolo sono state analizzate delle tecniche in grado di valutare le informazioni "utile", per poter fare una scelta tra più opzioni.

Nel prossimo capitolo parleremo di agenti intelligenti capaci di filtrare le informazioni disponibili sul web, in particolar modo per gli acquisti online. Discuteremo quindi delle diverse metodologie di filtraggio delle informazioni; descriveremo una tipologia di agenti chiamati anche "*crawler*" o "*spider*" utilizzati dai motori di ricerca quali *Google* per filtrare le informazioni; considereremo la possibilità di utilizzare negli acquisti online, sistemi di aiuto alle decisioni, meno intelligenti di un agente, perché incapaci di apprendere, ma molto efficaci per una rapida selezione delle opzioni.

## Capitolo 3

### Sistemi di supporto alle decisioni di acquisto

#### 3.1 Introduzione

Con la crescita delle possibilità di scelta, in termini sia di opzioni, sia di attributi che caratterizzano le diverse opzioni, è aumentata anche l'esigenza di sistemi sempre più sofisticati a supporto dei processi decisionali dei consumatori. In ambienti online, basta digitare una *query* in un qualsiasi motore di ricerca per ottenere una lista di pagine web rilevanti. Ma soprattutto, per facilitare gli utenti/consumatori nella scelta tra più prodotti e/o servizi, sono sempre più numerosi i motori di ricerca specializzati nella comparazione delle offerte disponibili nei diversi siti e-commerce. I comparatori online, mediante particolari sistemi di supporto alle decisioni di acquisto (*recommender systems*), aiutano i consumatori nei processi di acquisto.

La predisposizione di sistemi automatici di supporto ai processi decisionali, ha coinvolto in tutto il mondo ricercatori di diverse discipline (in particolare, economia, psicologia, informatica, matematica, statistica), operanti sia nelle università, sia in imprese, pubbliche e private. Numerosi sono infatti in questo campo, i contributi apportati dalle diverse discipline, principalmente a partire dalla metà del ventesimo secolo. In questo capitolo sono illustrati gli sviluppi realizzati nel filtraggio automatico delle informazioni, e in particolar modo i sistemi capaci di selezionare le opzioni di prodotto, per rendere più agevoli le scelte dei consumatori. Se il commercio online non riesce a decollare, soprattutto in Italia, una forte accelerata potrebbe essere data da questi sistemi di raccomandazione; applicazioni molto comode, perché utilizzabili anche dal computer di casa, e veloci, perché in pochi minuti consentono di effettuare scelte soddisfacenti.

#### 3.2 Il filtraggio automatico delle informazioni

Il termine “filtraggio delle informazioni” può riferirsi sia alla ricerca di informazioni che si desiderano (*filtering in*), sia all'eliminazione di informazioni non desiderate (*filtering out*); termini analoghi sono *information retrieval* (sistema di recupero delle informazioni) e *selective dissemination of information* (diffusione selettiva delle informazioni).

Mediante il filtraggio delle informazioni si cerca di estrarre da una grande quantità di informazioni quelle più importanti per una persona, escludendo quelle irrilevanti; in altre parole si cerca di selezionare alcune opzioni da un insieme più vasto, presentandole in ordine d'importanza. Tipicamente, l'insieme delle informazioni è dinamico, mentre le esigenze sono stabili e suscettibili di essere sintetizzate in un profilo personalizzato. Mediante il filtraggio si cerca di capire quali profili di preferenze possono, con più elevata probabilità, essere soddisfatti da una certa opzione, selezionata da un ampio insieme di opzioni.

Già Linch aveva messo in evidenza che internet ha innescato un processo di crescita esponenziale delle informazioni disponibili, poiché ognuno può diventare facilmente l'editore di se stesso; di conseguenza la quantità di informazioni disponibili è tendenzialmente pressoché infinita, e la loro qualità è estremamente variabile (*information overload* o sovraccarico di informazioni) [Linch, 1997].

Con la crescita esplosiva delle informazioni disponibili, la preoccupazione principale delle persone è diventata quella di come selezionare le informazioni rilevanti e corrette per ciascuno, anche in considerazione del fatto che informazioni molto importanti per una persona possono essere del tutto irrilevanti per un'altra; secondo Bowman già nei primi anni novanta del secolo scorso il 99 per cento delle informazioni disponibili non aveva alcuna rilevanza per il 99 per cento dei potenziali utenti [Bowman et al, 1994].

Le prime intuizioni sull'importanza prospettica del filtraggio delle informazioni può farsi risalire al 1945, quando Vanner Bush, costruttore della prima calcolatrice analogica e presidente del *Carnegie Institute* di Washington, pubblicò un saggio dal titolo “*as we may think*”, in cui descrisse le caratteristiche di un meccanismo per il recupero delle informazioni (*Memex*), che avrebbe potuto consentire di consultare i documenti presenti in un archivio secondo relazioni associative (non sequenziali), tipiche della mente umana, lungo sentieri caratterizzati da affinità tematica. Il progetto *Memex* non fu mai attuato concretamente, ma l'intuizione di Vanner Bush fu ripresa dopo 20 anni da Ted Nelson, che nel 1965 introdusse il termine “*ipertesto*” per indicare una modalità di accesso alle informazioni del tipo di quella ipotizzata da Vanner Bush, e nel 1967 progettò un ambiente letterario globale online consultabile da tutti con modalità associative e interattive (*Xanadu*). Il primo ipertesto online fu realizzato nel 1968 da Douglas Engelbart, un ricercatore americano che aveva inventato anche il *mouse*, che consente di passare da un documento all'altro, secondo le modalità associative tipiche della mente umana<sup>39</sup>.

Le reti di computer consentono la formazione di gruppi d'interesse che oltrepassano le frontiere geografiche; invece di inviare un articolo a un insieme specifico di persone, l'autore può collocarlo in un *newsgroup di usenet*, vale a dire una collocazione pubblica disponibile a chiunque sia interessato all'argomento. *Usenet* (abbreviazione di *user network* = rete di utenti) è una rete mondiale di server interconnessi in cui sono collocati articoli (*news*) inviati liberamente e consultabili da tutti gli abbonati; essa è articolata per temi in circa 50.000 *newsgroup*. Fondata nel 1979 da un neolaureato della Università americana di Duke, *Usenet* può essere considerata il precursore degli attuali web forum; gli articoli sono distribuiti sequenzialmente attraverso un insieme di moltissimi server interconnessi, Con lo sviluppo di internet, funzioni analoghe a quelle tradizionali di *Usenet* sono state gradualmente svolte dai “*community forum*”; tuttavia, mentre gli articoli collocati in *Usenet* vengono resi disponibili sui server di questa rete in tutto il mondo, gli articoli collocati in un *community forum* non sono di solito disponibili al di fuori di esso; anche se relativamente meno utilizzato che in passato, *Usenet* rimane ancora oggi un importante strumento di comunicazione.

L'esempio più importante di *information retrieval* è senza dubbio rappresentato dai motori di ricerca per il *World Wide Web*, dove un utente può inserire un quesito (*query*), e ottenere una lista di pagine rilevanti. Quindi un sistema di *information retrieval*, comprende una collezione di documenti (paragrafi, pagine, ecc.), una *query* che specifica ciò che l'utente vuole conoscere, un insieme di risultati, vale a dire il sottoinsieme di documenti che il sistema giudica rilevanti per il quesito, una modalità di presentazione dei risultati (una semplice lista di documenti, una mappa colorata, ecc.). Gran parte dei sistemi di *information retrieval* utilizzano modelli basati sulle statistiche dei conteggi delle parole; alcuni documenti hanno una maggiore probabilità di essere rilevanti per qualsiasi interrogazione, in virtù della loro qualità intrinseca; per esempio, per gli articoli accademici si può stimare la qualità in base al numero di citazioni da parte di altri autori; nel caso delle pagine web si può utilizzare come indicatore di qualità intrinseca il numero di collegamenti che puntano alle diverse pagine.

La performance di un sistema di *information retrieval* era spesso misurata in passato tenendo conto principalmente di due parametri: la copertura e la precisione; la precisione misura

---

<sup>39</sup>Un ipertesto è composto da documenti accessibili da un interfaccia elettronica in maniera non sequenziale, mediante collegamenti ipertestuali (*hyperlink*); il sistema di ipertesti di gran lungo più vasto è il *world wide web* di Internet, che impiega il linguaggio HTML (Hyper Text Mark-Up Language) per descrivere i documenti ipertestuali disponibili sul Web. Sviluppato verso la fine degli anni ottanta presso il CERN di Ginevra, HTML è un linguaggio pubblico, la cui sintassi è stata standardizzata e progressivamente migliorata dal *World Wide Web Consortium* (W3C); nonostante sia in corso la sua sostituzione con XHTML, che dovrebbe facilitare la trasmissione di file verso i telefoni cellulari, W3C, raccomanda ancora anche l'uso di HTML4, e ne sta anche sviluppando una versione più aggiornata (HTML5). Ogni documento ipertestuale scritto in HTML deve essere contenuto in un file di estensione *.htm* oppure *.html*, che inizia con l'indicazione della DTD, che dice al browser l'URL delle specifiche html utilizzate per il documento.

la percentuale di documenti effettivamente rilevanti fra quelli selezionati; la copertura misura la percentuale di documenti effettivamente rilevanti che sono presenti nell'insieme dei risultati; Precisione e copertura sono spesso in contrapposizione, selezionando un insieme molto grande aumenta in genere la percentuale di copertura ma diminuisce la precisione, e viceversa, se si punta a selezionare un insieme ristretto, aumenta probabilmente la precisione ma diminuisce la copertura. Attualmente gran parte delle ricerche di informazioni provengono da utenti di Internet<sup>40</sup>, ai quali spesso interessa molto anche l'immediatezza di una risposta rilevante; in questo caso una misura di efficacia può essere il reciproco della posizione del primo risultato rilevante (1 se il primo risultato che si ottiene è quello rilevante;  $\frac{1}{2}$  se il primo risultato non è rilevante ma lo è il secondo, e così via). Inoltre, un altro indicatore di efficacia può essere il tempo necessario per ottenere la risposta desiderata.

Sistemi raffinati di *information retrieval* cercano di tener conto delle correlazioni fra le parole, per esempio mediante il *case folding*, vale a dire la conversione automatica maiuscole/minuscole, oppure utilizzando gli algoritmi di *stemming* (estrazione delle radici delle parole); in questo modo la copertura tende ad aumentare, la precisione, però, di solito diminuisce; effetti analoghi si ottengono con le tecniche di riconoscimento dei sinonimi. Le tecniche di *information retrieval* possono essere potenziate anche considerando i "metadati", vale a dire dati esterni al testo del documento, come parole chiave e collegamenti ipertestuali forniti dalle persone.

Un criterio di presentazione dei risultati molto diffuso è costituito da liste ordinate per probabilità di rilevanza decrescente, eventualmente eliminando i risultati molto simili. Alcuni sistemi consentono all'utente di fornire "feedback sulla rilevanza", indicando quali documenti sono particolarmente rilevanti in un insieme inizialmente proposto, e proponendo poi un insieme di documenti simili a quelli indicati come rilevanti. I risultati a volte sono presentati nella forma di "alberi etichettati" anziché di liste ordinate; oppure classificati per argomenti; il *clustering* dei documenti comporta la creazione di un albero delle categorie per ogni insieme di risultati; la classificazione è più appropriata quando la collezione contiene un piccolo numero di argomenti, il *clustering* è più opportuno nel caso di insiemi di documenti molto ampi, come quelli tipici del *World Wide Web*.

La classificazione è un problema di apprendimento supervisionato che può essere impostato anche utilizzando gli alberi di decisione; dato un insieme di addestramento di documenti etichettati con le categorie corrette, è possibile costruire un singolo albero le cui foglie classificano ciascun documento; in questo modo è possibile raggiungere percentuali di accuratezza superiori al 90 per cento.

Il *clustering* dei documenti è più difficile, poiché lo spazio ha molte più dimensioni e non si sa in anticipo se i dati sono stati generati da un modello gaussiano. All'inizio, ogni documento è considerato come un *cluster* separato, poi si cercano i due cluster più vicini che vengono fusi, ripetendo il processo fino a quando rimane un solo cluster. La misura della distanza tra due cluster potrebbe corrispondere alla distanza fra le mediane, oppure considerare la distanza media tra i componenti dei *cluster*. Per la classificazione *clustering* dei documenti è necessario un algoritmo che tenga conto sia del loro contenuto, sia delle valutazioni degli utenti, e che possa fornire anche una misura del grado di similarità, in assenza di feedback, deve trattarsi di un algoritmo di apprendimento automatico *unsupervised*.

Le tipologie di algoritmi di *clustering* più studiate negli anni duemila sono state probabilmente le mappe auto-organizzantesi (*self-organizing map* o SOM), proposte da Kohonen [Kohonen, 1995, 1997, 2001], della Helsinki University of technology, la cui efficacia applicativa è anche dovuta alle loro ottime capacità di organizzazione e visualizzazione di masse di dati di elevate dimensioni. Le mappe auto-organizzantesi sono state utilizzate in diversi

---

<sup>40</sup> In una collezione di documenti molto grande come il *World Wide Web* è spesso praticamente impossibile calcolare con precisione la copertura; essa può essere soltanto stimata con un campionamento.

progetti industriali, e in diverse discipline scientifiche (statistica, analisi finanziarie, fisica, chimica, medicina), oltre che nell'organizzazione di vastissime quantità di documenti. Esse sono essenzialmente degli algoritmi reticolari neurali di apprendimento senza supervisione (*unsupervised*), e sono state proposte come uno dei modelli più realistici del funzionamento biologico del cervello. L'obiettivo dell'apprendimento è perseguito specializzando parti diverse della rete neurale artificiale a rispondere a particolari pattern d'ingresso, in analogia con il fatto che le informazioni sensoriali, visive, e uditive sono gestite da parti separate della corteccia cerebrale nel cervello umano.

Così come gran parte delle reti neurali artificiali, le mappe auto-organizzantesi hanno due modalità di funzionamento: 1) durante il processo di addestramento la rete neurale si organizza sulla base dei vettori d'ingresso ricevuti, che dovrebbero essere molti, e rappresentativi dei vettori attesi per la seconda fase; 2) durante il processo di *mapping*, i nuovi vettori d'ingresso vengono classificati; vincerà il neurone il cui vettore dei pesi è più vicino al vettore d'ingresso. Una mappa auto-organizzantesi può essere interpretata in due modi: 1) nella fase di addestramento i pesi di tutto il vicinato sono spostati nella stessa direzione, ed elementi simili tendono ad eccitare neuroni adiacenti; di conseguenza le mappe auto-organizzantesi formano una mappa semantica in cui opzioni simili sono vicine e quelle dissimili sono distanti; 2) i pesi neurali possono essere pensati come punti distribuiti nello spazio degli ingressi; più neuroni si dirigono verso regioni con una elevata concentrazione di campioni di addestramento e meno verso zone dove i campioni di addestramenti sono pochi.

L'algoritmo di apprendimento delle mappe auto-organizzanti elaborate da Teuvo Kohonen dell'Università Tecnologica di Helsinki è una brillante formulazione di apprendimento non supervisionato, ed è stato ampiamente utilizzato per la classificazione di elevate quantità di informazioni. Una mappa auto-organizzante si basa su una griglia di neuroni artificiali, i cui pesi sono continuamente adattati ai vettori forniti per l'addestramento, di solito di dimensione elevata; la dimensione dei vettori in uscita è invece di solito non superiore a 3, così da ottenere mappe visualizzabili a due o tre dimensioni. Le reti neurali lavorano in parallelo, e possono quindi elaborare moltissime informazioni in breve tempo; anche se alcune unità dovessero funzionare male, il sistema non dovrebbe bloccarsi. Una rete neurale è una sorta di "scatola nera", nel senso che le indicazioni da essa fornite difficilmente possono essere spiegate.

Le reti neurali sono di solito utilizzate quando i dati possono contenere diversi errori, oppure non sono disponibili modelli analitici adeguati; esse sono molto utilizzate nelle analisi di *data mining* (che saranno presentate più avanti), per le previsioni finanziarie e meteorologiche, nel campo della bio-informatica. Dopo una fase di apprendimento che può anche essere molto lunga, esse sono in genere in grado di indicare i risultati più probabili [Oja e altri, 2003]. Il principale motivo del successo delle mappe neurali auto-organizzantesi è che esse sono in grado di trasformare delle masse di informazioni multidimensionali in rappresentazioni visuali comprensibili; le opzioni simili sono raggruppate, così che esse appaiono vicine nella mappa dei risultati, così che esse evidenziano le relazioni esistenti fra opzioni caratterizzate da una elevata molteplicità di dimensioni anche quando la quantità dei dati è così elevata che le relazioni non possono essere evidenziate con altre metodologie. Poiché operano su vettori multidimensionali, esse possono essere facilmente utilizzabili per la classificazione dei documenti poiché sono in grado di sintetizzare le loro molteplici caratteristiche in una mappa a due dimensioni [Kohonen et al, 1996].

Le due strutture dati fondamentali in qualsiasi sistema di *information retrieval* sono il lessico, che elenca tutte le parole nella collezione di documenti, e l'indice inverso, che elenca tutti i punti in cui compare ogni parola nella collezione. Il lessico è una struttura dati che supporta una singola operazione: data una parola, restituisce la posizione dell'indice inverso che memorizza le sue occorrenze; il lessico deve essere implementato attraverso una struttura dati che consenta una ricerca veloce, eventualmente omettendo parole con poco contenuto informativo o "parole di stop" (il, di, a, e, un, ecc.). L'indice inverso, o semplicemente indice, è

costituito da un insieme di liste di occorrenze, che sono i punti in cui compare ogni parola; per il modello booleano a parole chiave una lista di occorrenze (“*hit list*”) consiste in una lista di documenti, nel modello a unigrammi è una lista di coppie documento-conteggio; per la ricerca di frasi intere la lista di occorrenze deve contenere le posizioni in cui la parola compare all’interno di ogni documento. Quando la *query* è composta da una singola parola (circa un quarto dei casi), l’elaborazione è molto veloce; con un solo accesso al lessico si ottiene l’indirizzo della lista di occorrenze. In concreto, i motori di ricerca che operano sul web suddividono l’indice in *k* segmenti, ognuno memorizzato su un computer diverso; le *query* sono inviate in parallelo a tutti i computer, e i *k* insiemi di risultati sono fusi in un unico insieme mostrato all’utente; i motori di ricerca gestiscono sul web diverse migliaia di *query* ogni secondo.

L’impresa che con maggior successo ha utilizzato gli agenti intelligenti per il filtraggio delle informazioni presenti su internet è stata probabilmente finora Google. Google è diventata leader della ricerca su internet facendo leva sulle valutazioni collettive degli utenti; la sua tecnologia *PageRank* è basata sull’idea che il modo migliore per trovare informazioni rilevanti sia di mettere in ordine di priorità i risultati di ricerca non in base alle caratteristiche dei documenti, ma al numero di siti che contengono un link ai diversi documenti. Secondo Tapscott e Williams “Questi esempi dimostrano come le persone possano “filtrare”, grazie alla collaborazione, la pletora di varietà e di scelte a loro disposizione sulla rete ... in assenza di una guida esperta [Tapscott, Williams, 2007] . Oggi, grazie alle iniziative pionieristiche di una serie di servizi web, le applicazioni dell’intelligenza collettiva stanno arrivando a cambiare le modalità tramite di organizzazione e classificazione del contenuto sulla Rete, attraverso l’utilizzo di un sistema di classificazione “popolare” chiamato *tagging*”. Questo sistema consente agli utenti di apporre “etichette” descrittive al contenuto di documenti, video, ecc.; probabilmente le persone che usano *tag* simili hanno interessi in comune.

### 3.3 Le principali metodologie di filtraggio delle informazioni

Le metodologie di filtraggio delle informazioni cercano di selezionare automaticamente le opzioni per ciascuno più interessanti [Good e altri, 1999]. Sebbene le prime ricerche su questo tema siano state effettuate sin dagli anni cinquanta del secolo scorso, le tecniche di “filtraggio” delle informazioni cominciarono a essere utilizzate diffusamente a partire dagli anni ottanta. Le principali metodologie di riferimento sono due: metodologie cognitive (*content based*) e metodologie “sociali” o “collaborative”: negli ultimi anni sono state avanzate anche delle proposte metodologiche di sintesi di questi due approcci.

Le metodologie cognitive (*content based*) selezionano le opzioni sulla base dei contenuti, per esempio la presenza o meno in un testo di alcune parole, oppure la presenza o meno in un telefonino di particolari caratteristiche; è anche possibile utilizzare gli operatori logici Booleani “and”, “or” e “not”, e i profili possono anche essere rappresentati da vettori di coefficienti di ponderazione che esprimono l’importanza relativa delle diverse caratteristiche. Alcune metodologie di filtraggio cognitivo aggiornano automaticamente i profili, tenendo conto delle valutazioni degli utenti sulle opzioni proposte sulla base dei profili correnti, le tecniche di aggiornamento dei profili utilizzano in genere il teorema di Bayes<sup>41</sup>, algoritmi genetici<sup>42</sup>, e altre

---

<sup>41</sup>Il teorema di Bayes indica in modo analiticamente rigoroso in che modo le nuove informazioni comportano un aggiustamento delle convinzioni; esso è comunemente attribuito al reverendo Thomas Bayes (1701-1761); Richard Price, esecutore testamentario di Bayes, scoprì fra le sue carte due saggi non pubblicati e li inviò alla Royal Society, che li pubblicò. Quei saggi non ricevettero tuttavia particolare attenzione, fino a quando lo stesso teorema non fu scoperto in modo indipendente dal grande matematico francese Laplace; i matematici inglesi riscoprirono allora rapidamente i lavori di Bayes. Secondo Stephen Stigler, un grande storico della statistica, il teorema di Bayes sarebbe stato scoperto in realtà da Nicholas Saunderson, un matematico cieco sin dal primo anno di vita, che fu professore di matematica nell’Università di Cambridge, su presentazione di Isaac Newton.

<sup>42</sup>Gli algoritmi genetici sono dei metodi euristici di ricerca, ottimizzazione e apprendimento automatico, basati sui principi della selezione naturale di Darwin, sviluppati principalmente negli Stati Uniti. Metodologie di simulazione

tecniche automatiche di apprendimento. Le metodologie di filtraggio automatico delle informazioni “basate sui contenuti, o cognitive”, risalgono al lavoro di Luhn in IBM, negli anni cinquanta del secolo scorso; Luhn propose di creare dei profili dei diversi utenti che potessero essere confrontati con le caratteristiche di diversi prodotti, per esempio libri, per segnalare a ciascuno dei libri da acquistare; le abitudini di lettura effettive potevano poi essere utilizzate per aggiornare continuamente i diversi profili [Oard, 1997]. Sin dall’origine queste metodologie cognitive differivano da quelle collaborative, in quanto facevano riferimento a ciascuna persona nella sua individualità e le segnalazioni si basavano sulle caratteristiche di opzioni effettivamente osservate; esse erano quindi applicabili pressoché esclusivamente a documenti testuali, poiché per gran parte degli altri prodotti era difficile implementare un’analisi computerizzata dei contenuti. Le prime applicazioni di queste metodologie erano essenzialmente dei motori di ricerca in grado di rintracciare i documenti che contenevano una certa parola [Li, Kim, 2003]; secondo Oard, dei circa 60 sistemi di questo tipo operanti nel 1969, soltanto quattro contenevano delle procedure automatiche di aggiornamento dei profili. Questi sistemi erano molto semplici, e fornivano la stessa lista di documenti a chiunque indicasse lo stesso termine di ricerca [Oard, 1997].

Lo sviluppo sia delle ricerche, sia delle possibilità computazionali dei computer, condusse a sviluppare la capacità dei motori di ricerca di analizzare i dati tenendo conto delle esigenze particolari di ciascuno, fornendo quindi delle indicazioni personalizzate; l’idea era che la “rilevanza” delle indicazioni potesse essere valutata soltanto su basi individuali; per avviare delle ricerche rilevanti non era quindi sufficiente fornire un semplice termine da cercare. La seconda generazione di sistemi di raccomandazione basati su contenuti erano caratterizzati dalla considerazione di semplici profili individuali di parole chiave, da cercare o da evitare; gli esempi più rilevanti furono le tecniche di filtraggio della posta elettronica e di USENET<sup>43</sup> come SIFT, Infoscope, e Information Lens [Oard, 1997]; il loro funzionamento si basava su semplici filtri basati su parole-chiave e quesiti; SIFT introdusse l’idea di raggruppare i profili degli utenti, così che gli articoli potessero prima essere selezionati con riferimento a diversi gruppi di utenti, e poi, da ciascun gruppo estrarre le segnalazioni per singole persone; questo sistema si rivelò più efficiente di quelli che trattavano separatamente ogni richiesta da singoli utenti; *Infoscope* introdusse degli automatismi, deducendo le parole chiave alle quali un utente poteva essere interessato dai documenti che aveva salvato, valutato interessanti, esaminati a lungo. Il lavoro di Malone e altri ricercatori coinvolti nel progetto *Information Lens* ebbe impatti più duraturi nel campo del filtraggio delle informazioni; essi misero in evidenza l’opportunità di tener conto anche di fattori sociali, suggerendo dei principi di filtraggio sociale che tenessero conto non soltanto delle caratteristiche dell’autore, ma anche dei riferimenti e delle raccomandazioni di altre persone [Malone et al, 1987].

---

computerizzata dei processi evolutivi cominciarono a essere sperimentate fin dal 1954 da Barricelli presso l’Università di Princeton; gli algoritmi genetici, in particolare, furono sviluppati presso l’università del Michigan da Holland (1975), che propose uno schema teorico per l’analisi dei mutamenti fra una generazione e quella successiva. Verso la metà degli anni ottanta l’Università di Pittsburgh organizzò la prima conferenza internazionale sugli algoritmi genetici, e negli anni successivi la General Electric propose la prima versione commerciale di un algoritmo genetico; nel 1989 fu commercializzato *evolver*, un algoritmo genetico per PC, le cui potenzialità applicative furono illustrate anche sul *New York Times*; una raccolta dei primi saggi sugli algoritmi genetici è stata curata da (Fogel, 1998).

<sup>43</sup> *Usenet* (abbreviazione di *user network* = rete di utenti) è una rete mondiale di server interconnessi in cui sono collocati articoli (*news*) inviati liberamente e consultabili da tutti gli abbonati; essa è articolata per temi in circa 50.000 *newsgroup*. Fondata nel 1979 da un neolaureato della Università americana di Duke, *Usenet* è il precursore degli attuali *web forum*; gli articoli sono distribuiti sequenzialmente attraverso un insieme di moltissimi server interconnessi; con lo sviluppo di internet funzioni analoghe a quelle tradizionali di *Usenet* sono state gradualmente svolte dai *community forum*; tuttavia, mentre gli articoli collocati in *Usenet* vengono resi disponibili sui server di questa rete in tutto il mondo, gli articoli collocati in un *community forum* non sono di solito disponibili al di fuori di esso; anche se relativamente meno utilizzato che in passato, *Usenet* rimane ancora oggi un importante strumento di comunicazione.

Man mano che l'attenzione delle nuove ricerche andò sempre più concentrandosi sulle informazioni in arrivo, aumentò l'uso di tecniche di *clustering* nell'ambito delle metodologie di filtraggio cognitivo, al fine di alleggerire il lavoro dei sistemi di raccomandazione; gli articoli venivano suddivisi in gruppi in base alle loro caratteristiche ritenute più rilevanti, e gli utenti fornivano delle indicazioni utili per ottenere un profilo delle loro preferenze. Nel caso di NewsWeeder di USENET i profili dei diversi utenti venivano ricavati sulla base delle loro valutazioni degli articoli, deducendo le loro parole-chiave preferite dal contenuto dei documenti su cui esprimevano valutazioni positive, NewsWeeder sollecita anche una partecipazione attiva degli utenti; invece di limitarsi a basarsi sul tempo che ciascun utente dedica ai diversi articoli, o sul fatto di salvarlo o meno, si chiede una sua valutazione in base a una scala da 1 a 5; in questo modo c'è la possibilità di effettuare delle previsioni sulle valutazioni di nuovi articoli, e quindi elencare in ordine di rilevanza prevista le diverse opzioni segnalate [Lang, 1995]; ciò naturalmente richiede un certo sforzo da parte degli utenti per creare e poi aggiornare i loro profili.

Il contributo più significativo di *NewsWeeder* alle ricerche sui sistemi di raccomandazione basati sui contenuti è stato il principio del “*minimum description length*” (MDL), applicato stimando la probabilità che ciascuna parola possa apparire in un documento sulla base della sua lunghezza, così da poter escludere i termini la cui capacità discriminatoria è ritenuta insufficiente. Una metodologia analoga è utilizzata attualmente dagli algoritmi di ispirazione Bayesiana per cercare di limitare l'accesso alla posta elettronica non desiderata. Applicando la regola di Bayes, il sistema stima la probabilità che una certa parola possa apparire in una categoria di documenti, piuttosto che in un documento specifico, e quindi predice la probabilità che il documento appartenga a una certa categoria sulla base delle parole in esso contenute. Lang mise in evidenza l'efficienza di questo metodo, le cui applicazioni si diffusero successivamente in diversi ambiti, tra cui la raccomandazione dei film [Ono e altri, 2005].

Un limite importante alle possibilità applicative dei metodi di filtraggio basati sui contenuti deriva dalle difficoltà di estendere le valutazioni a prodotti significativamente diversi da quelli valutati dall'utente, con il rischio di perdere importanti opportunità [Balabanovi'c, Shoham, 1997]; per alcuni articoli, inoltre, può essere difficile individuare indicatori significativi del loro contenuto intrinseco (ristoranti, ecc.).

A partire dagli ultimi anni del secolo scorso, su portali come Amazon, CDNow, ecc., cominciarono ad essere applicati concretamente con successo, alcuni sistemi di raccomandazione di tipo collaborativo rivolti soprattutto alle ricerche sugli algoritmi in grado di determinare meglio i raggruppamenti e le relazioni di similarità [Konstan, Riedl, 1999].

Le metodologie di filtraggio sociali o *collaborative* si basano sulle relazioni fra le valutazioni di persone diverse; l'applicazione più interessante è rappresentata dai filtri collaborativi, che cercano di aiutare le persone a decidere sulla base delle opinioni di altri, nell'ipotesi che le persone che hanno evidenziato preferenze analoghe in passato concorderanno anche in futuro

*GroupLens*, per esempio, è un sistema di filtraggio collaborativo che cerca di aiutare ciascuno a selezionare articoli di proprio interesse fra tutti quelli disponibili in rete. Nell'era della sovrabbondanza di informazioni, soprattutto online, uno dei problemi più rilevanti riguarda la selezione di opzioni utili e interessanti.

Il più rilevante fra i primi tentativi di implementazione di una metodologia di filtraggio delle informazioni di tipo collaborativo (*information Tapestry*) fu sperimentato nei primi anni novanta del secolo scorso nel centro di ricerca della Xeros, partendo dalla convinzione che il processo di selezione delle informazioni è più efficace quando più persone collaborano, aiutandosi vicendevolmente, annotando ciascuna le proprie valutazioni sugli articoli esaminati e condividendole [Goldberg, Nichols, Oki, Terry, 1992]. L'obiettivo immediato di quell'esperimento era di sviluppare una metodologia di filtraggio della posta elettronica, che consentisse agli utenti di evitare di essere inondati da un grande numero di messaggi

indesiderati; la metodologia era comunque applicabile ad altre categorie di documenti. In base a questo sistema, per esempio, un utente poteva decidere di vedere soltanto documenti che erano stati letti, ed eventualmente valutati positivamente, da almeno altri due utenti. L'esperimento *Tapestry* ebbe effetti limitati sullo sviluppo delle ricerche sulle metodologie di filtraggio delle informazioni, probabilmente perché esso era stato condotto in una impresa privata con obiettivi commerciali, e diversi suoi aspetti non furono adeguatamente pubblicizzati. Per evitare problemi analoghi, fu costituito presso l'Università di Minnesota un gruppo di ricerca sulle metodologie collaborative di filtraggio delle informazioni, con applicazioni ai documenti pubblicamente disponibili di USENET, con l'intento di rendere liberamente disponibili tutti risultati delle ricerche (*GroupLens project*) [Resnick et al, 1994]. Il principale contributo di *GroupLens* allo sviluppo delle metodologie di filtraggio collaborative è stata probabilmente la tecnica di aggregazione automatica delle richieste degli utenti; questa tecnica consente di correlare le valutazioni di tutti gli utenti al fine di individuare le similarità fra gli utenti, così che non è necessario che un consumatore sappia a quali altri utenti fare riferimento per il filtraggio, poiché il sistema automaticamente individua gli utenti simili ed effettua una previsione della valutazione dell'utente considerato sulla base delle valutazioni degli altri utenti [Li, Kim, 2003]; per alcuni aspetti analogo a *GroupLens* è "Ringo", un filtro collaborativo sviluppato da Shardanand e Maes [Shardanand, Maes, 1995]. L'architettura aperta di *GroupLens* ha contribuito a far sì che quel progetto diventasse il principale punto di riferimento delle ricerche sulle metodologie di filtraggio collaborative. Il postulato essenziale dei filtri collaborativi è che le persone che hanno espresso valutazioni analoghe in passato, continueranno a farlo in futuro; il processo di filtraggio collaborativo è articolato tipicamente in tre fasi: 1) diverse persone esprimono le loro valutazioni su un certo numero di opzioni; 2) vengono calcolati degli indici di correlazione fra le valutazioni delle diverse persone; 3) quando una nuova opzione viene valutata da una parte degli utenti, le valutazioni della stessa opzione da parte di un'altra persona viene predetta sulla base delle valutazioni delle persone le cui valutazioni passate sono risultate significativamente correlate con quelle della persona considerata. L'analisi di correlazione è opportuno che sia effettuata distintamente per tipi di valutazione, poiché non è detto che persone che esprimono valutazioni analoghe su alcuni tipi di valutazione si comportino in modo analogo anche per valutazioni di altra natura [Konstan et al, 1997].

La debolezza fondamentale delle metodologie di raccomandazione di tipo collaborativo è che esse non possono essere applicate né a nuovi utenti né a nuove opzioni; quando si presenta un nuovo utente non si sa con quali vecchi utenti sono correlate le sue valutazioni, e quindi non esistono elementi per consigliargli delle opzioni; le nuove opzioni non sono state ancora valutate da nessuno, e quindi non si sa a chi consigliarle; eventuali valutazioni di nuove opzioni da parte di nuovi potenziali consumatori non sono utili né per consigliare la nuova opzione ad altri potenziali acquirenti, né per consigliare qualche altra opzione al nuovo potenziale acquirente. In alcuni casi questo problema permane nel tempo, se in alcune aree il numero di valutazioni espresse rimane basso. Per una efficace implementazione di un sistema collaborativo di raccomandazione è necessario che sia raggiunta una massa critica di valutazioni; il sistema quindi funziona bene soltanto se c'è un gruppo sociale di persone che collaborano, e nel quale sono molte le persone con simili sistemi di preferenze.

Il principale vantaggio delle metodologie di filtraggio collaborative rispetto a quelle cognitive è la possibilità di fare a meno di analizzare direttamente le caratteristiche delle diverse opzioni; invece di perdere tempo nell'analisi dei documenti, e nello sviluppo di metodologie di analisi dei contenuti, chi si occupa di filtri collaborativi può concentrare l'attenzione sull'aspetto essenziale dei sistemi di filtraggio informativo: gli algoritmi per la definizione dei gruppi di utenti. Mentre l'implementazione di metodologie di filtraggio cognitivo comporta spesso l'esigenza di immagazzinare enormi quantità di dati sulle caratteristiche degli utenti e delle diverse opzioni, la progettazione di un sistema di filtraggio collaborativo è di solito abbastanza semplice: sono sufficienti profili di utenti definiti sulla base delle valutazioni effettuate, piuttosto

che stime di probabilità di presenza in un documento per tutte le parole. Soprattutto, le metodologie di filtraggio collaborative hanno possibilità applicative molto più vaste; non richiedendo l'analisi dei contenuti delle opzioni, esse possono essere applicate praticamente a qualunque tipo di scelta; esse sono state in effetti utilizzate per filtrare documenti e anche libri senza alcun riferimento al loro contenuto effettivo, e anche quadri, film, e una grande varietà di prodotti senza alcun riferimento al loro contenuto effettivo.

I sistemi di raccomandazione collaborativi sono molto popolari, sia per la loro semplicità, sia per la fiducia che i problemi possano essere superati, sia, soprattutto, perché con questi sistemi non è necessario considerare il contenuto, vale a dire le caratteristiche delle opzioni, per potere effettuare delle raccomandazioni. Per esempio, diverse ricerche hanno mostrato che esplicite valutazioni delle diverse opzioni possono a volte non essere necessarie, poiché possono essere sufficienti delle valutazioni implicite, come, per esempio, il tempo dedicato alla lettura di un articolo [Konstan et al, 1997]. Inoltre, la difficoltà della “partenza a freddo” può essere superata per le nuove opzioni utilizzando delle metodologie di filtraggio “ibride” o “multiple”.

Le metodologie di selezione ibride (o multiple) possono ovviare almeno in parte ai problemi di applicazione delle metodologie contenutistiche e collaborative [Balabanovi'c, Shoham, 1997].

Burke ha messo in evidenza cinque principali modalità di integrazione di diverse metodologie di filtraggio delle informazioni: 1) media ponderata dei risultati delle diverse metodologie; 2) passaggio da una metodologia all'altra a seconda delle informazioni disponibili; 3) presentazione simultanea dei risultati ottenuti applicando le diverse metodologie, lasciando la responsabilità della scelta agli utenti; 4) restrizione progressiva dell'insieme delle opzioni, mediante l'applicazione sequenziale di diverse metodologie; 5) utilizzazione dei risultati ottenuti con alcune metodologie come input addizionali nell'applicazione di altre metodologie [Burke, 2002].

*Tapestry* fu probabilmente il primo tentativo di combinare tecnologie di selezione contenutistiche e collaborative, anche se la metodologia collaborativa utilizzata era alquanto rudimentale, consistendo esclusivamente nella possibilità di tener conto in modo informale delle opinioni di altre persone.

Il progetto RAAP (*research assistant agent project*) fu uno dei primi sistemi ibridi di filtraggio delle informazioni; esso utilizza un algoritmo *tf-idf*<sup>44</sup> per proporre delle classificazioni dei documenti sulla base sia del loro contenuto sia dei profili degli utenti [Delgado et al, 1998].

Li e Kim proposero una metodologia che utilizza sia le valutazioni degli utenti, sia le informazioni sulle caratteristiche delle opzioni per determinare le relazioni di similarità; la loro metodologia è articolata in tre fasi: 1) raggruppamento dei documenti e creazione di una matrice di valutazione per gruppi, 2) determinazione delle relazioni di similarità, 3) effettuare delle previsioni sulle opzioni preferite dalle diverse persone sulla base delle relazioni di similarità [Li, Kim, 2003].

Secondo Breese [Breese et al, 1998], gli algoritmi più utilizzati possono essere distinti in due classi: 1) algoritmi che per le loro raccomandazioni utilizzano la media delle opinioni degli utenti, ponderate in base al grado di similarità fra gli utenti (esempi: algoritmo delle *k* opzioni più vicine (*k-nearest neighbor*<sup>45</sup>) e altri algoritmi utilizzati da *GroupLens*; 2) algoritmi che si basano sulla distanza fra le opzioni, usati per esempio nel progetto RAAP, oppure sulla similarità vettoriale come quelli utilizzati da *NewsWeeder* [Lemire e Maclachlan, 2005].

---

<sup>44</sup>Tf-idf (*term frequency – inverse document frequency*) è una tecnica spesso utilizzata per valutare l'importanza di una parola nell'ambito di un documento; l'importanza aumenta in proporzione alla frequenza con cui la parola appare in un certo documento, e diminuisce in proporzione alla frequenza con cui la parola appare nell'insieme dei documenti.

<sup>45</sup>Questo è il più semplice degli algoritmi di apprendimento automatico; ciascuna opzione è assegnata alla classe a cui appartengono il maggior numero delle *k* opzioni più vicine, *k* di solito è un numero abbastanza piccolo, possibilmente dispari nel caso di classificazioni binarie.

Secondo una ricerca svolta da Breese presso Microsoft i sistemi di filtraggio che meglio riescono a predire le decisioni degli utenti sono i metodi bayesiani e quelli che si basano su analisi di correlazione; quella ricerca è rimasta per molti anni un punto di riferimento per le verifiche delle performance di diversi sistemi di raccomandazione [Breese et al, 1998].

### **3.4 I Sistemi di estrazione delle informazioni**

Dalla relazione tra *information retrieval* e analisi statistica derivano i cosiddetti sistemi di estrazione delle informazioni. L'estrazione delle informazioni è il processo per cui vengono inseriti dati in un database esaminando un testo e cercando le occorrenze di una particolare classi di oggetti o eventi e le relazioni tra tali oggetti ed eventi.

Per facilitare l'estrazione di informazioni interessanti, sono stati realizzati particolari software, molto usati sul web, soprattutto dai motori di ricerca. Parliamo dei cosiddetti *crawler* (ma anche *spider* o *robot*), in grado di analizzare i contenuti di una rete (o di un database) in un modo metodico e automatizzato, in genere per conto di un motore di ricerca. Un crawler viene definito anche *bot*, perché è un programma o *script* che automatizza delle operazioni. I crawler solitamente acquisiscono una copia testuale di tutti i documenti visitati e le inseriscono in un indice.

Un uso estremamente comune dei crawler è nel Web. Sul Web, il crawler si basa su una lista di URL da visitare fornita dal motore di ricerca (il quale, inizialmente, si basa sugli indirizzi suggeriti dagli utenti o su una lista precompilata dai programmatori stessi). Durante l'analisi di un URL, identifica tutti gli hyperlink presenti nel documento e li aggiunge alla lista di URL da visitare. Il processo può essere concluso manualmente o dopo che un determinato numero di collegamenti è stato seguito.

Inoltre i crawler attivi su Internet hanno la facoltà di essere indirizzati da quanto indicato nel file *robots.txt* posto nella *root* del sito. All'interno di questo file, è possibile indicare quali pagine non dovrebbero essere analizzate. Il crawler ha la facoltà di seguire i consigli, ma non l'obbligo.

Le analisi empiriche realizzate all'interno di questo lavoro di tesi, espone nei capitoli 5 e 6, si basano soprattutto su dati estrapolati da pagine web; in particolare dati relativi alle caratteristiche delle diverse opzioni di prodotto e valutazioni rilasciate da utenti esperti o utenti che hanno provato il prodotto. Avvalendoci di uno di questi agenti *crawler* denominato *Web Content Extractor*, impostando come direttive quali tipologie di dati considerare, è stato possibile estrarre automaticamente dalle pagine web una grande quantità di informazioni rilevanti, che non sarebbe stato possibile ottenere con metodi tradizionali.

### **3.5 I sistemi di supporto alle decisioni di acquisto**

Anche negli ambienti online dedicati agli acquisti, l'eccessiva abbondanza di informazioni (*information overload*), risulta per l'utente/consumatore, un problema veramente difficile da superare. In questi casi, l'enorme varietà di opzioni di prodotto disponibili, e quindi anche di attributi che caratterizzano le diverse opzioni, tende spesso a rendere i processi di acquisto molto onerosi in termini sia di sforzi cognitivi sia di tempo impiegato, tanto da comportare una diminuzione della propensione a ricercare informazioni da parte degli acquirenti *online* [Alba, Linch, Weitz, Janiszewski, 1997]. Per ridurre la rilevanza di questi problemi possono essere utili dei sistemi di supporto alle decisioni di acquisto, come, per esempio, gli agenti di raccomandazione (*recommender system*), speciali software intelligenti in grado di considerare contemporaneamente un vasto numero di prodotti presenti in un *database* e di selezionarne un set più ristretto, sulla base delle preferenze di ciascun consumatore.

I *recommender system* sono delle applicazioni software, che hanno l'obiettivo di aiutare i consumatori/utenti durante i processi di acquisto, nella scelta del prodotto o servizio in grado di soddisfare meglio le proprie esigenze; in particolare questi sistemi cercano di capire le preferenze dei consumatori, di realizzare e successivamente mantenere aggiornato un profilo delle preferenze per ciascun consumatore, e, sulla base di questi risultati, formulare delle

proposte di acquisto personalizzate; più in generale, gli agenti di raccomandazione mettono a disposizione di ciascun consumatore ogni informazione potenzialmente utile per le sue decisioni di acquisto. Con la crescita delle possibilità di scelta dei consumatori, in termini sia di opzioni sia di attributi che caratterizzano le diverse opzioni, è aumentata anche l'esigenza di sistemi sempre più sofisticati a supporto dei processi decisionali dei consumatori.

Sono disponibili online molteplici siti che si caratterizzano per l'utilizzo di algoritmi e/o sistemi intelligenti di *comparison shopping*, e si differenziano soprattutto per le tipologie di informazioni utilizzate e per la natura delle elaborazioni effettuate. Una classificazione molto interessante dei sistemi di supporto alle decisioni di acquisto online (o "infomediari") è stata proposta da Guido, Peluso e Rizzato mediante una "matrice dei servizi di comparazione" (Tabella 3.1); la matrice è ottenuta incrociando due tipologie di ricerca delle informazioni, il *product brokering* (ricerca di informazioni sulle caratteristiche dei prodotti) e il *merchant brokering* (ricerca di informazioni sulle caratteristiche dei venditori), con due modalità di comparazione delle alternative d'acquisto (quantitative e qualitative) [Guido, Peluso, Rizzato, 2005].

Nel primo quadrante sono collocati i sistemi di comparazione dei prodotti su basi quantitative, suddivisi in agenti di raccomandazione *content-based filtering* che si basano sulle preferenze espresse da ciascun utente per i diversi attributi dei prodotti [Ariely, Linch, Aparicio, 2004], e semplici matrici di comparazione degli attributi dei diversi prodotti che consentono a ciascun utente di ordinare le diverse opzioni sulla base degli attributi da ciascuno ritenuti più importanti.

Nel secondo quadrante si trovano i sistemi di comparazione dei venditori su basi quantitative, vale a dire essenzialmente con riferimento ai prezzi di vendita; questi sistemi possono basarsi sui prezzi forniti direttamente dai venditori (sistemi non automatizzati di comparazione), oppure sui prezzi cercati in rete da *shopbots* (abbreviazione di *shopping robots*), vale a dire da agenti software che raccolgono e catalogano automaticamente informazioni provenienti da molteplici fonti su prezzi e/o altri attributi delle diverse opzioni (sistemi automatizzati di comparazione). Nel terzo quadrante rientrano i siti che offrono servizi di comparazione qualitativa dei prodotti sulla base delle preferenze espresse da altri consumatori o esperti (sistemi non automatizzati), oppure mediante sistemi automatizzati di *rating* che utilizzano agenti intelligenti del tipo *collaborative filtering*, in grado di selezionare i prodotti sulla base di valutazioni espresse da consumatori simili. Nel quarto quadrante si trovano i sistemi di comparazione dei venditori su basi qualitative, vale a dire attraverso indici di reputazione e credibilità ottenuti sulla base di giudizi forniti da consumatori o da esperti.

Tabella 3.1. - Classificazione dei sistemi di supporto alle decisioni di acquisto online (infomediari).

	<i>Product brokering</i>	<i>Merchant brokering</i>
Compa- razione Quanti- tativa	(I) <i>Product comparison</i> -Agenti di raccomandazione <i>content-based filtering</i> -Matrici di comparazione degli attributi	(II) <i>Price comparison</i> -Sistemi non automatizzati di comparazione -Sistemi automatizzati di comparazione ( <i>shopbots</i> )
Compa- razione Qualità- tativa	(III) <i>Rating</i> sui prodotti -Sistemi non automatizzati di <i>rating</i> , basati su opinioni di consumatori e esperti - Sistemi automatizzati di <i>rating</i> , basati su agenti di raccomandazione <i>collaborative filtering</i>	(IV) <i>Store comparison</i> - Opinioni e <i>rating</i> di consumatori - Giudizi e certificazioni di esperti

Fonte: Guido, Peluso e Rizzato (2005)

Algoritmi intelligenti creati per filtrare le informazioni online, sono diventati il *core business* di diversi portali di *e-commerce*; essi vengono anche chiamati *infomediari*, *cybermediari* o anche *digital intermediaries*, poiché svolgono simultaneamente, due funzioni molto importanti: 1) raccolgono e trasmettono efficientemente le informazioni ai consumatori, sulle varie offerte online, consentendo di ridurre costi e tempi di ricerca; 2) forniscono profili dettagliati degli stessi consumatori ai venditori online, consentendo a questi ultimi, nei limiti del rispetto della *privacy*, di disporre di informazioni, che da soli non riuscirebbero a reperire [Chircu e Kauffman, 1999]. Inoltre l'analisi delle preferenze dei consumatori e le loro opinioni post-acquisto possono consentire alle aziende produttrici di definire meglio le caratteristiche dei nuovi prodotti.

Tra i portali di *e-commerce* che usufruiscono di un sistema *collaborative filtering*, c'è il colosso eBay che ne ha fatto una bandiera verso altri siti d'aste *online*. A volte, però, basarsi sui giudizi di altri può portare a scelte non soddisfacenti; spesso vengono scoperte delle vere e proprie truffe sul web, derivanti da fenomeni di *crowdhacking*, vale a dire di processi di creazione artificiosa di feedback positivi basati su valutazioni di negozianti e acquirenti gonfiati artificialmente; curricula di società costruiti a tavolino per "accalappiare" ignari internauti che si fidano di giudizi positivi che credono siano stati inviati da altri utenti, e comprano l'opzione proposta.

Una percentuale molto elevata di giudizi positivi è un vantaggio in primo luogo, per il banditore dell'asta, che avrà ancora più clienti, ma anche per i clienti, che si sentiranno al riparo dalle truffe. La rivista "*Wired*", racconta di un colpo riuscito a Nancy Dreksler nel 2002-2003, che in questo modo ha "guadagnato" una piccola fortuna. Si è costruita una buona reputazione vendendo su eBay articoli di poco valore, come cd e dvd; a quel punto riuscì a vendere articoli inesistenti per un valore di oltre 100 mila dollari, lasciando più di 500 acquirenti a mani vuote. Rischi analoghi esistono in tutti i siti che fanno riferimento ai giudizi degli utenti, come Amazon.com, Price.com, Shopping.yahoo.com, Opinion.com, e tanti altri.

Di fronte ai rischi derivanti dall'utilizzo di sistemi *collaborative filtering*, l'utilizzo di un agente *content based filtering* può risultare molto più efficace e affidabile nella selezione di prodotti che corrispondano meglio alle preferenze di ciascun consumatore. Un sistema di raccomandazione content based, avvalendosi dei valori di performance degli attributi di ogni

opzione di prodotto, effettua delle comparazioni di prodotto con dati dimostrabili e non rilevati dalle indicazioni di un gruppo qualsiasi di opinionisti. Secondo Häubl e Trifts, infatti, le matrici di comparazione, ma soprattutto i sistemi *content-based filtering* consentono ai consumatori di prendere decisioni migliori con un minore sforzo cognitivo, allentando il trade-off tra grado di accuratezza delle decisioni e sforzo sostenuto [Häubl, Trifts, 2000].

### 3.5.1 Sistemi compensatori e non compensatori

Tra i sistemi di raccomandazione basati sui contenuti, ed inseriti nel primo quadrante della Tabella 3.1, rientrano anche particolari sistemi di aiuto alle decisioni, che si propongono di facilitare gli utenti nelle loro decisioni, attraverso processi di sfoltimento (*winnowing*) che consentono di ridurre le opzioni di scelta da dimensioni molto elevate a numeri più limitati e gestibili. Fasolo, Motta e Misuraca hanno evidenziato infatti, che una diversa tipologia di sistemi di aiuto alle decisioni (SAD), sono stati realizzati ed integrati in diversi siti, soprattutto negli Stati Uniti. Le tecniche decisionali utilizzate da queste applicazioni, sono definite “*compensatorie*”, se consentono di effettuare compensazioni fra valori positivi e negativi degli attributi delle diverse opzioni, oppure “*non-compensatorie*” [Edwards, 1992; Keeney, Raiffa, 1976; Zeleney, 1976; Fasolo, Motta e Misuraca, 2005].

Le tecniche non-compensatorie consentono agli utenti di utilizzare solo parte dell’informazione proposta, incoraggiandoli a concentrare la propria attenzione sugli attributi, uno alla volta. Le più comuni euristiche di scelta non-compensatorie sono l’*eliminazione per aspetti* [Tversky, 1972] e la *lessicografica* [Svenson, 1979]. Nel primo caso, il decisore stabilisce per ciascun attributo una soglia «accettabile». Se il valore di un attributo è al di sotto di tale soglia, l’alternativa viene scartata. Il processo continua fin quando non resta che una sola opzione. Nel secondo caso, invece, il decisore ordina gli attributi secondo l’importanza loro assegnata. Il processo di selezione inizia con il confronto di tutte le alternative in base all’attributo ritenuto più importante. L’alternativa con il valore più alto su tale attributo viene selezionata, mentre tutte le altre vengono eliminate. Se due alternative presentano lo stesso valore sull’attributo più importante, allora queste vengono confrontate sulla base del secondo attributo più importante. Il processo continua fin quando una sola opzione «supera» le altre nell’attributo considerato. Le strategie decisionali facilitate dai siti non compensatori sono più semplici e, di conseguenza, solitamente preferibilmente utilizzate, ma sono meno accurate di quelle compensatorie [Widing e Talarzyk, 1993].

Le strategie compensatorie sono generalmente ritenute procedure decisionali razionali, poiché comportano l’utilizzazione di tutte le informazioni disponibili, al fine di ottenere valutazioni complessive per le diverse opzioni, sulla base delle preferenze individuali per i diversi attributi. In questo modo le decisioni, se da un lato richiedono maggiori sforzi cognitivi, dall’altro consentono di selezionare le opzioni che massimizzano l’utilità attesa [Keeney e Raiffa, 1976].

Alcune indagini empiriche sul funzionamento dei sistemi di aiuto alle decisioni hanno sollevato però alcuni dubbi; questi sistemi, come tutti i sistemi *content based*, si basano sull’ipotesi che le preferenze delle persone siano note a priori e rimangano stabili nel corso dei processi decisionali; quindi la loro principale caratteristica consiste nella possibilità di facilitare specifiche euristiche decisionali, cioè strategie per acquisire informazioni e arrivare infine alla decisione [Payne et al, 1993]. Le euristiche sono delle vere e proprie scorciatoie cognitive che consentono di fronteggiare la complessità della realtà attraverso una riduzione delle informazioni considerate per prendere una decisione; sono strategie di pensiero che consentono di formulare rapidamente un giudizio sulle caratteristiche di un prodotto. L’applicazione di un procedimento decisionale basato su calcoli razionali, anche se dà garanzia di accuratezza, deve però fare i conti con le inevitabili complessità dei comportamenti umani [Russo, 2004]. Uno dei risultati fondamentali delle ricerche sui processi decisionali, è che le preferenze sono spesso «costruite» mediante aggiustamenti progressivi agli stimoli ambientali [Fischhoff, 1991; Payne, Bettman,

Coupey e Johnson, 1993]; in particolare, le decisioni possono essere influenzate dalle modalità di presentazione delle opzioni [Slovic, 1995], secondo Mandel e Johnson, anche il semplice colore dello sfondo del sito può influenzare le decisioni dell'acquirente; nel loro studio, i soggetti esploravano uno *shopping site* sperimentale e dovevano scegliere tra 2 poltrone [Mandel, Johnson, 2002]. Si osservò che i soggetti esposti ad uno sfondo con nuvole esprimevano una netta preferenza verso la poltrona che primeggiava sotto l'aspetto «comfort»; viceversa, i soggetti ai quali era stato mostrato lo sfondo con i *pennies* tendevano a scegliere sistematicamente la poltrona che primeggiava per il «prezzo». Ciò avverrebbe a causa di un *effetto priming*, meccanismo cognitivo in virtù del quale, se una data caratteristica (es. «comfort») è associata ad un certo oggetto (es. «soffici nuvole»), essa viene attivata con la semplice presentazione di uno sfondo blu a nuvole, portando perciò alla scelta dell'opzione migliore dal punto di vista della caratteristica considerata.

Può però capitare che, utilizzando strategie compensatorie, in presenza di opzioni con attributi inversamente correlati (dove attributi con performance alta compensano attributi con performance scarse), il consumatore potrebbe essere influenzato a scegliere una opzione di prodotto con attributi positivi per lui poco interessanti ed attributi negativi che invece erano per lui molto importanti.

Häubl e Trifts hanno infatti mostrato che le persone in genere hanno la tendenza a fidarsi delle raccomandazioni, e sono influenzati sistematicamente dalle performance degli attributi presenti nei prodotti selezionati; soprattutto quando la scelta è tra opzioni che hanno valori soddisfacenti per alcuni attributi, ma non soddisfacenti per altri (attributi negativamente correlati). Häubl e Trifts, ritengono però che con un numero più elevato di attributi, potrebbe migliorare significativamente la qualità delle decisioni (*inclusion effect*) [Häubl, Trifts, 2000]. Questo effetto è invece scarsamente significativo nei casi di correlazione positiva fra diversi attributi, vale a dire quando una opzione con valori soddisfacenti per un certo attributo, presenta valori soddisfacenti anche per gran parte degli altri attributi. Molto importanti sono anche le «proprietà collative» individuate da Berlyne (novità, sorpresa, complessità); queste caratteristiche influenzano la selezione degli stimoli, inducendo le persone a dedicare maggiore attenzione a certe caratteristiche rispetto ad altre [Berlyne, 1960]. I processi decisionali sono anche influenzati dalla natura dei prodotti; nelle comparazioni di prodotti omogenei l'assenza di informazioni su un attributo non pregiudica di solito la bontà della decisione, che in genere è influenzata principalmente dal prezzo. Le comparazioni di prodotti eterogenei sono in generale più complesse, comportano un maggior coinvolgimento del consumatore, e la mancanza di informazioni su alcuni attributi potrebbe compromettere la qualità della decisione [Chernev, 1997, Kivetz e Simonson 2000].

### **3.6 Decisioni di acquisto di prodotti con molte caratteristiche**

Le decisioni di acquisto sono particolarmente complesse nel caso di prodotti caratterizzati da molti attributi<sup>46</sup>, come per esempio i telefoni cellulari. Il fondamento teorico dell'analisi delle decisioni nel caso di prodotti caratterizzati da molti attributi è costituito dalla teoria del comportamento del consumatore di Lancaster, secondo cui il valore complessivo di un bene per un consumatore deriva dal valore delle sue diverse caratteristiche [Lancaster, 1966]. Le due principali applicazioni empiriche dell'approccio di Lancaster all'analisi delle decisioni dei consumatori sono la metodologia dei prezzi edonici, che cerca di stimare il valore delle diverse caratteristiche di un prodotto mediante un'analisi econometrica in cui il prezzo delle diverse versioni di un prodotto è la variabile dipendente e i valori delle diverse caratteristiche sono variabili esplicative, e la *conjoint analysis*, una metodologia di analisi proposta nell'ambito della

---

<sup>46</sup> Goode (2001) e Schwarts (2004) ha illustrato il “paradosso della scelta”, vale a dire il fatto che un aumento delle opportunità di scelta può provocare un peggioramento della qualità delle decisioni, o addirittura una sorta di “paralisi decisionale”.

psicologia matematica, e ampiamente utilizzata nelle ricerche quantitative di marketing [Johnson, 1974; Rosen, 1974; Green, Snirivasen, 1978; Thaler, 1985]. La *conjoint analysis* si basa sull'ordine di preferenza delle diverse versioni di un prodotto ottenuto attraverso indagini di mercato; essa è di solito utilizzata quando non sono disponibili stime affidabili dei prezzi di mercato per le diverse versioni del prodotto (per esempio quando si tratta di stimare il prezzo massimo che i consumatori potrebbero essere disposti a pagare per una nuova versione di un prodotto).

In teoria, un procedimento razionale di valutazione delle diverse opzioni richiederebbe la conoscenza della relazione fra l'intensità in cui i diversi attributi sono presenti in ciascuna opzione e il valore per il consumatore della opzione stessa; ottenuta in questo modo una stima monetaria del valore per il consumatore delle diverse opzioni, si potrebbero confrontare questi valori con i rispettivi prezzi, e decidere di acquistare l'opzione a cui corrisponde il maggior surplus netto del consumatore, o in termini assoluti (differenza fra il valore per il consumatore delle diverse opzioni e il loro prezzo), o in termini relativi (rapporto fra il valore per il consumatore delle diverse opzioni e il loro prezzo). In concreto gran parte delle decisioni dei consumatori sono adottate attraverso delle procedure molto semplificate ("euristiche") che tengono conto a volte soltanto di una piccola parte delle informazioni necessarie per una decisione ottimale<sup>47</sup>.

Un agente intelligente di tipo *content based filtering*, può essere molto utile per guidare il consumatore verso decisioni di scelta di alternative che meglio corrispondono alle sue esigenze e aspirazioni. Ma, soprattutto in presenza di numerosi attributi da considerare, è notevolmente difficile individuare delle opzioni di prodotto che soddisfano tutte le preferenze del consumatore; spesso, soltanto alcune alternative al top del mercato riescono a soddisfare tutte le esigenze del consumatore, ma a prezzi molto alti, ed includendo anche caratteristiche che magari non sono particolarmente importanti per il consumatore. In più, effettuare compensazioni tra attributi positivi e negativi, come accade con le tecniche compensatorie di supporto alle decisioni, non solo può portare ad escludere opzioni in grado di soddisfare al meglio le preferenze del consumatore, ma può portare anche a scegliere opzioni che presentano attributi con prestazioni più elevate di quelle in effetti richieste dal consumatore; quindi ad un prezzo più alto di quello che si dovrebbe pagare. La correlazione negativa tra attributi, spesso tutti importanti per il consumatore, solitamente è dovuto alle scelte aziendali dei produttori, che puntano a sviluppare alcune caratteristiche del prodotto a scapito di altre per differenziarsi dai concorrenti. Purtroppo però, mentre le opzioni disponibili per ciascun tipo di prodotto crescono a un ritmo esponenziale, le capacità cognitive dei consumatori rimangono sostanzialmente invariate, e quando si deve decidere quale prodotto acquistare risulta sempre più difficile tener conto di tutte le informazioni necessarie per la scelta.

### **3.6.1 Un esempio: le caratteristiche di un telefono cellulare**

La scelta tra migliaia di opzioni di prodotti disponibili, soprattutto in ambienti di acquisto online, diventa quasi impossibile, nel caso in cui si intende acquistare un telefono cellulare.

Il telefonino nell'ultimo decennio ha attraversato un processo continuo di evoluzione tecnologica, ed è diventato uno dei prodotti più complessi presenti sul mercato, essendo caratterizzato oramai da numerosi attributi. Per capire meglio quanto sia difficile oggi per un consumatore fare la scelta giusta, di seguito sono elencati 28 attributi relativi a un telefono cellulare di ultima generazione.

---

<sup>47</sup> Uno dei primi a mettere in evidenza le difficoltà di adottare in concreto processi decisionali ottimizzanti fu Herbert Simon (1947), premio Nobel per l'economia nel 1978, nonostante che la sua attività d'insegnamento universitario fosse prevalentemente nel campo della psicologia, dell'informatica, e dell'intelligenza artificiale.

Ad esempio, in presenza di *recommender system*, potrebbe essere chiesto a un utente online, intento a scegliere il suo cellulare ideale, di definire quanto sia interessato, in base ad una scala Likert che va da per niente interessato a molto interessato, alle seguenti caratteristiche:

*Dimensione (misurata in mm<sup>3</sup>):* la dimensione di un telefonino è data dal prodotto base\*altezza\*spessore.

*Spessore (misurato in mm):* oltre al volume complessivo, anche il solo spessore solitamente risulta molto importante quando si acquista un cellulare; alcune imprese come Samsung e Motorola hanno incentrato delle campagne di marketing principalmente sullo spessore ridottissimo di alcuni loro modelli.

*Peso (misurato in grammi) :* Il basso peso di un cellulare è una caratteristica abbastanza apprezzata dal consumatore; può variare dai 60 grammi ai 355 grammi di un palmare.

*Uscita sul mercato (misurata in numero di mesi di presenza sul mercato):* questa caratteristica considera da quanti mesi il telefono cellulare considerato è presente sul mercato.

*Autonomia della batteria (misurata in minuti):* l'autonomia della batteria è una caratteristica molto importante per gli utenti; L'inserimento di nuove applicazioni tecnologicamente avanzate all'interno di un cellulare porta comunque ad un maggiore dispendio di energia che non è facile soddisfare, soprattutto perché gli sviluppi tecnologici per le batterie ci sono stati, ma più lentamente rispetto alle esigenze energetiche delle nuove opzioni applicative; una batteria più grande, d'altronde, comporta un aumento del peso complessivo del cellulare. Soprattutto nei cellulari tecnologicamente più avanzati i produttori devono cercare un compromesso tra funzioni tecnologiche, autonomia della batteria, e peso complessivo del telefonino; diversi telefonini vengono venduti con doppia batteria per ridurre il rischio di rimanere con il cellulare scarico.

*Risoluzione del display (misurata in numero di pixel):* ogni pixel è una piccola cella dello schermo; solitamente maggiore è il numero di pixel, maggiore è la dimensione dello schermo.

*Qualità del display (misurata in numero di colori):* questa variabile definisce il numero di colori che è possibile visualizzare sullo schermo del cellulare; maggiore è il numero di colori, migliore sarà la visualizzazione di foto e video. Attualmente diversi cellulari vengono prodotti ancora con gli schermi in bianco e nero, o con pochissimi colori, perchè consumano molto meno energia. Al momento il massimo numero di colori visualizzabili su un display è circa 16 milioni, come lo schermo di un computer.

*Touch screen (può essere presente o assente):* questa caratteristica dicotomica valuta la presenza o meno sul telefonino di uno schermo tattile, molto utile per velocizzare le selezioni nei menu, e l'apertura di applicazioni presenti nel cellulare.

*Sistema operativo (Symbian, Microsoft, OS X):* i cellulari solitamente sono dotati di un minimo di capacità di programmazione per consentire all'utente di usufruire delle funzionalità essenziali di un cellulare (mandare un messaggio o una e-mail, fare una foto o un video, ascoltare la musica, etc.). Attualmente diversi cellulari sono dotati di un vero e proprio sistema operativo molto simile a quello dei personal computer; questi cellulari vengono denominati *smartphone* (telefonini intelligenti). I principali sistemi operativi realizzati ad hoc per *smartphone* o *palmari*, sono il Symbian OS, il *Microsoft mobile* e OS X, che è una versione leggera del sistema operativo Mac OS X di Apple utilizzata per *I-Phone*.

*Velocità di navigazione su internet (misurata in Kb/s):* questa caratteristica riguarda le prestazioni del cellulare nello scambio di dati, ciò è molto importante per una veloce navigazione sul web attraverso il cellulare. L'evoluzione tecnologica ha portato a grandi velocità di download nella trasmissione di dati; oramai attraverso la tecnologia Hsdpa integrata in un cellulare si può raggiungere la velocità in download di 10 Mb/s, simile all'Adsl. L'evoluzione in questo campo

nell'ultimo quinquennio è stata molto rapida: GSM (14,4 Kb/s), GPRS (40 Kb/s), EDGE (200 kb/s), UMTS (1920 kb/s), HSDPA (10 Mb/s).

*E-mail (può essere presente o assente):* questa variabile dicotomica riguarda la possibilità o meno di inviare mail attraverso il cellulare considerato; oramai circa l'80% dei cellulari disponibili sul mercato italiano consentono di inviare mail con il telefonino.

*Memoria disponibile (misurata in Mb):* questa variabile misura la quantità di memoria disponibile all'interno del telefonino; oramai non si memorizzano più soltanto numeri in rubrica, ma soprattutto file audio e video, applicazioni di genere diverso, etc... , quindi aumenta l'esigenza di memoria disponibile; l'I-Phone dispone di una memoria interna di 16 Gigabyte.

*Memoria espandibile (può essere presente o assente):* questa è una variabile dicotomica, che riguarda la presenza o meno all'interno di slot che consentano di aggiungere memoria a quella già presente all'interno del cellulare. Oramai circa il 66% dei cellulari nuovi acquistabili sul mercato italiano consentono di espandere la memoria interna attraverso memorie ausiliarie.

*Mp3 (può essere presente o assente):* Variabile dicotomica che definisce se è possibile o meno ascoltare file musicali codificati in Mp3, attraverso il telefono cellulare.

*Tastiera qwerty (può essere presente o assente):* questa caratteristica dicotomica indica se il cellulare dispone di una tastiera *qwerty*, molto simile a quella di un personal computer tradizionale per disposizione delle lettere, ovviamente in un formato molto più ridotto. Tenendo conto delle dimensioni di un cellulare, non è facile adoperare al meglio una tastiera *qwerty* integrata, o a scomparsa, sui recenti schermi tattili. Una vera e propria rivoluzione si avrà con l'uscita sul mercato di cellulari con tastiera virtuale, vale a dire cellulari capaci di proiettare una tastiera *qwerty* virtuale su qualsiasi superficie grazie ad un raggio laser; persino i tasti emetteranno il classico suono di una tastiera normale. Si tratterà di una invenzione molto pratica e comoda soprattutto per creare e modificare anche applicazioni office.

*Radio (può essere presente o assente):* variabile dicotomica che riguarda la presenza o meno della radio integrata nel cellulare.

*Risoluzione della fotocamera (misurata in numero di pixel):* questa variabile rileva la qualità delle foto realizzabili attraverso la fotocamera integrata del telefonino. La risoluzione della fotocamera si misura in pixel, come per la risoluzione del display, ma in questo caso maggiore è il numero di pixel, maggiore sarà il formato di stampa in grado di ottenere. La fotocamera con la più alta risoluzione tra quelle selezionate, ha una risoluzione oltre i 5 MegaPixel (oltre 5 milioni di pixel).

Tabella 3.3 - Risoluzione ideale della fotocamera in base alle dimensioni della stampa che si dovrà realizzare

Formato Foto	Risoluzione ideale in pixel ( $\approx$ MegaPixel)
10x13 cm	1920000 ( $\approx$ 2)
13x17 cm	3000000 ( $\approx$ 3)
15x20 cm	4320000 ( $\approx$ 4)
20x26 cm	5332000 ( $\approx$ 5)
30x40 cm	5591040 ( $\approx$ 6)
50x70 cm	7051800 ( $\approx$ 7)

*Flash (può essere presente o assente):* variabile dicotomica che riguarda la presenza o meno del flash integrato nella fotocamera del cellulare.

*Connettività senza fili (IrDA, Bluetooth o Wi-Fi):* questa caratteristica riguarda la possibilità di collegare il cellulare ad altri dispositivi tecnologici (cellulari, palmare, auricolare, computer, ..), senza l'utilizzo di scomodi cavetti. Al momento sono tre le tecnologie in grado di collegare

dispositivi senza fili, partendo dalla più datata e quella più recente e veloce per lo scambio di dati: Infrarossi, Bluetooth e la più moderna Wi-Fi.

*Gps o navigatore satellitare (può essere presente o assente):* questa variabile dicotomica indica se il cellulare incorpora anche un navigatore satellitare integrato, vale a dire un Gps (*Global Positioning System*); si tratta di un sistema di posizionamento su base satellitare, a copertura globale e continua, molto utile per rilevare la posizione in cui ci si trova, e quindi definire il tragitto migliore per arrivare in un'altra posizione.

*A-Gps (può essere presente o assente):* questa variabile dicotomica riguarda la presenza o meno dell'*Assisted GPS*, o *A-GPS*, un sistema che consente di abbattere i tempi necessari alla prima localizzazione durante l'uso di un terminale GPS; essa è utile soprattutto nei "canyon" urbani, quali vie strette o viali notevolmente alberati, in cui è difficile stabilire con precisione la lista di satelliti in vista al terminale. Questo sistema si è diffuso negli ultimi anni, ed è normalmente associato ai sistemi di localizzazione basati su telefonia cellulare; diversi dei telefonini considerati hanno Gps e A-Gps insieme, oppure uno solo dei due.

*Videocamera (può essere presente o assente):* variabile dicotomica che indica la presenza o meno della videocamera integrata; solitamente un cellulare che dispone di fotocamera, dispone anche di videocamera.

*Videochiamata (può essere presente o assente):* questa variabile dicotomica indica se è possibile o meno vedere l'interlocutore sullo schermo mentre si effettua una telefonata; questa opzione richiede una velocità di scambio dati non inferiore a quella ottenibile dall'Umts.

### **3.7 Valutare le relazioni tra le caratteristiche attraverso le tecniche di regressione lineare**

Per valutare meglio le relazioni esistenti tra le caratteristiche di un prodotto, ad esempio di un telefonino, posso essere utilizzate tecniche di regressione lineare.

La *regressione* formalizza il problema della stima dei parametri di una relazione funzionale tra variabili (*gli attributi del prodotto*), sulla base di dati campionari estratti da un'ipotetica popolazione (*le opzioni di prodotto*). Più formalmente, in statistica la *regressione lineare* rappresenta un metodo di stima del valore atteso condizionato di una variabile *dipendente*  $Y$  dati i valori di altre variabili *indipendenti*  $X_1, \dots, X_k$ .

Nel nostro caso la variabile dipendente è il prezzo delle diverse opzioni di prodotto, e i valori delle variabili indipendenti sono gli indicatori di performance degli attributi (es. le caratteristiche di un cellulare) di ogni opzione di prodotto.

E' importante prima di effettuare una regressione lineare:

- 1) verificare se è teoricamente fondata la relazione funzionale ipotizzata tra le variabili esaminate;
- 2) supposto che esista una relazione funzionale tra le variabili considerate, è necessario verificare che essa sia di tipo lineare.

Successivamente si procede con i seguenti passi:

- a) individuazione della relazione causale tra le variabili;
- b) stima dei parametri;
- c) verifica della significatività delle stime dei parametri.

Per relazione funzionale consideriamo direttamente una relazione di tipo lineare la cui forma è:

$$Y = \alpha + \beta X$$

Dove:

Y = variabile dipendente

$\alpha$  = intercetta della retta di regressione

$\beta$  = coefficiente angolare della retta di regressione

X = regressore (variabile indipendente)

L'equazione che abbiamo appena visto è denominata "equazione della retta di regressione".

La retta di regressione approssima la "nuvola di punti" data dalle osservazioni campionarie.

Può essere effettuata una regressione lineare utilizzando metodi differenti.

I principali metodi utilizzabili per effettuare regressione lineari sono il metodo dei minimi quadrati (*Ordinary Least Squares*) ed il metodo di regressione a passi (*Stepwise Regression*).

Il metodo *Least Squares* è una tecnica di stima dei parametri che permette di trovare una funzione che si avvicini il più possibile ad un'interpolazione di un insieme di dati (tipicamente punti del piano). In particolare la funzione stimata è quella che minimizza la somma dei quadrati delle distanze dai punti dati.

La *Stepwise regression* è invece una tecnica utilizzata soprattutto per scegliere le variabili, ovvero i termini, da includere in un modello di regressione multipla. In presenza di molti attributi, ovviamente non tutti sono indispensabili per effettuare una buona selezione delle opzioni (non tutte le variabili sono significative); attraverso la *stepwise* è possibile identificare quali sono le variabili indispensabili e quali sono le variabili o attributi che si può fare a meno di considerare.

La tecnica *Stepwise* utilizza una metodologia di selezione delle variabili che utilizza delle procedure di "forward selection" (selezione in avanti) e di "backward elimination" (eliminazione all'indietro).

La procedura di *forward selection* inizia includendo nell'equazione la variabile esplicativa o indipendente (predictor) che ha il più alto coefficiente di correlazione con la variabile da spiegare o dipendente (criterion variable). La procedura di *backward elimination* inizia con l'inclusione di tutte le variabili esplicative, ed elimina sequenzialmente quelle che hanno un più basso potere esplicativo.

*Step 1:* Viene individuata la variabile esplicativa (caratteristica dei cellulari) che presenta il più elevato coefficiente di correlazione con la variabile da spiegare (prezzo dei cellulari). Se la probabilità associata al test di significatività è minore o uguale di 0,05, questa variabile esplicativa è inclusa nell'equazione.

*Step 2:* Viene individuata una seconda variabile esplicativa che presenta il più elevato coefficiente di correlazione parziale con la variabile da spiegare; se la probabilità associata al test di significatività è minore o uguale di 0,05, anche questa variabile esplicativa è inclusa nell'equazione.

*Step 3:* Viene individuata la variabile esplicativa la cui eliminazione riduce meno il potere esplicativo del modello; se il criterio per eliminare la variabile è soddisfatto ( $P_{OUT} = .10$ ), la variabile viene eliminata.

*Step successivi:* Il programma prosegue poi alternando procedure di "forward selection" e di "backward elimination". Esso termina quando nessun'altra variabile esplicativa soddisfa i criteri per entrare o essere eliminata dall'equazione.

### 3.8 Recuperare informazioni rilevanti tra tante informazione superflue, attraverso le tecniche di *data mining*

L'*information overload* si avverte anche negli ambienti online dedicati agli acquisti. Possono essere tante le opzioni di prodotto, gli attributi di ogni prodotto, le informazioni sui venditori e sui produttori, senza dimenticare le valutazioni rilasciate sui siti specializzati da esperti e/o consumatori che hanno provato il prodotto.

Per selezionare informazioni rilevanti tra tante informazioni inutili, esistono dei particolari programmi statistici in grado di effettuare le cosiddette analisi di *Data Mining*; una di queste applicazioni si chiama *Microsoft Analysis Service*. La statistica può essere definita come "estrazione di informazioni utili da insiemi di dati", il *data mining* è analogo, ma a differenza della statistica, ha a che fare con insiemi enormi di dati. Il *data mining* è quindi una tipica applicazione informatica usata per rintracciare ed accorpare dati significativi sepolti sotto una montagna di informazioni irrilevanti. L'applicazione di apprendimento automatico statistico *Microsoft Analysis Service*, ad esempio, consente la progettazione, la creazione e la gestione di strutture multidimensionali contenenti dati aggregati provenienti da diverse sorgenti di dati, solitamente database relazionali; essa dispone di algoritmi di *data mining* che consentono di creare modelli di *data mining*. Il termine inglese *mining* fa proprio riferimento al lavoro di estrazione che viene fatto nelle miniere. Il *data mining* è ritenuta una delle attività cruciali per la comprensione e lo sfruttamento dei dati nella nuova era digitale; si tratta del processo automatico di scoperta e individuazione di strutture all'interno dei dati, vale a dire di *patterns*, modelli e relazioni. Questo processo, noto anche col nome KDD (*Knowledge Discovery in Databases*), consente di estrarre conoscenza, in termini di informazioni significative ed immediatamente utilizzabili, da grandi moli di dati, tramite l'applicazione di particolari tecniche ed algoritmi.

*Microsoft Analysis Service*, utilizza in particolare due algoritmi, per effettuare analisi di *data mining*: *Decision tree analysis*, in grado di realizzare alberi decisionali, e *Cluster analysis*, per segmentare l'insieme di cellulari venduti online in gruppi omogenei al loro interno e dissimili tra di loro, detti *clusters*.

Intendendo per progetto, in questo contesto, l'applicazione di tecniche di *data mining* in un ambito specifico e circoscritto, per il raggiungimento di uno specifico obiettivo, le fasi di un progetto di *data mining* sono le seguenti:

- individuazione delle fonti dei dati;
- estrazione / acquisizione dei dati (e integrazione, se provenienti da fonti o *data base* diversi);
- *pre-processing* (pulizia dei dati, analisi esplorative, selezione, trasformazione, formattazione);
- *data mining* vero e proprio (scelta dell'algoritmo, individuazione dei parametri, elaborazione, valutazione del modello);
- interpretazione / valutazione dei risultati;
- rappresentazione dei risultati.

L'algoritmo *Microsoft Clustering* viene definito anche algoritmo di segmentazione perché consente di dividere i dati in gruppi, o *cluster*, di elementi con proprietà simili. Data una popolazione di oggetti rappresentabili come punti in uno spazio multidimensionale in cui ciascuna dimensione corrisponde ad una caratteristica (*feature*) di interesse, e data una funzione che esprima la similarità o vicinanza tra due punti, effettuare un *clustering* significa raggruppare gli oggetti in un ridotto numero di insiemi (*cluster*) che caratterizzino al meglio la popolazione [Jain, 1988]. In generale, si può affermare che il *clustering* individua e mette in evidenza le somiglianze tra oggetti; poiché le categorie da individuare non sono note a priori, nella letteratura sul *pattern recognition* in intelligenza artificiale si parla anche di apprendimento non supervisionato. L'algoritmo utilizza tecniche iterative per raggruppare i casi di un set di dati in *cluster* con caratteristiche simili. Tali raggruppamenti sono utili per l'esplorazione dei dati, l'identificazione delle relative anomalie e l'effettuazione di stime. L'algoritmo di *clustering* si differenzia dagli altri algoritmi di *data mining*, per esempio dall'algoritmo *Microsoft Decision*

*Trees*, per il fatto che non è necessario designare una colonna stimabile per creare un modello di *clustering*.

L'Algoritmo *Microsoft Decision Trees*, è in grado invece di creare alberi decisionali in base agli attributi disponibili; viene definito anche *algoritmo di classificazione*, perché consente di stimare una o più variabili discrete, in base agli altri attributi del set di dati. Gli algoritmi decisionali vengono usati per la comprensione di un particolare fenomeno, poiché permettono di classificare in ordine di importanza le cause che portano al verificarsi di un evento. In un albero di decisione ciascun nodo intermedio modella una scelta, tipicamente attraverso una condizione che confronta un attributo con una costante, ciascun arco in uscita da un nodo rappresenta un possibile esito della condizione ivi contenuta, ciascuna foglia rappresenta una decisione finale. Il principale vantaggio di questo tipo di *pattern* è la facilità di interpretazione da parte dell'utente.

Per gli attributi discreti, l'algoritmo esegue le stime in base alle relazioni tra le colonne di input in un set di dati. I valori di tali colonne, noti come stati, vengono utilizzati per stimare gli stati di una colonna designata come stimabile. In particolare, l'algoritmo identifica le colonne di input correlate alla colonna stimabile. Se ad esempio, in uno scenario finalizzato alla stima dei clienti che probabilmente acquisteranno una bicicletta, nove su dieci tra i clienti più giovani acquistano la bicicletta mentre solo due su dieci tra i clienti meno giovani l'acquistano, l'algoritmo desume che l'età rappresenta un criterio di stima valido per l'acquisto di biciclette. L'albero delle decisioni esegue le stime in base alla tendenza verso un determinato risultato.

Per gli attributi continui, l'algoritmo utilizza una regressione lineare per determinare le divisioni dell'albero decisionale. Se più colonne vengono impostate come stimabili o i dati di input contengono una tabella nidificata impostata come stimabile, l'algoritmo crea un albero delle decisioni separato per ogni colonna stimabile. L'algoritmo *Microsoft Decision Trees* genera un modello di *data mining* creando una serie di divisioni, chiamate anche nodi, nella struttura. L'algoritmo aggiunge un nodo al modello ogni volta che rileva una colonna di input correlata in modo significativo alla colonna stimabile. Il modo in cui l'algoritmo determina una divisione varia a seconda che venga stimata una colonna continua o discreta

Qualunque sia l'algoritmo utilizzato, dopo aver effettuato l'elaborazione dei dati, il programma espone i risultati attraverso il *visualizzatore modelli di data mining*. Esso consente di visualizzare il contenuto di modelli di *data mining* dal punto di vista di un singolo attributo e delle relative relazioni. Questo strumento mostra il contenuto del modello di *data mining* per ogni nodo influenzato da un singolo attributo nonché i dati dell'istogramma per ogni nodo, visualizza i nodi del modello di *data mining* utilizzati nel modello, incluse le relazioni tra i nodi e le regole o gli attributi a essi assegnati, sotto forma di una serie di caselle collegate. Ogni casella rappresenta un nodo di una singola struttura decisionale o di un singolo *cluster*. A ogni nodo viene assegnato un colore indicante l'intensità dei dati di un attributo applicabile a un nodo selezionato in relazione al numero totale di casi elaborati dal nodo selezionato. Le assegnazioni dei colori e l'attributo selezionato possono essere modificati utilizzando l'elenco a discesa dei colori della struttura nel riquadro della legenda. I nodi vengono rappresentati in base alla classificazione dei fattori degli attributi, da sinistra verso destra, nel riquadro con i dettagli del contenuto. Quanto più una suddivisione è rappresentata verso la parte bassa della struttura, tanto minore sarà l'influenza del fattore che l'ha causata sul modello di *data mining*. Inoltre, il riquadro degli attributi consente di ordinare gli attributi in base al numero di casi o alla probabilità di occorrenza nel nodo selezionato, al fine di poter meglio comprendere la pertinenza di un dato attributo per un nodo.

I vantaggi offerti dalla visualizzazione del contenuto del modello di *data mining* mediante il *visualizzatore modelli di data mining* sono la possibilità di comprendere i modelli e le regole che caratterizzano un set di casi e la capacità di ottimizzare questi modelli e regole in base ai dati di *training*. Ad esempio, è possibile utilizzare le funzioni di visualizzazione del visualizzatore modelli di *data mining* per evitare un problema comune del *data mining* denominato "sovradattamento" (*overfitting*). Il sovradattamento si verifica quando il modello di *data mining*

inizia a creare regole specifiche di singoli casi. Il modello inizia così ad attribuire importanza a informazioni in realtà non rilevanti. Si supponga, ad esempio, che esista un insieme dei clienti per un modello di *data mining* relativo a un punto vendita contenente il cognome del cliente come campo di attributo. Il modello di *data mining* potrebbe creare una regola in base alla quale un cliente di nome Bianchi ha maggiori probabilità di acquistare alcuni utensili, soltanto perché un singolo cliente di nome Bianchi li ha acquistati. Questa regola si basa su un modello casuale, privo di contenuto significativo, e rappresenta un esempio di sovradattamento; la correlazione tra il cognome di un cliente e il tipo di prodotti acquistati in realtà non è significativa. Nella maggior parte dei casi il sovradattamento si verifica quando a un modello di *data mining* vengono aggiunti attributi che non hanno contenuto significativo; in tali casi, il modello tenta di creare regole laddove non ne esistono.

### 3.9 Conclusioni

In questo capitolo sono state presentate delle tecniche di *information retrieval*, vale a dire di recupero delle informazioni rilevanti. Sono stati evidenziati i vantaggi dell'uso dei motori di ricerca per il *World Wide Web*, dove un utente può inserire un quesito (*query*), e ottenere una lista di pagine rilevanti.

Sono state considerate alcune delle principali metodologie di filtraggio delle informazioni, che sono le metodologie cognitive (*content based*) e le metodologie "sociali" o "collaborative" oppure una sintesi dei due approcci. Le metodologie cognitive (*content based*) selezionano le opzioni sulla base dei contenuti, per esempio la presenza o meno in un testo di alcune parole, oppure la presenza o meno in un telefonino di particolari caratteristiche ma con lo svantaggio di non poter estendere le valutazioni a prodotti significativamente diversi da quelli valutati dall'utente, con il rischio di perdere importanti opportunità. Le metodologie di filtraggio collaborative si basano invece, sulle relazioni fra le valutazioni di persone diverse; l'applicazione più interessante è rappresentata dai filtri collaborativi, che cercano di aiutare le persone a decidere sulla base delle opinioni di altri, nell'ipotesi che le persone che hanno evidenziato preferenze analoghe in passato concorderanno anche in futuro. Ma queste ultime hanno lo svantaggio di non poter essere applicate né a nuovi utenti né a nuove opzioni; quando si presenta un nuovo utente non si sa con quali vecchi utenti sono correlate le sue valutazioni, e quindi non esistono elementi per consigliargli delle opzioni; oppure le nuove opzioni non sono state ancora valutate da nessuno, e quindi non si sa a chi consigliarle. Per una efficace implementazione di un sistema collaborativo di raccomandazione è necessario che sia raggiunta una massa critica di valutazioni, quindi il sistema funziona bene soltanto se c'è un gruppo sociale di persone che collaborano, e nel quale sono molte le persone con simili sistemi di preferenze. I sistemi di raccomandazione collaborativi, difatti, sono maggiormente utilizzati online, soprattutto per la valutazione dei prodotti o servizi, sia per la loro semplicità, sia per la fiducia che i problemi possano essere superati, sia, soprattutto, perché con questi sistemi non è necessario considerare il contenuto, vale a dire le caratteristiche delle opzioni, per potere effettuare delle raccomandazioni. Facendo riferimento alle principali tecniche di filtraggio delle informazioni sono stati realizzati diversi sistemi di raccomandazione; speciali software intelligenti in grado di considerare contemporaneamente un vasto numero di prodotti presenti in un *database* e di selezionarne un set più ristretto, sulla base delle preferenze di ciascun consumatore. Pur tenendo conto dei diversi vantaggi offerti dall'utilizzo di sistemi *collaborative filtering*, l'utilizzo di un agente *content based filtering*, può risultare molto più efficace e affidabile nella selezione di prodotti che corrispondano meglio alle preferenze di ciascun consumatore. Questo perché un sistema di raccomandazione *content based*, avvalendosi dei valori di performance degli attributi di ogni opzione di prodotto, effettua delle comparazioni con dati oggettivi e non derivanti dalle indicazioni di un gruppo qualsiasi di opinionisti. Secondo Häubl e Trifts, i sistemi *content-based filtering* consentono ai consumatori di prendere decisioni migliori con un minore sforzo

cognitivo, allentando il trade-off tra grado di accuratezza delle decisioni e sforzo sostenuto [Häubl, Trifts, 2000].

Infine per quel che riguarda le tecniche specifiche di filtraggio delle informazioni, l'attenzione è stata in particolare concentrata sugli algoritmi di *clustering*, e in particolare sulle reti neurali autoorganizzantesi, sulle tecniche di regressione, e sulle analisi di *data mining*.

## Capitolo 4

### Commercio elettronico e reputazione online

#### 4.1 Introduzione

Uno dei principali limiti degli acquisti online è rappresentato dalla mancanza di rapporto umano e contatto fisico, tipico invece dei negozi tradizionali dove i clienti non solo possono toccare con mano i prodotti ma possono contare sulla presenza fisica del venditore o del commesso. La scarsa fiducia dell'utente/consumatore unita alla normale diffidenza a ignorare ciò che non si conosce, rappresentano di fatto alcuni dei principali problemi nella vendita online. Per riuscire a ridurre il più possibile la distanza e la diffidenza, possono essere d'aiuto due strumenti spesso assenti nella maggior parte dei negozi online: il forum e la valutazione dei prodotti.

Il Forum, ovvero un luogo virtuale di discussione, è al tempo stesso uno strumento etichettabile come "*killer application*" ma può rischiare al tempo stesso di diventare un elemento fortemente negativo. I rapporti di fiducia che si instaurano nel forum, con conseguente effetto di appartenenza ad un gruppo, permettono di abbattere molte delle barriere mentali nei potenziali clienti, migliorando così i tassi di conversione. Un forum deserto e poco frequentato, però, rischia di dare un'idea negativa sul negozio, così come un locale vuoto non invoglia sicuramente ad entrare. Un altro strumento indispensabile è la valutazione dei prodotti, specialmente quando si adoperano agenti di raccomandazione *collaborative filtering*. La valutazione dei prodotti comporta un incremento delle informazioni a disposizione del consumatore, quindi può risultare vincente per convincere gli utenti incerti a decidere l'acquisto.

Spesso difatti in molte esperienze di e-commerce, sono stati molto frequenti i casi in cui gli utenti avrebbero comprato volentieri un determinato prodotto ma spesso la vendita non si concretizzava per diffidenza verso il prodotto stesso (14/06/06 <http://www.seotopic.com>).

#### 4.2 Le difficoltà di un sito e-commerce

In Italia nove siti e-commerce su dieci sono stati realizzati in economia, ovvero sfruttando le conoscenze tecniche del singolo imprenditore coadiuvato generalmente da un software *Open Source*, da un manuale e da qualche *tutorial* pescato online. Ciò comporta conoscenze approssimative e improvvisate, siti tecnicamente e strutturalmente ormai sorpassati e commercialmente dispersivi, scarsa o nulla considerazione dei potenziali clienti; per evitare tutto ciò risulta molto importante valutare la cosiddetta *usabilità* del sito<sup>48</sup>, perchè sussiste una completa assenza del concetto di *friendly search engine* (ricerca amichevole).

Dopo aver creato il sito, bisogna adoperarsi per effettuare tutte le operazioni necessarie per rendere il sito raggiungibile dai futuri clienti. Un utente che arriva nel sito e-commerce tramite una particolare ricerca su un motore di ricerca è già mentalmente predisposto ad una possibile conclusione positiva della sessione di *shopping*, diversamente invece accade, quando un utente arriva nel sito tramite sollecitazione esterna effettuata tramite un *banner* o una mail promozionale.

Inoltre chiunque volesse avviare un'attività online, deve comunque aver presente come opera l'ambiente circostante, nel quale si inserirà. Occorre infatti tener presente della fortissima crescita tecnica dei siti informativi, che incidono notevolmente nelle scelte dell'utente. I siti

---

<sup>48</sup> L'aspetto fondamentale che dovrebbe curare chi si avvicina all'usabilità come web designer, è quello di cercare di mettersi nei panni del potenziale utente quanto più possibile, quindi il prodotto che essi realizzano, dovrebbe essere soprattutto "*human oriented*".

informativi sono generalmente tecnicamente più avanzati rispetto ai tradizionali siti *e-commerce* e in linea con le necessità indotte dalla concorrenza e dalla evoluzione del mercato. I siti informativi spesso guadagnano attraverso la pubblicità, hanno quindi lo scopo principale di assorbire quanto più traffico possibile. Includono spesso sistemi interni di comparazione prezzi o vetrine dei vari *merchant* che risultano più appetibili agli occhi dei potenziali consumatori, quindi danno innanzitutto informazione, ma contemporaneamente cercano di monetizzare il loro lavoro inducendo il consumatore a fare una scelta che privilegi i siti di vendita associati (13/03/2006 <http://www.seotopic.com>).

La pubblicità comunque non basta, ci sono una miriade di siti internet che spendono decine di migliaia di euro in promozioni e *pay per click*, pensando che tante più visite ci saranno, tante più vendite si potranno ottenere, mettendo sempre in secondo piano l'usabilità del sito; un rallentamento dell'apertura di una pagina, può portare il consumatore a pensare che il sito sia carente, anche se potrebbe essere colpa del *server* e non del sito stesso.

Una delle cose più frequenti per chi possiede un *e-commerce* è che l'utente dopo aver visto il catalogo dei prodotti e aver inserito qualche prodotto nel carrello, ad un certo punto abbandona il tutto, senza che si riesca a capire quale potrebbe essere la motivazione logica.

I motivi per cui un utente può abbandonare la sessione di acquisto senza comprare, possono essere tanti, molti dei quali sono i seguenti:

- 1) Spesso l'abbandono è legato semplicemente all'umore dell'utente che all'improvviso può cambiare idea sull'acquisto da fare. Un buon aspetto grafico e comunicativo del sito può limitare questi abbandoni. Solitamente un utente valuta un sito nel giro di pochi secondi, e spesso la prima impressione rimane viva per lungo tempo.
- 2) Di solito il prezzo del prodotto viene giudicato troppo alto rispetto alle aspettative; l'utente si aspetta che on-line i prezzi siano sempre più bassi di quelli dei punti di vendita tradizionali, ma non è sempre così.
- 3) Prima di effettuare una scelta, il consumatore/utente verifica i prezzi proposti da più siti *e-commerce*, e quindi naviga in più siti senza l'intenzione reale di acquistare.
- 4) Tempi troppo lunghi e procedimenti complessi per acquistare un prodotto possono scoraggiare il consumatore, quindi è notevolmente importante facilitare l'utilizzo del carrello e semplificare la conferma dell'ordine.
- 5) Occorre limitare la richiesta di dati non necessari per l'evasione dell'ordine ( basterebbe nome, cognome, indirizzo e un recapito telefonico) per evitare che l'utente si intimorisca e abbandoni la fase di acquisto.
- 6) Può capitare che nell'utilizzo del carrello il sito si blocchi per problemi tecnici per via di *script* poco robusti o *server* poco affidabili, e questo crea nell'utente l'idea di sito poco affidabile.
- 7) Processi di pagamento non molto trasparenti, spese in eccesso non ben documentate, e forme di pagamento non usuali possono spaventare un cliente che prova per la prima volta ad acquistare online.
- 8) Altro problema sono le spese di spedizione, spesso troppo elevate rispetto al totale dell'ordine.
- 9) L'utente inoltre, solitamente è diffidente nei confronti delle cose che non può toccare; scarse o errate documentazioni sui prodotti possono quindi fare in modo che l'utente rinunci all'acquisto.
- 10) Informazioni insufficienti sull'azienda.
- 11) Infine, assistenza poco veloce, e-mail che tornano indietro per "*mail box full*", mancanza di numero verde o numeri telefonici fissi.

Gli accorgimenti che possono essere utili per ovviare a questi inconvenienti sono invece i seguenti:

- 1) Proporre prezzi in linea con il mercato; prezzi troppo bassi possono creare nell'utente preoccupazioni di "*fregatura*" legate a un mondo virtuale; prezzi troppo alti per alcuni prodotti faranno catalogare il sito come poco conveniente anche per tutti gli altri prodotti.

- 2) Visualizzare e specificare dall'inizio i prezzi per le spedizioni evitando di rimandarli a dopo la conferma dell'ordine.
- 3) Predisporre la possibilità di conservare il proprio carrello per un periodo di almeno 15-30 giorni in quanto spesso gli utenti ritornano sapendo di non dover rifare tutte le operazioni per inserire i prodotti nel catalogo, e comunque si tratta di un servizio utile per chi effettua spesso acquisti degli stessi prodotti.
- 4) Sono molto importanti le promozioni sui prodotti, che devono essere fatte spesso; in questo modo l'utente ritorna volentieri per vedere se un certo prodotto è in offerta, se ci sono occasioni e prodotti in liquidazione. Particolarmente stimolante possono essere le cessioni di prodotti, fatte gratuitamente, in base al valore del carrello ordinato; per avere qualcosa gratis i clienti aggiungono spesso prodotti che normalmente non avrebbero comprato.
- 5) Uno stimolo importante per l'acquisto è spesso la consegna gratuita, anche se un po' rischioso come sistema se non controllato bene; una buona idea potrebbe essere quella di collegarla all'importo dell'ordine oppure alla tipologia dei prodotti acquistati.
- 6) Evitare le registrazioni obbligatorie al sito; molti utenti sono restii a comunicare i propri dati personali preventivamente, e anche se acquistano, non per questo vogliono restare in qualche modo legati al sito.
- 7) Il catalogo dei prodotti deve essere sempre ben fornito e aggiornato; un catalogo misero non incentiva l'utente ad acquistare, e per evitare sorprese, occorre sempre indicare le quantità di merce giacente in magazzino.
- 8) Molto utile è il servizio di rintracciabilità dell'ordine on-line, con la visualizzazione della partenza della merce e la data di arrivo prevista; spesso si chiama l'assistenza per avere queste informazioni.
- 9) Limitare il numero di passaggi per confermare l'ordine; l'utente tende a spazientirsi molto facilmente e predilige compilare i campi su una singola pagina, o per lo meno non più di 4 passi consecutivi.
- 10) Proporre campagne promozionali a punti per fidelizzare il cliente, magari con buoni omaggio o sconti in percentuale dei punti raccolti; oppure generare un buono sconto in base agli acquisti fatti. L'idea di "guadagnare bonus mentre si spende" porta l'utente a spendere sempre qualcosa in più.
- 11) Comunicare sempre tutti i dati sull'azienda, titolare, responsabili di vendita e chi si occupa dell'assistenza, informazioni ridotte e poco chiare creano sfiducia. Usare i nomi reali e ancora meglio inserire una foto dello staff, perché rende meno freddo il rapporto cliente-venditore.
- 12) Importantissima è l'assistenza in tempo reale, eventuale numero verde, *chat* con responsabile vendite, per chiarire gli ultimi dubbi in tempo per completare il carrello.
- 13) Disponibilità di diverse modalità di pagamento, in modo da dare al cliente la possibilità di pagare con il sistema che preferisce.
- 14) Informare l'utente riguardo lo stato dell'ordine, con successive conferme (ricezione ordine, conferma presa in carico, conferma spedizione); in questo modo non soltanto si evitano chiamate al servizio clienti, ma si ricorda il sito all'utente, che può essere stimolato così a visitarlo nuovamente.
- 15) Conviene spendere qualche euro in più su una spedizione, piuttosto che rimetterci un cliente con relativa pubblicità negativa; i clienti online sono ancora relativamente pochi, ma generalmente sono persone piuttosto informate su Internet, spesso scrivono su blog, newsgroup, forum o parlano agli amici; un problema con un cliente può trasformarsi in una drastica diminuzione di nuovi potenziali clienti; un buon servizio si trasforma invece in pubblicità positiva (26/01/06 <http://www.seotopic.com>).

### 4.3 I siti “comparatori” di prodotti online

La nuova frontiera del commercio elettronico riguarda l'evoluzione dei siti di comparazione delle offerte dei prodotti venduti online, in modo da facilitare sempre più il processo decisionale del consumatore. Questi siti specializzati utilizzano dei *crawler*, (vedi capitolo precedente), per poter esplorare la rete alla ricerca delle offerte disponibili nei vari siti e-commerce. Gran parte dei siti di comparazione presenti online adottano ancora semplici matrici di comparazione per presentare agli utenti i vari prodotti, dove ogni colonna rappresenta una opzione di prodotto con l'offerta di prezzo del sito e-commerce e le righe mostrano i valori di performance degli attributi di ogni opzione di prodotto. Solitamente però, non è possibile comparare più di 5 opzioni per volta, pur essendoci decine e decine di opzioni disponibili sul web; inoltre risulta difficile valutare, tramite matrice, quale sia l'opzione migliore, poiché spesso, quando una opzione risulta soddisfacente per alcuni attributi, risulta insufficiente per altri attributi. Pertanto diversi siti comparatori oltre alle matrici di comparazione, utilizzano sistemi collaborativi di comparazione, consentendo agli utenti esperti e/o consumatori di lasciare non solo una opinione sul prodotto, ma anche un punteggio di valutazione (ad esempio da 1 a 10) che permette di ordinare le diverse opzioni in base alla popolarità ed alla preferenza. Come evidenziato nel capitolo precedente, sistemi di tipo collaborativo possono facilmente essere adottati per qualsiasi tipo di prodotto, e quindi sono molto più semplici da sfruttare, ma hanno diversi svantaggi. Sono invece ancora troppo pochi i siti di comparazione basati su sistemi di raccomandazione content based filtering, in Italia ad esempio sono del tutto assenti.

In ogni modo il web italiano è invaso da siti web che si definiscono comparatori di prezzo. Sono pochissimi quelli che stanno avendo un discreto successo. Tra i principali siti italiani, che si possono veramente definire motori di ricerca per l'acquisto online, segnaliamo Kelkoo.it, BuyCentral.it, Costameno.com, Twenga.it. Ognuno ha dei pregi e dei difetti, ma offrono comunque un buon servizio di comparazione dei prezzi esplorando la quasi totalità delle offerte disponibili sul mercato italiano. Kelkoo e Twenga, in particolare, non sono creazioni italiane, ma francesi, e successivamente sono state esportate negli altri paesi europei.

L'Italia, ovvero il web italiano incentrato sugli acquisti online, si è avvicinato tardi a questi nuovi servizi, senza pensare alle enormi opportunità che offrono. Ciò è stato un peccato, soprattutto perché secondo recenti sondaggi siamo tra gli ultimi in Europa nelle vendite e negli acquisti online. Facendo un esempio, sono 1108 i cellulari disponibili sul sito kelkoo.fr (Kelkoo francese) e per ogni prodotto sono presentate una media di almeno 8-10 offerte di vari siti e-commerce. Kelkoo.it ha un database di cellulari che supera di poco le 350 unità con una media di offerte di poco inferiore al Kelkoo francese. Invece dovrebbe essere il contrario, secondo una recente indagine della Banca mondiale che reputa l'Italia popolata di persone che fanno largo uso del telefono cellulare<sup>49</sup>.

---

<sup>49</sup> Secondo l'indagine della Banca mondiale, l'Italia vanta il primato del più alto numero di abbonamenti ai servizi di telefonia mobile, mentre a livello mondiale si trova al sesto posto. I dati relativi al 2006 riferiscono un tasso di penetrazione del 122 per cento (122 contratti di telefonia mobile su 100 cittadini). Nell'Unione Europea nessun mercato nazionale può vantare un così esteso bacino di utenza: alle spalle dello stivale si trovano la Gran Bretagna e il Portogallo (115), Spagna (105), Germania (102) e Grecia (100). Secondo la Banca mondiale, questo primato è dovuto alla convenienza delle tariffe applicate dagli operatori di telefonia mobile, precisando che il "paniere mensile" (ossia un'ipotesi media di 25 telefonate e 30 messaggi di testo) corrisponde ad una spesa di 14,1 dollari, una cifra ritenuta tra le più basse nel vecchio continente e battuta solo dalla cifra relativa alla Gran Bretagna (13,7 dollari). Secondo gli stessi parametri, nel mercato francese un numero inferiore di utenze sarebbe da imputare a tariffe di telefonia mobile più onerose (29,4 dollari, sempre per 25 telefonate e 30 SMS).

### 4.3.1 Cosa può fare un sito e-commerce per agevolare i motori di ricerca dedicati allo shopping

Spesso molti utenti online, entrano in un sito e-commerce dopo aver visionato le offerte selezionate da un comparatore online. È logico quindi che un sito e-commerce, non solo ha la necessità di accogliere nel migliore dei modi i suoi utenti, ma in primis deve adottare le giuste strategie per posizionarsi meglio sui migliori comparatori disponibili online.

I comparatori di prezzo, difatti, non sono tutti uguali, è quindi normale che molti *e-commerce* si affidino a più piattaforme di questo genere; utenza diversa, diversa usabilità e interfaccia, diversi contenuti, modalità e logica del motore e quindi diversa *SERP* (*search engines results page*), che non è altro che la pagina dei risultati di un motore di ricerca. La *SERP* di un comparatore di prezzo, ad esempio, è molto importante per un sito e-commerce, perché posizionarsi tra i primi in elenco, nei risultati di un motore di ricerca, in questo caso dello shopping, aumenta le possibilità di scelta da parte del consumatore. Per ottenere ciò il *webmaster* del sito deve adottare delle tecniche in grado di migliorare la posizione delle proprie pagine nelle *SERP* dei comparatori. Queste attività sono indicate con l'espressione "ottimizzazione" o "SEO" (*Search Engine Optimization*), mentre l'operazione in sé è chiamata "posizionamento nei motori di ricerca". Per ottenere un buon posizionamento sui comparatori online, un sito e-commerce dovrebbe seguire taluni accorgimenti:

- 1) non portare lo scontro soltanto sul *fattore prezzo*, anche se questo è un suggerimento rivolto anche ai merchant che non utilizzano i comparatori. "E-commerce" non è solo sinonimo di "prezzi stracciati"; cercare di battere i propri concorrenti sul solo *fattore prezzo* spesso significa innescare delle dinamiche deleterie per il proprio mercato di riferimento e, nel medio periodo, anche per il proprio business; vendite con margini eccessivamente contenuti, nessuna differenziazione tra il proprio e-shop e quello dei concorrenti, scarsa fidelizzazione del cliente che sarà dunque portato sempre e comunque a scegliere di acquistare solo dove paga meno; il danno provocato da una concorrenza eccessiva sui prezzi è inversamente proporzionale ai margini del settore. Se osserviamo meglio l'elenco dei cellulari immessi sul mercato dai produttori, notiamo che molti modelli non sono stati nemmeno presi in considerazione dai venditori online, quindi gran parte di essi puntano su pochi modelli, arrivando ad una deleteria guerra dei prezzi.
- 2) Fornire un buon *feed prodotti* ai comparatori, perché questi aggiornano quotidianamente i loro database con le informazioni presenti nei siti e-commerce clienti (prezzo, dato sul prodotto e immagini disponibili, spese di spedizione) attraverso spiders o robot o crawler<sup>50</sup>. Sappiamo che queste applicazioni, mentre analizzano il contenuto del sito, interrogano un particolare file messo a disposizione dal venditore per portare a conoscenza del comparatore i dati sui prodotti venduti. Di solito questo file è in formato "txt" o "csv", quindi più è completo ed aggiornato questo file e più efficace ed efficiente sarà la comparazione con i concorrenti.
- 3) L'assenza di un'immagine associata al prodotto chiaramente non spronerà l'utente a visitare l'offerta di quel negozio (è come avere una vetrina senza niente esposto); senza contare poi che molti motori di comparazione offrono all'utente la possibilità di effettuare delle ricerche tra i soli prodotti ai quali è associata un'immagine.
- 4) E' opportuno raccogliere, analizzare e intervenire sui dati che riguardano la navigazione degli utenti; gli utenti infatti, lasciano una scia sui siti e-commerce, anche se poi decideranno di non comprare; attraverso ad esempio strumenti come *Google Analytics* è possibile definire provenienza, interessi, e incertezze del consumatore durante la navigazione, ma anche le preferenze per quanto concerne la logistica e le modalità di pagamento, i contenuti delle

---

<sup>50</sup> Ecco alcuni crawler presenti online, specificati dal nome del Motore di ricerca e dal "nome spider": Google "googlebot", Altavista "scooter" e "mercator", Teoma "teoma\_agent", Yahoo "slurp".

pagine web sui quali focalizzare l'attenzione e quali sono i link che seguono, le risorse esterne dalle quali arrivano, i tempi di permanenza medi e il numero di pagine visitate da ogni singolo utente. Un'analisi dei dati raccolti porterà a correggere molti aspetti della propria attività: come la presenza e il posizionamento sui motori di ricerca, le campagne *cpc* (*costo per click*), i programmi di affiliazione, l'usabilità del sito, l'efficacia del proprio *customer care*, l'allineamento dei prezzi rispetto alla concorrenza, le partnership esterne, etc. Il *CPC* ("*costo per click*", vale a dire un importo calcolato su ogni visita generata) può variare in modo molto sensibile in funzione della categoria merceologica.

I prossimi sottoparagrafi descrivono alcuni comparatori online che si differenziano dagli altri per alcune specificità molto apprezzate dai consumatori che decidono di acquistare online.

#### 4.3.2 Phonescoop.com: un portale che seleziona i cellulari con un sistema *content based*

Phone scoop, è un portale dedicato esclusivamente al mondo della telefonia; esso è molto utile sia per i consumatori intenzionati ad acquistare un cellulare, sia per gli esperti interessati a informazioni dettagliate sui prodotti. E' stato lanciato sul web nel dicembre del 2001, e si trova fra i primi 100 siti più interessanti del web nella classifica dei TOP 100 della rivista *PC Magazine*. Tuttavia, dato che è un sito tutto USA, è possibile fare il confronto solo dei cellulari disponibili nel mercato americano, con annessi "carriers" (gestori telefonici).

La pagina più innovativa di questo sito è la pagina "*Phone Finder*", che aiuta l'utente seguendo due criteri, *simple* oppure *weighted* per trovare il cellulare più adatto. La comparazione dei cellulari viene fatta tenendo conto dell'importanza assegnata da ciascun consumatore ad ogni caratteristica disponibile in un cellulare.

Figura 4.1 – La pagina "*Phone Finder*" di *Phonescoop.com*.

The screenshot shows the Phonescoop.com website interface. At the top, there's a blue header with the 'phone scoop' logo and a search bar for jumping to a phone. Below the header, there's a featured advertisement for the BlackBerry Bold PDA. The main navigation menu on the left includes links for News, In Depth, Phones, Phone Finder, Carriers, Forums, Glossary, and Links. The central 'Phone Finder' section is currently in 'Simple' mode and provides instructions on how to use the search filters. A 'Carrier' dropdown menu is open, showing options like 'Don't Care', 'Show Me', 'Preferred', 'Important', 'Very Important', and 'Require'. The right sidebar contains several 'Ads by Google' for various mobile-related topics.

La scala di valutazione utilizzata per definire l'importanza è la seguente: *Don't Care*, *Show Me*, *Preferred*, *Important*, *Very important*, *Require*; si può definire il cellulare desiderato nei minimi dettagli, poiché *Phone Scoop* presenta un database dei cellulari notevolmente più completo dei diversi siti *e-commerce* che danno la possibilità di effettuare una comparazione dei prodotti tenendo conto al massimo di 10-15 caratteristiche disposte su una matrice di comparazione, e comunque, pur essendo poche caratteristiche, risulta sempre molto difficile valutare quale sia il prodotto più adatto.

Per le caratteristiche che possono presentare differenti valori di *performance*, ad esempio il numero di pixel della fotocamera, l'autonomia minima che deve avere il cellulare, il tipo di memoria espandibile preferito, etc., l'utente seleziona la *performance* che lo soddisferebbe, e successivamente definisce l'importanza da attribuire alla caratteristica in base alla classifica vista precedentemente. Invece, per le caratteristiche dicotomiche, che possono essere semplicemente presenti o assenti, quali, ad esempio, bluetooth, Wi-Fi, Registrazione Video, MMS, il consumatore deve solo definire quale importanza attribuire alla presenza della caratteristica considerata.

La figura seguente illustra come si presentano alcune caratteristiche:

Figura 4.2 – La pagina delle caratteristiche di *Phonescoop.com*.

The screenshot displays a web interface for selecting mobile phone characteristics. The categories and their options are as follows:

- Battery Life: Talk Time**: Importance dropdown set to "Don't Care". Below it, a "Minimum:" input field is followed by "hours" and statistics: "(Min: 0 Avg: 4.2 Max: 18)".
- Battery Life: Standby Time**: Importance dropdown set to "Don't Care". Below it, a "Minimum:" input field is followed by "hours" and statistics: "(Min: 0 Avg: 222.7 Max: 770)".
- Display Resolution (approx.)**: Importance dropdown set to "Don't Care". Below it, several resolution options are listed with checkboxes: 96 x 64, 101 x 80, 128 x 128, 128 x 160, 176 x 220, 240 x 240, 240 x 320 (QVGA), and > QVGA.
- Display Type**: Importance dropdown set to "Don't Care" (with a dropdown menu open showing options: Don't Care, Show Me, Preferred, Important, Very Important, Require). Below it, several display technology options are listed with checkboxes: LCD (Black & White), LCD (Grayscale), LCD (Color), LCD (Color STN), LCD (Color TFT/TFD), OLED (Full-Color), and Electrophoretic.
- Video Capture**: Importance dropdown set to "Don't Care".
- Video Sharing**: Importance dropdown set to "Don't Care".
- Integrated PDA**: Importance dropdown set to "Don't Care".
- Software**: Importance dropdown set to "Don't Care" (with a dropdown menu open showing options: Don't Care, Show Me, Preferred, Important, Very Important, Require, Exclude, Don't Care).
- BREW**: Importance dropdown set to "Preferred".
- Games**: Importance dropdown set to "Require".
- Java (J2ME)**: Importance dropdown set to "Don't Care".

*PhoneScoop* non solo offre all'utente la possibilità di confrontare i cellulari tenendo conto delle performance degli attributi, ma l'utente se vuole può anche considerare prima di scegliere anche il cosiddetto "Average User Rating", vale a dire il voto dato dagli altri utenti esperti o acquirenti, che quindi hanno conosciuto il prodotto precedentemente. Per fare ciò, il

consumatore stabilisce il minimo *rating* che lo soddisferebbe, e poi definisce l'importanza da attribuire al voto dei navigatori, come si può vedere dalla figura successiva.

Figura 4.3 – La pagina “Average User Rating” di *Phonescoop.com*.

L'applicazione di PhoneScoop è stata testata stabilendo le *performance* minime e l'importanza da attribuire ad alcune caratteristiche ritenute importanti.

La figura successiva mostra la pagina web con i cellulari filtrati ed elencati in modo decrescente secondo il punteggio ottenuto; se la performance dell'attributo è superiore a quella minima che era stata inserita, viene visualizzata in grassetto altrimenti no; neanche il cellulare che ha ottenuto il punteggio più elevato soddisfa tutte le richieste.

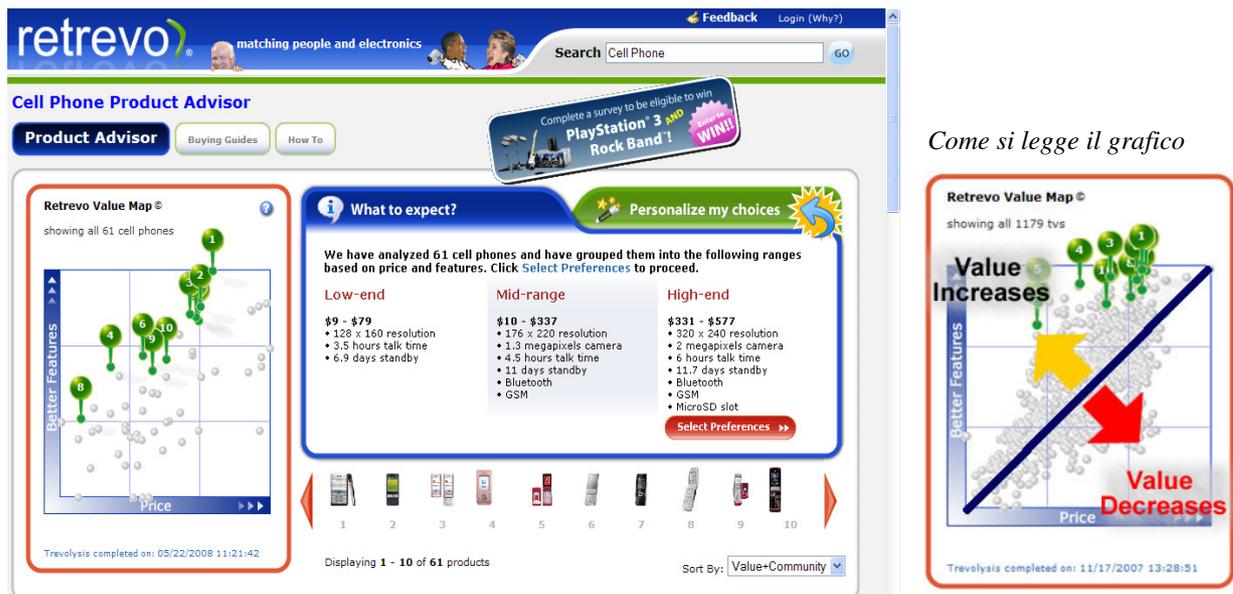
Figura 4.4 – La pagina dei risultati di *Phonescoop.com*

CLICK TO SORT	Manufacturer & Model (sorted)	Talk Time	Standby Time	Display Resolution	Bluetooth	Camera	Email Client	Hard Drive	Memory Card Slot	MMS	Packet Data	Video Capture	Wi-Fi	Compare
	<b>Samsung SGH-D900 / Black Carbon</b>	6.5000000	260.0000000	240 x 320	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	EDGE	Yes	-	<input type="checkbox"/>
	<b>Nokia 6500 slide</b>	6.0000000	320.0000000	240 x 320	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	EDGE	Yes	-	<input type="checkbox"/>
	<b>Nokia 5610</b>	6.0000000	311.9975040	240 x 320	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	EDGE	Yes	-	<input type="checkbox"/>
	<b>Samsung SGH-P850 / P858</b>	3.0000000	300.0000000	240 x 320	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	EDGE	Yes	-	<input type="checkbox"/>
	<b>Samsung Ultra Edition 10.9 / U600</b>	4.5000000	345.4000000	240 x 320	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	EDGE	Yes	-	<input type="checkbox"/>
	<b>Nokia E90</b>	5.8000000	336.0000000	800 x 352	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	HSDPA 3.6	Yes	Yes	<input type="checkbox"/>
	<b>HP iPAQ 610 Business Navigator</b>	4.0000000	250.0000000	240 x 320	Yes	3+ megapixel	Yes	-	microSD (TransFlash)	Yes	HSDPA 7.2	Yes	Yes	<input type="checkbox"/>

### 4.3.3 Retrevo.com: con un sistema *content based*, mette in relazione prezzo e performance delle caratteristiche per diverse tipologie di prodotto.

Un interessante sito di comparison shopping è *retrevo.com.*; questo sito è unico nel suo genere, perché fornisce immediatamente al consumatore una valutazione dei prodotti su un diagramma cartesiano, dove l'asse delle y identifica la prestazione complessiva delle caratteristiche di ogni prodotto, mentre l'asse delle x misura il prezzo ( Figura 5.5).

Figura 4.5 – Cellulari selezionati mettendo in relazione il prezzo con la performance complessiva delle caratteristiche del prodotto nel sito *Retrevo.com.* (Per accedere alla seguente pagina web: *Retrevo.com (Home Page) → Cell Phone*)



La figura 4.5 mostra come i cellulari che si trovano sotto la diagonale, che divide in due l'asse cartesiano, sono da scartare, a differenza di quelli che si trovano sopra la diagonale; i primi 10 cellulari si trovano infatti sopra la diagonale. Occorre però far presente che a differenza di altri comparatori online, questo sito dispone nel proprio database delle informazioni di soli 61 cellulari, quindi viene rappresentata una minuscola fetta di quello che potrebbe essere il mercato complessivo. Inoltre tramite la rappresentazione con diagramma cartesiano l'applicazione riesce a delineare 3 raggruppamenti di prodotti, vale a dire i cellulari di fascia bassa (\$9 - \$79), quelli di fascia media (\$80 - \$337) e i cellulari di fascia alta (\$338 - \$577).

L'ordine di preferenza dei cellulari può anche essere fatto tenendo conto delle opinioni rilasciate dalla *community* di *Retrevo.com*, vale a dire le opinioni degli utenti/consumatori e degli utenti/esperti registrati al sito. Le figure 4.6, 4.7, 4.8 mostrano i risultati ottenibili, seguendo 3 diversi metodi, vale a dire, *Valori + Community*, oppure *solo Valori* oppure *solo Community*.

Figura 4.6 – Cellulari selezionati in base al prezzo, alla performance complessiva, e alle opinioni degli utenti, nel sito *Retrevo.com*

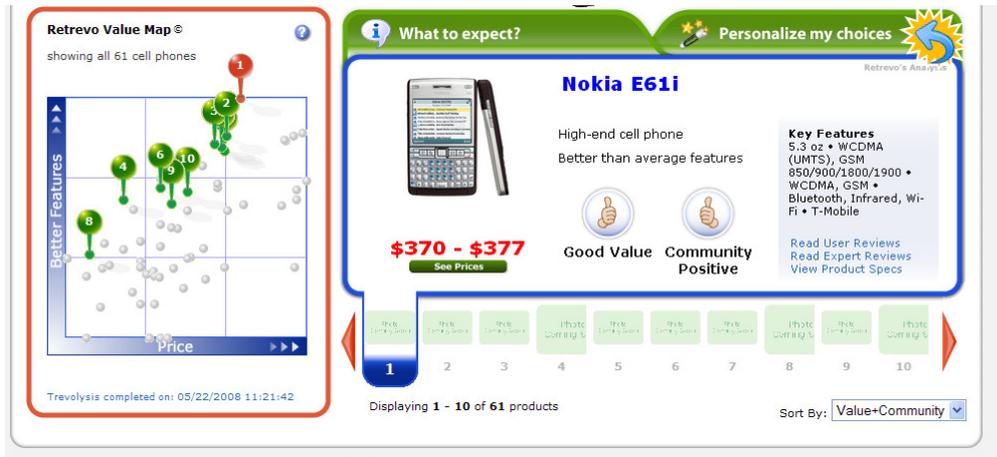


Figura 4.7 – Cellulari selezionati in base al prezzo e alla performance complessiva, nel sito *Retrevo.com*

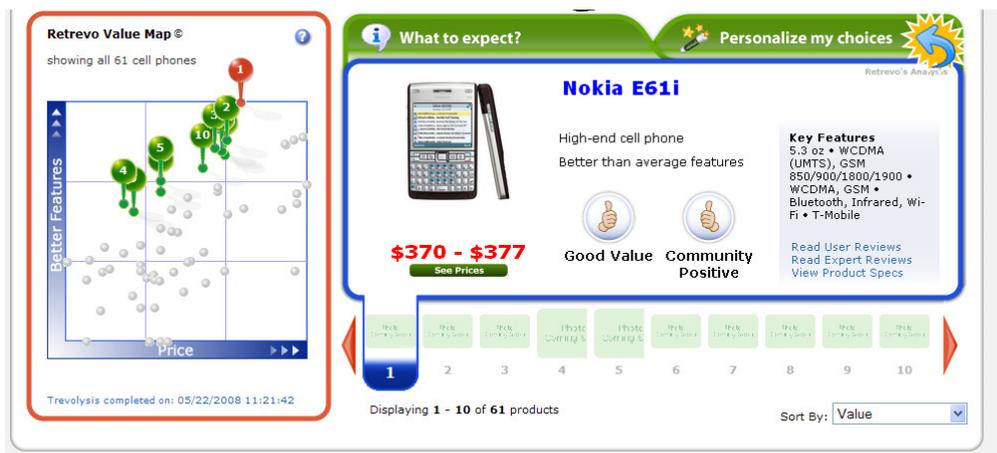


Figura 4.8 – Cellulari selezionati in base alle opinioni degli utenti, nel sito *Retrevo.com*.



Retrevo.com è stato fondato nel 2005 a *Sunnyvale*, in California ed opera esclusivamente all'interno del mercato USA.

#### 4.3.4 Twenga.it: uno dei più completi *price comparison*

Sono numerosi i siti online che si occupano prevalentemente di comparare i prezzi dei prodotti venduti in differenti siti e-commerce, il cosiddetto “*price comparison*”, quelli di maggior successo presentano la comparazione dei prodotti di alcune centinaia di siti preposti alla vendita online, e più o meno danno al consumatore/utente una panoramica generale di quello che offre il mercato e quanto verrebbe la spesa. Quindi al consumatore per poter effettuare una buona scelta non basterebbe accedere a un solo motore di ricerca, ma prima di acquistare dovrebbe conoscere le offerte di più motori di ricerca e quindi effettuare una comparazione tra i diversi motori. Alla fine otterrà una migliore panoramica del mercato, ma rimarrà sicuramente all’oscuro delle offerte di qualche sito non considerato dai motori di ricerca sottoposti a confronto. Una soluzione al problema che abbatte di molto la ricerca del consumatore nel confronto dei risultati di differenti motori di ricerca è stata ottenuta dal portale web “*Twenga*”.

*Twenga* è un motore di ricerca per lo shopping tra i più completi al mondo; rispetto ai concorrenti potenziali, il suo principale punto di forza è dato dal vastissimo catalogo di prodotti che possiede. Solo il portale italiano, “*Twenga.it*”, offre un catalogo di 4 milioni di prodotti disponibili in 2.500 negozi che vendono online nei confini italiani. *Twenga*, nel dialetto africano Swahili significa « dritto all’obiettivo », quindi propone una vasta selezione di prodotti e opinioni, per guidare il consumatore nell’universo dell’acquisto on-line. *Twenga* è nato in Francia, a Parigi, nel 2004, per iniziativa di Bastien Duclaux e Cédric Anès; attualmente esso è presente in tutta Europa, con differenti versioni del sito in Inghilterra, Francia, Germania, Spagna, Italia, Polonia, Olanda, Belgio e Svizzera.

Cos’è *Twenga*, e dove vogliono arrivare i suoi fondatori, si può capire leggendo la loro *mission* presente sul sito: “*offrire una guida allo shopping esaustiva e imparziale considerando tutti i venditori, le marche e i prodotti disponibili sul web*”. Questa grande disponibilità è stata ottenuta attraverso alcuni algoritmi sviluppati internamente, e conosciuti nel mondo informatico come *crawl*; si tratta di software che scovano automaticamente i prodotti e i negozi on-line e li integrano nei cataloghi quotidianamente. Il *crawl* realizzato da *Twenga* è stato chiamato *Twengabot*, un software che identifica sul web i siti di commercio elettronico, acquisisce le informazioni di pubblico dominio, relative ai prodotti venduti sul sito e-commerce, e le invia al database presente su *Twenga*; per mantenere aggiornate le informazioni dei prodotti venduti sui vari siti e-commerce, il robot visita quotidianamente le pagine web interessate.

Figura 4.9– La pagina di Twenga.it che guida gli utenti nel mondo della telefonia



La figura precedente mostra la pagina che consente di accedere all'interno del portale Twenga, alla comparazione dei prodotti che fanno parte della Telefonia Mobile; vengono mostrate le differenti categorie che si possono visitare (cellulare, smartphone, ...), il numero di negozi online del settore inclusi nel database con il relativo numero di prodotti, e tanto altro. Molto interessante è l'informazione relativa al numero di prodotti che hanno subito una diminuzione di prezzo, vale a dire se il giorno prima lo stesso prodotto costava di più in quel negozio, questo viene segnalato al consumatore, in modo che se non l'avesse ancora acquistato, oggi potrebbe approfittarne.

Figura 4.10 – Cellulari selezionati in base alla popolarità nel sito *Twenga.it*  
 (Per accedere alla seguente pagina web: *Twenga.it* ( Home Page ) → *Telefonia* → *Cellulare*)

The screenshot displays the Twenga.it website interface for mobile phone searches. At the top, there is a search bar labeled 'Ricerca Twenga' with an 'OK' button. Below the search bar, navigation links include 'Home', 'Telefonia', and 'Cellulare'. The main heading is 'Cellulare'. A filter section titled 'Restringi i tuoi risultati' includes dropdown menus for 'Marca', 'Negozio', 'Prezzo', and 'Diminuzione prezzi'. Below this, another filter section includes 'Opzioni Multimedia', 'Peso', 'Tecnologia', 'Interfaccia wireless', and 'Design'. The search results are displayed as 'Risultati 1-30 su 11 853 presso 255 negozi'. A sorting menu at the top right of the results area offers options: 'Pertinenza', 'TwengaScore', 'Prezzo crescente', 'Prezzo decrescente', and 'Popolarità'. The product listings are arranged in a grid. Each listing features a 'Confronta !' checkbox, a product image, the model name, a price range, a star rating (represented by five asterisks), and a 'Confronta offerte' button. The products shown are: Motorola RAZR V3i (€ 90 - € 249), Nokia 6300 (€ 99 - € 217), Samsung SGH E250 (€ 59 - € 130), Nokia E65 (€ 149 - € 1.400), Nokia N73, Nokia 5200, Nokia N95, and Samsung SGH U700 Ultra. A yellow banner on the right side of the grid contains the text 'L'Espresso'.

Selezionando la categoria Cellulare si accede alla pagina web mostrata in Figura 4.10. I cellulari sono stati disposti in base alla popolarità, quindi in base alle scelte fatte dai consumatori/utenti. Twenga, infatti, considera nella selezione i commenti lasciati dagli utenti/consumatori che hanno acquistato il prodotto ed i commenti lasciati dagli utenti/esperti che valutano le performance delle caratteristiche del prodotto. La disposizione dei prodotti inoltre può essere effettuata in relazione alla pertinenza, oppure tenendo conto dei *Twenga Scores* ottenuti, oppure in base al prezzo dei prodotti ordinati in modo crescente o decrescente.

Come si arriva alla determinazione dei *Twenga Scores* non è chiaro, ma presumibilmente si tiene conto delle *performance* delle caratteristiche del prodotto (si veda Figura 4.11). La grande ampiezza del database di Twenga.it si vede leggendo il numero dei risultati ottenuti; 11.853 cellulari trovati in 255 negozi online. Così come tanti altri motori di ricerca, Twenga.it dà la possibilità al consumatore di restringere il campo di ricerca definendo alcune caratteristiche del prodotto; ciò è molto utile, ma risulta insufficiente per guidare l'utente verso la scelta del prodotto più adatto alle sue esigenze.

Figura 4.11 – Cellulari Selezionati in base al “Twenga-Score”

(Per accedere alla seguente pagina web: *Twenga.it* ( Home Page ) → *Telefonia* → *Cellulare* → *TwengaScore*”)

The screenshot shows the Twenga.it website interface. At the top, there's a navigation menu with categories like Marche, Foto, Audio - Video, **Telefonia**, Informatica, Elettrodomestici, Casa, Mobili, Forniture, Giardinaggio - Bricolage, and Moda. Below this is a search bar with the text "Ricerca Twenga" and an "OK" button. The main content area is titled "Cellulare" and displays a grid of mobile phone listings. Each listing includes a product image, the model name (e.g., Motorola MOTORAZR2 V8, Sony Ericsson W610i Walkman, Nokia E90 Communicator, Samsung SGH ZV40), a price range (e.g., € 181 - € 324), a star rating (\*\*\*\*\*), and a "Confronta" button with the number of offers (e.g., "Confronta 42 offerte"). The page also shows filters for "Restringi i tuoi risultati" and "Opzioni Multimedia".

#### 4.4 La reputazione online

Come abbiamo visto, la maggior parte dei motori di ricerca che si occupano di *price comparison* effettuano la comparazione dei prodotti, non solo in base al prezzo, ma soprattutto in base alle opinioni lasciate sul sito da utenti/consumatori e utenti/esperti; la stessa *Ebay.it* ha investito molto in questo campo. Nel 2007 infatti, il principale motore di aste online ha comprato per 75 milioni di dollari *StumbleUpon*, sito che fornisce la valutazione dei prodotti acquistabili in rete. *StumbleUpon* è una grande comunità online in cui gli utenti esprimono le proprie preferenze valutando prodotti, siti web, software e qualsiasi altro oggetto si possa trovare in rete. Ogni iscritto può pubblicare la propria recensione e condividerla con tutti gli altri visitatori; la base di utenti superava nel 2007 i 2 milioni di unità, per oltre 5 milioni di raccomandazioni al giorno, con un tasso di crescita del 150% all'anno; *StumbleUpon* esercita quindi una significativa influenza nella formazione delle opinioni degli utenti di Internet (*Pim.it*, Paolo Iasevoli, 31 maggio 2007).

La reputazione è una caratteristica dei soggetti che deriva dalle azioni del gruppo in cui sono inseriti, e in particolare dalle azioni di trasmissione di valutazioni sociali [Conte, Paolucci, 2002]; si tratta di un argomento che soprattutto negli ultimi tempi viene molto dibattuto da economisti, e psicologi. La reputazione viene utilizzata online già da 10 anni sia per attività commerciali che per usi di socializzazione e di ricerca di informazioni specifiche; attraverso la creazione di una reputazione, si cerca di superare il problema delle asimmetrie informative particolarmente rilevanti sul web [Akerlof, 1970]. Nel processo di valutazione sono almeno quattro gli agenti implicati: i *valutatori* (set E), che portano avanti l'attività di valutazione del

target in base ad esperienze dirette; i *target* (set T), oggetto della valutazione; i *beneficiari* (set B) dei comportamenti prescritti dalle norme sociali su cui si basa la valutazione; gli agenti *memetici*<sup>51</sup> (set M), o terze parti, che trasmettono valutazioni sociali. Una persona può avere ruoli diversi nella dinamica reputazionale; a volte si riscontra durante la fase di valutazione una tendenza ottimistica (migliore del reale), si ipotizza l'applicazione di una *regola della cortesia* nella trasmissione di reputazione, che si instaura nella situazione in cui  $set E = set T = set B$  (con  $E = M$  come ipotesi semplificatrice). Qui la benevolenza dei valutatori è verso i target; la regola della cortesia ha una forma del tipo: "passa buona reputazione sempre, la cattiva solo se certa". Nel caso di tendenza prudente, invece, si ipotizza l'applicazione di una *regola della prudenza*, derivante da una situazione in cui  $set E = set B$  and  $set E \neq set T$  (con  $E = M$  per ipotesi). Qui la benevolenza dei valutatori è verso i beneficiari e non verso il target. La regola della prudenza è del tipo: "passa cattiva reputazione sempre, la buona solo se certa". Un esempio, viene dal giornalismo, che secondo W. Lippmann, dovrebbe essere il "cane da guardia della democrazia"; dovrebbe cioè interpretare e valutare le azioni dei governanti (e dei potenti in genere) al fine di riportarle al pubblico in generale, di cui il giornalista dovrebbe adottare il punto di vista. La responsabilità influenza la circolazione di reputazione; più è grande la responsabilità percepita dal pettegolo (set M) riguardo gli effetti della propria azione memetica, minore sarà l'intensità della stessa; nell'interfaccia di un numero crescente di siti internet di varia natura si possono ritrovare diversi tipi di meccanismi reputazionali. L'applicazione maggiormente studiata è quella fornita dalla casa d'aste e-Bay, dove lo scopo dell'introduzione di un meccanismo reputazionale a livello utente è creare fiducia, garantire un alto livello di qualità negli scambi, superare il problema della forte asimmetria informativa intrinseca al mezzo.

La reputazione online è particolarmente rilevante nelle aste elettroniche, nei *social networks*, nei fora di discussione, nei sistemi recensori. In qualsiasi ambiente, la reputazione serve a creare fiducia nel venditore (o nel professionista) al quale si paga il bene (o a cui si assegna l'esecuzione del progetto); nei *social networks*, ad esempio, (sistemi deputati al trasferimento in rete della propria rete di connessioni sociali, allo scopo di condivisione del quotidiano), le reputazioni sono dedicate all'allargamento della propria rete sociale, con la ricerca di nuovi partner d'affari, o di attività ricreative, o per nuove relazioni affettive. In questo contesto quindi, la reputazione di un individuo è riconducibile al suo grado di connessione con gli altri utenti, oltre che a quel che gli altri dicono di lui sul sito. Ciò accade anche nei fora di discussione, luoghi dove si possono porre domande su argomenti specifici, sui quali si ricevono di solito diverse risposte da altri utenti più o meno esperti. La reputazione serve pertanto a discernere rapidamente le informazioni rilevanti; attraverso i sistemi recensori gli utenti possono esprimere un loro giudizio su prodotti o servizi, oppure su esperti o professionisti.

Esistono però, diversi studi che sollevano forti dubbi sul funzionamento effettivo del sistema reputazionale di eBay (il cosiddetto "feedback forum"). Resnick e altri affermano, in base ai dati raccolti, che su tale sistema circa la metà degli utenti rilasciano valutazioni e che le valutazioni fatte, sono nel 99 per cento dei casi di segno positivo. Questo fatto, incrociato coi dati sulle denunce di frodi perpetrate su Internet, e in particolare nelle aste elettroniche, fa riflettere sulla effettiva utilità dei meccanismi reputazionali attualmente in uso [Resnick et al., 2002].

Questo avviene perché la logica del meccanismo reputazionale tende a inibire le voci critiche e a creare un ambiente cortese, dove non si danno valutazioni negative per paura di riceverne altrettante negative come ritorsione. Ciò tuttavia riduce di molto l'utilità di queste informazioni e l'ottimismo reputazionale crea ambienti fortemente positivi, utili come *advertising*, ma a rischio di scorrettezze per l'ingenuità o la fiducia eccessiva generate [Marmo, 2006].

---

<sup>51</sup> Per meme si intende un elemento in grado di determinare i comportamenti degli individui di una collettività, fondato sull'imitazione dei modelli e consuetudini impostisi nella specie di appartenenza.

Quindi, fino a poco tempo fa, su ebay anche il venditore doveva lasciare un giudizio positivo o negativo sulla controparte (acquirente), e quest'ultima era abbastanza condizionata a evitare di lasciare un giudizio negativo. Ora invece, come soluzione al problema, solo l'acquirente viene invitato a lasciare un commento, e il venditore non può più opporsi. Ma in questo modo, il venditore rischia di non essere pagato, a grande vantaggio degli acquirenti che possono fare il bello ed il cattivo tempo, anche quando il venditore è onesto.

Inoltre è possibile "drogare" i feedback su ebay, per esempio acquistando molti beni di poco prezzo in grandi quantità, oppure vendendosi fra amici beni a basso valore. In questo modo il punteggio sale velocemente in pochi giorni; Nel 2006 la polizia postale ha denunciato 4 cittadini italiani di 40 anni residenti a Torino per aver truffato oltre 150 persone in tutta Italia attraverso Internet per un valore complessivo di 70 mila euro in sei mesi; Durante il periodo pre-natalizio i malviventi avevano messo in vendita materiale informatico ed elettronico a prezzi molto competitivi; una volta ottenuti i pagamenti si sono resi irreperibili senza spedire il materiale ordinato. I soggetti si erano costruiti una fama di venditori on-line validi e onesti, in questo modo è stato più facile truffare i clienti di e-bay che si affidavano a loro sulla base dei giudizi positivi emessi dai precedenti acquirenti (La stampa.it , 2007).

Altri difetti dei sistemi recensori adottati nelle valutazioni online sembrano essere i seguenti:

- Spesso sono pochi, i soggetti che quotidianamente danno la loro opinione per le diverse tipologie di prodotti presentate nel sito. Come in tutte le indagini, quando il campione utilizzato è troppo piccolo e non corrisponde all'identità generale della popolazione, i risultati non possono essere abbastanza attendibili. Per esempio, su doyou.it sono sempre le stesse persone che fanno le valutazioni per più prodotti, per avere un campione di opinioni ampio su un particolare prodotto, occorre aspettare molto tempo; in questo modo però una maggiore affidabilità nella scelta, potrebbe comportare l'acquisto di un prodotto oramai antiquato.
- In genere, esprimono la propria opinione sui siti e-commerce che presentano sistemi recensori, persone o molto insoddisfatte o molto soddisfatte del prodotto acquistato. Non c'è quindi la sicurezza che un prodotto con un basso indice di reputazione sia effettivamente da scartare.
- Spesso in alcuni siti si creano gruppi di utenti "esperti" che commentano i prodotti entrando molto nel dettaglio e criticando magari attributi che raramente risultano utili per un comune consumatore: è possibile quindi che le preferenze del consumatore meno esperto vengano modificate verso richieste che non corrispondono più ai suoi bisogni.
- Difficilmente i siti di opinione presentano le review di tutti i prodotti presenti sul mercato, quindi un consumatore difficilmente terrà conto nella sua scelta, di alternative di prodotto che non presentano valutazioni.
- Di solito è molto facile che il consumatore venga condizionato da commenti lasciati da *reviewer* (revisori) ingannatori. Il sito *Amazon.com*, ad esempio, è uno dei principali siti e-commerce con sistemi recensori; ogni revisore che inserisce le sue opinioni sul sito, viene al tempo stesso valutato dagli utenti che leggono il suo commento. Il *rank* è il punteggio ottenuto da ogni *reviewer*; non esiste un modello matematico per stabilirlo, e nemmeno viene pubblicato online il sistema utilizzato per definirlo. Approssimativamente un revisore ottiene un punto quando tre utenti si dimostrano soddisfatti del suo commento, ne perde uno quando un solo utente non è soddisfatto. Infatti sotto ogni commento lasciato sul sito, vi sono due *botton click* (*soddisfatto* o *non soddisfatto*), quindi anche se ci fossero le migliori intenzioni da parte dei revisori, non esiste comunque alcuna affidabilità scientifica nel calcolo dei *ranks*.

#### 4.4.1 L'influenza dei blog nelle decisioni di acquisto dei consumatori

La Hotwire-Ipsos Mori, una società di ricerche di mercato e pubbliche relazioni che opera nel mercato dell'*information technology*, ha recentemente realizzato una interessante ricerca di mercato sul comportamento degli utenti europei nell'uso dei *blog*, sulla loro confidenza con questi spazi online, le loro scelte<sup>52</sup>. Un monito alle imprese che non hanno ancora compreso il nuovo paradigma del *blog* e la rivoluzione strisciante provocata dalla blogosfera e dai nuovi strumenti WEB 2.0 con le loro caratteristiche di collaborazione, partecipazione e condivisione di informazioni, di contenuti ( *www.youtube.com*, *www.myspace.com*, *ecc.* ). Un monito anche per quelle imprese che si stanno lasciando prendere dalla frenesia del *blogging*, ma continuano a snobbarlo come uno spazio minore e comunque secondario rispetto alle pratiche di marketing, di comunicazione, e commerciali consuete.

La ricerca, è stata condotta in Gran Bretagna, Francia, Germania, Spagna e Italia. Dai risultati della ricerca sembra che la blogosfera sia diventata un luogo di influenza sempre più importante sui comportamenti degli utenti dei blog consumatori. Venticinque milioni di europei hanno infatti affermato di aver modificato il loro modo di pensare e la loro percezione di un'azienda e/o dei suoi marchi e prodotti, dopo aver letto i commenti e le informazioni fornite da altri utenti su un blog. Il 34% degli intervistati hanno affermato di avere deciso di non acquistare un prodotto dopo avere letto su un blog commenti da parte di altri consumatori. Ancora più interessante è il fatto che i blog vengano ritenuti la seconda fonte di informazione percepita come più fidata (24%) dopo i giornali (30%) e prima della pubblicità in televisione (17%) e l'e-mail marketing (14%). Secondo la ricerca fra le persone che spendono più di 145 euro al mese online, sale al 30% la percentuale che da maggiore fiducia ai blog. In Europa sei utenti intervistati su dieci conoscono i blog e un utente su sei ha letto almeno un blog. La nazione con il maggior numero di blog è la Francia con una percentuale del 90% di intervistati che conoscono e praticano la blogosfera. Invece l'Inghilterra è risultata essere la nazione meno propensa all'uso dei blog con una percentuale del 50%. In Germania conosce i blog il 55% del campione intervistati, mentre in Italia e Spagna il 51%. Sempre secondo la ricerca, i blog influenzano sempre di più i comportamenti e i processi decisionali all'acquisto dei consumatori; il 52% degli europei intervistati affermano di acquistare preferibilmente prodotti di cui hanno già sentito parlare bene dalle persone di loro conoscenza o in rete. I blog sono anche capaci di bloccare la decisione all'acquisto; 40 milioni di europei lo fanno dopo aver letto commenti negativi su un prodotto o una marca. Il potere delle opinioni delle persone nelle decisioni di acquisto è da sempre conosciuta da chi si occupa di Marketing. Risulta però molto interessante che questo fenomeno si manifesti oggi anche in rete e con le stesse caratteristiche di influenza sul comportamento dei consumatori. Questo perché gli effetti e le influenze di questo nuovo metodo di passa parola cadranno su milioni di persone, quindi non avviene più localmente ma globalmente, non riguarderà più un mercato e un'offerta limitata ma globale. Se un prodotto venduto in gran parte del mondo delude un consumatore americano, renderà dubbioso nell'acquisto un consumatore tedesco. Ma non tutti i blog sono uguali e molte aziende potrebbero buttare il loro denaro dalla finestra ricorrendo al blog come nuovo strumento di marketing senza essersi attentamente preparate per farlo.

Che le aziende si preoccupino molto degli sviluppi di questi nuovi metodi di comunicazione si carpisce dalle continue soluzioni che si realizzano per farlo diventare un vantaggio nei confronti dei concorrenti; la società *Complexlab* ha realizzato un'applicazione software, denominata *Coolhunting*, utile per indagare e scoprire nuove tendenze (*trends*) in rete

---

<sup>52</sup> La ricerca condotta da Ipsos Mori Europa ha coinvolto un campione rappresentativo di 2.214 adulti di età superiore ai 15 anni in Gran Bretagna (526), Francia (440), Germania (485), Italia (378) e Spagna ( 385). Il campione era composto da persone adulte che usano Internet da casa e/o dal lavoro. Le interviste sono state compiute con incontri faccia a faccia durante il mese di settembre del 2007 usando i servizi di Ipsos Capibus.

e chi sono i loro originatori (*trendsetters*). La rete è ricca di dati e di informazioni che possono essere utilizzati per scoprire dove si stanno manifestando in rete fenomeni di innovazione e quali siano gli innovatori protagonisti. Queste informazioni possono essere utili per tutte quelle realtà imprenditoriali che hanno messo il cliente al centro delle loro strategie e che vogliono costruire nel tempo relazioni durature con una clientela che cambia costantemente e che è sempre meno fedele perché immersa in un mercato sempre più competitivo, ricco e globale.

La soluzione di *Complexlab* permette di convertire le nuove idee, le nuove tendenze scoperte in rete e le conoscenze ad esse associate, in prodotti e servizi più corrispondenti ai loro bisogni e alle loro aspirazioni. Le nuove tendenze possono essere seguite nella loro nascita, nel loro crescere, spostarsi e moltiplicarsi nei vari luoghi e spazi di incontro che caratterizzano la rete. I dati così individuati e raccolti vengono poi utilizzati per visualizzare una mappa grafica di questi luoghi e delle relazioni che li caratterizzano, una mappa in grado di evidenziare ciò che il mercato vorrebbe essere oggi, rispetto a quello che invece è. *Coolhunting* quindi, utilizza tecnologie e soluzioni software che permettono di catturare nuove informazioni là dove esse vengono create, nei luoghi della rete dove ha luogo l'innovazione e si manifesta la creatività delle persone nelle loro vesti di utenti di servizi e consumatori di marchi e di prodotti.

#### **4.4.2 Le attività di correzione della reputazione online**

Siti, forum, blog, motori di ricerca, social network, video online, etc., possono risultare utili, per farsi una idea su un prodotto, servizio, venditore o professionista, ma è necessario tener presente la possibilità di trovarsi davanti a dei tranelli. Stanno infatti nascendo online, sempre più siti che si offrono come “guardiani della reputazione” (*reputation guardian*), e offrono i propri servizi a tutti coloro che sono interessati ad avere una corretta reputazione online, o anche a nascondersi dai motori di ricerca.

Questi siti hanno ideato un'applicazione chiamata *reputation manager*, un software avanzato di ricerca automatica delle informazioni, che, grazie all'interazione fra *keywords* e analisi semantica, si occupa di cercare le citazioni in tutto il web, canali informali compresi, riguardanti il brand, le figure aziendali, i prodotti, i servizi, le situazioni o i contesti; *reputation manager* riesce a sondare tutti i canali del web e a valutare gli impatti che questi hanno nell'immagine aziendale. Anche in Italia esistono alcuni siti di *reputation guardian*, il più famoso è [Reputazioneonline.it](http://Reputazioneonline.it); [Reputazioneonline.it](http://Reputazioneonline.it) si rivolge a enti pubblici e privati, organizzazioni profit e non-profit, aziende farmaceutiche, istituzioni, operatori turistici, agenzie di moda, operatori del settore alberghiero e della ristorazione, politici e altri personaggi pubblici, insomma tutti quelli che sono consapevoli che ogni giorno online si parla di loro, della loro immagine, delle loro azioni, della qualità dei servizi offerti, dei loro prodotti.

Come affermano i realizzatori di [Reputazioneonline.it](http://Reputazioneonline.it), l'intercettazione, l'analisi e l'interpretazione dei giudizi, pregiudizi e opinioni presenti sulla rete costituisce un'attività indispensabile per la tutela della visibilità, della reputazione su internet e della credibilità di ogni azienda, del suo brand, dei suoi servizi, del suo management; è necessario quindi che il monitoraggio sia tempestivo per essere realmente efficace e per consentire di mettere in campo tutti gli strumenti capaci di fronteggiare una crisi e di dominarla prima che le ondate di opinione si diffondano senza controllo nel web.

Tutto ciò può tuttavia suscitare preoccupazioni e perplessità; inserendo le parole *reputation manager* su Google appaiono esclusivamente links dove si parla soltanto bene del servizio offerto; queste applicazioni infatti, non eliminano le informazioni negative, ma le nascondono. Le strategie per “correggere” la reputazione possono essere tante; secondo quanto scritto sul sito [Reputazioneonline.it](http://Reputazioneonline.it), rilevata una pagina web che necessita di correzione, viene definita di volta in volta una strategia che consenta di limitare gli effetti negativi di quanto rilevato. Prendiamo il caso ad esempio di quello che offre [ReputationDefender.com](http://ReputationDefender.com), che per soli 30 dollari al mese farà in modo che nella ricerca di informazioni riguardo un'azienda su un motore di ricerca, i risultati in cima saranno positivi e quelli negativi in fondo (*Ansa.it*,

24/05/2008); è probabile che molti personaggi pubblici, già lo adoperino per nascondere notizie negative. Purtroppo, man mano che aumenta il numero di utenti che utilizzano questo servizio, sarà sempre più difficile la ricerca di informazioni obiettive; un'azienda potrebbe anche adoperare questi nuovi *crawls* per screditare i concorrenti, facendo in modo che le notizie negative riguardanti i concorrenti si posizionino ai primi posti dei risultati dei motori di ricerca. I rischi di questo meccanismo reputazionale sono molti, e sarà più difficile per gli utenti avere "ancore corrette" capaci di indirizzarli verso le scelte giuste.

#### **4.4.3 Monitorare le preferenze dei consumatori**

Piuttosto che basarsi sulle opinioni lasciate online, che come abbiamo visto precedentemente, non sono del tutto affidabili, la *startup* californiana *Baynote*, ha messo a punto un software in grado di monitorare le preferenze dei consumatori online basato sulle azioni invece che sulle opinioni; è del tutto evidente l'importanza strategica di comprendere preferenze ed esigenze dei potenziali acquirenti online; la soluzione giusta, secondo *Baynote*, è di registrare direttamente le azioni degli utenti, attraverso i loro *click*. Sembrerebbe un'idea banale, dal momento che il monitoraggio dei click è una pratica in uso fin dagli albori di Internet; interessante e innovativo è il modo in cui questa tecnica viene adoperata. *Baynote* ha infatti sviluppato un software che, una volta integrato in un sito di e-commerce, è in grado di individuare i prodotti più desiderati, dando così l'opportunità all'azienda di puntare maggiormente su di essi. Il meccanismo è semplice: più utenti cliccano su un determinato prodotto, più questo guadagna visibilità; ciò significa che se un consumatore ha dimostrato particolare interesse per un nuovo *mouse* ottico, alla sua prossima visita quel *mouse* sarà automaticamente posizionato nell'area più visibile della pagina, pronto per essere acquistato. Si tratta dunque di un sistema in grado di creare soluzioni personalizzate per ogni singolo utente, che rappresenta un'alternativa (o, meglio ancora, un complemento) alla modalità precedentemente più utilizzata: il monitoraggio delle opinioni dei consumatori. Spesso infatti molte strategie di marketing si formano esclusivamente basandosi sui commenti lasciati da altri, e ciò potrebbe risultare dannoso per l'azienda; a giudicare la qualità di un prodotto sono soltanto una piccolissima parte degli effettivi acquirenti, e il giudizio può essere sporcato da numerosi fattori estranei, per esempio una personale disposizione negativa nei confronti dell'azienda. Anche se le opinioni danno comunque un'idea di come la pensano in generale gli utenti, le parole possono condurre fuori strada, mentre le azioni concrete compiute dagli utenti sarebbero una "scienza esatta". E non c'è motivo di non crederlo, visto che il software di *Baynote* è stato già scelto da giganti del commercio elettronico come eBay, Glam.com e Motorola (Pmi.it, Paolo Iasevoli, 2 agosto 2007).

#### **4.5 Indagini sugli acquisti online nei principali paesi europei**

La *European Interactive Advertising Association* (EIAA) è un'associazione che si occupa di esaminare i trend e modelli di spesa degli acquirenti online europei; nel mese di settembre 2007 essa ha condotto un'indagine effettuata attraverso più di 7000 interviste telefoniche in Gran Bretagna, Germania, Francia, Spagna, Italia, Paesi Baltici, Belgio e Olanda. L'indagine ha evidenziato che, dopo una ricerca sul Web, il 40% degli acquirenti online europei cambia idea sui marchi da acquistare. Gli inglesi sono più influenzati dalle informazioni disponibili su Internet (49%), mentre gli italiani sono tra quelli che più difficilmente cambiano idea (27%). Oltre la metà (59%) degli acquirenti online considera i siti Web di marchi noti, una fonte importante di informazioni per la ricerca o la valutazione di un prodotto o servizio, mentre i motori di ricerca sono considerati più utili (76%) delle segnalazioni personali (72%). Il 61% considera i siti Web con i confronti di prezzi un'utile fonte di informazioni, mentre il 57% legge le recensioni dei clienti dei siti Web per agevolare la scelta. Secondo i risultati dell'indagine, dopo una ricerca online i consumatori europei tendono a cambiare idea soprattutto quando si

tratta di acquistare prodotti elettronici (il 59% ha cambiato idea sul marchio), mentre per le prenotazioni di pacchetti vacanze e viaggi ha cambiato idea il 43%.

L'80% degli utenti europei di Internet ha acquistato online un prodotto o servizio, con un aumento del 3% rispetto al 2006 e del 100% rispetto al 2004. In effetti, in appena sei mesi, questi acquirenti online europei hanno effettuato 1,3 miliardi di acquisti, spendendo mediamente € 747 a testa. Nel complesso, gli acquisti online più diffusi sono rappresentati dai biglietti di viaggio (54% degli acquisti complessivi), seguiti da pacchetti vacanze (42%), libri (40%), biglietti per concerti/teatri/festival (38%), abbigliamento (33%) e prodotti elettronici (33%). Dalle indagini risulta che le donne sono particolarmente orientate all'acquisto di abbigliamento (40% contro il 27% degli uomini) e pacchetti vacanze (46% contro il 38% degli uomini), mentre i maschi sono più orientati all'acquisto di dispositivi high-tech (39% contro il 27% delle donne) e CD (25% contro il 20% delle donne).

Tabella 4.1 - Utenti europei che cambiano idea sul marchio da acquistare dopo una ricerca su Internet in una indagine del 2007

Regno Unito	49%
Danimarca	48%
Norvegia	44%
Paesi Bassi	44%
Belgio	42%
Svezia	42%
Media europea	40%
Francia	36%
Germania	37%
Spagna	34%
Italia	27%

Fonte: *European Interactive Advertising Association*

Tabella 4.2 – Principali beni acquistati online nel 2007 nei principali paesi europei (acquisti online in percentuale degli acquisti complessivi)

Biglietti di viaggio	54%
Pacchetti vacanze	42%
Libri	40%
Biglietti per concerti/teatri/festival	38%
Abbigliamento	33%
Prodotti elettronici	33%
CD	23%
Biglietti cinematografici	23%
DVD	21%
Download di musica	19%

Fonte: *European Interactive Advertising Association*

#### 4.5.1 Indagini sugli acquisti online italiani

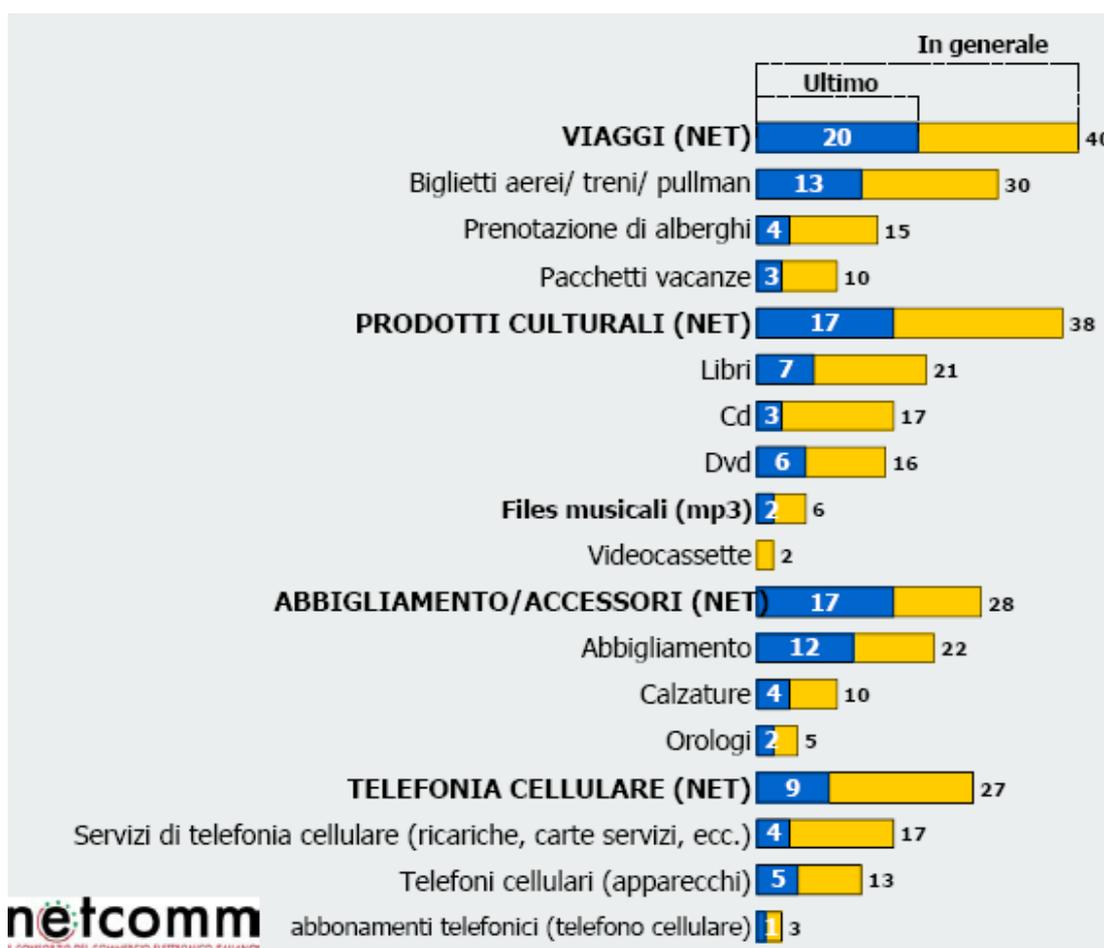
Ogni anno il Consorzio del Commercio Elettronico Italiano NETCOMM effettua delle indagini sull'evoluzione del commercio elettronico in Italia insieme all'Università Bocconi e a GfK Eurisko. Per l'annuale appuntamento NETCOMM eCOMMERCE FORUM del 14 e 15 maggio 2008 tenutosi a Milano sono stati presentati i risultati di una interessante indagine alla quale hanno partecipato 400 soggetti che hanno effettuato almeno un acquisto online. Dai primi risultati si evince che ben oltre il 90% degli intervistati intende ripetere l'acquisto online e quasi il 100% di questi, esprimono un giudizio da buono a ottimo sull'esperienza; si tratta di un chiaro segnale dell'aumento di fiducia tra quanti hanno sperimentato l'e-commerce; rimane il problema di portare verso gli acquisti online coloro che non vi sono mai ricorsi. Lo stesso presidente di NETCOMM Roberto Liscia che ha avviato i lavori del Forum, ha parlato del cosiddetto *info-commerce*, un fenomeno sociologico che riguarda un numero molto elevato di persone che usano direttamente o indirettamente Internet per acquisire informazioni prima di un acquisto online o sui canali tradizionali. Sarebbe opportuno riconoscere che internet è entrato ormai nella vita quotidiana degli italiani, non soltanto per le crescenti opportunità di fare acquisti a condizioni migliori rispetto agli altri canali, ma anche per le enormi possibilità informative che mette a disposizione.

Secondo la ricerca di NETCOMM, a fronte di circa 18 milioni di utenti Internet che hanno navigato negli ultimi tre mesi in Italia, coloro che almeno una volta hanno fatto un acquisto online sono poco più di 5 milioni. Da non sottovalutare invece il fatto che i clienti abbandonino nel 38,8% dei casi dopo aver ricercato informazioni (forte impatto dell'*info-commerce*), nel 16,8% dei casi nel corso della registrazione, nel 24% dei casi durante l'acquisto e nel 16,4% durante il pagamento.

L'acquirente online è molto oculato e attento; egli si informa e confronta tra diverse offerte prima di comprare; il 47% degli intervistati girano il web per informarsi prima sul prodotto, poi sul modello e successivamente sulla marca. La decisione di acquisto è autonoma nel 76% dei casi, proviene dal passaparola di un familiare nel 23%, per il 16,9% dal web e solo per il 3% dalla pubblicità tradizionale.

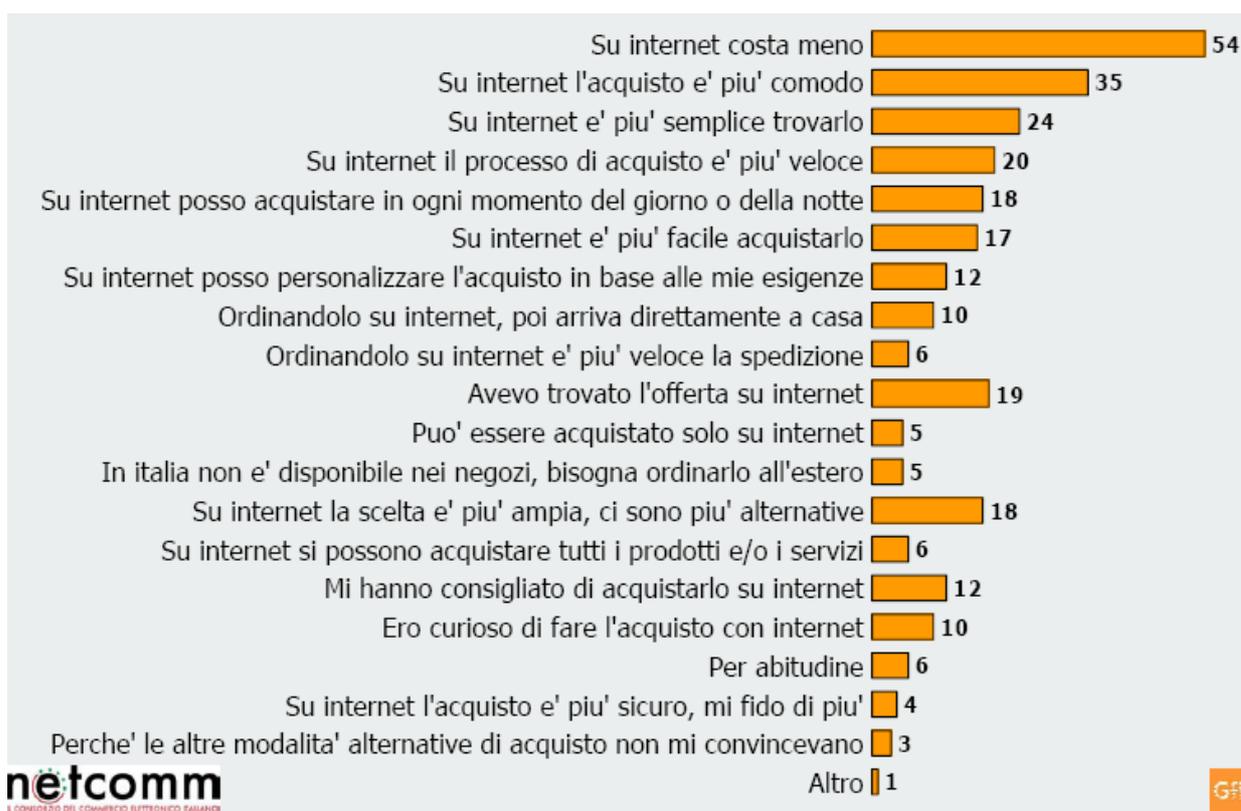
La figura seguente mostra i prodotti che più si acquistano online: in Italia nel 2007 sono stati acquistati on-line il 30 per cento dei biglietti di treni, aerei e pullman, il 22% dei prodotti per abbigliamento, il 21% dei libri, il 17% delle ricariche telefoniche, il 15% delle prenotazioni di alberghi, il 13% dei telefoni cellulari.

Figura 4.12 – Prodotti o servizi maggiormente acquistati online in Italia nel 2007



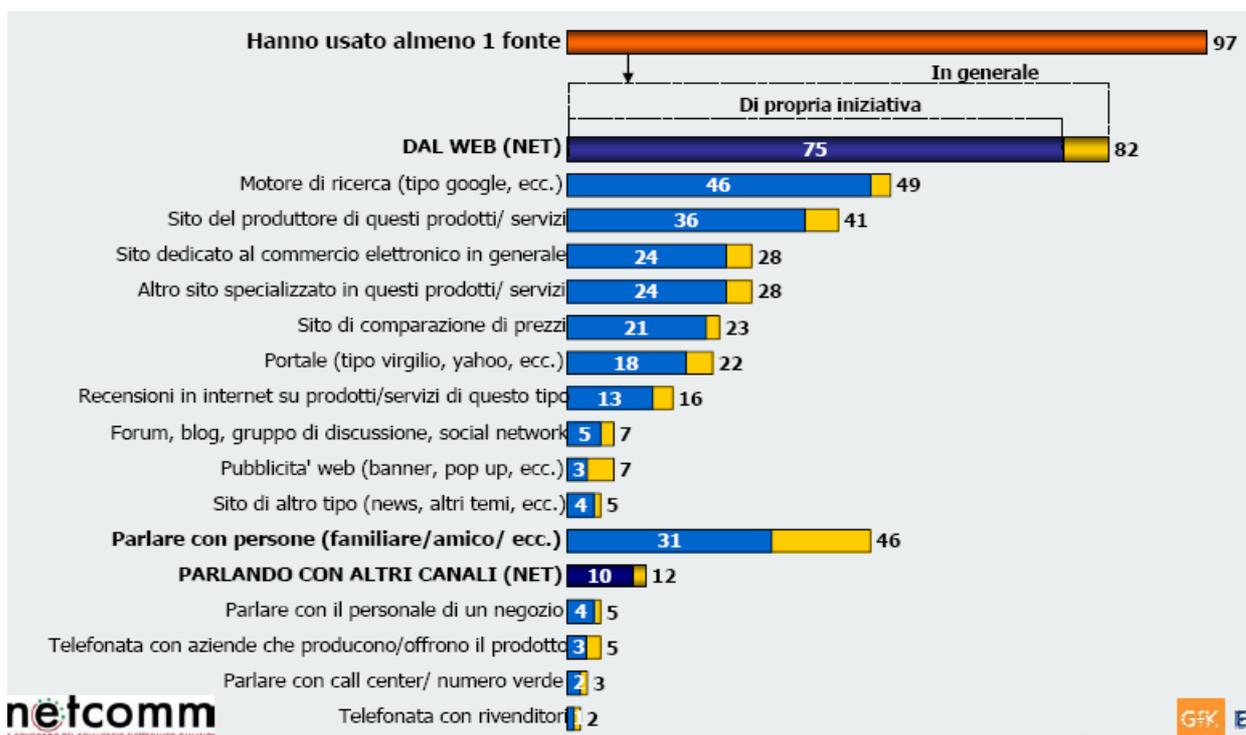
La figura seguente evidenzia le ragioni che spingono principalmente ad acquistare online in Italia. Secondo il 54% degli intervistati da Netcomm, su internet si riesce a risparmiare di più, è più comodo l'acquisto online per il 35%, è più facile la ricerca del prodotto o servizio da acquistare per il 20%, e così via.

Figura 4.13 – Le ragioni che spingono ad acquistare online in Italia



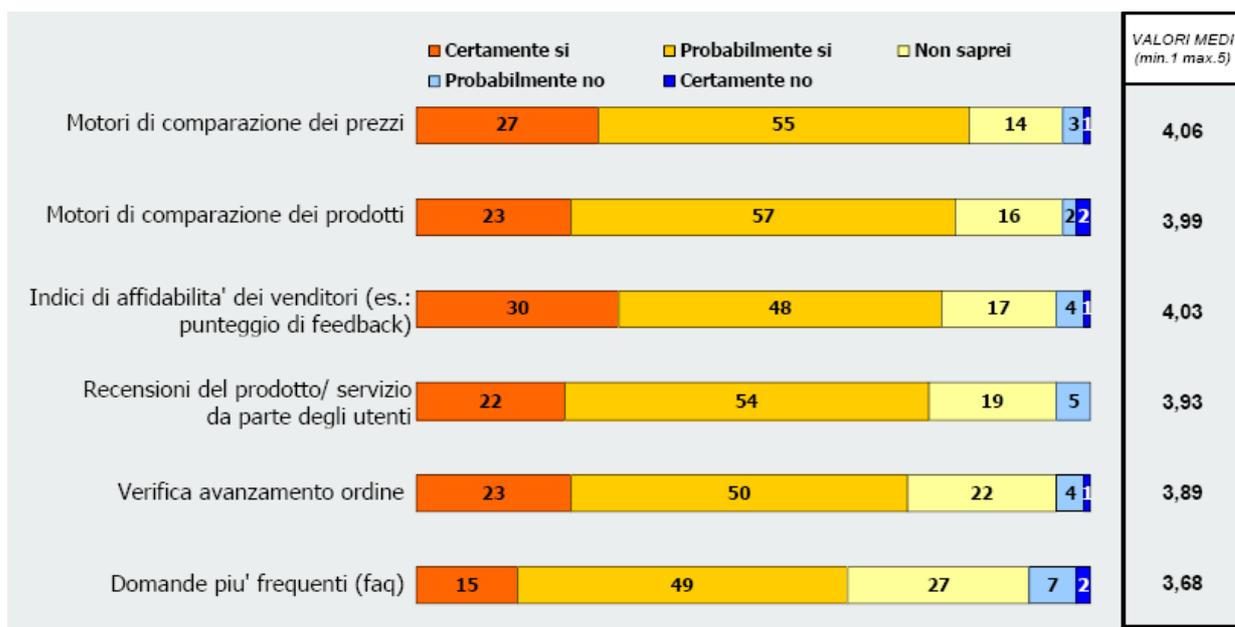
La figura successiva evidenzia le fonti di informazioni maggiormente utilizzate per il processo decisionale: il 49% degli intervistati acquisisce informazioni tramite motori di ricerca, il 41% dal sito del produttore, il 28% da un sito di commercio elettronico, il 28% da altri siti specializzati, il 23% da un sito comparatore, il 22% da un portale, il 16% da recensioni su internet, ecc.); in complesso l'82 per cento degli intervistati ottiene informazioni dal WEB, e il 46 per cento da familiari, amici, ecc. (passaparola) (Il totale delle percentuali è maggiore di 100 poiché molti utenti utilizzano più fonti di informazione).

Figura 4.14 – Le fonti di informazioni maggiormente utilizzate per i processi decisionali



Molto significativa è la propensione ad usufruire degli strumenti e servizi di supporto offerti dalla Rete (Figura 4.15); gli intervistati hanno dichiarato di accogliere positivamente i diversi strumenti disponibili online per aiuti durante il processo decisionale. Molto apprezzati sono i motori di comparazione dei prodotti e dei prezzi. Ma anche gli indici di affidabilità dei venditori online e le recensioni dei prodotti/servizi da parte degli utenti.

Figura 4.15 – Propensione ad usufruire di diversi strumenti e servizi di supporto online



#### 4.6 Sviluppi nelle tecnologie dell'informazione

Secondo il Prof. Wolfgang Wahlster, direttore del Centro di ricerca tedesco per l'intelligenza artificiale (DFKI)<sup>53</sup>, intervenuto il 19 ottobre 2007 in occasione del simposio di ITEA 2<sup>54</sup> a Berlino, le tecnologie dell'informazione e delle telecomunicazioni (TIC) sono il principale motore delle attività innovative e il software rappresenta il carburante di questo motore<sup>55</sup>; la diffusione di queste tecnologie può contribuire ad aumentare la sicurezza, a migliorare la salute, ad offrire migliori opportunità di apprendimento e di divertimento. Gradualmente si arriverà a un "accesso ambientale alla rete", vale a dire a una connessione sempre presente ed efficace, reti ad hoc, reti locali individuali, reti mobili sull'aereo e sul treno, reti domestiche e così via. Il professor Wahlster immagina la casa del 2012 come un luogo completamente connesso mediante diverse schede SIM, che consentiranno di organizzare l'intrattenimento, monitorare la salute, collegare in rete i diversi elettrodomestici, controllare gli accessi, riciclare i rifiuti, mantenere le memorie digitali dei prodotti; dei robot intelligenti si occuperanno della pulizia degli ambienti.

La connessione sempre e ovunque non è l'unica evoluzione dei prossimi anni: anche internet cambierà; secondo il professor Wahlster, si passerà gradualmente a servizi internet integrati, alla comunicazione da macchina a macchina e infine al web semantico (contenuti web comprensibili ai computer)<sup>56</sup>; gradualmente si arriverà a una combinazione tra l'internet delle cose, una rete tra oggetti autoconfigurante senza fili, l'internet integrato e l'internet dei servizi; mettendo insieme questi elementi si passerà dal web 2.0 al web 3.0. Secondo il professor Wahlster, le grandi sfide tecnologiche della società dell'informazione sono l'auto sicura, il dispositivo multilingue sempre a portata di mano, il robot di servizio, il computer che si automonitora e autoripara, l'agente di polizia su internet, il simulatore dei farmaci e della malattia a base cellulare, il potenziamento dei ricordi personali, la giacca per la comunicazione pervasiva, il visualizzatore utilizzabile ovunque, il negozio intelligente. Tutte queste sfide hanno una dimensione sociale: la salute e la società che invecchia, la connettività sempre e ovunque, nuovi valori e servizi, sicurezza, privacy e fiducia; la chiave di tutto ciò è garantire l'interoperabilità fra i diversi sistemi.

I sistemi di comunicazione di quarta generazione dovrebbero essere in grado di fornire una soluzione IP complessiva, in cui voce, dati e flussi multimediali possano essere disponibili per gli utenti ovunque e in qualsiasi momento. L'obiettivo principale è avere un sistema IP pienamente integrato, mediante una convergenza delle tecnologie con fili e senza fili in grado di

---

<sup>53</sup> Il centro di ricerca tedesco per l'intelligenza artificiale, costituito nel 1988, è uno dei maggiori istituti di ricerca del mondo nel campo delle tecnologie del software innovativo basate sulle metodologie dell'intelligenza artificiale. Tra gli azionisti figurano anche grandi imprese come Microsoft, SAP, BMW, DaimlerChrysler. Un altro importantissimo centro di ricerca internazionale sull'intelligenza artificiale è l'*MIT Computer Science and Artificial Intelligence Laboratory* (CSAIL), costituito nel 2003 attraverso la fusione fra il laboratorio di Intelligenza Artificiale e il laboratorio di Informatica. Nell'ambito del CSAIL svolge attività di ricerca Tim Berners-Lee, inventore nel 1989 del World Wide Web, e Direttore della sua fondazione nel 1994 del World Wide Web Consortium (W3C). Presso l'MIT opera anche, in ambiti connessi con l'intelligenza artificiale, il *Media Lab*, concepito nel 1980 da Negroponte, e in particolare il *Media Lab's Ambient Intelligence research group*, fondato e diretto da Pattie Maes.

<sup>54</sup> ITEA 2 (*Information technology for European advancement*) è un programma strategico europeo per attività avanzate di ricerca e sviluppo precompetitiva nel software per sistemi e servizi *software-intensive*. Il programma mira a mobilitare in 8 anni 20.000 annualità di lavoro in attività di ricerca e sviluppo, anche al fine di contribuire a ridurre il gap rispetto all'obiettivo di investimenti in ricerca e sviluppo pari al 3 per cento del prodotto interno lordo concordato nel 2000 al vertice europeo di Lisbona. Al programma partecipano università, centri di ricerca e imprese pubbliche e private. Il progetto ITEA 2 è la prosecuzione del progetto ITEA lanciato nel 1999 (e concluso nel 2007), su iniziativa delle imprese e con il sostegno dei governi europei. Il suo obiettivo è stimolare la competitività delle imprese europee, contribuendo a superare il gap rispetto agli Stati Uniti nei sistemi *software-intensive*.

<sup>55</sup> CORDIS.EUROPA.eu, 25/10/2007

<sup>56</sup> Sulle opportunità offerte dal "web semantico" si veda, fra gli altri, Berners-Lee e altri (2001, 2006), e Feigenbaum e altri (2007).

fornire delle velocità di trasmissione delle informazioni dell'ordine di 100 Mbit in movimento e da un Gbit da fermi, sia all'interno che all'esterno delle costruzioni, con elevata qualità e sicurezza.

#### 4.7 Conclusioni

Quando un utente abbandona la sessione di acquisto di un sito e-commerce senza comprare, i motivi sono tanti. La rinuncia può dipendere dall'umore dell'utente, dall'aspetto grafico e comunicativo del sito, dal prezzo troppo alto del prodotto, dai tempi troppo lunghi e/o procedimenti complessi per acquistare un prodotto, dai processi di pagamento non molto trasparenti, oppure dalle spese di spedizione troppo elevate. Qualcosa si può fare, ad esempio, applicare dei prezzi in linea con il mercato, visualizzare e specificare dall'inizio i prezzi per le spedizioni, predisporre la possibilità di conservare il proprio carrello per un periodo di almeno 15-30 giorni, effettuare delle promozioni sui prodotti, oppure la consegna gratuita, avere un catalogo dei prodotti ben fornito e aggiornato, limitare il numero di passaggi per confermare l'ordine ed importantissima è l'assistenza in tempo reale, eventuale numero verde, *chat* con responsabile vendite, per chiarire gli ultimi dubbi in tempo per completare il carrello.

La nuova frontiera del commercio elettronico è l'evoluzione dei siti di comparazione delle offerte dei prodotti venduti online, che permettono di facilitare il processo decisionale del consumatore. Molti di questi siti, al momento, utilizzano semplici matrici di comparazione attributi x opzioni di prodotto, oppure sistemi di collaborative filtering. I sistemi che consentono scelte più razionali, perché più vicine alle preferenze del consumatore/utente, sono i sistemi di raccomandazione content based filtering, ma sono ancora pochi i siti price comparison che li adoperano, e soprattutto sono made in USA, mentre in Italia sono ancora quasi del tutto assenti.

I sistemi collaborative filtering utilizzati da molti siti che si occupano di vendite sul web, rientrano nel complesso sistema delle reputazioni online. Difatti basta fare una veloce ricerca su un qualsiasi motore di ricerca, per farsi una idea sulla reputazione acquisita da un prodotto o servizio, venditore o produttore oppure da un professionista. Le valutazioni online vengono fatte nelle aste elettroniche, nei *social networks*, nei fora di discussione, nei sistemi recensori; in qualsiasi ambiente, la reputazione serve a creare fiducia nel venditore (o nel produttore o nel professionista) al quale si paga il bene (o a cui si assegna l'esecuzione del progetto). Esistono però, diversi studi che sollevano forti dubbi sul funzionamento effettivo del sistema reputazionale di eBay (il cosiddetto "feedback forum"), oppure sul sistema adottato da Amazon che si basa principalmente sui *ranks*, vale a dire i punteggi ottenuti da ogni *reviewer* (revisore).

In più, stanno nascendo online sempre più siti che si offrono come "guardiani della reputazione" (*reputation guardian*), e offrono i propri servizi a tutti coloro che sono interessati ad avere una corretta reputazione online, o anche a nascondersi dai motori di ricerca. Quindi siti, forum, blog, motori di ricerca, social network, video online, etc., possono risultare utili, per valutare un prodotto, servizio, venditore o professionista, ma è necessario tener presente la possibilità di trovarsi, sempre più, davanti a degli inganni.

Nel capitolo successivo saranno esposte alcune indagini realizzate per poter osservare più in profondità le difficoltà riscontrate dai consumatori negli ambienti online dedicati agli acquisti. Prendendo come prodotto di indagine il telefono cellulare, sono stati considerati attributi, opzioni, ed offerte disponibili online, sul prodotto esaminato. In particolar modo, tenuto conto di quanto detto, soprattutto in questo capitolo, è stata effettuata una analisi delle opinioni rilasciate dagli utenti (consumatori e/o esperti) su uno dei principali price comparison italiani.

## Capitolo 5

### **Analisi empiriche e indagini esplorative sulle decisioni di acquisto nel mercato della telefonia mobile**

#### **5.1 Introduzione**

In questo capitolo sono riportati i risultati ottenuti mediante tre diverse indagini empiriche, realizzate per analizzare in modo più approfondito gli ambienti online dedicati agli acquisti, e quindi come si comporta il consumatore in tali ambienti, nell'ambito dei processi decisionali di acquisto. I dati utilizzati per queste indagini sono stati presi dal web, utilizzando un crawler chiamato *Web Content Extractor*.

Nelle indagini si è fatto riferimento a una particolare tipologia di prodotto, il telefono cellulare, in quanto prodotto complesso caratterizzato da una rapida evoluzione tecnologica, per il quale al momento dell'acquisto occorre considerare numerosi attributi; si tratta di uno dei prodotti maggiormente acquistati online, e molti dei suoi attributi possono essere facilmente misurati quantitativamente.

La prima indagine è stata realizzata con l'obiettivo di esaminare i telefoni cellulari venduti online in Italia. I dati considerati sono i valori di performance degli attributi di diverse opzioni di prodotto e le offerte di prezzo disponibili nelle centinaia di siti e-commerce italiani. Per una comodità di selezione, sono state considerate le offerte presentate su uno dei principali comparatori di prezzi del web italiano: Kelkoo.it; questo sito dispone di migliaia di offerte, provenienti da decine e decine di negozi online che vendono telefoni cellulari.

L'obiettivo della seconda indagine è stato quello di esaminare l'intero mercato dei telefoni cellulari venduti non solo online ma anche offline (negozi tradizionali), nel decennio dal 1998 al 2008, un periodo nel quale i telefoni cellulari sono stati caratterizzati da rapidi sviluppi tecnologici, passando da semplici apparecchi per comunicare, a veri e propri minicomputer tascabili. La terza indagine ha consentito di analizzare le opinioni espresse da esperti o consumatori che hanno acquistato dei telefoni cellulari. In particolare, sono state analizzate le numerose opinioni lasciate dagli utenti di Kelkoo.it, con riferimento a telefoni cellulari venduti negli ultimi cinque anni. Sono stati considerati i punteggi attribuiti dagli utenti per diversi aspetti dei prodotti (valutazione complessiva, valutazione dello stile, della durata della batteria, della facilità d'uso, della qualità, e del rapporto qualità-prezzo).

Sono state utilizzate sia metodologie di apprendimento supervisionato, come l'analisi di regressione multivariata semplice e l'analisi di regressione *stepwise*, sia metodologie di apprendimento non supervisionato, come il *data mining*, mediante algoritmi di *Decision tree analysis* e di *Cluster analysis*<sup>57</sup>.

#### **5.2 Analisi sui cellulari venduti online sul sito *kelkoo.it***

Sono numerosissimi i siti di e-commerce online dove è possibile acquistare un telefonino, e non è facile quindi confrontare le diverse offerte, se non attraverso i comparatori di prezzo, motori di ricerca che confrontano le offerte provenienti dai diversi siti e-commerce. Uno dei più famosi motori di ricerca dedicati agli acquisti è Kelkoo.it, con migliaia di offerte provenienti da numerosi negozi online.

Kelkoo non è una creazione italiana ma francese, successivamente esportata negli ambienti online di altri paesi europei. Il web italiano centrato sugli acquisti online si è avvicinato tardi a questi nuovi servizi, sottovalutando le grandi opportunità che essi offrono. Ciò

---

<sup>57</sup> Analisi approfondite di queste metodologie sono presentate nel terzo capitolo di questa tesi, e in particolare nei paragrafi 3.4, 3.5 e 3.6.

è stato un peccato, soprattutto perché secondo recenti sondaggi l'Italia è tra gli ultimi paesi in Europa nelle vendite e negli acquisti online.

In questo paragrafo sono presentate le analisi effettuate in base ai dati relativi alle opzioni di cellulari disponibili sul sito Kelkoo.it a luglio del 2008; sono state utilizzate sia metodologie di analisi di regressione multivariata semplice, sia metodi di analisi di regressione *stepwise*, sia metodologie di *cluster analysis*.

Accedendo al sito <http://www.kelkoo.it>, e successivamente linkando all'indirizzo web <http://www.kelkoo.it/c-100010713-cellulari.html>, si accede alla pagina dedicata alle offerte dei diversi siti e-commerce italiani raggruppate per opzione di prodotto; il consumatore ha la possibilità, per ogni opzione di prodotto, di accedere immediatamente all'offerta più conveniente, oppure visualizzare per ogni opzione, tutte le diverse offerte disponibili.

Nel mese di luglio 2008, alla pagina dedicata ai cellulari del sito kelkoo.it, erano disponibili 298 cellulari delle principali marche di telefonini: 2 cellulari Apple, 5 cellulari Hewlett Packard, 19 cellulari Htc, 13 cellulari i-Mate, 24 cellulari Lg, 35 cellulari Motorola, 69 cellulari Nokia, 5 cellulari Rim, 89 cellulari Samsung e 47 cellulari Sony Ericsson.

Per gran parte dei telefonini si disponeva di diverse decine di offerte provenienti da diversi siti, ma nell'indagine sono state considerate, per ogni opzione di prodotto, l'offerta con il prezzo più basso. Per ogni cellulare sono state considerate 28 caratteristiche, di cui alcune dicotomiche (presenti o assenti), alcune discrete e altre continue.

### 5.2.1 Statistiche descrittive e matrice di correlazione

La tabella 5.1 illustra alcune statistiche descrittive del campione di cellulari considerato. Le caratteristiche dalla numero 1 alla numero 11 sono misurate da variabili continue, quelle dalla numero 12 alla numero 14 da variabili discrete (lo zoom può assumere 8 valori discreti fra zero e 100; Internet può essere *Assente*, *Gprs*, *edge*, *Umts*, *Hsdpa*; la connettività senza fili può essere *Assente*, *Infrarossi*, *Bluetooth*, *Wi-Fi*). Le caratteristiche dalla numero 15 alla numero 28 sono dicotomiche (presenti o assenti). Da rilevare che più del 75% dei cellulari considerati dispone di Mp3, di videocamera, e quindi anche di fotocamera, e offre la possibilità di inviare e-mail.

La tabella 5.2 presenta la matrice di correlazione tra le variabili relative alle caratteristiche dei cellulari venduti online su Kelkoo.it inclusi nell'analisi. Sinteticamente, osservando la matrice di correlazione si può rilevare come la qualità del display (il numero di colori del display) sia correlata positivamente con la risoluzione della fotocamera (scatti di qualità, hanno bisogno di display di qualità); come la connettività senza fili è correlata positivamente con le caratteristiche: sistema operativo Windows, mail, connessione a Internet, memoria espandibile, mp3, peso, prezzo, risoluzione del display e della fotocamera, *touch screen* e videocamera. Dalla matrice si può osservare inoltre che gran parte dei cellulari con sistema operativo Windows presentano anche una buona connettività senza fili e schermo *Touch Screen*, ma hanno peso e dimensioni elevate.

Le variabili velocità di connessione a internet (internet), risoluzione della fotocamera e risoluzione del display (dimensione del display) costituiscono *l'essenza tecnologica del telefonino*, perché sono quelle sulle quali i produttori hanno puntato maggiormente, facendosi una fortissima concorrenza, che ha determinato una evoluzione tecnologica molto rapida dei telefoni cellulari, senza consistenti aumenti di prezzo. Ecco perché, come dimostra la matrice di correlazione, la velocità di navigazione (Internet) e la risoluzione della fotocamera siano entrambe correlate con le stesse variabili: colori e risoluzione del display, connettività senza fili, Gps, mail, flash e zoom della fotocamera, risoluzione della fotocamera, presenza di memoria espandibile, mp3, videocamera, e funzione di videochiamata. Lo stesso vale per la risoluzione del display che non si discosta molto dai coefficienti di correlazione delle altre due variabili; infine, tutte e tre le variabili sono correlate negativamente con l'uscita sul mercato.

Tabella 5.1 - valori minimi, valori massimi, valori medi, e deviazione standard per le caratteristiche misurate da variabili continue, e percentuali di presenza per le variabili dicotomiche.

	Min	Max	Media	Deviazione Standard	% presenza variabili dicotomiche	
1	Prezzo in euro	24	1277	207	174	
2	Mesi presenza sul mercato	-2	67	18	11	
3	Autonomia conversazione in ore	60	600	286	102	
4	Autonomia in standby in ore	100	720	297	102	
5	Risoluzione display	0	307200	57395	43854	
6	Colori display	0	16000000	2136672	5229346	
7	Dimensioni in mm <sup>3</sup>	26977	393536	84780	30658	
8	Spessore in mm	6	43	16	4	
9	Memoria interna	0	16000	330	1632	
10	Peso in grammi	65	355	104	30	
11	Risoluz. fotocamera	0	5038848	1641371	1263041	
12	zoom	0	100	19	22	
13	Internet	0	100	54	31	
14	connettività	0	100	62	28	
15	flash	0	100	29	45	29%
16	videocamera	0	100	75	43	75%
17	Memoria espandibile	0	100	66	48	66%
18	mp3	0	100	87	34	87%
19	touch	0	100	14	35	14%
20	email	0	100	79	41	79%
21	videochiamata	0	100	35	48	35%
22	d_so_mac	0	100	1	8	1%
23	d_so_symbian	0	100	11	32	11%
24	d_so_windows	0	100	15	36	15%
25	a-gps	0	100	4	20	4%
26	gps	0	100	7	26	7%
27	qwerty	0	100	3	17	3%
28	radio	0	100	42	49	42%

	A_GPS	AUTONOM. CONVERSAZ.	AUTONOMIA IN STANBY	COLORI DISPLAY	CONNETTIVITA	D_SO_MAC	D_SO_SYMBIAN	D_SO_WINDOWS	DIMENSIONI	EMAIL	FLASH	GPS	INTERNET	MEMORIA ESPAND.	MEMORIA INTERNA	MP3	PESO	PREZZO	QWERTY	RADIO	RISOLUZ. DISPLAY	RISOLUZ. FOTOCAMERA	SPESSORE	TOUCH	PRESENZA SUL MERCATO	VIDEOCAMERA	VIDEOCHIAMATA	ZOOM
A_GPS	1.00	0.02	0.01	0.28	0.23	0.19	0.25	0.10	0.10	0.11	0.25	0.27	0.27	0.08	0.22	0.08	0.15	0.23	0.16	0.07	0.11	0.29	-0.01	0.11	-0.19	0.12	0.21	0.27
AUTONOM. CONVERSAZ.	0.02	1.00	0.45	-0.13	-0.11	0.01	-0.01	-0.11	-0.15	-0.05	0.04	0.14	-0.02	0.02	0.01	-0.12	-0.08	-0.11	-0.02	0.09	-0.01	0.04	-0.14	-0.02	-0.24	-0.08	0.00	0.00
AUTONOMIA STANBY	0.01	0.45	1.00	0.15	-0.16	0.03	0.10	-0.09	-0.03	0.00	0.00	0.08	-0.03	-0.07	0.03	-0.17	-0.05	0.01	0.07	0.16	-0.04	-0.03	-0.04	0.03	-0.11	-0.17	0.02	-0.06
COLORI DISPLAY	0.28	-0.13	0.15	1.00	0.22	0.22	0.48	-0.16	0.01	0.18	0.26	0.20	0.30	0.09	0.27	0.13	0.19	0.25	0.11	0.18	0.24	0.36	-0.02	-0.10	-0.26	0.23	0.28	0.33
CONNETTIVITA	0.23	-0.11	-0.16	0.22	1.00	0.11	0.25	0.41	0.23	0.63	0.30	0.29	0.58	0.53	0.21	0.62	0.53	0.52	0.21	-0.01	0.55	0.55	-0.10	0.38	-0.15	0.66	0.35	0.25
D_SO_MAC	0.19	0.01	0.03	0.22	0.11	1.00	-0.03	-0.03	0.00	0.04	-0.05	0.13	0.05	-0.11	0.79	0.03	0.08	0.14	-0.01	-0.07	0.18	0.02	-0.10	0.20	-0.09	0.05	-0.06	-0.07
D_SO_SYMBIAN	0.25	-0.01	0.10	0.48	0.25	-0.03	1.00	-0.15	0.12	0.19	0.38	0.26	0.35	0.21	0.13	0.14	0.23	0.21	0.12	0.23	0.25	0.31	0.04	0.06	-0.14	0.14	0.36	0.48
D_SO_WINDOWS	0.10	-0.11	-0.09	-0.16	0.41	-0.03	-0.15	1.00	0.31	0.20	-0.02	0.20	0.15	0.28	0.03	0.17	0.50	0.37	0.20	-0.26	0.37	0.11	-0.05	0.57	-0.04	0.22	0.03	-0.17
DIMENSIONI	0.10	-0.15	-0.03	0.01	0.23	0.00	0.12	0.31	1.00	0.08	0.05	0.19	0.07	0.15	0.09	0.03	0.61	0.29	0.07	-0.05	0.25	0.03	0.68	0.31	0.21	0.04	0.06	0.04
EMAIL	0.11	-0.05	0.00	0.18	0.63	0.04	0.19	0.20	0.08	1.00	0.28	0.15	0.51	0.48	0.10	0.56	0.30	0.34	0.04	0.06	0.40	0.45	-0.15	0.21	-0.16	0.54	0.30	0.25
FLASH	0.25	0.04	0.00	0.26	0.30	-0.05	0.38	-0.02	0.05	0.28	1.00	0.11	0.36	0.33	0.06	0.23	0.21	0.23	0.02	0.16	0.32	0.58	0.00	0.10	-0.16	0.33	0.35	0.45
GPS	0.27	0.14	0.08	0.20	0.29	0.13	0.26	0.20	0.19	0.15	0.11	1.00	0.33	0.12	0.25	0.11	0.39	0.35	0.18	0.10	0.41	0.31	-0.02	0.25	-0.23	0.13	0.22	0.14
INTERNET	0.27	-0.02	-0.03	0.30	0.58	0.05	0.35	0.15	0.07	0.51	0.36	0.33	1.00	0.55	0.15	0.51	0.36	0.47	0.16	0.04	0.57	0.66	-0.15	0.22	-0.38	0.58	0.81	0.40
RISOLUZ. FOTOCAMERA	0.29	0.04	-0.03	0.36	0.55	0.02	0.31	0.11	0.03	0.45	0.58	0.31	0.66	0.52	0.12	0.46	0.35	0.53	0.11	0.12	0.54	1.00	-0.20	0.24	-0.44	0.68	0.55	0.48
RISOLUZ. DISPLAY	0.11	-0.01	-0.04	0.24	0.55	0.18	0.25	0.37	0.25	0.40	0.32	0.41	0.57	0.42	0.33	0.37	0.63	0.55	0.26	-0.04	1.00	0.54	-0.15	0.45	-0.26	0.45	0.42	0.19
MEMORIA ESPAND.	0.08	0.02	-0.07	0.09	0.53	-0.11	0.21	0.28	0.15	0.48	0.33	0.12	0.55	1.00	-0.11	0.48	0.32	0.26	0.13	0.06	0.42	0.52	-0.13	0.24	-0.28	0.61	0.38	0.27
MEMORIA INTERNA	0.22	0.01	0.03	0.27	0.21	0.79	0.13	0.03	0.09	0.10	0.06	0.25	0.15	-0.11	1.00	0.08	0.27	0.29	-0.02	0.01	0.33	0.12	-0.08	0.27	-0.13	0.09	0.05	0.05
MP3	0.08	-0.12	-0.17	0.13	0.62	0.03	0.14	0.17	0.03	0.56	0.23	0.11	0.51	0.48	0.08	1.00	0.25	0.30	0.07	0.05	0.37	0.46	-0.16	0.16	-0.26	0.63	0.29	0.26
PESO	0.15	-0.08	-0.05	0.19	0.53	0.08	0.23	0.50	0.61	0.30	0.21	0.39	0.36	0.32	0.27	0.25	1.00	0.68	0.20	-0.11	<b>0.63</b>	0.35	0.17	0.52	0.00	0.31	0.26	0.09
PREZZO	0.23	-0.11	0.01	0.25	0.52	0.14	0.21	0.37	0.29	0.34	0.23	0.35	0.47	0.26	0.29	0.30	0.68	1.00	0.16	-0.03	0.55	0.53	-0.05	0.47	-0.18	0.39	0.33	0.09
QWERTY	0.16	-0.02	0.07	0.11	0.21	-0.01	0.12	0.20	0.07	0.04	0.02	0.18	0.16	0.13	-0.02	0.07	0.20	0.16	1.00	-0.07	0.26	0.11	-0.07	0.04	-0.13	0.10	0.16	0.00
RADIO	0.07	0.09	0.16	0.18	-0.01	-0.07	0.23	-0.26	-0.05	0.06	0.16	0.10	0.04	0.06	0.01	0.05	-0.11	-0.03	-0.07	1.00	-0.04	0.12	0.05	-0.11	-0.21	-0.01	0.05	0.29
SPESSORE	-0.01	-0.14	-0.04	-0.02	-0.10	-0.10	0.04	-0.05	0.68	-0.15	0.00	-0.02	-0.15	-0.13	-0.08	-0.16	0.17	-0.05	-0.07	0.05	-0.15	-0.20	1.00	-0.07	0.38	-0.19	-0.04	0.05
TOUCH	0.11	-0.02	0.03	-0.10	0.38	0.20	0.06	0.57	0.31	0.21	0.10	0.25	0.22	0.24	0.27	0.16	0.52	0.47	0.04	-0.11	0.45	0.24	-0.07	1.00	-0.13	0.15	0.08	-0.11
PRESENZA SUL MERCATO	-0.19	-0.24	-0.11	-0.26	-0.15	-0.09	-0.14	-0.04	0.21	-0.16	-0.16	-0.23	-0.38	-0.28	-0.13	-0.26	0.00	-0.18	-0.13	-0.21	-0.26	-0.44	0.38	-0.13	1.00	-0.38	-0.30	-0.23
VIDEOCAMERA	0.12	-0.08	-0.17	0.23	0.66	0.05	0.14	0.22	0.04	0.54	0.33	0.13	0.58	0.61	0.09	0.63	0.31	0.39	0.10	-0.01	0.45	0.68	-0.19	0.15	-0.38	1.00	0.42	0.33
VIDEOCHIAMATA	0.21	0.00	0.02	0.28	0.35	-0.06	0.36	0.03	0.06	0.30	0.35	0.22	0.81	0.38	0.05	0.29	0.26	0.33	0.16	0.05	0.42	0.55	-0.04	0.08	-0.30	0.42	1.00	0.43
ZOOM	0.27	0.00	-0.06	0.33	0.25	-0.07	0.48	-0.17	0.04	0.25	0.45	0.14	0.40	0.27	0.05	0.26	0.09	0.09	0.00	0.29	0.19	0.48	0.05	-0.11	-0.23	0.33	0.43	1.00

Tabella 5.2 - Matrice di Correlazione tra le variabili relative alle caratteristiche dei cellulari (*I coefficienti di correlazione maggiori o uguali a 0,30 sono evidenziati in rosso*).

### 5.2.2 Un'analisi di regressione multivariata, dei prezzi edonici dei cellulari

Kelvin Lancaster introdusse negli anni sessanta una innovazione fondamentale nell'analisi del comportamento dei consumatori, mettendo in evidenza che ciò che i consumatori spesso desiderano non è tanto un prodotto specifico, ma un insieme di caratteristiche; per esempio un uomo d'affari che viaggia spesso richiede che il suo computer portatile sia leggero, abbia una buona autonomia e una elevata potenza di calcolo, ecc. piuttosto che un certo logo sulla confezione; ovviamente, spesso è necessario stabilire dei compromessi (*trade-off*) fra diverse caratteristiche [Lancaster, 1966]. L'analisi dei prezzi edonici discende dalla teoria di Lancaster del comportamento dei consumatori in termini di caratteristiche dei prodotti. Le stime dei coefficienti in un'analisi di regressione multivariata, in cui il prezzo di diverse specificazioni di un prodotto rappresenta la variabile dipendente (da spiegare) e gli indicatori delle diverse caratteristiche sono variabili indipendenti o esplicative (*predictor*), possono essere interpretate come indicatori del valore delle diverse caratteristiche per i consumatori<sup>58</sup>. L'analisi dei prezzi edonici fu inizialmente applicata alle case per abitazione; un'applicazione ai telefoni cellulari sul mercato tedesco fu realizzata da Dewenter [Dewenter et al., 2007].

Il prezzo riassume in sintesi le prestazioni di un prodotto; è importante quindi valutare le relazioni esistenti tra le performance delle varie caratteristiche del prodotto con il prezzo di mercato, perché ciò offre la possibilità di evidenziare quanto influisce singolarmente ciascuna caratteristica, sul prezzo del prodotto.

Attraverso l'applicativo statistico Eviews 6.0, sono state effettuate delle regressioni lineari con metodo *Least Square* (metodo dei minimi quadrati), ponendo come variabile dipendente il prezzo e variabili esplicative le diverse caratteristiche dei cellulari. La tabella 5.3 mostra i primi risultati ottenuti mediante l'analisi di regressione.

Dai risultati ottenuti dalla regressione si può vedere quali caratteristiche evidenziano delle relazioni statisticamente significative con il prezzo; in particolare, sono significativamente diverse da zero, a un livello di probabilità inferiore all'1 per cento, le stime dei coefficienti delle variabili autonomia di conversazione, memoria espandibile, peso, risoluzione della fotocamera, zoom; è significativamente diversa da zero a un livello di probabilità inferiore al 5 per cento la stima del coefficiente della variabile autonomia in standby; sono significativamente diverse da zero, a un livello di probabilità inferiore al 10 per cento, le stime dei coefficienti delle variabili dimensioni, flash, e connessione a internet; è significativamente positivo il segno dei coefficienti delle variabili: autonomia in standby, connessione a internet, peso, risoluzione fotocamera; è significativamente negativo il segno dei coefficienti delle variabili autonomia in conversazione, dimensione, flash, memoria espandibile. Sembra strano avere un coefficiente positivo altamente significativo per quanto riguarda il peso del telefonino; ciò perché, essendo il peso una caratteristica negativa per l'acquirente, esso dovrebbe risultare inversamente proporzionale al prezzo; ciò deriva probabilmente dal fatto che i telefonini cosiddetti *smartphone* perché ultratecnologici, caratterizzati da prezzi molto alti, utilizzano molta energia e quindi devono essere dotati di grandi batterie che fanno aumentare di molto il peso complessivo del telefonino. In modo analogo si può probabilmente spiegare il coefficiente significativamente negativo per la variabile autonomia di conversazione: diversi telefonini di elevate prestazioni (in particolare per il display), e prezzo elevato, consumano molta energia, e hanno quindi una bassa autonomia di conversazione.

---

<sup>58</sup> Un'altra applicazione empirica della teoria di Lancaster del comportamento dei consumatori è la *conjoint analysis*, elaborata nell'ambito della psicologia matematica (Johnson, 1974; Green and Srinivasan, 1978; Thaler, 1985). La *conjoint analysis* si applica in genere quando non sono disponibili prezzi affidabili dei prodotti, e si basa principalmente su interviste a potenziali acquirenti di diverse specificazioni di un prodotto.

Tabella 5.3 – Stime dei parametri di una relazione lineare tra i prezzi e le caratteristiche dei cellulari ottenute con il metodo dei minimi quadrati (*least square*).

Dependent Variable: PREZZO

Method: Least Squares

Date: 07/19/08 Time: 09:24

Sample: 1 298

Included observations: 298

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-174.8783	52.33539	-3.341492	0.0010
A_GPS	0.415507	0.387959	1.071008	0.2851
AUTONOM_CONV	-0.208245	0.079397	-2.622844	0.0092
AUTONOMIA_STAND-BY	0.157734	0.078498	2.009398	0.0455
COLORI_DISPLAY	-9.34E-07	1.74E-06	-0.537985	0.5910
CONNETTIVITA	0.291431	0.428353	0.680354	0.4969
D_SO_MAC	-0.177970	1.410256	-0.126197	0.8997
D_SO_SYMBIAN	0.193506	0.280103	0.690839	0.4903
D_SO_WINDOWS	0.219846	0.276304	0.795666	0.4269
DIMENSIONI	-0.000682	0.000412	-1.655736	0.0989
EMAIL	0.088608	0.220636	0.401602	0.6883
FLASH	-0.319986	0.189682	-1.686962	0.0928
GPS	-0.173520	0.300996	-0.576485	0.5648
INTERNET	0.799299	0.478409	1.670743	0.0959
MEMORIA_ESP	-0.570541	0.201483	-2.831713	0.0050
MEMORIA_INTERNA	0.005790	0.007359	0.786744	0.4321
MP3	-0.063610	0.275155	-0.231177	0.8174
PESO	3.337111	0.403332	8.273855	0.0000
QWERTY	-0.130643	0.420771	-0.310486	0.7564
RADIO	0.124084	0.148531	0.835408	0.4042
RISOLUZ_DISPLAY	-6.49E-05	0.000244	-0.266505	0.7901
RISOLUZ_FOTOC	5.70E-05	9.91E-06	5.749127	0.0000
SPESSORE	1.225751	2.635190	0.465147	0.6422
TOUCH	0.243003	0.277489	0.875721	0.3820
USCITA_MERCATO	-0.437963	0.775603	-0.564674	0.5728
VIDEOCAMERA	-0.034106	0.273922	-0.124510	0.9010
VIDEOCHIAMATA	-0.154314	0.251872	-0.612669	0.5406
ZOOM	-1.085827	0.395630	-2.744552	0.0065
<i>R-squared</i>	<i>0.641587</i>	<i>Mean dependent var</i>		<i>206.7315</i>
<i>Adjusted R-squared</i>	<i>0.605746</i>	<i>S.D. dependent var</i>		<i>174.0130</i>
<i>S.E. of regression</i>	<i>109.2622</i>	<i>Akaike info criterion</i>		<i>12.31463</i>
<i>Sum squared residuals</i>	<i>3223321.</i>	<i>Schwarz criterion</i>		<i>12.66200</i>
<i>Log likelihood</i>	<i>-1806.879</i>	<i>F-statistic</i>		<i>17.90078</i>
<i>Durbin-Watson statistics</i>	<i>0.899455</i>	<i>Prob(F-statistic)</i>		<i>0.000000</i>

La maggior parte delle variabili relative alle 26 caratteristiche considerate sono direttamente proporzionali all'utilità derivabile dall'acquisto del prodotto; quanto più alto è il loro valore, tanto più alta sarà la soddisfazione dell'utente e tanto più alto sarà anche il prezzo del prodotto. Alcune variabili come peso, dimensione, spessore, e uscita sul mercato (mesi di presenza sul mercato), sono invece, inversamente proporzionali all'utilità attesa. Si può ottenere anche per queste caratteristiche una variabile positivamente correlata alla soddisfazione dell'utente, ponendo uguale a 100 il valore della variabile quando il peso (o la dimensione o lo spessore, o l'uscita sul mercato) è minimo, un valore zero quando il peso è massimo, un valore 50 quando il peso ha il valore mediano, e ottenendo per interpolazione lineare tutti gli altri valori intermedi. In questo modo viene anche uniformato l'ambito di variabilità all'intervallo 0–100. Le stime ottenute utilizzando i valori standardizzati di queste variabili sono riportati nella Tabella 5.4.

Tabella 5.4 – Stime dei parametri di una relazione lineare tra i prezzi e le caratteristiche dei cellulari ottenute con metodo dei minimi quadrati (*least square*), con le variabili peso, dimensione, spessore e mesi di presenza sul mercato standardizzate.

Dependent Variable: PREZZO

Method: Least Squares

Sample: 1 298

Included observations: 298

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	300.7131	58.96103	5.100202	0.0000
A_GPS	0.418581	0.418336	1.000585	0.3179
AUTONOM_CONV	-0.213185	0.085677	-2.488249	0.0134
AUTONOMIA_STAND-BY	0.174214	0.084925	2.051395	0.0412
COLORI_DISPLAY	7.67E-07	1.84E-06	0.417140	0.6769
CONNETTIVITA	0.373582	0.459163	0.813615	0.4166
D_SO_MAC	-2.211567	1.506028	-1.468477	0.1431
D_SO_SYMBIAN	0.146657	0.302271	0.485182	0.6279
D_SO_WINDOWS	0.428773	0.299769	1.430344	0.1538
DIMENSIONI Standard.	0.369621	1.342762	0.275270	0.7833
EMAIL	-0.059728	0.240304	-0.248552	0.8039
FLASH	-0.326496	0.203340	-1.605661	0.1095
GPS	0.119875	0.323026	0.371099	0.7109
INTERNET	0.518747	0.516357	1.004630	0.3160
MEMORIA_ESP	-0.615687	0.216135	-2.848626	0.0047
MEMORIA_INTERNA	0.015335	0.007791	1.968301	0.0501
MP3	0.009128	0.295935	0.030844	0.9754
PESO Standardizzato	-3.649510	0.770510	-4.736487	0.0000
QWERTY	-0.035121	0.456159	-0.076993	0.9387
RADIO	0.066598	0.159497	0.417552	0.6766
RISOLUZ_DISPLAY	0.000380	0.000252	1.509961	0.1322
RISOLUZ_FOTOC	5.57E-05	1.08E-05	5.165431	0.0000
SPESSORE Standardizzato	-0.166672	0.958281	-0.173928	0.8621
TOUCH	0.575231	0.294271	1.954769	0.0516
USCITA sul MERCATO Standard.	0.123209	0.503519	0.244696	0.8069
VIDEOCAMERA	-0.121146	0.297588	-0.407094	0.6843
VIDEOCHIAMATA	-0.070046	0.272299	-0.257240	0.7972
ZOOM	-1.238678	0.426917	-2.901450	0.0040
<i>R-squared</i>	<i>0.583792</i>	<i>Mean dependent var</i>	<i>206.7315</i>	
<i>Adjusted R-squared</i>	<i>0.542171</i>	<i>S.D. dependent var</i>	<i>174.0130</i>	
<i>S.E. of regression</i>	<i>117.7426</i>	<i>Akaike info criterion</i>	<i>12.46413</i>	
<i>Sum squared residuals</i>	<i>3743093.</i>	<i>Schwarz criterion</i>	<i>12.81150</i>	
<i>Log likelihood</i>	<i>-1829.155</i>	<i>F-statistic</i>	<i>14.02642</i>	
<i>Durbin-Watson statistics</i>	<i>0.729125</i>	<i>Prob(F-statistic)</i>	<i>0.000000</i>	

Le principali differenze fra i risultati delle due regressioni, a parte ovviamente l'inversione di segno dei coefficienti delle variabili peso, dimensione, e spessore, sono le seguenti: il segno dei coefficienti delle variabili risoluzione display, numero colori del display, e funzione GPS cambia da negativo a positivo, ma le loro stime rimangono non significativamente diverse da zero.

Ogni volta che si stimano i parametri di una relazione mediante un'analisi di regressione si ottengono dei valori cosiddetti "*residui*", che il programma *Eviews* inserisce in una tabella apposita; si ottiene un residuo per ciascun telefonino. I residui indicano di quanto si discostano i prezzi effettivi dei diversi telefonini dai prezzi stimati in base all'equazione di regressione. I valori dei residui possono essere utilizzati per cercare di individuare i cellulari più convenienti (quelli con prezzo effettivo nettamente inferiore a quello stimato in base all'equazione di

regressione) e i cellulari meno convenienti (quelli il cui prezzo effettivo è nettamente maggiore di quello stimato in base all'equazione di regressione).

La tabella 5.5 presenta una sintesi dei risultati ottenuti osservando i residui di entrambe le equazioni stimate; essa mostra, ordinati in senso decrescente di convenienza, i dieci telefonini più convenienti e i dieci telefonini meno convenienti sulla base dei residui dalle equazioni stimate. In blu sono indicati i telefonini più convenienti secondo i residui dall'equazione di regressione stimata utilizzando i valori non standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato; in rosso sono indicati i telefonini meno convenienti secondo i residui dall'equazione di regressione stimata utilizzando i valori non standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato. In nero sono indicati i telefonini più convenienti, e quelli meno convenienti, secondo i residui dall'equazione di regressione stimata utilizzando i valori standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato.

La differenza tra le due classifiche è notevole soprattutto per alcuni cellulari; per esempio, l'HTC X7500 nella regressione con variabili non standardizzate si trova tra i primi 10, mentre in quella con alcune variabili standardizzate si trova al 278° posto; ciò si verifica presumibilmente perché, essendo l'HTC X7500 il cellulare con il peso maggiore, la standardizzazione della variabile peso ha influenzato maggiormente la relazione fra il suo peso effettivo e quello stimato in base alla regressione lineare.

Tabella 5.5 Cellulari più convenienti e meno convenienti secondo i risultati dell'analisi di regressione fra i prezzi e le caratteristiche dei cellulari (in blu sono indicati i telefonini più convenienti secondo i residui dall'equazione di regressione stimata utilizzando i valori non standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato; in rosso sono indicati i telefonini meno convenienti secondo i residui dall'equazione di regressione stimata utilizzando i valori non standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato; in nero sono indicati i telefonini più convenienti, e quelli meno convenienti, secondo i residui dall'equazione di regressione stimata utilizzando i valori standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato).

	marchio	modello	prezzo	residui con LS semplice	residui ols con alcune standardizz.
1°	Samsung	SGH-G600	180	-172	-198
2°	Samsung	SGH-U300	108	-180	-198
3°	i-Mate	JAMA	139	-119	-185
4°	Samsung	SGH-P200	76	-154	-180
5°	Nokia	N95 8GB	369	-107	-179
6°	Samsung	SGH-F700	279	-179	-177
7°	Motorola	V600	58	-192	-173
8°	Nokia	6500 Classic	200	-105	-161
9°	HTC	Touch Diamond	490	53	-158
10°	Nokia	3120 Classic	139	-122	-156
11°	i-Mate	JAQ3	233	-197	-151
14°	LG Electronics	KG920	204	-172	-136
15°	Motorola	RIZR Z3	89	-157	-130
18°	Nokia	E61i	215	-157	-117
39°	HTC	P4350	320	-179	-83
278°	HTC	X7500	911	-186	124
287°	RIM	Blackberry 8310 Curve	378	206	165
288°	Samsung	SGH-Z400	399	188	178
289°	Hewlett Packard	iPAQ h6345	544	112	181
290°	HTC	P6500	660	38	193
291°	RIM	Blackberry 8800	352	218	219
292°	Hewlett Packard	iPAQ hw6915	605	153	239
293°	Samsung	SGH-P520 Giorgio Armani	538	292	261
294°	Samsung	SGH-E800	358	289	270
295°	Nokia	8800 Arte	969	435	507
296°	Nokia	8800 Sapphire Arte	1033	500	572
297°	Nokia	8800 Sirocco Edition	1155	742	827
298°	Nokia	8800	1277	764	833

Il programma statistico Eviews offre anche la possibilità di realizzare dei grafici in grado di visualizzare l'andamento dei residui dalle relazioni stimate mediante l'analisi di regressione tenendo conto del prezzo dei telefonino. I prodotti sono stati disposti in ordine decrescente dei prezzi dei cellulari. in modo da facilitare la lettura (la curva dei prezzi è evidenziata in rosso).

Figura 5.1– L'andamento dei residui per la relazione stimata con le variabili non standardizzate

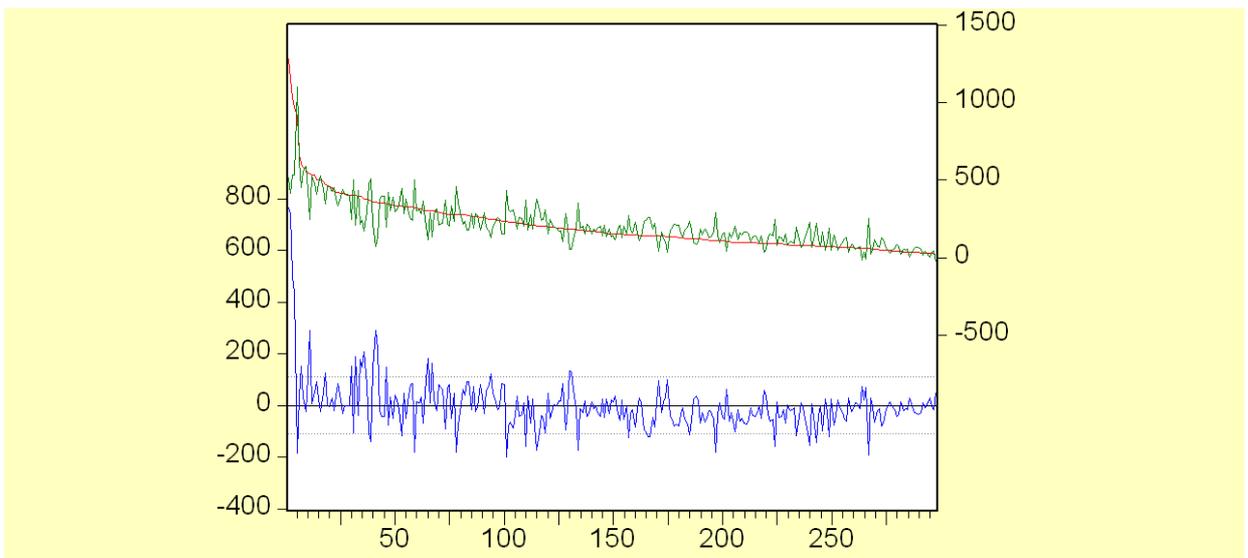
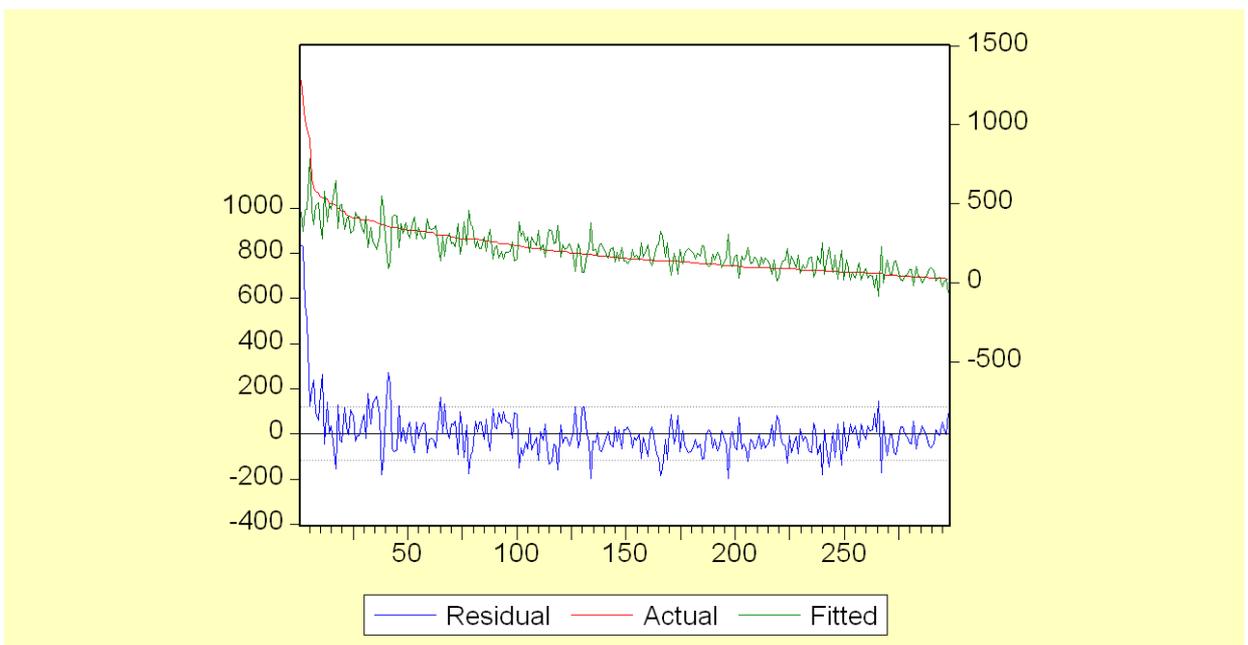


Figura 5.2– L'andamento dei residui dalla relazione stimata utilizzando i valori standardizzati delle variabili peso, dimensione, spessore, e mesi di presenza sul mercato.



### 5.2.2.1 Analisi di regressione con eliminazione progressiva delle variabili

La regressione multivariata che include tutte le 28 caratteristiche dei cellulari sopra elencate consente di spiegare il 64 per cento della variabilità dei prezzi dei cellulari; le stime di alcuni coefficienti hanno tuttavia un segno opposto rispetto a quello atteso; il caso più clamoroso è quello della variabile peso, il cui coefficiente è maggiore di zero con un livello di significatività statistica molto elevata. La spiegazione più plausibile di questo risultato è la presenza di una elevata correlazione statistica fra il peso e alcune caratteristiche qualitative dei cellulari che non è possibile rilevare in modo soddisfacente; in particolare diversi “cellulari intelligenti” (*smartphone*) con prestazioni particolarmente sofisticate hanno peso e prezzi particolarmente elevati. Esempi di caratteristiche positive dei cellulari per le quali il coefficiente stimato è risultato essere negativo con un elevato livello di significatività statistica sono l'autonomia in conversazione, la presenza o meno di memoria espandibile, la presenza o meno dello zoom. Si è ritenuto utile stimare i coefficienti della relazione fra il prezzo e le caratteristiche dei cellulari escludendo progressivamente le caratteristiche i cui coefficienti stimati avevano un segno opposto rispetto a quello atteso.

Alla fine di questo processo di successive eliminazioni di variabili sono state ottenute le stime riportate nella tabella 5.6. Eliminando 16 caratteristiche, sono rimaste nell'equazione di regressione 11 variabili esplicative delle originarie 27 (meno della metà). Ciò nonostante, il valore della percentuale della variabilità dei prezzi spiegata dal modello (misurata dal valore di R-quadro) è diminuita soltanto dal 58 al 49 per cento, e il valore di R-quadro aggiustato per i gradi di libertà è diminuito soltanto da 0,54 a 0,487. Contrastanti sono state le variazioni degli indicatori più utilizzati per valutare i modelli tenendo conto sia del loro potere esplicativo sia della loro complessità: il criterio di informazione di Akaike e il criterio di informazione bayesiano o di Schwarz; il valore del criterio di informazione di Akaike è aumentato da 12,46 a 12,56, mentre il valore del criterio di informazione di Schwarz è diminuito da 12,81 a 12,71. Considerato che i modelli migliori sono quelli a cui corrispondono valori più bassi dei valori di questi due criteri di informazione, gli effetti sulla bontà del modello esplicativo della variabilità dei prezzi dei cellulari della drastica riduzione delle caratteristiche considerate da 27 a 11 è ambiguo; tenendo conto nella valutazione sia della capacità esplicativa che della complessità del modello, secondo il criterio di Akaike la riduzione delle variabili comporterebbe un peggioramento della bontà complessiva del modello, mentre secondo il criterio di Schwarz la riduzione delle variabili comporterebbe addirittura un miglioramento della qualità del modello. Queste valutazioni contrastanti derivano dal fatto che il criterio di informazione bayesiano o di Schwarz penalizza maggiormente del criterio di Akaike l'aumento delle variabili, e quindi della complessità del modello.

Tabella 5.6 - Stime dei parametri di una relazione lineare tra i prezzi e le caratteristiche dei cellulari ottenute con metodo dei minimi quadrati (*least square*) senza peso, uscita sul mercato, flash, zoom, videocamera, mp3, email, radio, memoria espandibile, videochiamata, presenza sistema operativo Mac o Symbian, autonomia in conversazione, a\_gps, spessore e dimensioni

Dependent Variable: PREZZO

Method: Least Squares

Sample: 1 298

Included observations: 298

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18.11823	32.43715	0.558564	0.5769
RISOLUZ_FOTOCAM	3.59E-05	8.68E-06	4.131033	4.75E-05
COLORI_DISPLAY	3.07E-06	1.68E-06	1.828105	0.068575
TOUCH	0.903951	0.289362	3.123946	0.001968
D_SO_WINDOWS	0.76014	0.280938	2.70572	0.007225
CONNETTIVITA	0.570556	0.382549	1.491458	0.136943
GPS	0.318046	0.319812	0.994476	0.320832
INTERNET	0.279132	0.350261	0.796925	0.426156
QWERTY	0.153073	0.464356	0.329645	0.74191
AUTONOMIA_STANBY	0.051496	0.076519	0.67298	0.501503
MEMORIA_INTERNA	0.010809	0.005242	2.061965	0.040114
RISOLUZ_DISPLAY	0.000462	0.000253	1.826643	0.068796
<i>R-squared</i>	<i>0.489561</i>	<i>Mean dependent var</i>		<i>206.7315</i>
<i>Adjusted R-squared</i>	<i>0.469929</i>	<i>S.D. dependent var</i>		<i>174.0130</i>
<i>S.E. of regression</i>	<i>126.6919</i>	<i>Akaike info criterion</i>		<i>12.56083</i>
<i>Sum squared residuals</i>	<i>4590536.</i>	<i>Schwarz criterion</i>		<i>12.70970</i>
<i>Log likelihood</i>	<i>-1859.563</i>	<i>F-statistic</i>		<i>24.93657</i>
<i>Durbin-Watson statistics</i>	<i>1.785867</i>	<i>Prob(F-statistic)</i>		<i>0.000000</i>

Considerando le 11 caratteristiche rimaste, quelle che risultano essere associate in misura statisticamente più significativa con le variazioni dei prezzi dei cellulari sono la risoluzione della fotocamera, il *touch*, il sistema operativo Windows, la memoria interna, i colori del display, e la risoluzione del display.

### 5.2.2.2 Analisi di *stepwise regression*

In questo paragrafo sono presentati i risultati di un'analisi di regressione multivariata con il metodo *stepwise* (Tabella 5.7), per identificare quali sono le variabili indispensabili e quali sono le variabili o attributi che si può fare a meno di considerare.

Tabella 5.7 – Stime dei parametri di una relazione lineare tra i prezzi e le 28 caratteristiche che definiscono i 298 telefoni cellulari venduti online, in base alla metodologia *Stepwise Regression*.

Dependent Variable: PREZZO  
 Method: Stepwise Regression  
 Sample: 1 298  
 Included observations: 298  
 No always included regressors  
 Number of search regressors: 28  
 Selection method: Stepwise forwards  
 Stopping criterion: p-value forwards/backwards = 0.5/0.5

	Coefficient	Std. Error	t-Statistic	Prob.
PESO	3.250791	0.339183	9.584181	0.0000
DIMENSIONI	-0.000558	0.000272	-2.046645	0.0416
RISOLUZ_FOTOCAMERA	5.35E-05	8.17E-06	6.555018	0.0000
C	-165.3343	36.48194	-4.531951	0.0000
ZOOM	-1.049537	0.360437	-2.911845	0.0039
MEMORIA_ESPAND	-0.525992	0.179965	-2.922736	0.0037
INTERNET	0.586605	0.301767	1.943901	0.0529
AUTONOM_CONVERSAZ	-0.200890	0.069502	-2.890409	0.0041
AUTONOMIA_STANBY	0.150516	0.071074	2.117736	0.0351
TOUCH	0.384015	0.221439	1.734181	0.0840
FLASH	-0.297095	0.175805	-1.689910	0.0921
A_GPS	0.435958	0.347681	1.253901	0.2109
CONNETTIVITA	0.351047	0.328246	1.069465	0.2858
RADIO	0.128743	0.136066	0.946185	0.3449
MEMORIA_INTERNA	0.003752	0.004311	0.870354	0.3848
<i>R-squared</i>	0.638467	<i>Mean dependent var</i>		206.7315
<i>Adjusted R-squared</i>	0.620582	<i>S.D. dependent var</i>		174.0130
<i>S.E. of regression</i>	107.1866	<i>Akaike info criterion</i>		12.23604
<i>Sum squared residuals</i>	3251378.	<i>Schwarz criterion</i>		12.42214
<i>Log likelihood</i>	-1808.171	<i>Hannan-Quinn criter.</i>		12.31054
<i>F-statistic</i>	35.69845	<i>Durbin-Watson statistics</i>		1.773436
<i>Prob(F-statistic)</i>	0.000000			

Utilizzando il metodo di *stepwise regression* risulta che 14 caratteristiche delle 27 inizialmente considerate riescono a spiegare la maggior parte della variabilità; 13 caratteristiche (fra cui risoluzione del display, email, mp3, qwerty, spessore, videocamera, videochiamata, ecc.) secondo questo metodo possono essere rimosse dal modello senza ridurre significativamente il suo potere esplicativo. In effetti, eliminando queste 13 caratteristiche la percentuale della variabilità dei prezzi spiegata dal modello diminuisce soltanto dal 64,1 al 63,8 per cento, il valore di R-quadro aggiustato per i gradi di libertà aumenta da 0,606 a 0,621, il valore del

criterio di informazione di Akaike diminuisce da 12,315 a 12,236, il valore del criterio di informazione bayesiano o di Schwarz diminuisce da 12,662 a 12,422<sup>59</sup>. Le caratteristiche che risultano essere associate positivamente in misura statisticamente significativa a un livello di probabilità inferiore all'1 per cento con i prezzi dei telefoni cellulari sono il peso e la risoluzione della fotocamera; risultano essere associate positivamente con il prezzo a un livello di probabilità inferiore al 5 per cento l'autonomia in standby e la connessione a Internet. Il valore significativamente maggiore di zero del coefficiente della variabile peso riflette la presenza di diversi telefonini che pesano molto perché hanno prestazioni avanzate che non è stato possibile misurare in modo adeguato, e che hanno un prezzo elevato, non perché pesano molto, ma per le funzioni avanzate che li caratterizzano. Le caratteristiche che risultano essere associate negativamente in misura statisticamente significativa a un livello di probabilità inferiore all'1 per cento con i prezzi dei telefoni cellulari sono lo zoom e la memoria espandibile; risulta essere associata negativamente con il prezzo a un livello di probabilità inferiore al 5 per cento la dimensione dei cellulari.

Fasolo, McClelland e Todd (2007, *pagg. 13-14*) hanno evidenziato che “la possibilità di scelta é una lama a doppio taglio. Da un lato le persone sono naturalmente attratte dalla possibilità di scegliere .... Dall'altro lato, però, più possibilità di scelta non costituiscono sempre un fatto positivo.” Fin dagli anni settanta diverse ricerche hanno evidenziato che una sovrabbondanza di attributi dei prodotti può originare un problema di sovraccarico di informazioni (*information overload*). Un caso esemplare in cui può presentarsi in concreto un problema di sovraccarico di informazioni riguarda l'acquisto di un telefono cellulare. I risultati qui presentati mostrano che poche caratteristiche sono in grado di spiegare una proporzione elevata della variabilità dei prezzi dei telefoni cellulari, e quindi, presumibilmente, del valore per la media dei consumatori di diverse specificazioni dei telefonini.

## 5.2.3 Un'analisi di Data Mining

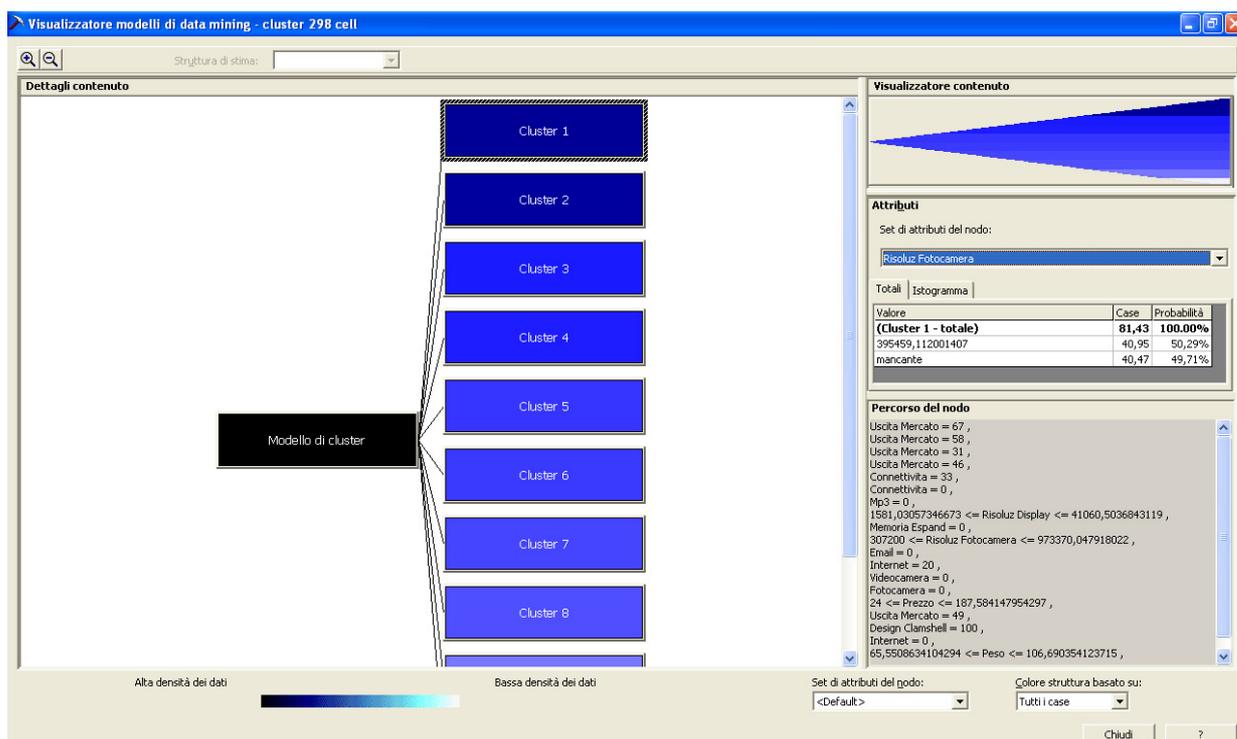
### 5.2.3.1 Analisi di clustering

L'applicazione *Microsoft Analysis Manager*, attraverso l'elaborazione dei dati disponibili sui cellulari venduti online, ha creato un modello di *cluster* (Figura 5.3) che raggruppa i *cluster* ottenuti. Per evitare una dispersione di informazioni, è stato imposto al programma di raggruppare le opzioni di prodotti in non più di 8 *cluster*. La Figura 5.3 presenta la pagina di output del programma in cui si evidenzia il modello di *cluster* realizzato.

---

<sup>59</sup> Si ricorda che una diminuzione dei valori dei criteri di informazione di Akaike e di Schwarz segnala un miglioramento del modello, poiché la perdita in termini di potere esplicativo (modestissima in questo caso) è più che compensata dalla minore complessità del modello per la diminuzione delle variabili esplicative da 27 a 14. La più forte diminuzione dei valori del criterio di informazione di Schwarz rispetto a quello di Akaike riflette il fatto che il criterio di informazione di Schwarz penalizza maggiormente di quello di Schwarz gli aumenti di complessità del modello.

Figura 5.3 – La pagina di output contenuta in *visualizzatore modelli di data mining*, ottenuta avvalendosi dell’algoritmo di *clustering*.



Come evidenzia la Figura 5.3, il modello di *cluster* raggruppa i gruppi, nel nostro caso 8 *clusters*, che sono stati composti dal programma. Ad esempio, avendo selezionato il primo *cluster*, nella colonna a destra della Figura 5.3 compaiono i dati relativi ai cellulari che fanno parte del *cluster* considerato. In particolare, nel riquadro “Attributi”, selezionando un attributo, compaiono, i valori medi che quell’attributo ha all’interno del *cluster*, il numero di casi (prodotti) che compongono il *cluster* e la probabilità (la percentuale di casi che sono vicini ai valori indicati). Ad esempio, in questo caso è stata selezionata la variabile risoluzione della fotocamera, mostrando che di 81 casi (cellulari) che compongono il primo *cluster*, il 50,29% (41 casi) dei cellulari presenti nel primo *cluster* hanno una risoluzione media di circa 0,3 Mpixel (395.459 pixel) e il restante 49,71% (40 casi) non hanno la fotocamera. Invece nel riquadro più in basso “percorso del nodo” sono evidenziati i valori delle variabili che soddisfano il *cluster*. Dall’esempio raffigurato, si deduce che la gran parte dei cellulari che compongono il primo *cluster* non hanno l’Mp3, non hanno connettività senza fili (“=0”), o al massimo solo tecnologia infrarossi (“=33”), sono usciti sul mercato tanto tempo fa (31, 46, 58, 67 mesi fa), e tante altre informazioni.

Per poter esaminare più facilmente la composizione dei diversi *cluster*, le informazioni più importanti sono state raggruppate nella tabella 5.8, che presenta le caratteristiche utilizzate dal programma per effettuare il *clustering*, e quali valori le variabili assumono mediamente in ogni *cluster*.

La prima colonna sulla destra della tabella 5.8 evidenzia le caratteristiche utilizzate dal programma per valutare le dissimilarità tra i gruppi di cellulari. Il programma ha raggruppato i cellulari in 8 *cluster* ben definiti.

Di seguito, in base alla tabella 5.8, viene presentata una ragionevole interpretazione dei risultati, in modo da definire le similarità all’interno dei gruppi e le dissimilarità tra i gruppi.

Il primo *cluster* raggruppa 81 cellulari con le funzionalità essenziali per telefonare usciti sul mercato italiano intorno agli anni 2003 e 2004; essi hanno fotocamera di base, e un display di media qualità.

Il secondo *cluster* raggruppa invece 64 cellulari con funzionalità di base, che a differenza di quelli del primo *cluster*, sono usciti di recente sul mercato: essi sono più evoluti dei cellulari presenti nel primo *cluster*, hanno in più connettività mediante *bluetooth*, velocità di navigazione *Edge* e *Umts*, possono inviare *mail*, hanno memoria espandibile, sono in grado di riprodurre file musicali (mp3), hanno una buona risoluzione della fotocamera, etc. .

Il terzo *cluster* raggruppa 36 cellulari che sono una via di mezzo per quelli che vogliono un cellulare tecnologicamente evoluto ma senza esagerare; una via di mezzo tra un cellulare base ed uno *smartphone*. La maggior parte dei cellulari presenti in questo *cluster* sono usciti sul mercato nel 2007 e nel 2008 e sono molto più evoluti di quelli dei *cluster* precedenti, per la risoluzione della fotocamera, la velocità di navigazione (*Umts* o *Hsdpa*), e la presenza di videochiamata.

Nei cluster dal quarto al settimo rientrano “cellulari intelligenti” o *smartphone*; essi contengono il top della tecnologia disponibile in un cellulare, sono molto simili ad un mini pc, disponendo di un sistema operativo molto simile a quello dei personal computer. Il quarto *cluster* raggruppa difatti 29 cellulari, i più sottili e leggeri tra gli *smartphone* venduti online. I produttori principali di questi cellulari sono Samsung e Sony Ericsson; essi dispongono di memoria espandibile, videochiamata, un’ottima fotocamera, e una connessione a internet super veloce *Hsdpa*.

Il quinto *cluster* raggruppa invece 26 cellulari con una buona performance delle caratteristiche, ma quello che li differenzia soprattutto dagli altri *smartphone* è l’ampio schermo *touch screen*. Tra questi rientra il cellulare della Apple *I-phone*, dotato di un ampio display con una evoluta tecnologia *touch screen*, che i concorrenti stanno cercando, con molte difficoltà, di imitare. Il sesto *cluster* comprende 25 cellulari *smartphone*, simili a dei minipc tascabili; essi sono ultratecnologici, con una mini tastiera *qwerty*; ciò comporta un aumento rilevante delle dimensioni e del peso; principali produttori di questi cellulari sono Nokia e Htc. Nel settimo *cluster* rientrano 22 cellulari che si possono definire “cellulari di lusso”, perché hanno buone basi tecnologiche e prezzi molto elevati. L’ottavo *cluster* raggruppa i 16 cellulari economici, con una discreta base tecnologica, e diversi dagli altri perché gran parte hanno un *design* con sportellino.

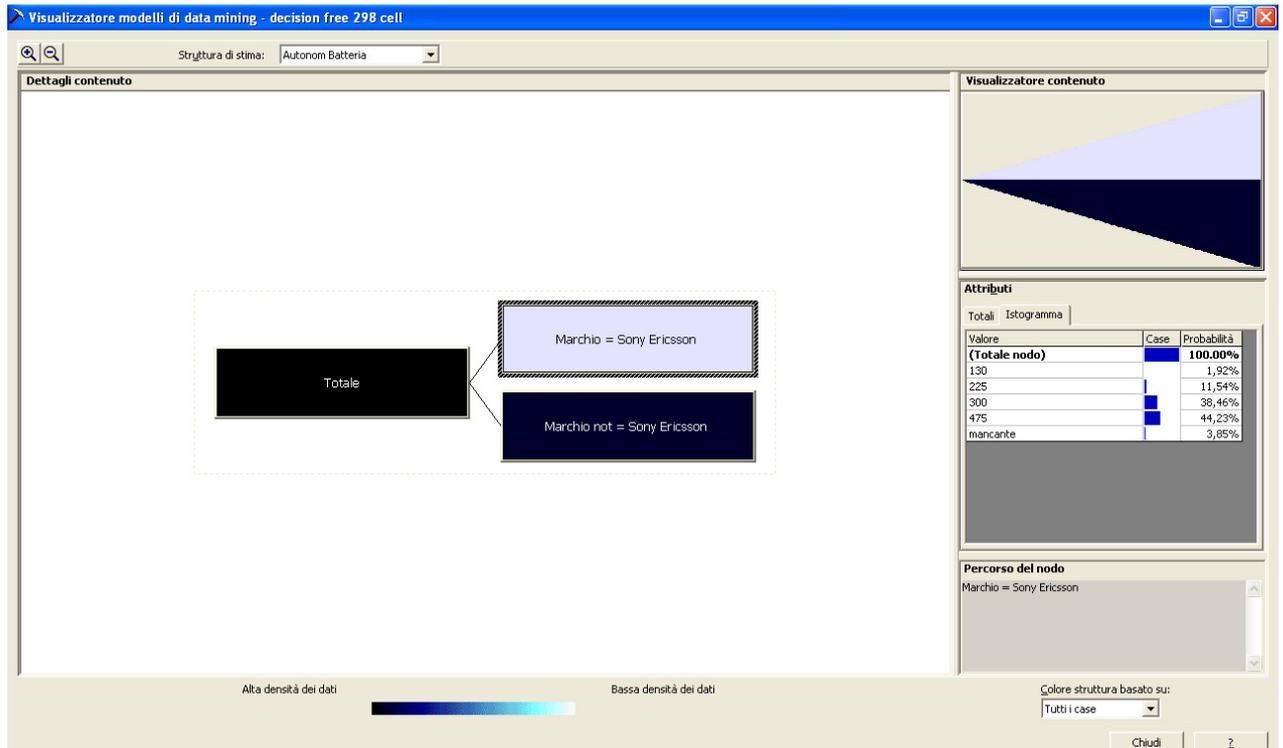
Tabella 5.8– Sintesi dei risultati ottenuti dalla *cluster analysis* su 298 cellulari venduti online.

	Cellulari vecchi modelli con tecnologia di base	Cellulari nuovi modelli con tecnologia di base	Cellulari nuovi modelli di media tecnologia	Cellulari di media tecnologia, Ultrasottili e leggerissimi	Smartphone con ampio touchscreen e discreta base tecnologica	Smartphone ultratecnologici con Gps e tastiera Qwerty	smartphone costosi ma leggeri, sottili e con discreta base tecnologica	Cellulari economici con discreta base tecnologica e design a sportellino
Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
<b>Case (num. opzioni)</b>	<b>81</b>	<b>64</b>	<b>36</b>	<b>30</b>	<b>26</b>	<b>25</b>	<b>22</b>	<b>16</b>
<b>Prezzo (in Euro)</b>	da 24 a 171	da 24 a 310	da 93 a 366		da 103 a 588		Prezzo > 972	35<= Prezzo <= 136
<b>Uscita</b>	2003 e 2004	fine 2006	Fine 2007 a 2008	2007 e 2008	2005 e 2006	Fine 2007 e 2008	2008	
<b>Marchio</b>		Samsung	Nokia	Samsung , Sony Ericsson	Apple, i-Mate, Hewlett Packard, HTC	Nokia, HTC	RIM , Sony Ericsson	Nokia
<b>Peso (in grammi)</b>	da 65 a 100	da 78 a 124	da 80 a 130	da 69 a 111	da 84 a 203	> 253	da 85 a 154	
<b>Spessore</b>				da 7 a 17		> 31	da 10 a 19	
<b>Dimensioni (in mm3)</b>		da 52015 a 110675		da 38786 a 88950		> 256647		da 63760 a 108850
<b>Sistema Operativo</b>	Assente	Assente		Assente	Presente	Presente	Presente	Assente
<b>Connettività</b>	Assente e infrarossi	Bluetooth	Bluetooth	Bluetooth	Wi-Fi	Wi-Fi	Wi-Fi	Infrarossi
<b>Mp3</b>	Assente	Presente	Presente					
<b>Risoluzione Fotocamera (in pixel)</b>	307200	da 974688 a 1920000	da 1920000 a 5038848	3145728	1920000	2073600 , 3981312	3145728	da 307200 a 1228800
<b>Flash</b>	Assente		Presente	Presente		Presente	Presente	
<b>Videocamera</b>	Assente	Presente	Presente	Presente		Presente		Assente
<b>Internet</b>	Assente e gprs	Edge, Umts	Umts, Hsdpa	Edge, Hsdpa	Edge	Hsdpa	Edge	Gprs, Edge ,
<b>Email</b>	Assente	Presente	Presente				Presente	
<b>Memoria espan</b>	Assente	Presente	Presente	Presente	Presente		Presente	Assente
<b>Risoluzione Display (in pixel)</b>	da 1100 a 39039	da 10292 a 98248	76800	da 34598 a 104106	da 21362 a 151969		da 50166 a 94677	da 4798 a 43192
<b>Colori Display (in num. colori)</b>	<= 261096	da 118219 a 359846		<= 2224818				<=420268
<b>Videochiamata</b>	Assente		Presente		Assente	Presente		Assente
<b>Design</b>		Slide		Slide	Bar	Slide	Bar e Flip	Sportellino
<b>Radio</b>			Presente		Assente		Presente	
<b>Touch screen</b>					Presente	Presente	Presente	
<b>Autonomia Batteria (in minuti di conversazione)</b>					da 104 a 390			
<b>Gps</b>						Presente		
<b>Tastiera Qwerty</b>						Presente	Presente	

### 5.2.3.2 Analisi di *decision trees*

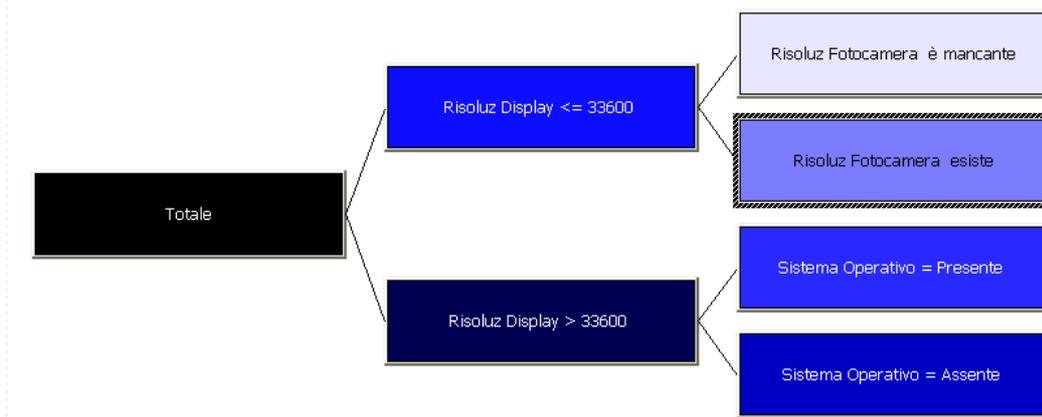
L'Algoritmo *Microsoft Decision Trees* di *Microsoft Analysis Services* offre la possibilità di creare alberi decisionali tenendo conto dei valori assunti dagli attributi per ogni opzione di prodotto considerato. La Figura 5.4 mostra una pagina di output ottenuta seguendo la procedura di questo algoritmo.

Figura 5.4 – La pagina di output contenuta in visualizzatore modelli di *data mining*, ottenuta avvalendosi dell'algoritmo *Microsoft Decision Trees* e prendendo in considerazione l'attributo *autonomia della batteria*.



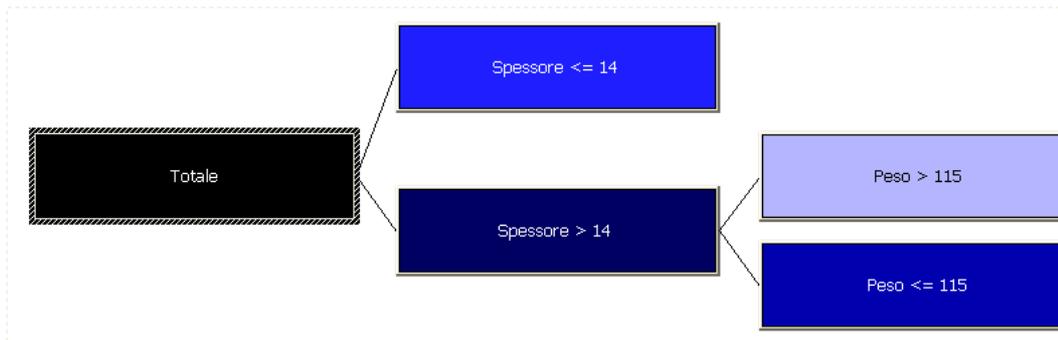
La Figura 5.4 mostra i risultati dell'albero decisionale ottenuto considerando l'*autonomia della batteria*. In base a questa selezione l'applicazione ha individuato come principale discriminante il produttore del cellulare. Nel riquadro a destra compaiono nel riquadro *attributi* delle fasce di durata della batteria (130, 225, 300, 475 ore), e selezionando il nodo *marchio=Sony Ericsson*, compaiono le percentuali di *case* aventi marchio Sony Ericsson per le diverse fasce di autonomia della batteria. Si può infatti osservare che più dell'80% dei telefonini aventi marchio Sony Ericsson hanno un'autonomia altissima, oltre le 300 ore, quindi si può affermare che solitamente essi hanno la migliore autonomia tra tutti i cellulari considerati. Nelle figure successive per una comodità di lettura, vengono riportati solo gli alberi decisionali più significativi.

Figura 5.5 – albero decisionale ottenuto prendendo in considerazione l'attributo *connettività*



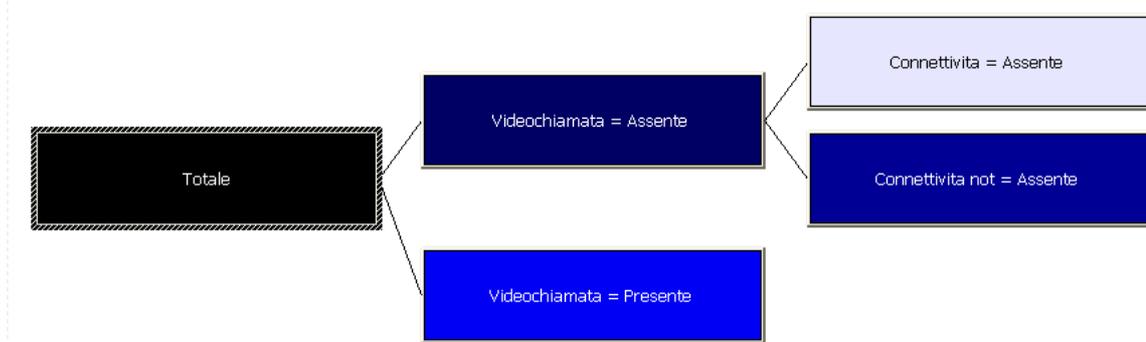
Un albero decisionale ottenuto partendo dalla connettività senza fili considera la relazione esistente tra Sistema Operativo e Risoluzione del display. Secondo la seguente stratificazione: in presenza di sistema operativo, spesso il cellulare presenta un ampio display, in particolare, una risoluzione maggiore di 33600 pixels, e una buona connessione senza fili, in particolar modo Wi-Fi.

Figura 5.6 – albero decisionale ottenuto prendendo in considerazione l'attributo *dimensioni*



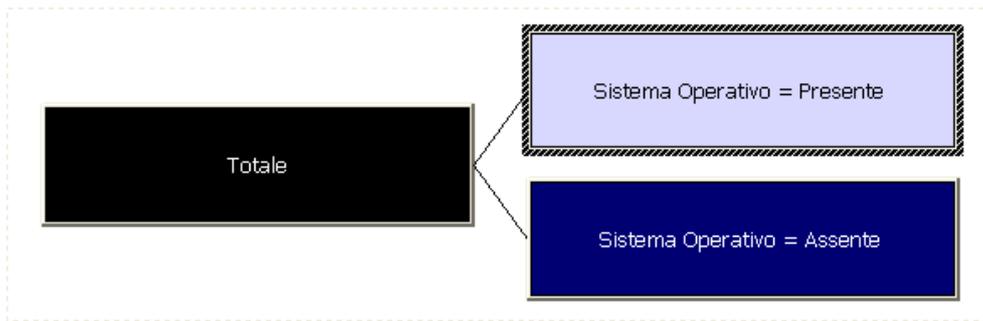
Partendo invece dalle dimensioni del cellulare, secondo la seguente disposizione dei nodi, spesso quando il peso è elevato, lo spessore ha un valore al di sopra della media, ed allo stesso modo le dimensioni

Figura 5.7 – albero decisionale ottenuto prendendo in considerazione l'attributo *internet*



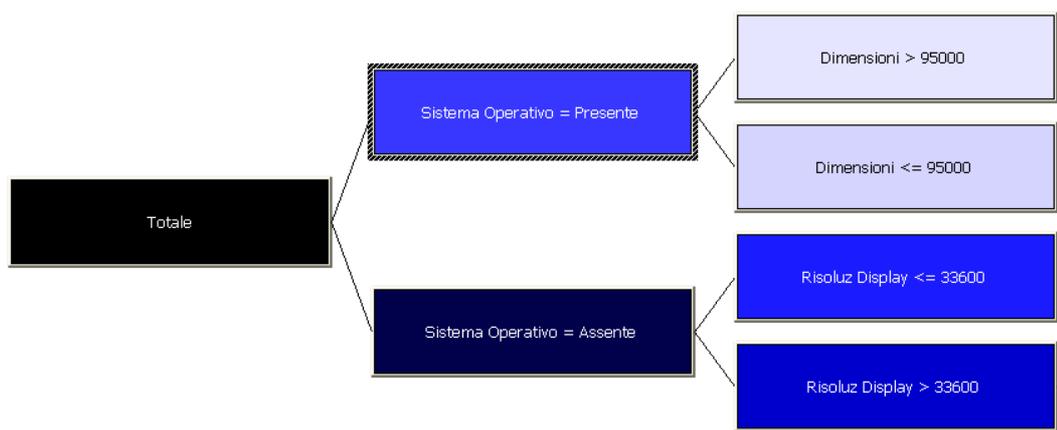
Il precedente albero decisionale considera la velocità di scambio dei dati, l'informazione più importante è che quando il cellulare presenta la videochiamata, si è in presenza di una veloce tecnologia nello scambio dei dati; ciò è dovuto al fatto che la videochiamata ha bisogno di almeno una connessione Umts per poter funzionare.

Figura 5.8 – albero decisionale ottenuto prendendo in considerazione l'attributo *marchio*



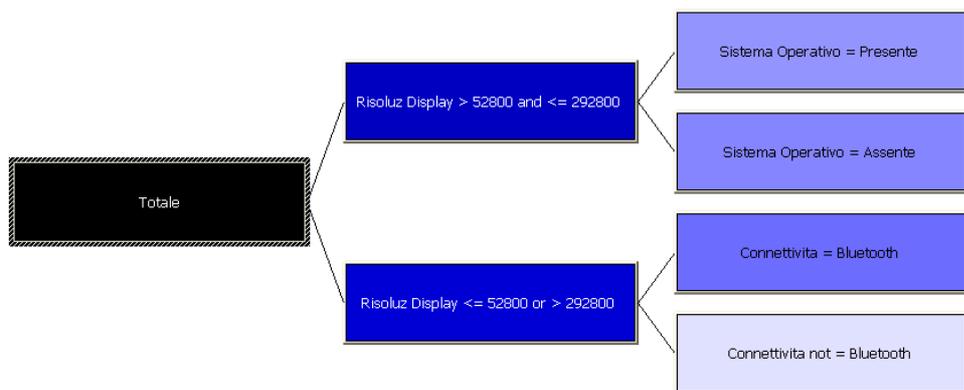
Il precedente albero decisionale, ottenuto partendo dal marchio, evidenzia come il sistema operativo sia molto legato al produttore del cellulare. In particolar modo la Apple con il suo sistema operativo OS X, Gli smartphone Htc, e I-Mate dispongono del sistema operativo di casa Microsoft (windows mobile), gli smartphone della Nokia, hanno integrato il sistema operativo Symbian OS, mentre gli smartphone della Samsung hanno alcuni Windows Mobile ed alcuni il sistema Symbian OS.

Figura 5.9 – albero decisionale ottenuto prendendo in considerazione l'attributo *peso*



L'albero decisionale ottenuto prendendo in considerazione l'attributo peso, definisce come 81 cellulari con sistema operativo su 217 cellulari hanno in gran parte maggior peso e dimensioni mentre i cellulari che non sono *smartphone* sono più leggeri e con una risoluzione del display maggiore.

Figura 5.10 – albero decisionale ottenuto prendendo in considerazione l'attributo *prezzo*



L'albero decisionale ottenuto partendo dalle stime ottenibili considerando la variabile prezzo, considera i cellulari *smartphone*, quindi dotati di sistema operativo, i cellulari con

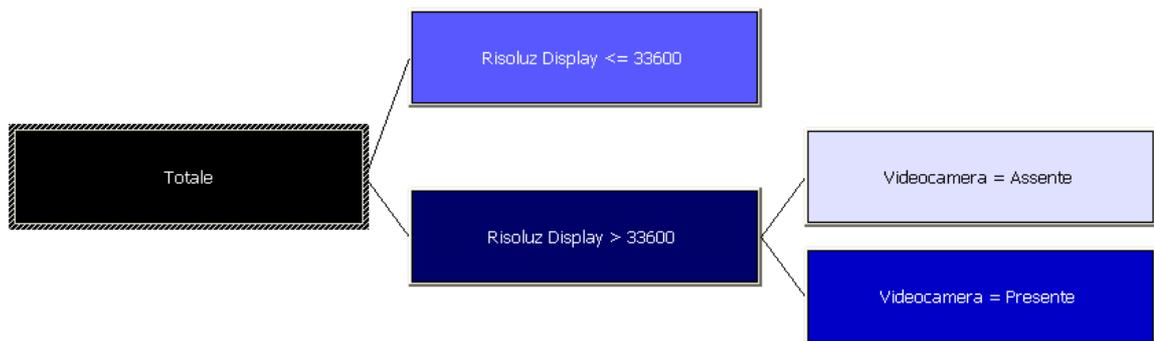
maggiore peso, una maggiore dimensione del display e quindi un prezzo maggiore; per i cellulari “non *smartphone*”, invece, la caratteristica maggiormente legata alle variazioni di prezzo è la presenza o meno della connettività *Bluetooth*.

Figura 5.11 – albero decisionale ottenuto prendendo in considerazione l’attributo *prezzo*



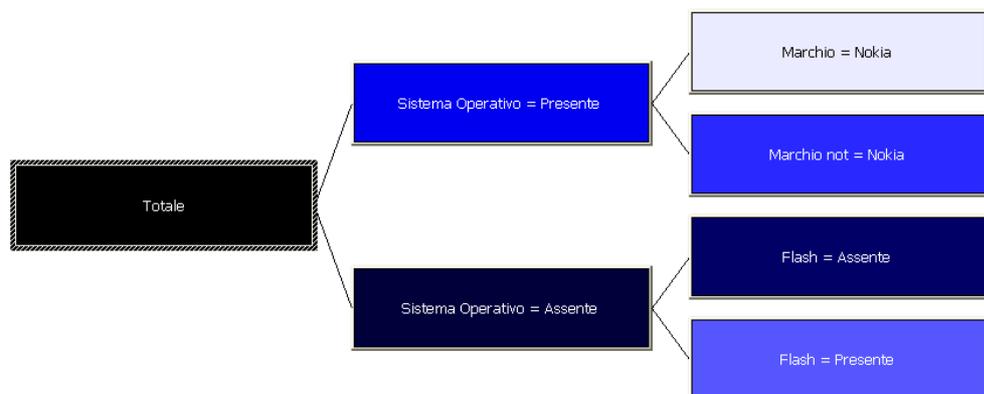
L’albero decisionale ottenuto considerando la caratteristica marchio evidenzia come su 124 cellulari che hanno la radio, 52 cellulari sono Nokia mentre 29 sono Sony Ericsson.

Figura 5.12 – albero decisionale ottenuto prendendo in considerazione l’attributo *risoluzione fotocamera*



L’albero decisionale ottenuto partendo dalla risoluzione della fotocamera mostra che i cellulari con videocamera hanno una migliore risoluzione del display e una migliore risoluzione della fotocamera.

Figura 5.13 – albero decisionale ottenuto prendendo in considerazione l’attributo *touch screen*



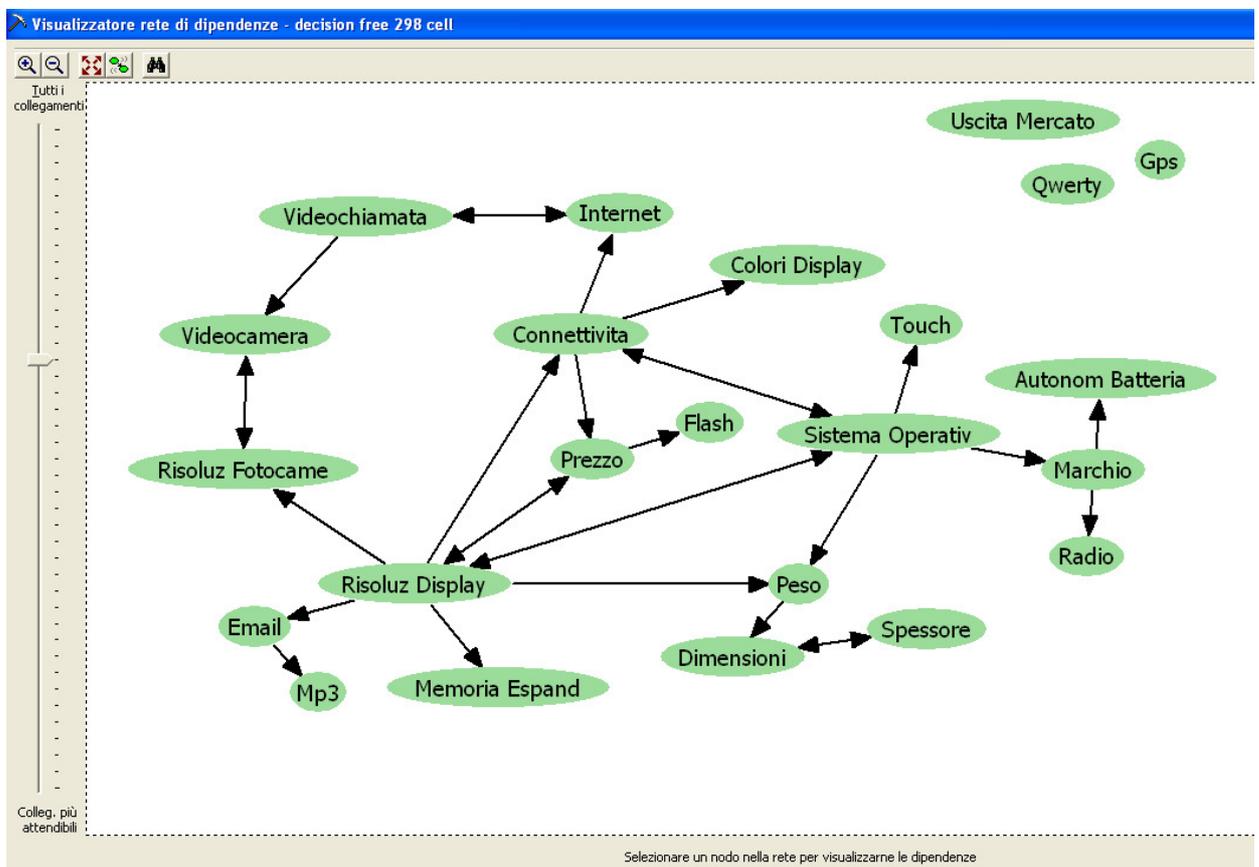
La principale informazione ottenibile osservando il precedente albero decisionale basato sulle stime relative alla presenza o meno dello schermo *touch screen* evidenzia come 37 cellulari su 43 che hanno lo schermo *touch screen* hanno anche il sistema operativo e solo uno di questi è

marchiato Nokia. La Nokia, infatti, pur essendo il principale produttore di telefonini al mondo, rimane indietro nell'ingresso dei telefonini con lo schermo tattile, sempre più richiesti dall'utenza.

### 5.2.3.3 La rete delle dipendenze

Mentre il *visualizzatore modelli di data mining* consente di visualizzare le informazioni sulla distribuzione e le relazioni dal punto di vista di un singolo attributo, il *visualizzatore rete di dipendenze* offre la possibilità di visualizzare il modello di *data mining* dal punto di vista di tutti gli attributi utilizzando le sole informazioni sulle relazioni esistenti tra gli attributi, in modo da presentare una prospettiva più ampia dell'intero modello di *data mining*. Per esempio, l'attributo fotocamera consente di stimare la memoria totale, la predisposizione per memoria aggiuntive (memoria espandibile), la risoluzione del display, la videocamera, e il tipo di connettività senza fili.

Figura 5.14 - Visualizzazione delle dipendenze tra gli attributi utilizzati per spiegare i dati di 298 cellulari venduti online sul sito Kelkoo.it a luglio del 2008



La figura 5.14 fornisce significative informazioni aggiuntive; il *visualizzatore rete di dipendenze*, mostra gli attributi identificandoli come nodi; le frecce indicano la direzione della stima. Per esempio, la relazione esistente tra prezzo e flash sta a significare che il prezzo consente di stimare l'attributo flash; allo stesso modo, la risoluzione e il prezzo secondo la rete di dipendenze possono stimarsi reciprocamente. tra le tante relazioni, si può sottolineare come la risoluzione del display consente di stimare il peso del cellulare ( display più grande ha bisogno di più energia quindi maggiore autonomia della batteria e quindi maggior peso); il peso consente poi di stimare le dimensioni abbastanza correlate con lo spessore. Il marchio consente di stimare l'autonomia della batteria e la presenza o meno nel cellulare della radio integrata; secondo le informazioni disponibili i cellulari Sony Ericsson presentano le batterie più durature, mentre gran

parte dei cellulari Nokia dispongono di radio integrata. Per quanto riguarda il nodo sistema operativo, esso è correlato con la funzione *touch screen*, perché quasi tutti i cellulari con lo schermo tattile hanno un sistema operativo; esso è anche correlato con il peso, perché i cellulari con sistema operativo richiedono una maggiore autonomia della batteria, e con il marchio perché solo pochi produttori come Htc, Hp, Samsung, e Nokia, hanno presentato sul mercato cellulari *smartphone*. La risoluzione del display consente di stimare la risoluzione della fotocamera, proprio perché solitamente un cellulare con un grande display consente di visualizzare meglio le foto e, in particolare quelle realizzate con una buona fotocamera integrata. Altra relazione da sottolineare è quella tra funzione di videochiamata e velocità di scambio dati (Internet), questo perché non si può effettuare una videochiamata se il telefono non dispone almeno di una velocità di scambio dati di tipo Umts.

### **5.3 Analisi sui cellulari venduti in Italia**

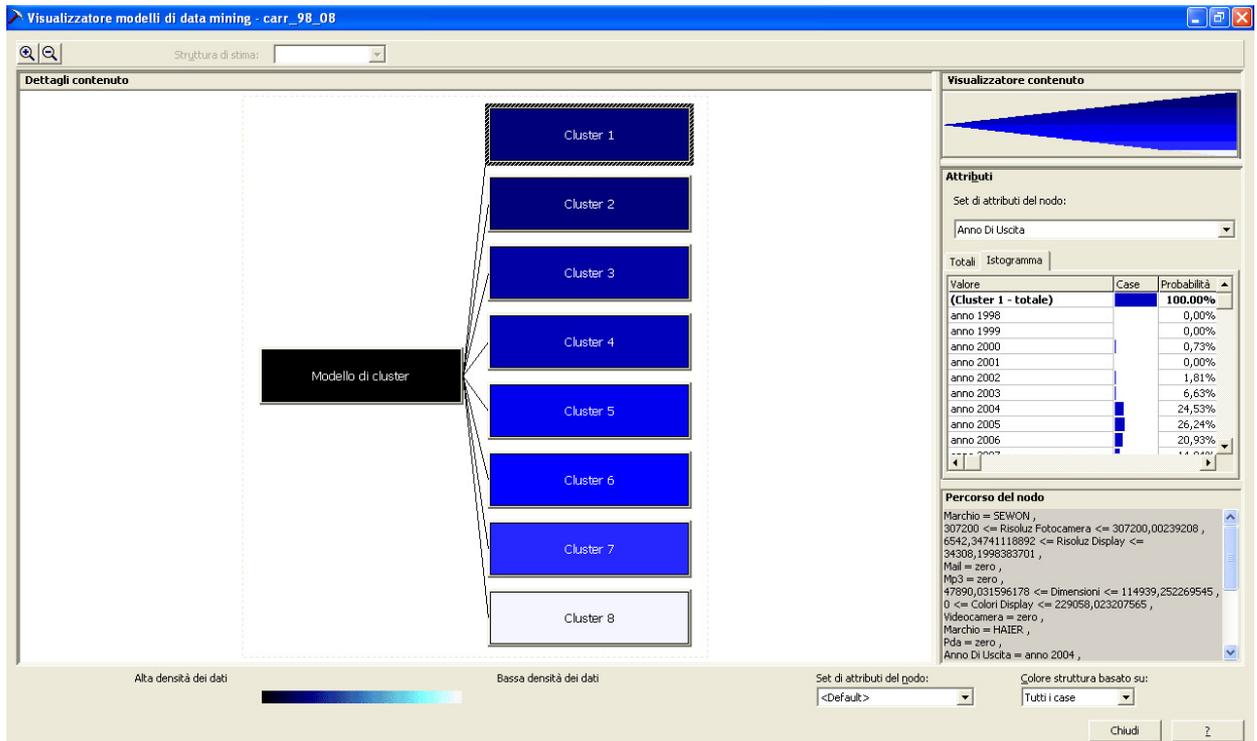
Moltissimi sono i cellulari usciti sul mercato nell'ultimo decennio; molti dei quali hanno avuto enorme successo, per le novità introdotte, e quindi riprodotti sul mercato con nuove edizioni aggiornate tecnologicamente, e molti sono stati dei veri e propri flop, che non sono durati nemmeno una stagione. Gran parte di questi cellulari, per la precisione 2120 presenti sul mercato italiano dal 1998 al 2008, sono stati esaminati, tenendo conto di 28 caratteristiche in grado di rappresentarli dettagliatamente (da un semplice calcolo si pensi che sono stati raccolti 59.360 dati, vale a dire, *2120 cellulari x 28 caratteristiche*). Si presentano i risultati ottenuti attraverso un'analisi di *data mining* con algoritmi di *Cluster analysis* e *Decision tree analysis*.

#### **5.3.1 Analisi di *data mining***

##### **5.3.1.1 Analisi di *clustering***

Disponendo dei dati relativi alla gran parte di cellulari presentati sul mercato italiano nell'ultimo decennio (1998-2008), attraverso l'applicativo *microsoft analysis service* si è pensato di effettuare un'analisi *Clustering*, utile in particolare per valutare la composizione dei gruppi definiti dall'applicativo. I risultati della *cluster analysis* sono presentati nella Figura 5.15.

Figura 5.15 – Risultati della *cluster analysis* su 2120 cellulari presenti sul mercato italiano nel periodo 1998-2008



Il programma *Microsoft analysis service* ha raggruppato i cellulari in 10 *cluster*. La figura 5.15 mostra come selezionando il primo *cluster* è possibile osservare in basso a destra la performance delle caratteristiche dei cellulari del primo *cluster*, e spostandoci più in alto, viene evidenziato il set di attributi utilizzato nella stima, (nel nostro caso sono state utilizzate 30 caratteristiche). Selezionando un attributo viene evidenziato, sempre per il primo *cluster*, la distribuzione delle performance di ogni variabile; per esempio si nota facilmente che il *cluster 1* presenta un gruppo di 415 cellulari che, secondo la caratteristica “anno di uscita”, include cellulari introdotti nel periodo 2003-2005.

La tabella 5.9 presenta una sintesi dei risultati ottenuti mediante la *cluster analysis* su 2120 cellulari presenti sul mercato nell’intervallo di tempo tra il 1998 ed il 2008

Tabella 5.9 – Sintesi dei risultati ottenuti mediante la *cluster analysis* sui 2120 cellulari presenti sul mercato italiano nel periodo fra il 1998 ed il 2008

	Cellulari dal 2004 al 2005 con tecnologia di base	Cellulari dal 1998 al 2003 con tecnologia di base	Cellulari dal 2006 al 2007 con una buona base tecnologica	Cellulari del 2005 con una buona base tecnologica	cellulari con grande display e non molto recenti (2004-2005)	cellulari con la più ampia memoria interna e risoluz. del display	smartphone con connettività Wi-Fi	I primi smartphone, sul mercato, quindi di grandi dimensioni e discreta base tecnologica
Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
Case (num. opzioni)	545	469	302	278	175	137	129	64
Uscita	2004-2005	1998-2003	2006-2007	2005	2004-2005	2007-2008	2006-2007	2001-2003
Marchio			Lg	Samsung	Gigabyte, Qtek, I-Mate	Nokia	Asus, Hp, Htc, Rim	Htc, Rim, Gigabyte, Eten, Qtek
Spessore (in millimetri)	media di 19	media di 21	media di 17	media di 20	media di 22	media di 15	media di 15	media di 21
Dimensioni (in millimetri <sup>3</sup> )	media di 81463	media di 93048	media di 82475	media di 88976	media di 125545	media di 78228	media di 96204	media di 148005
Sistema Operativo	Assente	Assente	Assente	Assente	presente per metà	presente per metà	Presente	Presente
Connettività	Assente	Assente	Bluetooth	Bluetooth	Bluetooth	Bluetooth	Wi-Fi	Bluetooth
Mp3	Assente	Assente	Presente	Presente	Presente	Presente	Presente	
Risoluzione Fotocamera (in pixel)	307200	Assente	in media 1,7 milioni	in media 1,2 milioni	in media 1,3 milioni	in media 3 milioni	in media 1,5 milioni	in media 2 milioni
Flash								
Videocamera	Presente	Assente	Presente	Presente	Presente	Presente	Presente	Presente
Internet						Umts	Umts/Hsdpa	
Email	Assente	Assente	Presente	Presente	Presente	Presente	Presente	Presente
Memoria totale (in Mb)	media di 19	media di 4	media di 42	media di 28	media di 104	media di 850	media di 430	media di 86
Memoria espandibile	Assente	Assente	Presente	Assente	Presente	Presente	Presente	Presente
Risoluzione Display (in pixel)	18857	14184	52636	31934	93432	76800	68525	56670
Colori Display (in num. colori)	63452	2433	211003	111803	91327	5954529	79331	2015639
Videochiamata	Assente	Assente	presente per metà	Assente		presente	presente per metà	Assente
Radio			presente per un terzo			presente		
Autonomia Batteria (in minuti di conversazione)	236	226	243	247	210	294	285	200
Gps						in parte Gps	in parte Gps	

Il primo *cluster* raggruppa 545 cellulari con le funzionalità essenziali per telefonare usciti sul mercato italiano intorno agli anni 2004 e 2005; si tratta in genere di cellulari che possono inviare messaggi MMS, hanno una fotocamera di base, un display di media qualità, e una discreta autonomia della batteria.

Il secondo *cluster* raggruppa 469 cellulari con funzionalità di base che, a differenza di quelli del primo *cluster*, sono apparsi sul mercato fra il 1998 e il 2003; si tratta di cellulari tecnologicamente meno evoluti, privi di fotocamera e di videocamera, che hanno il peggior display rispetto all'insieme di cellulari considerato (display di piccole dimensioni e scarsa qualità delle immagini visualizzabili, risoluzione del display bassa, numero di colori del display basso).

Il terzo *cluster* comprende cellulari che, usciti sul mercato, fra il 2006 e il 2007, presentano una discreta base tecnologica; essi hanno il *Bluetooth*, la funzione Mp3, una buona risoluzione della fotocamera, la videocamera, possono inviare email, una discreta memoria anche espandibile; circa la metà di essi hanno funzioni di videochiamata e un terzo dispone della radio.

Il quarto *cluster* raggruppa 278 cellulari con una discreta base tecnologica rispetto agli anni di uscita sul mercato (gran parte di questi sono usciti nel 2005). Basta confrontarli con i cellulari presenti nel primo cluster per accorgersi che questi cellulari hanno in più il *Bluetooth*, la funzione Mp3, una migliore risoluzione della fotocamera, la possibilità di inviare mail e un *display* più grande e di qualità migliore).

Gli ultimi 4 *cluster* sono composti dai cellulari più evoluti tecnologicamente, la maggior parte di essi ha infatti un sistema operativo (cosiddetti *smartphone*). Il quinto *cluster* presenta 175 cellulari; metà di essi sono dotati di sistema operativo, quindi potrebbero definirsi anche *smartphone*. La maggior parte di questi cellulari hanno comunque il display più grande tra i 2100 cellulari considerati, e sono conseguentemente ingombranti per via delle dimensioni e dello spessore. Usciti sul mercato nel periodo 2004-2005, presentano connettività *Bluetooth*, funzione Mp3, discreta risoluzione della fotocamera. Il sesto *cluster* comprende cellulari tecnologicamente alquanto evoluti, che presentano un'ottima memoria interna anche espandibile e un'ottima risoluzione del display, che consente la visualizzazione ottimale delle foto; la maggior parte di questi cellulari sono dotati di ottime fotocamere, che presentano le risoluzioni tra le migliori sul mercato; essi sono inoltre dotati di tecnologia Umts, molto utile per navigare velocemente online. Il settimo *cluster* raggruppa cellulari dotati tutti di sistema operativo, quindi *smartphone* che tecnologicamente sono molto simili ai cellulari del cluster precedente, pur essendo meno recenti (usciti sul mercato nell'anno 2006-2007), ma in più dispongono di connettività Wi-Fi, la migliore disponibile sul mercato. Questo *cluster* include cellulari creati appositamente per applicazioni *office mobile*; gran parte di questi cellulari sono marchiati Asus, Hp, Htc e Rim che producono esclusivamente cellulari *smartphone*.

L'ultimo *cluster* costituito da 62 cellulari, include i primi *smartphone* creati da Htc, Rim, Gigabyte, Eten, Qtek nel periodo 2001-2003, sono quindi meno evoluti tecnologicamente rispetto a quelli del *cluster* precedente e hanno dimensioni piuttosto grandi.

La Tabella 4.14 presenta un confronto tra due diverse *cluster analysis* effettuate sui dati relativi ai telefoni cellulari, una relativa a 298 cellulari venduti online a luglio del 2008, l'altra relativa a 2101 cellulari venduti sul mercato italiano nel periodo dal 1998 al 2008.

Sostanziali differenze tra le due analisi in *cluster* si notano in particolare negli ultimi *clusters* che in entrambe le analisi presentano i cellulari più evoluti tecnologicamente. La *cluster* sui 298 cellulari ha definito ben 3 gruppi di *smartphone* (*cluster* 5, 6 e 7), mentre nella *cluster* con 2101 cellulari sono definiti in particolar modo solo nel 7° e nell'8° *cluster*. Ciò si può spiegare in base al gran numero di *smartphone* usciti sul mercato negli ultimi mesi, mentre la *cluster* sui 2101 cellulari considera un intervallo di tempo molto più ampio, e quindi in questo caso gli *smartphone* si trovano in una percentuale molto più piccola rispetto alla prima *cluster analysis*. Inoltre, confrontando le 2 *cluster* si nota come, pur essendoci oggi una fortissima evoluzione tecnologica delle funzioni disponibili in un cellulare, ci sono comunque ancora molti cellulari con performance simili a quelle presenti nei cellulari usciti sul mercato nei primi anni del 2000. Questo perché sono molti i consumatori, che, ancora oggi, sono interessati ad acquistare un telefonino che abbia il minimo di funzioni necessarie per telefonare e non sono attirati da altri servizi. Ovviamente, a differenza di quelli usciti sul mercato prima del 2000, i nuovi modelli sono comunque più leggeri e meno ingombranti, hanno una migliore autonomia della batteria, e costano molto meno.

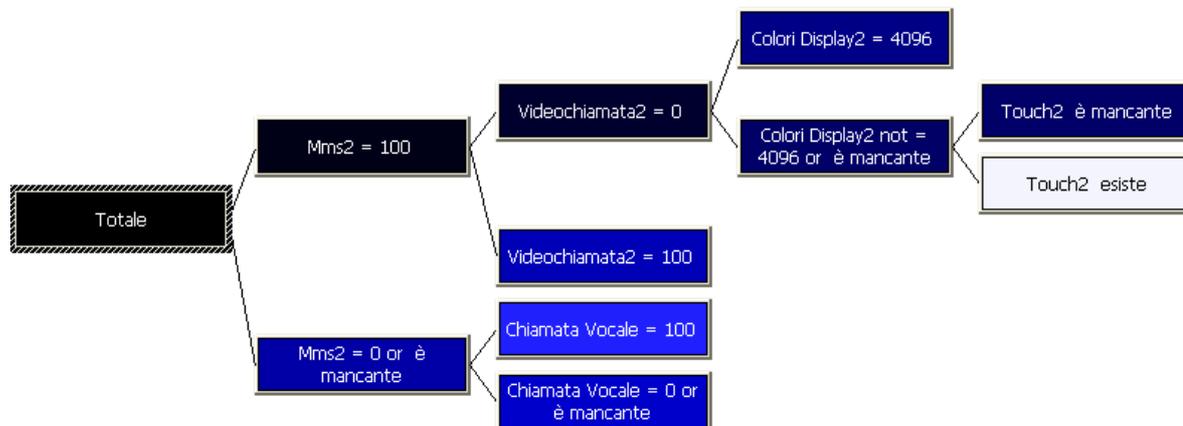
Tabella 5.10 Confronto tra la *cluster analysis* effettuata sui 298 cellulari disponibili a luglio del 2008 sul sito Kelkoo.i e la *cluster analysis* effettuata sui 2120 telefoni cellulari presenti sul mercato italiano dal 1998 ad oggi.

Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
<b>Cluster Analysis con i 298 cellulari venduti online a luglio del 2008</b>	Cellulari vecchi modelli con tecnologia di base	Cellulari nuovi modelli con tecnologia di base	Cellulari nuovi modelli di media tecnologia	Cellulari di media tecnologia, Ultrasottili e leggerissimi	Smartphone con ampio touchscreen e discreta base tecnologica	Smartphone ultratecnologici con Gps e tastiera Qwerty	smartphone costosi ma leggeri, sottili e con discreta base tecnologica	Cellulari economici con discreta base tecnologica e design a sportellino
<b>Cluster Analysis con i 2120 cellulari venduti sul mercato italiano dal 1998 al 2008</b>	Cellulari dal 2004 al 2005 con tecnologia di base	Cellulari dal 1998 al 2003 con tecnologia di base	Cellulari dal 2006 al 2007 con una buona base tecnologica	Cellulari del 2005 con una buona base tecnologica	cellulari con grande display e non molto recenti (2004-2005)	cellulari con la più ampia memoria interna e risoluz. del display	smartphone con connettività Wi-Fi	I primi smartphone, sul mercato, quindi di grandi dimensioni e discreta base tecnologica

### 5.3.1.2 Analisi di *decision trees*

La figura 5.16 mostra la rappresentazione dei dati relativi alle performance delle caratteristiche di ogni cellulare venduto sul mercato italiano nell'ultimo decennio attraverso un albero decisionale stimato partendo dalla caratteristica "anno di uscita sul mercato".

Figura 5.16 - Albero decisionale ottenuto prendendo in considerazione l'attributo *uscita sul mercato*



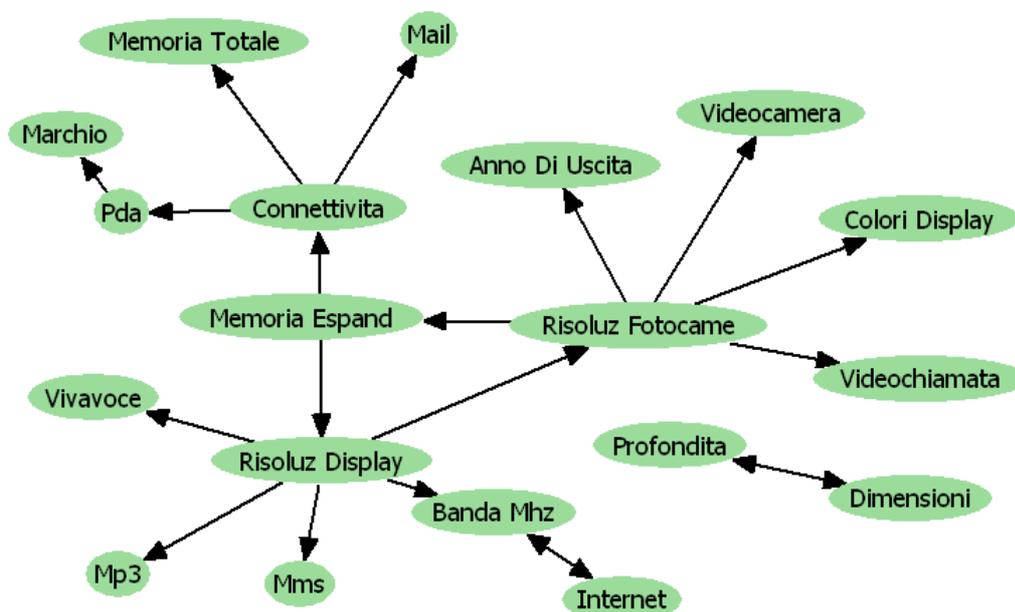
L'albero decisionale mostrato nella Figura 5.16 presenta al primo livello di nodi l'attributo Mms, questa caratteristica dal 2002 in poi è stata aggiunta progressivamente a tutti i cellulari venduti sul mercato. Allo stesso modo, dal 2004 in poi, sono progressivamente aumentati i cellulari con la funzione videochiamata. Da ciò si deduce che, considerando l'anno di uscita sul mercato, le caratteristiche che hanno avuto uno sviluppo più rapido nel tempo sono la funzione MMS e la funzione videochiamata.

### 5.3.1.3 La rete delle dipendenze

La figura 5.17 evidenzia che le relazioni tra le variabili sono diverse da quelle ottenute prendendo in esame solo uno spicchio della popolazione, e cioè, i 298 cellulari venduti online sul sito kelkoo.it a luglio del 2008. Solo in particolari casi, come per la relazione esistente tra risoluzione del display che riesce a stimare la risoluzione della fotocamera, che a sua volta riesce

a stimare la presenza o meno della fotocamera, oppure il sistema operativo (chiamato in questo caso Pda) che riesce a stimare il produttore del cellulare.

Figura 5.17 - Visualizzazione delle dipendenze tra gli attributi utilizzati per spiegare i dati di 2120 cellulari presenti sul mercato italiano dal 1998 al 2008.



### 5.3.2 La penetrazione sul mercato di diverse caratteristiche dei telefoni cellulari

Le figure 5.18 e 5.19 sono utili per comprendere come si sono evolute nell'ultimo decennio le diverse caratteristiche che definiscono un telefono cellulare. Sono state considerate in particolare 21 caratteristiche, suddivise in 2 grafici al fine di rendere più facile la lettura che altrimenti risulterebbe molto più difficile.

La figura 5.18 mostra come siano migliorate nel corso del periodo 1998-2008 le principali caratteristiche che maggiormente incidono sul prezzo finale del telefonino sulla base dei risultati ottenuti dalle regressioni presentate nei precedenti capitoli. È stato posto pari a 100% il picco massimo raggiunto dalla caratteristica (la più alta penetrazione ottenuta sul mercato). Ad esempio le caratteristiche Gps, videochiamata, quantità di colori del display, memoria totale e *touch screen* hanno avuto una forte accelerazione negli ultimi 3 anni. In particolar modo per la funzione *touch screen*, quasi sconosciuta agli inizi del 2007, è ormai disponibile in quasi tutti i cellulari con ampi display. C'è stato comunque un forte aumento della penetrazione sul mercato dei telefonini anche per le caratteristiche memoria totale, memoria espandibile, e tecnologia Umts per la connessione a Internet. Quest'ultima ha avuto addirittura una diminuzione di presenza nell'ultimo anno, forse perché stanno sempre più ritornando numerosi sul mercato i cellulari che consentono esclusivamente di telefonare. Invece la crescita della penetrazione sul mercato dei cellulari con alte performance per alcune caratteristiche come connettività senza fili e risoluzione della fotocamera è stata meno rapida rispetto alla crescita della penetrazione delle caratteristiche analizzate precedentemente. La caratteristica autonomia della batteria è invece la caratteristica che è cresciuta di meno nell'ultimo decennio, e in particolare negli anni duemila. Ciò si spiega col fatto che, pur essendo le batterie indispensabili per far funzionare un telefonino tecnologicamente avanzato, poiché con l'aggiunta di nuove funzioni sul cellulare occorre sempre più energia, la ricerca nel settore dello sviluppo dell'autonomia delle batterie riesce a stento a stare al passo con le richieste.

Figura 5.18 – Analisi della penetrazione sul mercato fra il 1998 e il 2008 delle prime 11 fra le principali caratteristiche di un telefono cellulare

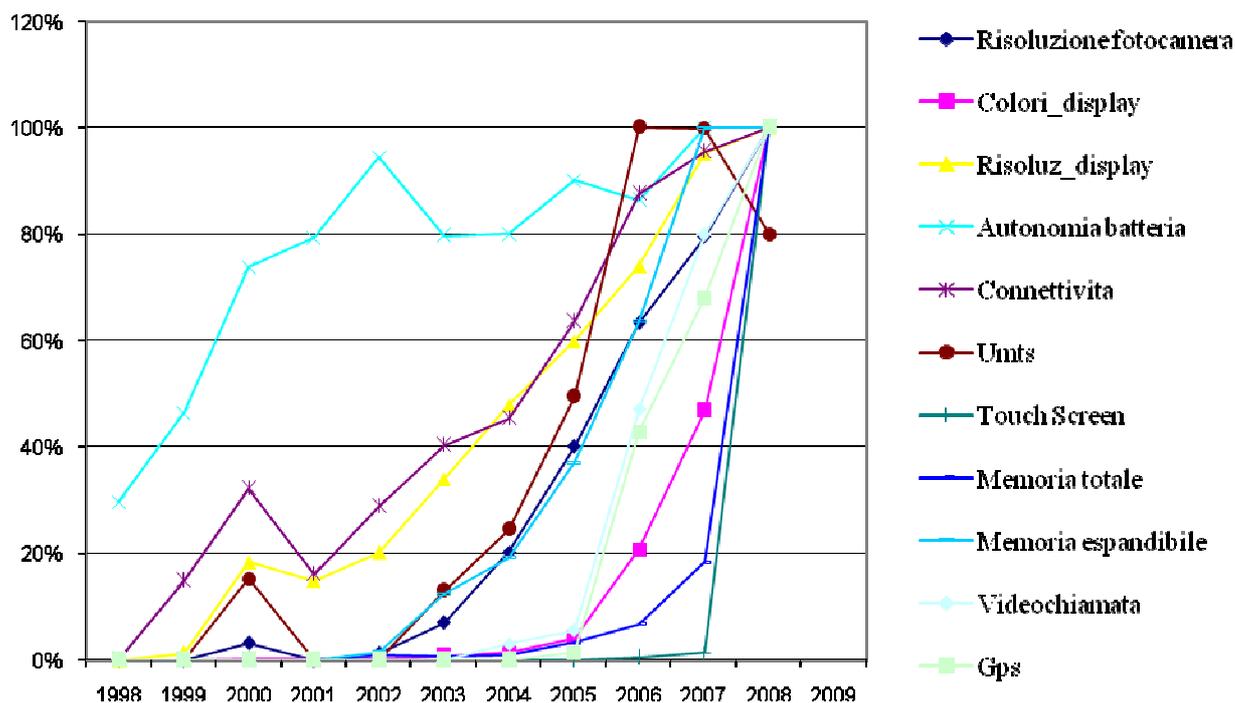
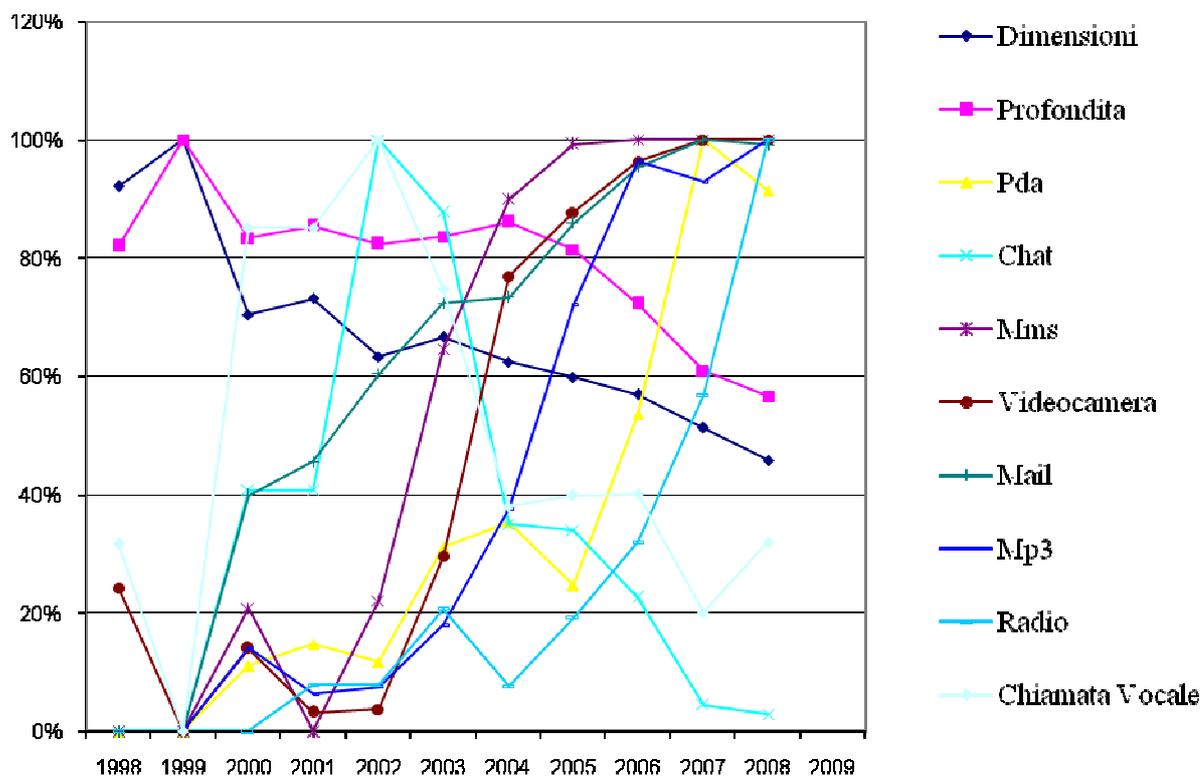


Figura 5.19 – Analisi della penetrazione sul mercato fra il 1998 e il 2008 delle altre 10 principali caratteristiche di un telefono cellulare



La figura 5.19 mostra le restanti 10 principali caratteristiche dei telefoni cellulari considerati. Le caratteristiche radio e funzione di Pda (*personal digital assistant*), sono cresciute rapidamente negli ultimi anni. Altre caratteristiche quali ad esempio Mms, Mp3, Mail e videocamera, hanno avuto una forte crescita dal 2002 al 2005, e poi si sono stabilizzate, anche

perché oramai si trovano in quasi tutti i cellulari. Alcune caratteristiche come *Chat* e chiamata vocale hanno avuto una forte crescita della penetrazione sul mercato fino al 2002, seguita da una ripida discesa non ancora conclusasi perché sempre meno richieste dai consumatori. Le variabili continue dimensioni e profondità hanno avuto invece un andamento completamente opposto, dal 1998 ad oggi hanno avuto una lenta e graduale discesa, ciò sta a significare che nel complesso i telefonini pur essendo tecnologicamente sempre più avanzati, stanno diventando sempre più piccoli.

#### **5.4 Analisi delle opinioni degli utenti del sito *Kelkoo.it***

In questo paragrafo vengono presentati i risultati di un'indagine condotta online tenendo conto delle opinioni lasciate dagli utenti del sito *Kelkoo.it* sui cellulari presentati sul sito. Obiettivo dell'indagine è stato quello di scoprire, partendo dalle opinioni rilasciate sul sito, quali siano state le caratteristiche dei cellulari valutati che non hanno corrisposto alle aspettative degli utenti, quali siano stati i cellulari che hanno ottenuto un maggior numero di opinioni, e quali tra questi quelli maggiormente consigliati.

L'importanza delle opinioni delle persone nelle decisioni di acquisto è da sempre conosciuta da chi si occupa di Marketing. Risulta però molto interessante che questo fenomeno si manifesti oggi anche in rete e con le stesse caratteristiche di influenza sul comportamento dei consumatori. Questo perché gli effetti e le influenze di questo nuovo metodo di passa parola cadranno su milioni di persone, quindi non avviene più localmente ma globalmente, non riguarderà più un mercato e un'offerta limitata ma globale. Se un prodotto venduto in gran parte del mondo delude un consumatore americano, potrà rendere dubbioso nell'acquisto anche un consumatore tedesco. La maggior parte dei motori di ricerca che si occupano di *price comparison* effettuano la comparazione dei prodotti, non solo in base al prezzo, ma soprattutto in base alle opinioni lasciate sul sito da utenti/consumatori e utenti/esperti; la stessa *Ebay.it* ha investito molto in questo campo. Per l'indagine è stato scelto il comparatore di prezzi *Kelkoo*, perché è uno dei principali motori di ricerca europei dedicato alla comparazione dei prezzi dei prodotti venduti online. *Kelkoo* è una società che dal 2004 fa parte del gruppo *Yahoo*, e la sua vocazione, secondo quanto si legge sul suo sito, è: *Rendere sempre più semplice, più rapido e più attrattivo lo shopping online*.

In genere gli utenti, avendo un po' paura di acquistare in un mercato dove non si ha un contatto fisico con gli interlocutori e nemmeno con il prodotto da acquistare, apprezzano moltissimo qualsiasi servizio o strumento utile a ridurre i loro dubbi. Solitamente, più si parla di un prodotto o di un venditore, meglio è; quindi prima di acquistare sarebbe utile anche leggere i blog e le risposte dei forum, oppure le valutazioni lasciate dagli utenti che hanno acquistato il prodotto. In questo modo ci si ritrova però con un altro problema: l'eccessiva informazione. Se si considerano, per esempio, le opinioni lasciate dai consumatori, quando sono in tanti a valutarlo, anche se il prodotto da acquistare è perfetto, può capitare di trovare qualcuno che si lamenta, generando quindi dei dubbi in chi potrebbe acquistarlo.

Il sito *Kelkoo.it* offre agli utenti non solo la possibilità di lasciare un giudizio sui prodotti offerti, ma suggerisce di definire meglio la soddisfazione ottenuta attraverso la valutazione di alcune caratteristiche ritenute maggiormente importanti (si veda Figura 5.20). Questa valutazione è effettuata usufruendo di una scala *Likert* che va da 0 a 5. Dalla somma dei voti si ottiene la valutazione complessiva di ogni opinionista, e la media delle valutazioni complessive per un prodotto definirà la sua posizione rispetto agli altri presenti sul sito. Chiunque acceda al sito può visualizzare i prodotti non soltanto in base ad un ordine crescente o decrescente dei prezzi o del nome, ma anche in base al punteggio ottenuto in base alle valutazioni degli utenti.

Figura 5.20 - La valutazione di un cellulare sul sito Kelkoo.it

**buono**  
Di Anonimo (25/12/2007)

**Vantaggi:** completissimo  
**Svantaggi:** batteria

eccellente prodotto, peccato per la batteria e per le mappe a pagamento [Leggi l'intero giudizio »](#)

E' stato utile per te leggere questo giudizio? **Sì / No**  
([Segnala abuso](#))

**Valutazione Complessiva:** ★★★★★

Stile:	★★★★☆
Durata della Batteria:	★★★★☆
Facilita' d'Uso:	★★★★☆
Qualita':	★★★★☆
Rapporto Qualita'-Prezzo:	★★★★☆
<b>Consigli il Prodotto:</b>	<b>Sì</b>

Questo tipo di valutazione è più facile da confrontare con altre; sarebbe stato comunque interessante scoprire quali sono le caratteristiche che non soddisfano gli utenti, quali li soddisfano maggiormente, quali sono i prodotti maggiormente consigliati, quali quelli che hanno avuto più opinioni; tuttavia kelkoo.it mostra soltanto un punteggio totale complessivo e non spiega altro. Quindi per valutare meglio le opinioni lasciate bisognava in primis raccogliere queste valutazioni; per fare ciò è stato opportuno utilizzare uno *spider software* in grado di analizzare velocemente diverse pagine web per poter selezionare le informazioni ritenute più rilevanti. Nel nostro caso, lo *spider* è stato configurato stabilendo quali erano le pagine web che contenevano le opinioni degli utenti di Kelkoo.it e quali informazioni erano utili per la nostra indagine; lo *spider* quindi inseriva le informazioni raccolte in una tabella più comoda da interpretare. Lo *spider* utilizzato si chiama *Web Content Extractor*; esistono online diversi software simili, ma questa applicazione ha una usabilità simile a quella di un'applicazione *office*, e quindi non è stato molto difficile configurarlo. Sono tanti i prodotti che hanno raccolto valutazioni sul sito kelkoo.it; ma per limitare il campo di indagine ci siamo soffermati esclusivamente sulle opinioni lasciate sui cellulari.

#### 5.4.1 Sintesi dei risultati

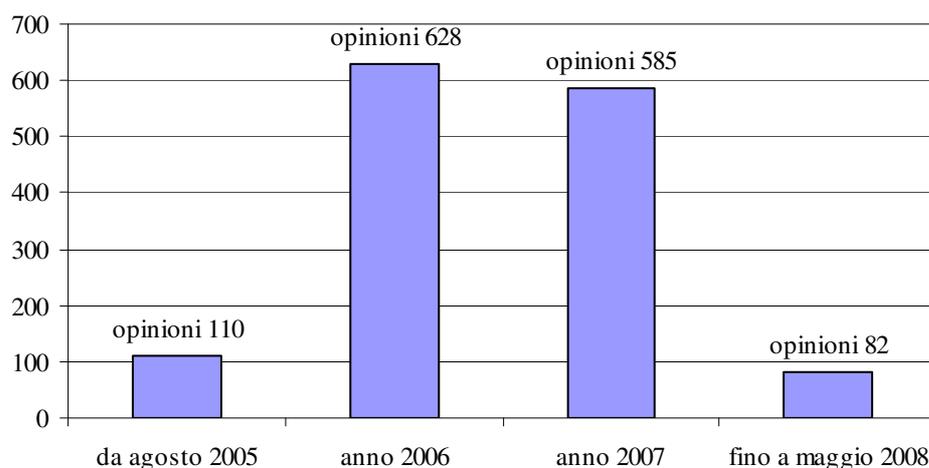
In base alle informazioni raccolte, da una prima analisi si può evidenziare che sono state rilevate 1405 opinioni lasciate sul sito kelkoo.it in un periodo di riferimento che va da agosto 2005 a giugno 2008.

Più del 68% delle opinioni sono date da opinionisti che hanno rilasciato una sola opinione e solo il 5,5% delle opinioni sono date da opinionisti che hanno rilasciato più di una opinione, anche se circa il 19% delle opinioni sono state rilasciate da persone rimaste in anonimato.

Tra le 1405 rilasciate, sicuramente il 72% erano positive, mentre le rimanenti, vale a dire il 28%, risultavano negative, nei confronti del prodotto. Ciò si spiega osservando la figura precedente (Figura 7.1) dalla quale si nota che quando un utente lasciava una opinione doveva concludere rispondendo alla domanda: "Consigli il Prodotto?"; quasi i due terzi degli opinionisti sono stati soddisfatti del prodotto acquistato.

Osservando la figura successiva (Figura 5.21), si vede che il 2006 è l'anno in cui sono state rilasciate più opinioni, mentre dal 2007 in poi si rileva una diminuzione delle opinioni rilasciate sul sito Kelkoo.it, forse per l'avvento di nuovi motori di ricerca per la comparazione dei prezzi, e forse anche a causa della flessione delle vendite di cellulari.

Figura 5.21 - Le opinioni degli utenti di Kelkoo.it divise per anno

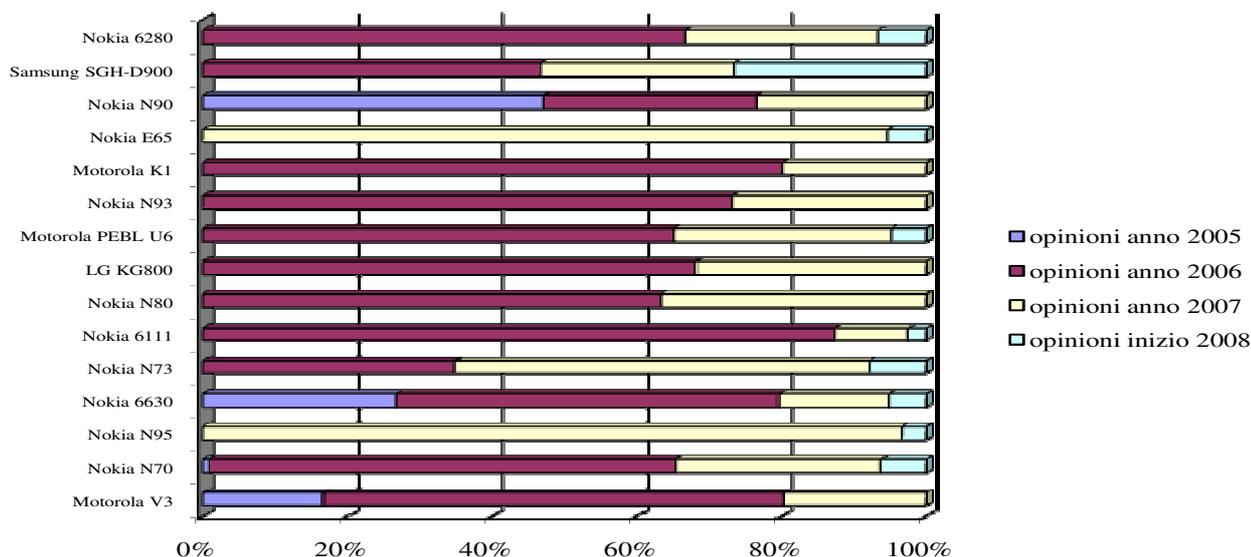


Pur essendo tanti i cellulari che hanno avuto almeno una valutazione soggettiva da parte degli utenti del sito, sono soltanto 15 i cellulari che hanno avuto almeno 15 opinioni a testa. La tabella 5.11 mostra i dati relativi ai cellulari che hanno ottenuto più opinioni; i più criticati sono stati il Motorola V3, il Nokia N70 ed il Nokia N95.

Tabella 5.11 - Telefoni cellulari che hanno avuto un maggior numero di opinioni sul sito Kelkoo.it

	Marca e Modello	Opinioni Tot Kelkoo.it	Uscita sul Mercato	opinioni anno 2004	opinioni anno 2005	opinioni anno 2006	opinioni anno 2007	opinioni inizio 2008
1	Motorola V3	20,85%	2004	-	49	186	58	0
2	Nokia N70	6,76%	2005	-	1	61	27	6
3	Nokia N95	6,33%	2006	-	-	-	86	3
4	Nokia 6630	5,62%	2004	-	25	49	14	5
5	Nokia N73	4,48%	2006	-	-	22	36	5
6	Nokia 6111	2,78%	2005	-	-	34	4	1
7	Nokia N80	2,14%	2006	-	-	19	11	0
8	LG KG800	1,57%	2006	-	-	15	7	0
9	Motorola PEBL U6	1,42%	2005	-	-	13	6	1
10	Nokia N93	1,42%	2006	-	-	19	7	0
11	Motorola K1	1,42%	2006	-	-	16	4	0
12	Nokia E65	1,28%	2007	-	-	-	17	1
13	Nokia N90	1,21%	2005	-	8	5	4	0
14	Samsung SGHD900	1,07%	2006	-	-	7	4	4
15	Nokia 6280	1,07%	2005	-	-	10	4	1
	<i>Altri Cellulari</i>	<i>41%</i>						

Figura 5.22 - L'andamento delle opinioni dei 15 cellulari che hanno avuto maggior numero di opinioni sul sito Kelkoo.it. negli anni dal 2005 al 2008.



Anche dalla figura 5.22 si rileva come il 2006 sia stato l'anno con maggior numero di opinioni rilasciate; ma è importante inoltre notare come i cellulari vengano maggiormente apprezzati i primi mesi di uscita sul mercato e come il numero delle opinioni vada rapidamente diminuendo nei mesi successivi. Cellulari quali ad esempio il Motorola K1, il Nokia N93, il Nokia N80, e il cellulare LG KG800 hanno avuto un periodo di valutazione massimo di due anni (anni 2006 e 2007), mentre cellulari quali i Nokia N70 e 6630 hanno ricevuto giudizi in tutti gli anni dal 2005 al 2008.

Il Nokia 6111 e il Nokia N95 sono stati quelli che hanno ottenuto il maggior numero di opinioni nell'arco di un anno avendo quindi suscitato maggiore curiosità da parte degli utenti; per esempio, il Nokia N95 è stato da tutti riconosciuto come il cellulare più completo del 2007. Agli utenti che lasciano una opinione su un prodotto il sito Kelkoo.it offre la possibilità di precisare meglio la loro valutazione, definendo il loro gradimento su una scala *likert* che va da 0 a 5, per 5 caratteristiche ritenute importanti in un cellulare: lo stile, la durata della batteria, la facilità d'uso, la qualità, il rapporto tra la qualità e il prezzo.

La tabella 5.12 riporta la media dei voti lasciati dagli utenti rispetto alle suddette caratteristiche, sempre per i primi 15 cellulari che hanno ottenuto un maggior numero di opinioni. Per una comodità di lettura, i risultati delle valutazioni sono state trasformate in percentuali.

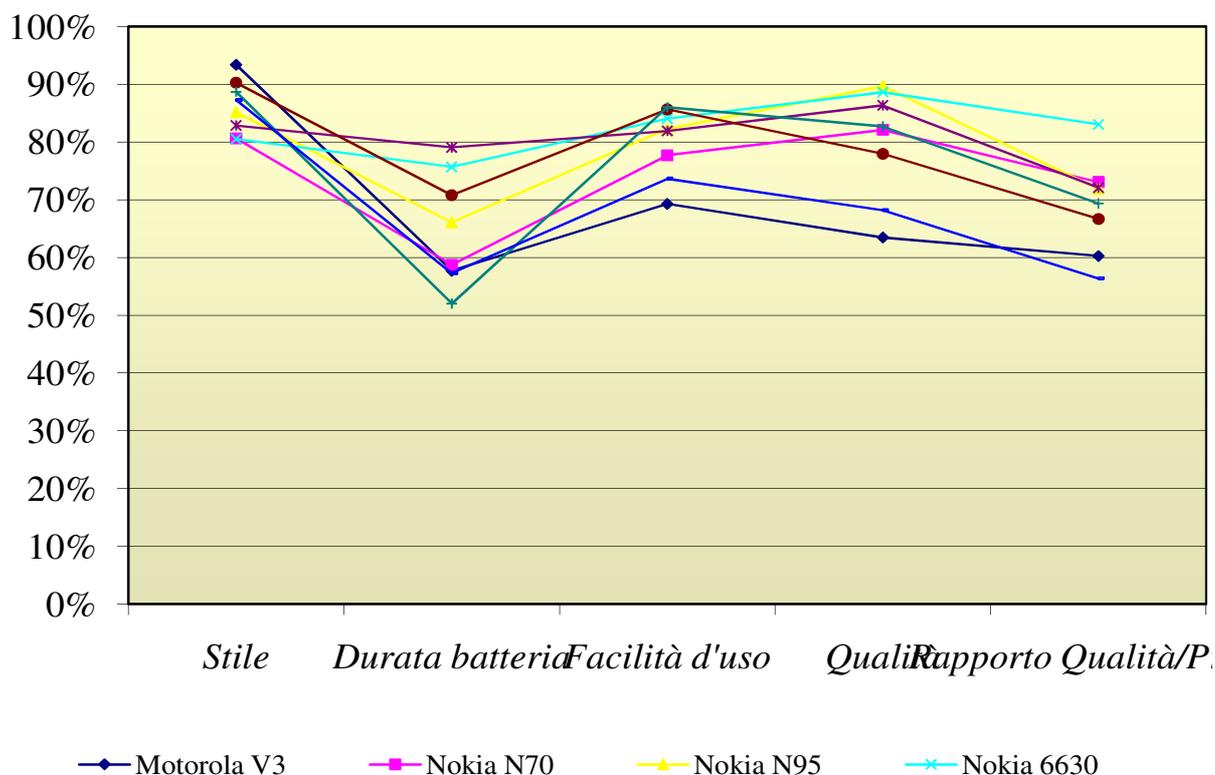
Tabella 5.12 - Valutazione di alcune caratteristiche dei cellulari che hanno ottenuto un maggior numero di opinioni

	<i>Stile</i>	<i>Durata batteria</i>	<i>Facilità d'uso</i>	<i>Qualità</i>	<i>Rapporto Qualità/Prezzo</i>
Motorola V3	93/100	58/100	69/100	63/100	60/100
Nokia N70	81/100	59/100	78/100	82/100	73/100
Nokia N95	85/100	66/100	82/100	90/100	72/100
Nokia 6630	81/100	76/100	84/100	89/100	83/100
Nokia N73	83/100	79/100	82/100	86/100	72/100
Nokia 6111	90/100	71/100	86/100	78/100	67/100
Nokia N80	89/100	52/100	86/100	83/100	69/100
LG KG800	87/100	57/100	74/100	68/100	56/100

La tabella 5.12 mostra che i cellulari con maggior numero di opinioni hanno destato maggiore curiosità da parte degli utenti subito dopo l'uscita sul mercato; difatti, osservando i dati riferiti allo stile, una caratteristica che si apprezza prima dell'acquisto, tutti e 15 i cellulari hanno ottenuto valori molto alti. Addirittura il Motorola V3 ha ottenuto quasi il massimo da tutti gli opinionisti; Motorola infatti, come si capiva dalla pubblicità, puntò moltissimo sul formato rivoluzionario dei suoi prodotti, e in particolare sull'innovativa tastiera sulla quale bastava sfiorare il carattere per comporre un numero di telefono o scrivere un messaggio. Questa innovazione ha avuto molto successo, ma, come si può notare dall'opinione degli utenti riguardo il rapporto qualità/prezzo, il prezzo è stato ritenuto piuttosto alto, rendendo alquanto dubbia la convenienza dell'acquisto.

Purtroppo, come spesso accade, l'apparenza inganna, e quindi i consumatori, attratti dal design innovativo, si aspettavano anche performance superiori; invece i punteggi rilasciati per l'autonomia della batteria, la facilità d'uso, e la qualità in generale non superano di molto la sufficienza.

Figura 5.23 Confronto tra i cellulari con maggior numero di opinioni rilasciate, in base alle caratteristiche valutate.

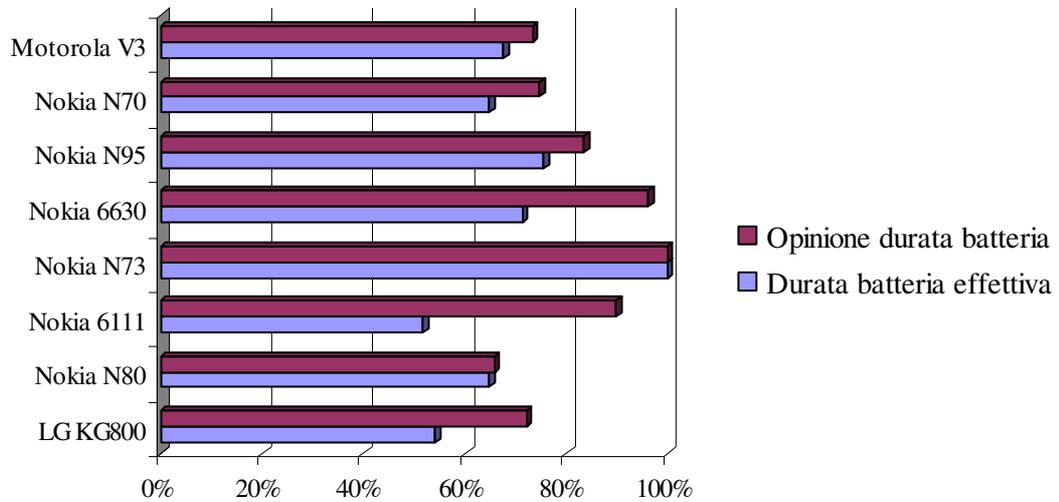


Dalla Figura 5.23 si nota come i 15 cellulari considerati abbiano ottenuto indici più bassi per quanto riguarda l'autonomia della batteria ed il rapporto qualità/prezzo; la batteria è un problema per quasi tutti i cellulari di ultima generazione, con tante funzioni avanzate che consumano molta energia; però se la tecnologia sta facendo passi da gigante, i progressi nell'efficienza delle batterie sono stati minori di quelli necessari a far fronte all'inserimento di funzioni tecnologicamente sempre più avanzate.

L'inserimento di nuove funzioni comporta inoltre rilevanti aumenti di prezzo; considerato che alcune delle nuove funzioni spesso non vengono utilizzate, si manifesta quindi insoddisfazione da parte di utenti, perché si rendono conto di avere pagato significative maggiorazioni di prezzo per funzioni aggiuntive di fatto scarsamente utilizzate.

La figura 5.24 mostra come effettivamente la valutazione dei consumatori spesso sia corretta. Il grafico mette infatti a confronto le valutazioni date sulla durata della batteria con le autonomie delle batterie riportate dalle casa produttrici per i primi 10 cellulari che hanno avuto un maggior numero di opinioni; le differenze fra autonomia effettiva e percepita sono minime per tutti i cellulari; ciò dimostra l'affidabilità dei giudizi degli utenti/consumatori durante il processo decisionale.

Figura 5.24 - Confronto fra autonomia della batteria percepita e autonomia effettiva per i 10 cellulari maggiormente valutati.



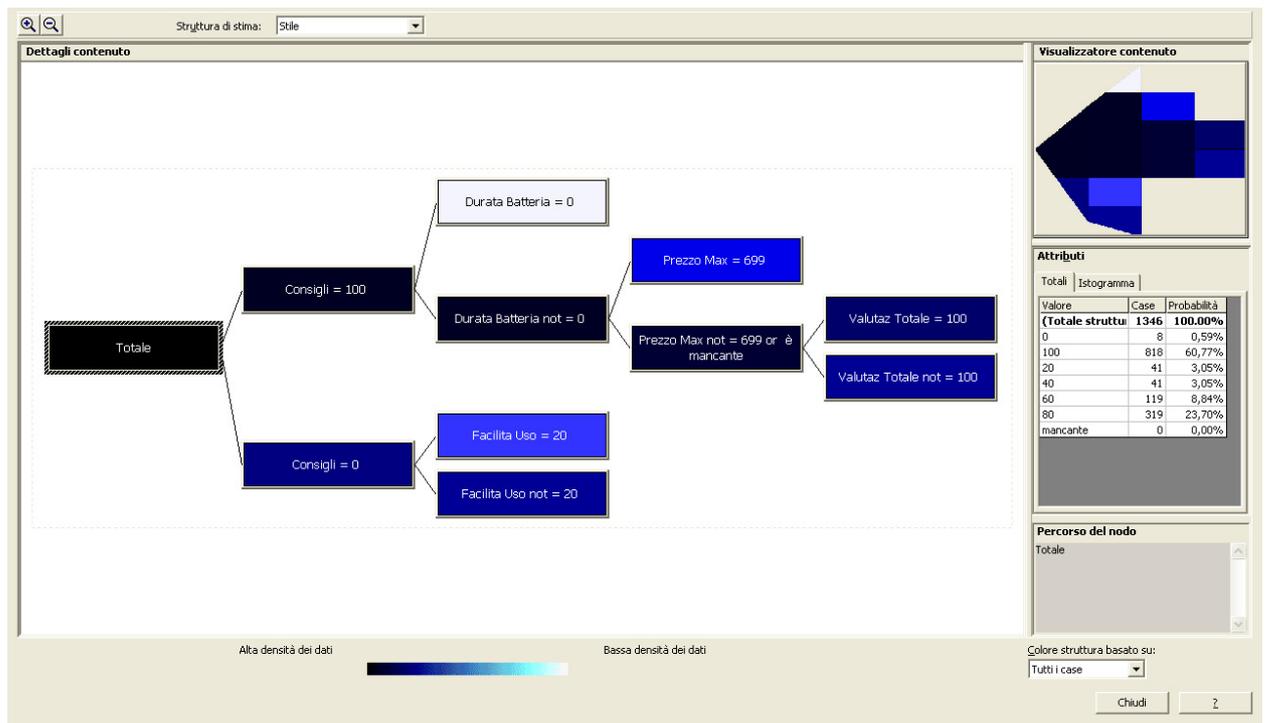
## 5.4.2 Un'analisi di Data Mining

### 5.4.2.1 Analisi di decision trees

Sono stati realizzati alberi decisionali tenendo conto dei punteggi lasciati dagli utenti opinionisti sulle cinque caratteristiche scelte dal sito Kelkoo.it

Particolarmente interessante è l'albero decisionale che si ottiene utilizzando come variabile di stima le opinioni lasciate sulla caratteristica "stile"; un albero decisionale di questo tipo è illustrato nella Figura 5.25.

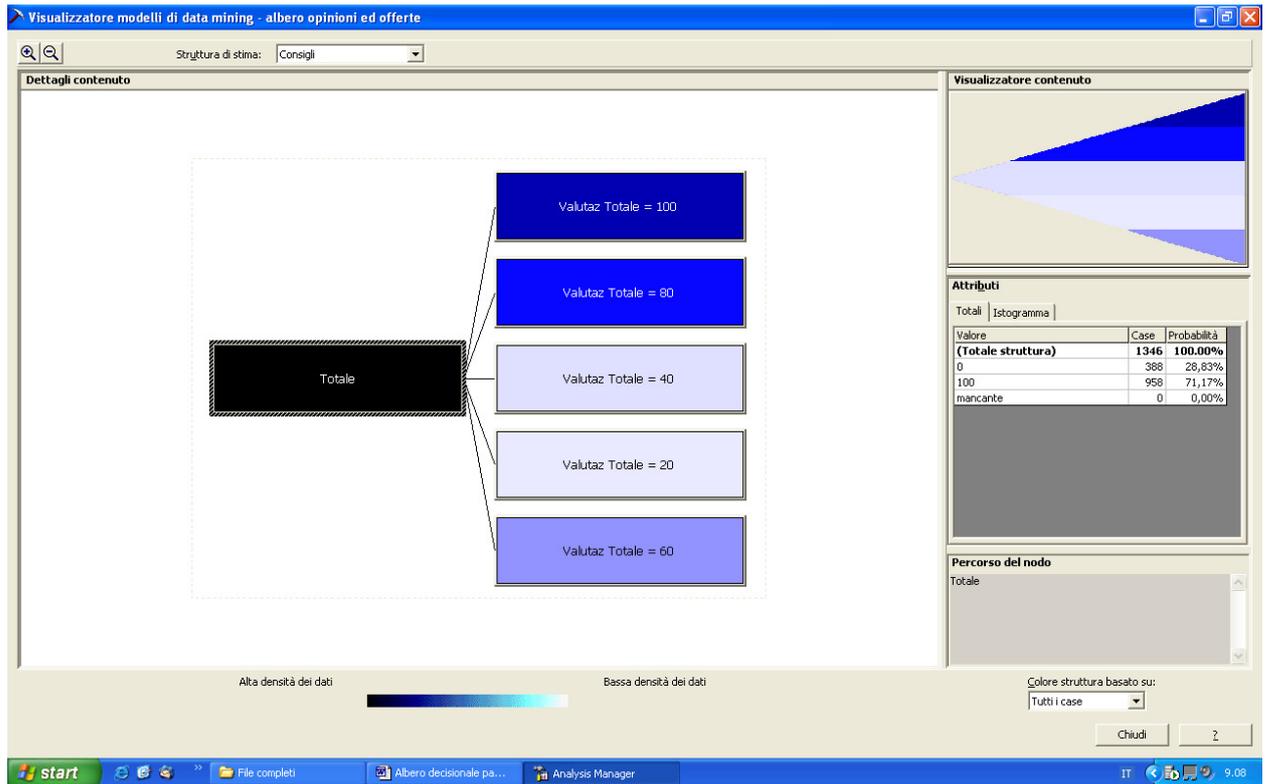
Figura 5.25 - Albero decisionale ottenuto partendo da valutazioni sullo stile del prodotto



Analizzando meglio i risultati ottenuti dall'albero decisionale, risulta che tra coloro i quali consigliavano di acquistare il prodotto, il 68% giudicava molto positivo lo stile e la durata della batteria del cellulare considerato. Ciò può voler dire che nell'acquisto di un cellulare è molto importante non soltanto lo stile, ma anche la durata della batteria.

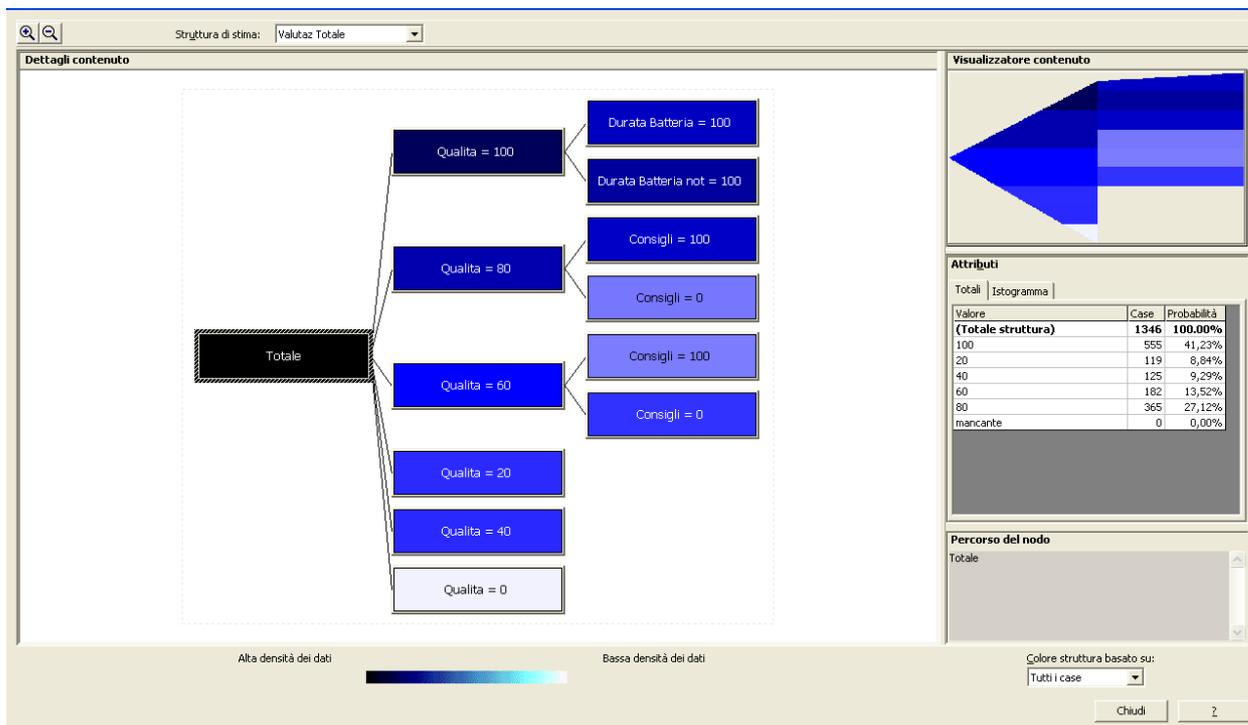
Quest'ultima considerazione si evidenzia meglio osservando la figura 5.26 in cui viene analizzata la valutazione totale delle opinioni lasciate dagli utenti del sito Kelkoo.it.

Figura 5.26 - Albero decisionale ottenuto partendo da valutazioni complessive dei prodotti, quindi dalla domanda: "Consiglia di acquistare il prodotto?".



La figura 5.26 mostra l'albero decisionale derivante dalla domanda "consiglia di acquistare il cellulare?", (Risposta: *Si* – *No*). Si nota una forte relazione tra l'attributo e la valutazione totale ottenuta dalla media dei voti lasciati per ogni cellulare giudicato; data l'elevata correlazione tra le due variabili, sarebbe opportuno valutare l'albero decisionale che si ottiene stabilendo come variabile di stima la valutazione totale dei cellulari (Figura 5.27).

Figura 5.27 - Albero decisionale ottenuto stabilendo come variabile di stima la caratteristica “valutazione totale del prodotto”



Secondo l’output del programma, il 74% degli utenti che hanno dato una valutazione massima della qualità del prodotto (qualità = 100) considerano importantissima anche la qualità della batteria; si può quindi dedurre che quando si acquista un cellulare si tiene conto dello stile del prodotto, anche perché è quella che più colpisce l’utente, soprattutto quando si acquista in un negozio tradizionale; tuttavia, si rimane soddisfatti dell’acquisto fatto, e quindi si consiglia di acquistarlo, quando si rimane soddisfatti della qualità del prodotto, in particolare per quel che riguarda l’autonomia della batteria.

## 5.5 Conclusioni

In conclusione presentiamo una sintesi dei principali risultati ottenuti dalle diverse indagini presentate in questo capitolo. Dalla matrice di correlazione tra le variabili relative alle caratteristiche dei cellulari venduti online su Kelkoo.it (Tabella 5.2), si rileva che la qualità del display è correlata positivamente con la risoluzione della fotocamera (scatti di qualità, hanno bisogno di display di qualità), che la connettività senza fili è correlata positivamente con le caratteristiche: sistema operativo Windows, Mail, connessione ad internet, memoria espandibile, mp3, peso, prezzo, risoluzione del display e della fotocamera, touch screen e videocamera.

Dalla regressione multivariata (Tabella 5.3), si rilevano quali caratteristiche presentano delle relazioni statisticamente significative con il prezzo; in particolare, sono significativamente diverse da zero, le stime dei coefficienti delle variabili autonomia di conversazione, memoria espandibile, peso, risoluzione della fotocamera, zoom, autonomia in standby, dimensioni, flash, e connessione a internet; è significativamente positivo il segno dei coefficienti delle variabili: autonomia in standby, connessione a internet, peso, risoluzione fotocamera, invece è significativamente negativo il segno dei coefficienti delle variabili autonomia in conversazione, dimensione, flash, memoria espandibile. Il peso è una caratteristica negativa per l’acquirente, ma risulta direttamente proporzionale al prezzo; ciò deriva probabilmente dal fatto che i telefonini cosiddetti *smartphone* perché ultratecnologici, caratterizzati da prezzi molto alti,

utilizzano molta energia e quindi devono essere dotati di grandi batterie che fanno aumentare di molto il peso complessivo del telefonino. In modo analogo si può probabilmente spiegare il coefficiente significativamente negativo per la variabile autonomia di conversazione: diversi telefonini di elevate prestazioni (in particolare per il display) e prezzo elevato consumano molta energia, e hanno quindi una bassa autonomia di conversazione.

Esaminando da diverse angolazioni il mercato della telefonia mobile, dapprima basandoci nella prima indagine solo sui cellulari venduti a luglio del 2008 sul sito Kelkoo.it e poi nella seconda indagine, sull'intera popolazione di cellulari venduti nell'ultimo decennio, i risultati ottenuti mediante la *cluster analysis* (Tabelle 5.8 e 5.9), sono alquanto concordanti. In entrambe le analisi sono stati ottenuti dei *cluster* con cellulari raggruppati per vecchio o nuovo modello, per tecnologia di base, media o alta, oppure per dimensione, quindi cellulari ultrasottili, di dimensioni ridotte, o *smartphone* di grandi dimensioni. Sostanziali differenze sono dovute soprattutto alla velocità di evoluzione tecnologica delle caratteristiche che definiscono un cellulare. Mentre tra i telefonini venduti a luglio 2008 una percentuale elevata dispongono di display touch screen, Gps, tastiera qwerty, tecnologia Wi-Fi e connessione ad internet superveloce tramite Hsdpa, considerando invece l'intera popolazione di cellulari usciti sul mercato dal 1998 ad oggi, i cellulari che hanno queste caratteristiche rappresentano una percentuale molto bassa del totale, mentre moltissimi dispongono soltanto di funzione MMS, connettività ad infrarossi, e velocità di connessione ad internet Gprs o al massimo Edge. Tutto ciò si vede meglio osservando le Figure 5.18 e 5.19, che mostrano la penetrazione sul mercato delle principali caratteristiche di un telefono cellulare. Ad esempio, le caratteristiche Gps, videochiamata, quantità di colori del display, memoria totale e *touch screen* hanno avuto un forte sviluppo negli ultimi 3 anni. In particolare, la funzione *touch screen*, quasi sconosciuta agli inizi del 2007, è oramai disponibile in quasi tutti i cellulari con ampi display. Mentre per le caratteristiche connettività senza fili e risoluzione della fotocamera c'è stata una crescita meno rapida rispetto a quelle analizzate precedentemente. Altre caratteristiche quali ad esempio mms, mp3, mail e videocamera, hanno avuto una forte crescita dal 2002 al 2005, e poi si sono stabilizzate, anche perché oramai si trovano in quasi tutti i cellulari.

La terza indagine che aveva l'obiettivo di valutare il comportamento degli opinionisti online, ha portato ad alcuni interessanti risultati. Alcuni di questi sono evidenziati nella Tabella 5.12, che mostra come i 15 cellulari che hanno avuto maggior numero di opinioni su Kelkoo.it, hanno destato maggiore curiosità subito dopo l'uscita sul mercato; lasciando punteggi alti per quanto riguarda lo *stile* del telefonino, una qualità che si apprezza prima dell'acquisto. Ma come spesso accade l'apparenza inganna, e quindi i consumatori, attratti dal design innovativo, si aspettavano anche performance superiori; invece i punteggi rilasciati per le caratteristiche: autonomia della batteria, facilità d'uso, e qualità in generale non superano di molto la sufficienza; lo stesso vale per il rapporto qualità/prezzo, rendendo quindi, alquanto dubbia la convenienza dell'acquisto. Ciò si deduce anche dall'albero decisionale raffigurato in Figura 5.27; prendendo come variabile di stima la valutazione totale del prodotto, si rileva che il 74% degli utenti che hanno valutato massima la qualità del prodotto (qualità = 100), considerano importantissima anche la qualità della batteria; si può quindi ipotizzare, che quando si acquista un cellulare ovviamente si tiene conto molto dello stile del prodotto, anche perché è quello che più colpisce l'utente, soprattutto quando si acquista in un negozio tradizionale, ma si rimane soddisfatti dell'acquisto fatto, e quindi si consiglia di acquistare lo stesso telefonino, soltanto quando si valuta alta la qualità del prodotto, che dipende molto anche dall'autonomia della batteria.

In conclusione, basandoci sui risultati ottenuti dalle ricerche effettuate nei settori dell'intelligenza artificiale, della psicologia cognitiva, della statistica e dell'informatica (capitoli 1, 2 e 3), e dopo aver esaminato la situazione degli ambienti online dedicati agli acquisti, italiani ed esteri (capitolo 4), sono state presentate in questo capitolo alcune indagini empiriche che aiutano a comprendere meglio le decisioni dei consumatori, in particolare per quel che riguarda

gli acquisti online. Il capitolo successivo (Capitolo 6) presenta una proposta per alcuni aspetti innovativa di sistema di aiuto alle decisioni di acquisto, reso disponibile anche online. Esso si basa su una metodologia decisionale e su una interfaccia utente che consentono di orientare il processo decisionale verso scelte in grado di corrispondere meglio alle esigenze e aspirazioni dei consumatori.

## Capitolo 6

### Progettazione e sviluppo di un sistema di supporto alle decisioni

#### 6.1 Introduzione

In questo capitolo si presenta un nuovo sistema di supporto decisionale, con una metodologia di filtraggio diversa rispetto a quella utilizzata dai principali sistemi di raccomandazione *content based filtering*, come ad esempio i sistemi di aiuto alle decisioni di tipo *compensatori*, chiamati così, perchè per ogni opzione, gli attributi negativi vengono compensati dagli attributi positivi (vedi paragrafo 3.5.1); la metodologia decisionale di un sistema compensatorio si basa sulla Teoria della massimizzazione dell'utilità attesa [Keeney e Raiffa, 1976]. Nel paragrafo 6.2 si evidenzia come, seguendo questo metodo, si corra il rischio di selezionare opzioni sbagliate quando gli attributi sono inversamente correlati. La metodologia del sistema qui proposto, in sintesi, tende a minimizzare, per ogni attributo di ogni opzione di prodotto, il gap tra performance richiesta dal consumatore e performance effettiva, selezionando le opzioni di prodotto con i più bassi valori complessivi di gap negativi.

L'applicazione proposta è stata testata, all'inizio, per la selezione del cellulare più adatto tra quelli presenti nel reparto di telefonia di un centro commerciale della provincia di Cosenza, con risultati in complesso positivi. L'applicazione è disponibile anche online sul sito SceltaGiusta.it.

L'applicazione può essere utilizzata per la scelta di qualsiasi prodotto "complesso" (multi-attributo), quando si è in presenza di molte opzioni di scelta; essa può quindi facilitare l'acquisto di una lavatrice, un televisore, un'auto, e persino un servizio, come l'organizzazione di un viaggio.

#### 6.2 Una diversa metodologia decisionale

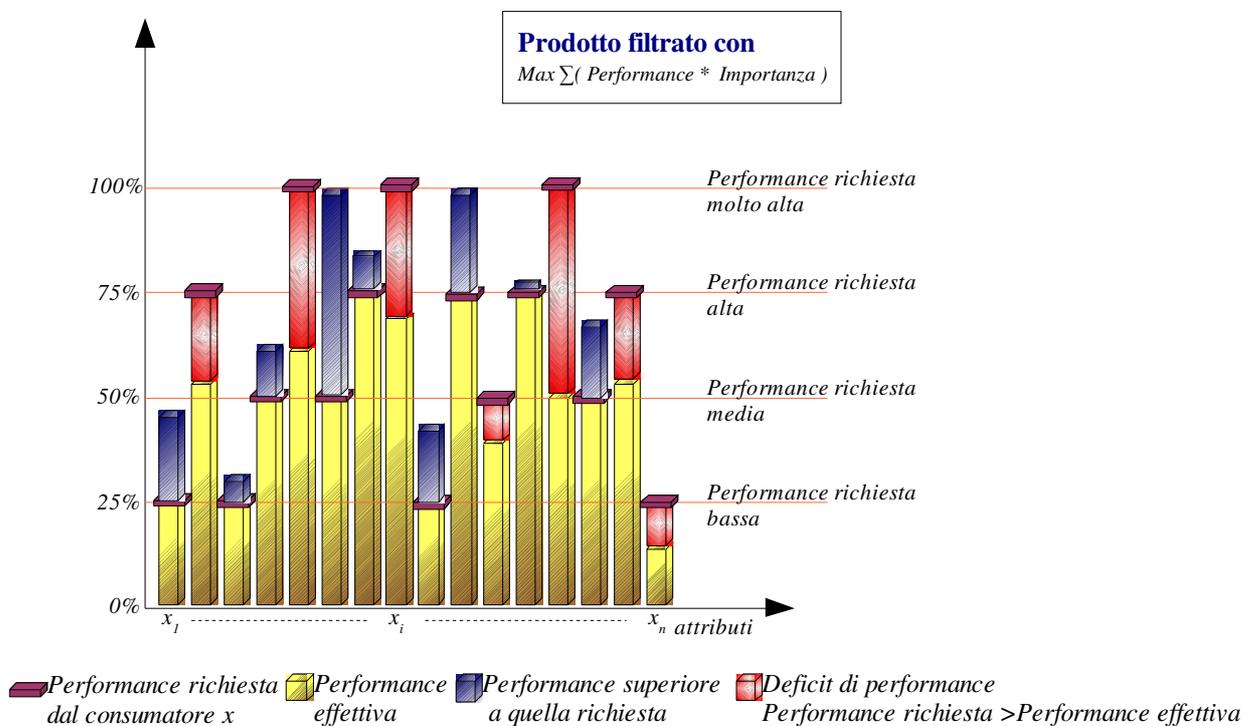
I sistemi di supporto alle decisioni di acquisto facilitano le scelte dei consumatori, specialmente i sistemi *content-based filtering*, capaci di selezionare le opzioni più vicine alle preferenze di ciascuno. Si corre però il rischio di ottenere selezioni poco soddisfacenti, soprattutto, quando i sistemi *content based* utilizzano metodologie compensatorie.

Questi sistemi selezionano le opzioni di prodotto in base al metodo che massimizza la somma dei prodotti dei valori degli indicatori di performance dei diversi attributi per gli indici di importanza dei diversi attributi, facendo riferimento alla teoria della massimizzazione dell'utilità attesa [Keeney e Raiffa, 1976]. Le strategie compensatorie sono ritenute procedure decisionali razionali, poiché comportano l'utilizzazione di tutte le informazioni disponibili, al fine di ottenere valutazioni complessive per le diverse opzioni, sulla base delle preferenze individuali per i diversi attributi. Le strategie compensatorie non sono però affidabili, in presenza di opzioni con attributi inversamente correlati, dove attributi con performance alta compensano attributi con performance scarse; in questo caso, il consumatore potrebbe essere influenzato a scegliere una opzione di prodotto con caratterizzate da performance molto buone per alcuni attributi, e performance molto negative per altri aspetti. Häubl e Trifts hanno mostrato che le persone, avendo la tendenza a fidarsi delle raccomandazioni, soprattutto quando la scelta è tra opzioni che hanno valori soddisfacenti per alcuni attributi, ma non soddisfacenti per altri (attributi negativamente correlati), vengono influenzate sistematicamente da performance alte degli attributi dei prodotti selezionati, anche se per esse, questi attributi sono poco interessanti.

La logica delle diverse metodologie decisionali è illustrata dalle Figure 4.1 e 4.2, che evidenziano le performance dei diversi attributi dell'opzione selezionata in base a differenti criteri decisionali.

La Figura 4.1 evidenzia ciò che può accadere compensando performance positive e negative. Nella figura viene esaminata la specificazione di prodotto selezionata da un agente intelligente che compensa performance positive con performance negative delle caratteristiche. Le righe orizzontali individuano l'importanza attribuita dal consumatore ad ogni attributo del prodotto; i parallelepipedi rossi mostrano il divario esistente tra la performance effettiva di una caratteristica e la performance minima richiesta dal consumatore per quella caratteristica. Applicando il metodo che massimizza la somma dei prodotti dei valori degli indicatori di performance dei diversi attributi per gli indici di importanza dei diversi attributi si ottiene un cellulare in cui diverse caratteristiche presentano delle performance insufficienti rispetto a quelle richieste, e altre caratteristiche hanno performance molto più elevate rispetto a quelle richieste dal consumatore. Le carenze di performance per alcuni attributi sono compensate da performance molto elevate per altri attributi.

Figura 6.1 - Rappresentazione delle performance degli attributi del prodotto selezionato secondo il metodo che massimizza la somma dei prodotti dei valori degli indicatori di performance dei diversi attributi per gli indici di importanza dei diversi attributi.



Un'azienda che è interessata ad immettere un nuovo prodotto sul mercato, per individuare dove posizionarsi, vale a dire la tipologia di consumatori che potrebbero essere interessati a quel prodotto, utilizza diversi metodi, tra i quali la *conjoint analysis*, per scoprire l'importanza attribuita dai consumatori alle diverse caratteristiche di un prodotto. Qui abbiamo a che fare con prodotti in cui le performance delle caratteristiche sono stabilite a priori e non si possono modificare (tranne per alcuni attributi, quali la memoria del cellulare se è espandibile, la batteria se è sostituibile, ecc.).

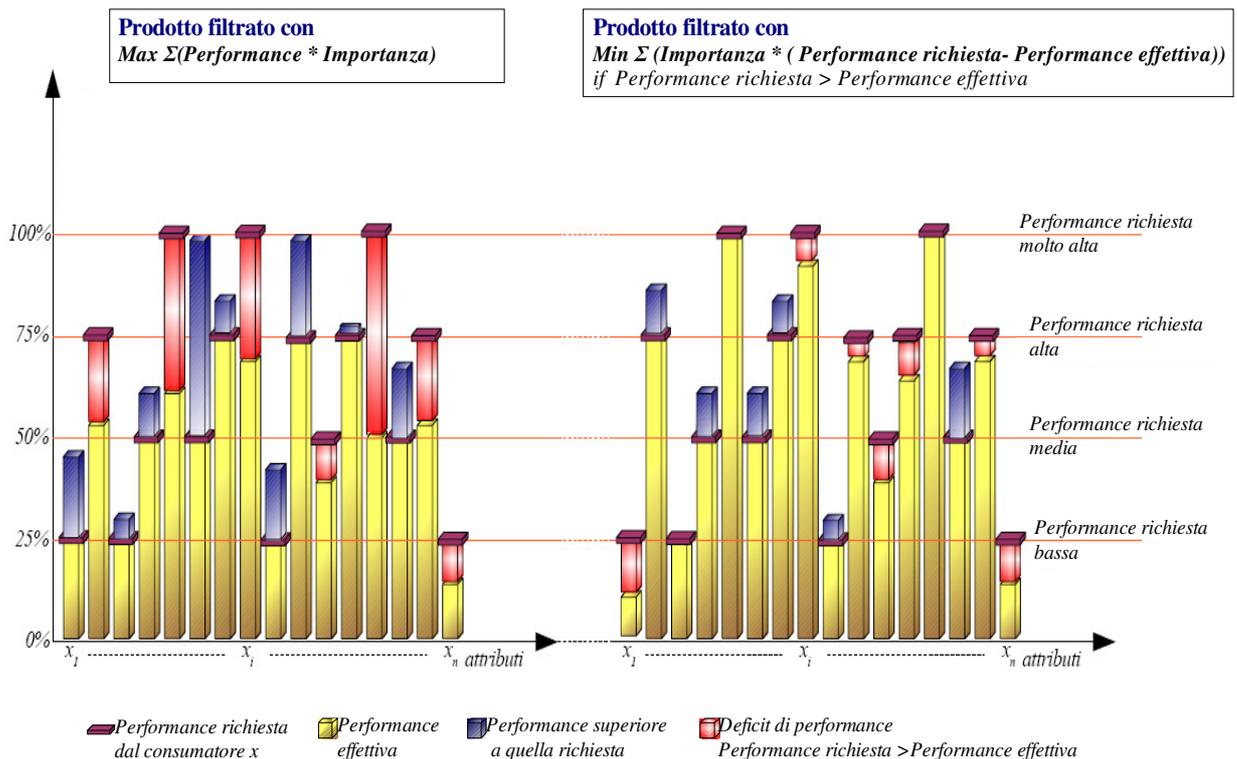
Una soluzione a questo problema potrebbe essere quella di selezionare tra le opzioni di prodotto presenti sul mercato quelle che si avvicinano maggiormente al "prodotto ideale" del consumatore  $x$ , ovvero, minimizzare il divario tra il prodotto ideale e i prodotti reali.

Questa differenza è stata definita deficit di performance ( $I_i - P_i$ ). Quindi considerare in particolar modo le caratteristiche che non raggiungono i valori di performance desiderati; perchè

un prodotto con performance migliori di quelle richieste, che lasciano comunque conveniente il prezzo, può far solo piacere al consumatore.

Il prodotto ottimale ( $j$ ) sarà quindi quello a cui corrisponde il valore minore della sommatoria dei deficit di performance effettiva rispetto a quella minima richiesta.

Figura 6.2 - Rappresentazione delle performance degli attributi del prodotto selezionato secondo il metodo che massimizza la somma dei prodotti dei valori degli indicatori di performance dei diversi attributi per gli indici di importanza dei diversi attributi, *versus* rappresentazione delle performance degli attributi del prodotto selezionato secondo il metodo che minimizza la somma degli scarti negativi dei valori degli indicatori di performance dei diversi attributi rispetto ai valori di performance indicati come soddisfacenti.



La Figura 6.2 illustra i risultati che si ottengono attraverso la massimizzazione della sommatoria dei prodotti tra performance e Importanza, e la minimizzazione della sommatoria dei gap tra performance richiesta e performance effettiva, ponderati per i coefficienti di importanza dei diversi attributi. Il secondo prodotto presenta attributi con performance vicine a quelle richieste, soprattutto per quanto riguarda le caratteristiche reputate più importanti. Inoltre, siccome i deficit di performance vengono minimizzati tenendo conto dell'importanza attribuita (priorità alle caratteristiche che suscitano maggiore interesse nel consumatore), sarebbe molto difficile ottenere come output una opzione caratterizzata da performance particolarmente negative per attributi particolarmente importanti.

### 6.3 La progettazione dell'applicazione

Come risultato della ricerca portata avanti per scoprire come si comporta il consumatore nei processi decisionali, e in particolar modo negli ambienti online, è stata realizzata un'applicazione denominata "Acquisto Facile", che si propone di ridurre al minimo gli sforzi cognitivi dei consumatori, selezionando un gruppo limitato di prodotti che, in base a criteri predefiniti, sono in grado di corrispondere meglio alle aspirazione ed esigenze di ciascun

consumatore. L'applicazione è disponibile online sul sito [www.SceltaGiusta.it](http://www.SceltaGiusta.it)<sup>60</sup>; al momento la versione "beta", aiuta i consumatori/utenti a individuare il telefono cellulare più vicino alle sue richieste, successivamente potrà essere impiegata per altre tipologie di prodotti caratterizzati da molti attributi. L'applicazione, ovviamente, si basa sulla metodologia decisionale qui proposta, che minimizza la somma degli scarti negativi tra performance richieste dal consumatore e performance effettive degli attributi delle opzioni considerate, ponderati per i coefficienti di importanza dei diversi attributi<sup>61</sup>.

Il codice interno dell'applicazione è stato scritto in *php*<sup>62</sup>, (un linguaggio di *scripting* interpretato, con licenza *open source*<sup>63</sup>); l'algoritmo che, secondo le preferenze dell'utente, seleziona dal database i telefonini più adatti alle sue esigenze, si basa su questo metodo decisionale. Viene quindi immediatamente offerta ai soggetti interessati la possibilità di valutare la validità del metodo proposto con dati reali ed in ambienti online.

L'interfaccia grafica del sistema "Acquisto Facile", che interagisce con l'utente online, è stata realizzata attraverso l'applicativo *Adobe Flash CS3 Professional*, l'ambiente di *authoring*<sup>64</sup> più avanzato del settore per la creazione di contenuti interattivi di alto livello per piattaforme digitali. Il software *Adobe Flash* consente di creare animazioni vettoriali<sup>65</sup> principalmente per il web; esso viene utilizzato anche per creare giochi o interi siti web; grazie all'evoluzione delle ultime versioni, esso è divenuto un potente strumento per la creazione di *Rich Internet Application*<sup>66</sup>. I contenuti creati con *Adobe Flash*, non solo possono essere utilizzati per Internet, ma sono anche compatibili con un'ampia gamma di telefoni cellulari e dispositivi elettronici di consumo; quindi il contenuto viene creato una sola volta e può essere successivamente riutilizzato per più piattaforme. Integrando l'applicazione all'interno del dispositivo, sarà possibile individuare il prodotto che meglio corrisponde alle esigenze e aspirazioni dell'utente anche quando si è davanti alla vetrina di un negozio.

---

<sup>60</sup> Attraverso la pagina [http://www.sceltagiusta.it/Acquisto%20facile\\_.html](http://www.sceltagiusta.it/Acquisto%20facile_.html), si accede immediatamente alla pagina iniziale dell'applicazione.

<sup>61</sup> La principale caratteristica di questo sistema di aiuto alle decisioni è di evitare le compensazioni tra *performance* positive e negative, a differenza dei sistemi che massimizzano la somma dei prodotti tra i valori degli indicatori di performance dei diversi attributi e i valori degli indicatori di importanza degli stessi attributi.

<sup>62</sup> Originariamente concepito per la realizzazione di pagine web dinamiche, *php* attualmente è utilizzato principalmente per sviluppare applicazioni web lato server; esso può essere utilizzato anche per scrivere *script* a linea di comando o applicazioni *stand-alone* con interfaccia grafica. Il suo nome è un acronimo ricorsivo che sta per *Hypertext Pre-processor* (*pre-processore di ipertesti*).

<sup>63</sup> *Open source*, (termine inglese che significa "sorgente aperta") indica un software rilasciato con un tipo di licenza per la quale il codice sorgente è lasciato alla disponibilità di eventuali sviluppatori, in modo che con la collaborazione (in genere libera e spontanea) di tantissimi utenti il prodotto finale possa raggiungere una complessità maggiore di quanto potrebbe ottenere un singolo gruppo di programmazione.

<sup>64</sup> Gli applicativi d'autore (*authoring*) sono quei software verticali che consentono la realizzazione di una comunicazione multimediale, articolata e riproducibile su personal computer. L'intento è quello di poter produrre e veicolare contenuti (immagini statiche, animazioni grafiche, filmati video, commenti sonori, effetti audio e altro) via web e su supporti come cd-rom, DVD, oppure attraverso un circuito, chiuso o aperto, di display distribuiti e connessi tra loro in rete.

<sup>65</sup> La grafica vettoriale è una tecnica utilizzata in computer grafica per descrivere un'immagine. Un'immagine descritta con la grafica vettoriale è chiamata immagine vettoriale. Nella grafica vettoriale un'immagine è descritta mediante un insieme di primitive geometriche (semplici figure geometriche come punti, linee, triangoli) che descrivono punti, linee, curve e poligoni ai quali possono essere attribuiti colori e anche sfumature.

<sup>66</sup> Le *Rich Internet Application* (RIA) sono applicazioni web che possiedono le caratteristiche e le funzionalità delle tradizionali applicazioni per computer, senza però necessitare dell'installazione sul disco fisso. Le RIA si caratterizzano per la dimensione interattiva e per la velocità di esecuzione. Infatti la parte dell'applicazione che elabora i dati è trasferita a livello *client* e fornisce una pronta risposta all'interfaccia utente, mentre la gran parte dei dati e dell'applicazione rimane sul server remoto, con notevole alleggerimento per il computer utente. Anche l'interazione con una RIA avviene in remoto, tramite un comune *web browser*. In un certo senso le RIA rappresentano una generazione di applicazioni che permettono un'interazione totalmente rinnovata, fondata sulle migliori caratteristiche funzionali e progettuali che finora erano prerogative, alternativamente, del web o delle applicazioni desktop.

L'applicazione "Acquisto Facile" si appoggia su un *data-base* che contiene i dati relativi ad oltre un migliaio di cellulari disponibili attualmente sul mercato italiano, definiti da una ventina di caratteristiche. Il sistema di supporto alle decisioni, al termine della procedura iterativa, seleziona un gruppo limitato di prodotti in grado di soddisfare le esigenze dell'utente, presentandogli le migliori offerte disponibili sui principali siti e-commerce italiani, con il link al sito del venditore, e offrendo quindi la possibilità di acquistare immediatamente uno dei prodotti selezionati. Nel web italiano sono centinaia i siti e-commerce che vendono telefoni cellulari; individuare quale specificazione di prodotto acquistare, vale a dire le offerte più vantaggiose, è in genere difficilissimo per il consumatore; i comparatori di prezzo, motori di ricerca per gli acquisti online, sono utili perché raccolgono le offerte di moltissimi siti e-commerce e le presentano agli utenti, ordinandole in base al prezzo, alla popolarità, al produttore, e anche in base alle valutazioni lasciate da altri utenti o da esperti. Attraverso i principali siti di comparazione (*kelkoo.it*, *BuyCentral.it*, *Twenga.it*, *TrovaPrezzi.it*, *MigliorScelta.it*) sono state raccolte migliaia e migliaia di offerte disponibili online. Per poter facilmente raccogliere tutti questi dati nel minor tempo possibile, e aggiornarli quotidianamente (perché le offerte possono variare ogni giorno), è stato utilizzato un particolare tipo di software chiamato *Web Content Extractor*, in grado di estrapolare dalle pagine web esclusivamente le informazioni richieste. *Web Content Extractor* è un *crawler*; i *crawler*, chiamati anche *spider* o *softbot*, sono applicazioni che analizzano i contenuti di una rete (o di un database) in modo metodico e automatizzato; in genere essi vengono utilizzati online per conto di un motore di ricerca<sup>67</sup>; ad esempio lo *spider* utilizzato da Google si chiama *googlebot*, quello di Yahoo si chiama *slurp*.

L'applicazione "Acquisto Facile" è un sistema di supporto alle decisioni di acquisto completo, perché offre la possibilità all'utente di valutare dettagliatamente le diverse alternative di prodotto sulla base dei valori di performance delle caratteristiche, e successivamente acquistare l'alternativa di prodotto selezionata sul sito e-commerce italiano che presenta l'offerta più conveniente.

Il paragrafo successivo spiega come, attraverso dei brevi e semplici passaggi, l'utente può definire come dovrebbe essere il suo telefonino ideale; l'interazione termina con la pagina dei risultati, in cui sono classificati in ordine decrescente i telefonini che meglio soddisfano le esigenze dell'utente.

L'applicazione qui proposta si articola in cinque fasi.

Nella prima fase l'utente definisce il suo interesse per i diversi attributi del prodotto, raggruppandoli per grado di interesse (Figura 6.3).

Utilizzando il mouse, l'utente si posiziona su una caratteristica e, tenendo premuto il tasto sinistro, la sposta in una delle aree delimitate per definire il grado di performance ritenuto soddisfacente per ogni attributo del prodotto, secondo la scala di misurazione seguente: performance "nulla", "bassa", "media", "elevata", "molto elevata". Per facilitare l'utente nella comprensione della funzionalità di ogni caratteristica, è stata inserita una "legenda" posizionata in alto a destra della pagina web. Passando il mouse su una caratteristica, all'interno della legenda, viene visualizzata una breve spiegazione dell'attributo, accompagnata da una immagine esplicativa; cliccando sull'icona in basso a destra il sistema passa alla seconda fase.

---

<sup>67</sup> Un uso estremamente comune dei *crawler* è nel Web. Sul Web, il *crawler* si basa su una lista di *url* da visitare fornita dal motore di ricerca (il quale, inizialmente, si basa sugli indirizzi suggeriti dagli utenti o su una lista precompilata dai programmatori stessi). Durante l'analisi di un *url* il *crawler* identifica tutti gli *hyperlink* presenti nel documento e li aggiunge alla lista di *url* da visitare; il processo può essere concluso manualmente o dopo che un determinato numero di collegamenti è stato seguito. I *crawler* attivi su Internet hanno la facoltà di essere indirizzati secondo quanto indicato nel file "*robots.txt*" posto nella *root* del sito; all'interno di questo file è possibile indicare quali pagine non dovrebbero essere analizzate; il *crawler* ha la facoltà, ma non l'obbligo, di seguire i consigli.

Figura 6.3 – Definizione dei livelli d’interesse per le diverse caratteristiche di un telefono cellulare

Definisci quanto sei interessato alle seguenti caratteristiche di un cellulare  
 tenendo premuto il tasto sinistro del mouse sposta la caratteristica in uno dei rettangoli vuoti, posti in basso

**Legenda**  
**Sistema Operativo:**  
 utilizzare il cellulare come se fosse un minipc

Radio Am/Fm	Videochiamata	Autonomia Batteria	Lettores Mp3
Internet	Conness. senza fili	Sistema Operativo	Touch Screen
E-mail			
Qualità Display		Spessore	Memoria interna
Dimens. Display	Videocamera		Memoria espandibile



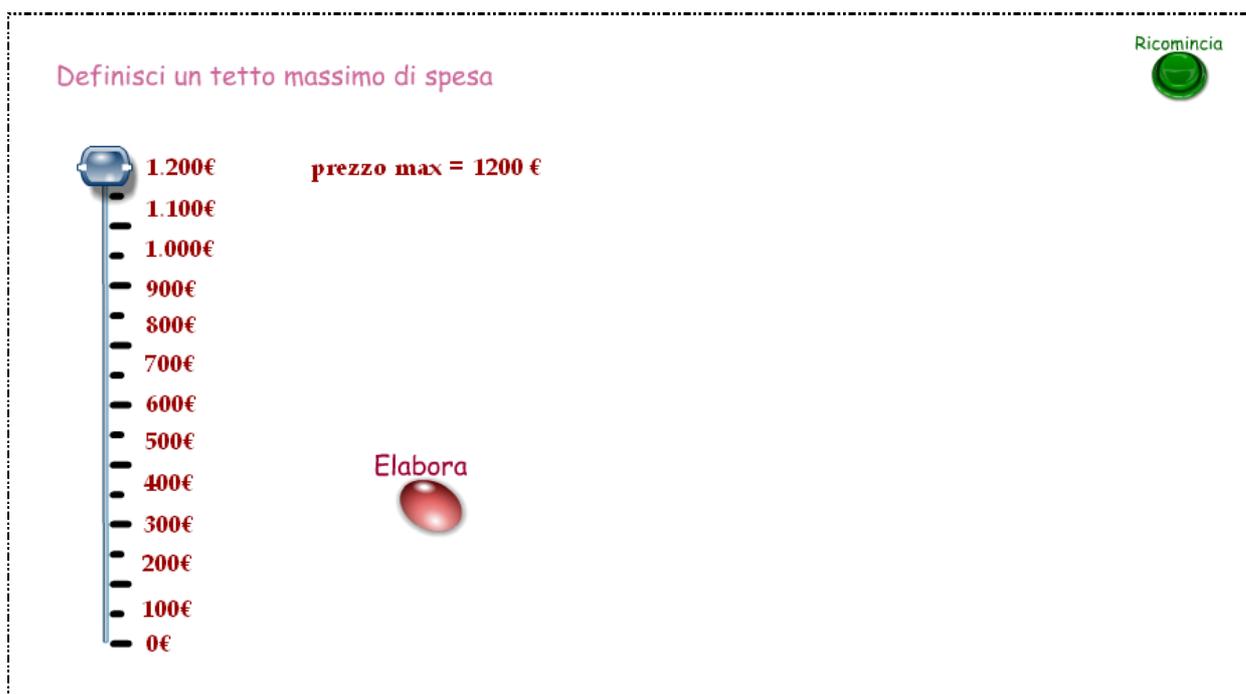
Per niente interessato      Poco interessato      Mediamente interessato      Abbastanza interessato      Molto interessato

		Tastiera Qwerty Risoluz. Fotocamera	Peso	Navigatore Dimensione
--	--	--	------	--------------------------

▶▶

Nella terza fase l’utente stabilisce il suo tetto massimo di spesa con la possibilità quindi di escludere alternative di prodotto troppo onerose (Figura 6.4).

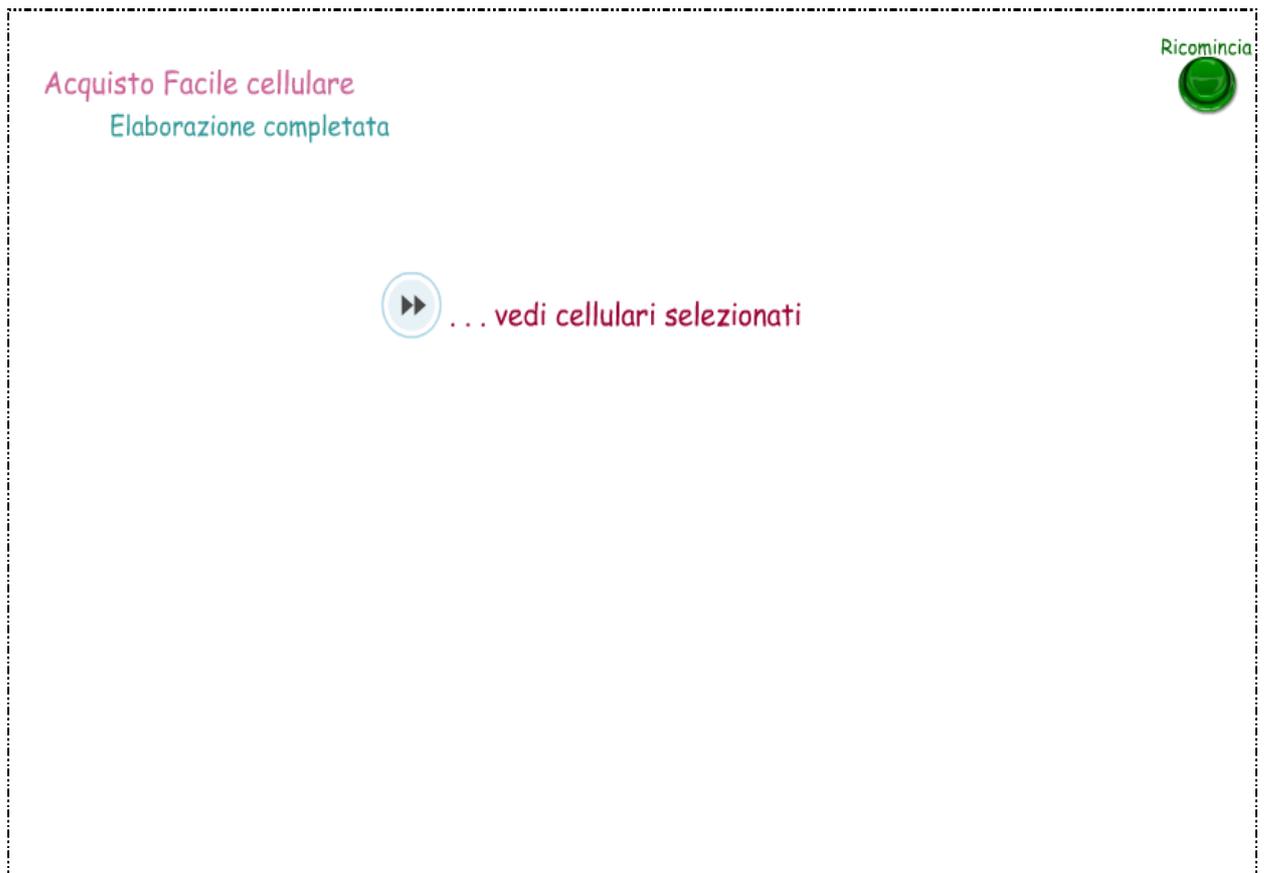
Figura 6.4 – Definizione tetto massimo di spesa



Nella terza fase l'applicazione "Acquisto Facile", procede all'elaborazione, tenendo conto del grado di *performance* ritenuto soddisfacente per l'utente per ogni caratteristica del prodotto e delle *performance* delle caratteristiche ritenute rilevanti dei cellulari presenti nel *dataset*; l'applicazione minimizza la somma degli scarti negativi tra *performance* effettive delle caratteristiche di ogni specificazione di prodotto e livelli di *performance* richiesti dall'utente. Questo sistema riduce in modo radicale la possibilità di compensazione di significative carenze in alcune caratteristiche con *performance* particolarmente elevate da altri punti di vista.

La *performance* minima richiesta dall'utente viene valutata dall'applicazione secondo i seguenti valori: 0 (nulla), 25 (bassa), 50 (media), 75 (elevata), 100 (molto elevata). Gli indicatori percentuali di *performance* di ogni caratteristica, invece, vengono valutati attribuendo il valore 100 alla *performance* migliore, il valore 50 alla *performance* mediana, il valore zero alla *performance* peggiore; gli altri valori sono ottenuti per interpolazione lineare. Nel caso dell'indicatore di dimensione, per tener conto del fatto che, a parità di altre condizioni, la *performance* è inversamente proporzionale alla dimensione, l'indicatore percentuale di *performance* è ottenuto attribuendo il valore 100 al cellulare di dimensione minima, il valore zero al cellulare di dimensione massima, il valore 50 al cellulare di dimensione mediana, calcolando tutti gli altri valori per interpolazione lineare; in modo analogo sono ottenuti i valori degli indicatori percentuali di *performance* relativi al peso dei cellulari. La Figura 6.5 mostra la pagina web che avvisa l'utente quando l'elaborazione è completata.

Figura 6.5 - Pagina web che avvisa l'utente quando l'elaborazione è completata.



Nella fase conclusiva l'applicazione “*Acquisto Facile*” seleziona un gruppo ristretto di cellulari che presentano caratteristiche con *performance* prossime a quelle richieste dal consumatore.

La Figura 6.6 mostra i cellulari selezionati mettendo a confronto le performance per ogni attributo. La presente pagina dei risultati evidenzia inoltre, anche un indicatore di convenienza all'acquisto (*Differenza Qualità/Prezzo*), ottenuto sottraendo per i diversi telefonini, l'indice standardizzato di prezzo dall'indice standardizzato di performance.

Figura 6.6 – Elencazione dei cellulari che meglio corrispondono alle esigenze e aspirazioni dell'utente



**ACQUISTO FACILE  
CELLULARE**

[mostra tabella senza dettagli](#)

Stabilisci un prezzo max: .00 €

Ricevi questa pagina dopo aver risposto ad un breve questionario **VEDI**

**Importanza data:**

per niente ★★★★★

poco ★★★★★

mediamente ★★★★★

abbastanza ★★★★★

molto ★★★★★

**Performance caratteristiche:**

0% ★★★★★ <60% ★★★★★

<10% ★★★★★ <70% ★★★★★

<20% ★★★★★ <80% ★★★★★

<30% ★★★★★ <90% ★★★★★

<40% ★★★★★ <100% ★★★★★

<50% ★★★★★

						
Prezzo max stabilito: 400 €	Immagini non disponibili					
Marchio	Samsung	Samsung	Samsung	Nokia	Sony Ericsson	Asus
Modello	SGH-i620	SGH I560	SGH-i710	E61i	P990i	M530w
Performance	91.2%	88.8%	87.1%	81.5%	81.5%	81.5%
Prezzo standard	38.6876%	39.667%	43.9765%	45.8374%	42.7032%	41.1851%
 Differenza Qualità/Prezzo	OK 1.584	OK 1.50352	OK 1.91625	OK 1.67334	OK 1.26617	OK 1.06894
Prezzo	★★★★★ 400€	★★★★★ 380€	★★★★★ 292€	★★★★★ 254€	★★★★★ 318€	★★★★★ 349€
★★★★★ Tastiera Qwerty	✓	✗	✗	✓	✓	✓
★★★★★ Internet	★★★★★ UMTS	★★★★★ GPRS	★★★★★ GPRS	★★★★★ UMTS	★★★★★ GPRS	★★★★★ UMTS
★★★★★ E-mail	✓	✓	✓	✓	✓	✓
★★★★★ Qualità display	★★★★★ 65000	★★★★★ 262144	★★★★★ 262000	★★★★★ 16000000	★★★★★ 262000	★★★★★ 65000
★★★★★ Dimensione display	★★★★★ 76800	★★★★★ 76800	★★★★★ 76800	★★★★★ 76800	★★★★★ 76800	★★★★★ 76800
★★★★★ Touch screen	✗	✗	✓	✗	✗	✓
★★★★★ Connettività	★★★★★ Wi-Fi	★★★★★ Wi-Fi	★★★★★ Wi-Fi	★★★★★ Wi-Fi	★★★★★ Wi-Fi	★★★★★ Wi-Fi
★★★★★ Risoluz. fotocamera	★★★★★ 1920000	★★★★★ 3145728	★★★★★ 1920000	★★★★★ 1920000	★★★★★ 1920000	★★★★★ 1920000
★★★★★ Antenna GPS	✓	✓	✓	✗	✗	✗

Nota: Gli indicatori percentuali di *performance* sono ottenuti attribuendo il valore 100 alla *performance* migliore, il valore 50 alla *performance* mediana, il valore zero alla *performance* peggiore; gli altri valori sono ottenuti per interpolazione lineare. Nel caso dell'indicatore di dimensione, per tener conto del fatto che, a parità di altre condizioni, la *performance* è inversamente proporzionale alla dimensione, l'indicatore percentuale di *performance* è ottenuto attribuendo il valore 100 al cellulare di dimensione minima, il valore zero al cellulare di dimensione massima, e il valore 50 al cellulare di dimensione mediana; in modo analogo sono ottenuti i valori degli indicatori percentuali di *performance* relativi al peso dei cellulari.

#### 6.4 Valutazione del sistema di aiuto alle decisioni “Acquisto Facile”

L'applicazione, prima di essere pubblicata online, è stata testata nella scelta del cellulare più adatto tra quelli disponibili presso il reparto di telefonia di un centro commerciale della provincia di Cosenza. Su Internet si hanno tantissime opzioni di scelta provenienti dai diversi siti e-commerce; in questo ambito è quindi potenzialmente più efficace un agente di raccomandazione. In un processo decisionale online è più facile che prevalga l'aspetto cognitivo rispetto a quello emotivo, poiché l'utente si pone in genere un maggior numero di domande sulla convenienza del proprio acquisto alle quali l'agente dovrà trovare risposta. I risultati dell'indagine qui presentata mostrano che anche off-line è possibile ottenere buoni risultati per il consumatore. Un vantaggio da non sottovalutare, rispetto ad un acquisto on-line, è la possibilità

di vedere e provare immediatamente le alternative filtrate dal *database*, e quindi valutare meglio la convenienza dell'acquisto effettuato.

Il test dell'agente intelligente qui proposto all'interno del reparto di telefonia di un grande centro commerciale in provincia di Cosenza (Emmezeta) ha consentito di scegliere il prodotto multi-attributo più popolare e tecnologicamente evoluto (un cellulare), in presenza di una molteplicità di opzioni.

#### **6.4.1 Metodologia e risultati dell'indagine**

È stato chiesto ai consumatori intenti ad osservare i cellulari all'interno delle vetrine, di indicare attraverso un indice di gradimento che va da "per niente interessato" a "molto interessato", la loro preferenza per ogni caratteristica relativa ai cellulari disponibili all'interno dello stand. Alla fine il programma individuava un gruppo limitato di opzioni elencate in ordine di valore complessivo decrescente. Dopodichè i consumatori avevano la possibilità di eliminare tra le alternative filtrate i cellulari che non piacevano esteticamente e quelli che superavano il costo massimo sostenibile. Circa il 41% dei consumatori che avevano intenzione di acquistare un cellulare, acquistarono uno dei prodotti filtrati dal programma. La tabella 4.1 mostra le 40 caratteristiche presentate ai consumatori e le relative percentuali di intervistati che si sono pronunciati sull'interessamento personale per ogni attributo del prodotto. L'ultima colonna della tabella mostra quali sono le caratteristiche preferite dagli intervistati (*smile* positiva e colore verde), mentre non sono per niente interessati alla visualizzazioni e modifica di documenti di testo e fogli elettronici (*PDA/organizer*), alle mascherine intercambiabili e disinteressati alle offerte dei gestori telefonici (Tim, Vodafone, Tre, mentre la Wind non presenta offerte negli stand MZ). Sono stati inoltre esaminati i risultati, verificando la presenza o meno del cosiddetto *inclusion effect* (la presenza di prodotti che presentano alte performance per alcune caratteristiche per le quali gli intervistati indicano valori minimi di preferenza potrebbe pregiudicare la bontà dei risultati).

Tabella 6.1 – Percentuale di intervistati che si sono pronunciati sul grado d'interesse per i diversi attributi di un telefono cellulare.

	Per niente interesa- sato	Poco interesa- sato	Mediamen- te interessato	abbastanza interessato	molto interesa- sato
Ricezione/Invio messaggi MMS	20%	5%	25%	30%	20%
Videomessaggio	20%	10%	10%	40%	20%
Funzione <i>chat</i>	<b>75%</b>	15%	0%	10%	0%
Connessione a Internet	<b>45%</b>	5%	10%	15%	25%
Invio/Ricezione e-mail	35%	10%	5%	25%	25%
La TV sul cellulare	<b>65%</b>	15%	10%	5%	5%
N° colori display interno	5%	0%	10%	<b>50%</b>	35%
Risoluz. display interno in n° pixel	5%	0%	5%	<b>50%</b>	<b>40%</b>
Connettività tramite USB	15%	10%	5%	30%	40%
connettività senza fili	15%	0%	5%	<b>45%</b>	35%
Risoluz. max fotocamera in n° pixel	5%	5%	5%	<b>40%</b>	<b>45%</b>
Zoom digitale	5%	5%	5%	<b>60%</b>	25%
Flash	15%	5%	10%	<b>60%</b>	10%
Foto per numeri in rubrica	20%	20%	10%	30%	20%
Funzione videocamera	15%	10%	5%	<b>45%</b>	25%
Lettore Multimediale	10%	10%	10%	<b>50%</b>	20%
Musica in formato MP3	10%	5%	10%	<b>45%</b>	30%
PDA/organizer	<b>55%</b>	5%	15%	15%	10%
Gusci e mascherine intercambiabili	<b>80%</b>	15%	5%	0%	0%
Videochiamate	20%	15%	10%	25%	30%
Radio stereo AM/FM	<b>40%</b>	0%	30%	15%	15%
Vivavoce integrato	0%	0%	30%	25%	<b>45%</b>
Chiamata a vibrazione	5%	5%	15%	25%	<b>50%</b>
Tipo di telefono in Mhz	30%	10%	10%	30%	20%
Dimensioni in mm3	0%	0%	15%	<b>40%</b>	<b>45%</b>
Peso con batteria in grammi	10%	0%	20%	30%	<b>40%</b>
Tempo di durata in stand-by in ore	5%	0%	5%	25%	<b>65%</b>
Tempo di durata in conversaz. in ore	5%	0%	5%	30%	<b>60%</b>
Chiamata a comando vocale	<b>40%</b>	15%	15%	25%	5%
<i>Voice Note</i> (registratore vocale)	15%	15%	20%	25%	25%
Memoria totale in Mb	5%	0%	15%	<b>40%</b>	<b>40%</b>
Memoria espandibile	10%	5%	10%	<b>40%</b>	<b>35%</b>
<i>Design</i> a conchiglia	30%	5%	20%	30%	15%
N° colori display esterno	30%	5%	15%	<b>40%</b>	10%
Risoluz. display esterno in n° pixel	25%	5%	15%	<b>45%</b>	10%
Apertura a scorrimento	30%	5%	15%	25%	25%
Gestore Tim	<b>80%</b>	5%	5%	0%	10%
Gestore Vodafone	<b>55%</b>	5%	5%	15%	20%
Gestore Tre	<b>85%</b>	5%	5%	0%	5%

### 6.4.2 Il Questionario

Successivamente è stato chiesto ai clienti del centro commerciale di rispondere ad un breve questionario articolato in 4 sezioni. Nella prima sezione si chiedeva di indicare il grado di soddisfazione per i risultati del sistema di supporto decisionale proposto; i risultati sono riportati nella tabella 6.2.

Tabella 6.2 – Valutazione dei risultati del sistema di supporto alle decisioni di acquisto da parte gli intervistati.

<b>1.a) Come valuta il sistema di acquisto proposto?</b>	<i>per niente soddisfatto</i>	<i>poco soddisfatto</i>	<i>mediamente soddisfatto</i>	<i>abbastanza soddisfatto</i>	<i>molto soddisfatto</i>
Comprensione delle caratteristiche presenti nel prodotto	0%	0%	0%	54%	46%
Organizzazione grafica delle pagine del programma	0%	0%	0%	58%	42%
Dimensione dei caratteri	0%	0%	8%	46%	46%
Visualizzazione grafica dei prodotti filtrati	0%	0%	13%	63%	25%
Utilità del programma nella fase di pre-acquisto	0%	0%	8%	<b>42%</b>	<b>50%</b>
Soddisfazione per i prodotti selezionati	0%	4%	8%	<b>33%</b>	<b>54%</b>

Da una prima analisi si vede una notevole soddisfazione dei consumatori rispetto ai diversi aspetti del sistema di supporto decisionale a soggetti indecisi nella scelta del prodotto più adatto e conveniente. Infatti, se si sommano i risultati ottenuti per quanto riguarda le risposte più favorevoli (abbastanza soddisfatto, molto soddisfatto), si raggiunge, in media, circa il 93% degli intervistati. Ciò evidenzia che i consumatori sono consapevoli delle difficoltà esistenti nella scelta di un prodotto complesso, per il quale è necessario valutare numerose caratteristiche. Essi quindi ritengono utile l'aiuto di un assistente, che può essere anche un software, in grado di guidarli verso l'alternativa più adatta, al prezzo più vantaggioso. Anche le risposte relative alla "soddisfazione per i prodotti selezionati" sono molto favorevoli (si raggiunge l'87% degli intervistati), e ciò conferma la validità del sistema di supporto decisionale qui proposto.

Nella seconda parte del questionario (*Sezione b*) sono riportate le risposte dei consumatori ad alcune domande inerenti il prodotto esaminato.

#### *Sezione b:*

Tabella 6.3 – Numero di cellulari posseduti da ciascuna persona ( % di intervistati ) :

	0	1	2	3
	0%	38%	58%	4%

Secondo la tabella 6.3), il 62% degli intervistati possiede almeno 2 cellulari, mentre nessuno dichiara di non possedere alcun cellulare.

Tabella 6.4 – Principali marche di cellulari ( % di intervistati ) :

<i>Nokia</i>	<i>Motorola</i>	<i>Samsung</i>	<i>Motorola e Nokia</i>
38%	46%	25%	21%

Nella Tabella 6.4 sono presentate le risposte inerenti le marche di cellulari maggiormente acquistate. Modelli Motorola e Nokia sono posseduti dalla maggioranza degli intervistati; il 21% degli intervistati possiedono un cellulare Nokia e un cellulare Motorola. In genere gli intervistati si trovano in difficoltà nel ricordare il modello di cellulare posseduto, ma ricordano benissimo la marca; ciò conferma l'importanza della *brand image*. Nella *brand image* è possibile rintracciare una componente cognitiva, il *brand territory*, che distingue ciascuna marca dalle concorrenti sul piano della qualità e dei *benefit*, una componente antropomorfa, la *brand personality*, che fa considerare il *brand* come fosse un individuo dotato di personalità, ed una componente relazionale, la *brand relationship* che indica il vissuto del rapporto con la marca [Rutelli, Siri, 2004].

Tabella 6.5 – Grado di soddisfazione per il cellulare posseduto ( % di intervistati ) :

E' soddisfatto/a del suo/suoi cellulare/i?		
<i>Si</i>	<i>No</i>	<i>Non risponde</i>
83%	8%	8%

Alla terza domanda della sezione b l'83% degli intervistati risponde di essere soddisfatto dei propri cellulari (Tabella 6.5). Questo risultato si può spiegare forse con il fatto che il cellulare serve essenzialmente per telefonare e alla fine ognuno si accontenta di quello che ha. Invece l'8% degli intervistati sono insoddisfatti, perché hanno avuto problemi di malfunzionamento del telefono e di lentezza nell'assistenza; ciò ovviamente inciderà negativamente sulla *brand relationship*.

Tabella 6.6 – Intenzioni di acquisto ( % di intervistati ) :

Aveva in mente di acquistare un cellulare in particolare?	
<i>Si</i>	<i>No</i>
71%	29%

La tabella 6.6 evidenzia che il 71% degli intervistati aveva in mente di acquistare un cellulare in particolare.

Tabella 6.7 – Motivazioni per l'acquisto di un telefono cellulare ( % di intervistati ) :

Per quale motivo intende acquistare un telefonino?			
<i>il mio cell non ha . .</i>	<i>cell aggiuntivo</i>	<i>per cambiarlo</i>	<i>regalo</i>
44%	39%	11%	6%

In base alla Tabella 6.7, il 44% di quelli che hanno intenzione di acquistare un telefonino sono interessati ad un cellulare che presenti alcune caratteristiche assenti nei telefonini posseduti, mentre il 39% necessita di in altro cellulare. E' interessante inoltre osservare che l'11% degli

intervistati, interessati alle novità ed alla curiosità di ispezionare gli ultimi modelli, desiderano sostituire il vecchio telefonino.

Tabella 6.8 – Altri prodotti elettronici posseduti ( % di intervistati ) :

Quali altri prodotti elettronici possiede?	
<i>Notebook</i>	38%
<i>Agenda elettronica</i>	8%
<i>Lettore Mp3</i>	63%
<i>Palmare</i>	4%
<i>Lettore Cd</i>	58%
<i>Telecamera</i>	42%
<i>Fotocamera digitale</i>	63%
<i>Navigatore satellitare</i>	17%

La seconda parte del questionario si conclude chiedendo all'intervistato se possiede altri prodotti elettronici. La maggior parte degli intervistati possiede un lettore Mp3, un lettore Cd, ed una fotocamera digitale.

La terza parte (Sezione c) del questionario contiene domande connesse con Internet.

#### **Sezione c:**

La tabella 6.9 evidenzia che il 46% degli intervistati usa internet più di 15 ore al mese, mentre un 30% di intervistati naviga saltuariamente (da 2 a 10 ore al mese).

Tabella 6.9 – Ore mensili di utilizzo di internet ( % di intervistati ) :

Quante ore naviga mensilmente su internet?					
<i>Mai</i>	<i>0-2</i>	<i>&gt;2-5</i>	<i>&gt;5-10</i>	<i>&gt;10-15</i>	<i>&gt;15</i>
13%	8%	13%	17%	4%	46%

Tabella 6.10 – Acquisti su internet ( % di intervistati ) :

Cosa acquista su internet?		
<i>Nessun prodotto</i>	60%	
<i>elettronica</i>	44%	tra quelli che acquistano
<i>libri</i>	11%	/// ///
<i>musica</i>	11%	/// ///
<i>biglietti concerti/festival</i>	22%	/// ///
<i>biglietti viaggio</i>	33%	/// ///
<i>abbigliamento</i>	0%	/// ///
<i>arredamento</i>	11%	/// ///
<i>altro</i>	11%	/// ///

Una ricerca condotta da *marketingroutes.com*, avente come target di riferimento i soggetti che avevano comprato, almeno una volta, un qualsiasi prodotto online, ha mostrato che essi

hanno incontrato delle difficoltà a trovarli. Ancora oggi moltissime persone hanno difficoltà a comprare online. Molti non si pongono neppure problema, (rispondendo “non mi interessa” o “non ci ho mai pensato”) e concludono dicendo “Perché comprare on-line quando ci sono i negozi con delle belle vetrine e commesse/i disponibili a darmi una mano?”. Non c’è quindi da stupirsi se il 60% degli intervistati (tabella 6.10) afferma di non aver mai acquistato on-line. Tra gli intervistati che acquistano online, la maggior parte compra prodotti elettronici e biglietti di vario genere.

Le principali difficoltà incontrate negli acquisti online sono evidenziate nella tabella 6.11.

Tabella 6.11 - Principali ostacoli agli acquisti online( % di intervistati) :

Principali ostacoli agli acquisti on line (massimo 3 risposte)	
<i>rischio nei pagamenti online</i>	77%
<i>non si può toccare con mano il prodotto</i>	64%
<i>poca informazione</i>	18%
<i>Scarsa fiducia nei venditori</i>	36%
<i>prezzi dei prodotti elevati</i>	5%
<i>spese di trasporto elevate</i>	18%
<i>altro</i>	0%

Secondo gli intervistati, le principali difficoltà riscontrabili quando si acquista online sono la rischiosità percepita dei pagamenti e l’impossibilità di toccare con mano i prodotti. Per molte persone è fondamentale guardare, toccare, farsi trasportare dai sensi, anche perché si ha la sensazione di aver un maggior controllo sulla decisione di acquisto, pensando che così sarà più difficile venir ingannati. Una cospicua parte degli intervistati è scettica sull’affidabilità dei venditori sul web; per cercare di ridurre questi timori, Ebay sta investendo molto per cercare di migliorare i rapporti tra venditori e acquirenti, riducendo il rischio di truffe.

Tabella 6.12 – Principali vantaggi dell’e-commerce ( % di intervistati) :

Vantaggi dell’e-commerce (massimo 3 risposte)	
<i>acquisto quando voglio</i>	41%
<i>acquisto da dove voglio, basta un terminale</i>	32%
<i>facilita il confronto tra prodotti simili</i>	36%
<i>risparmio di tempo, rapidità nell'acquisto</i>	32%
<i>migliori offerte di acquisto</i>	27%
<i>il prodotto arriva a domicilio</i>	27%
<i>altro</i>	0%

I vantaggi di acquistare online sono ben riconosciuti dagli intervistati, ma non c’è una netta preferenza tra le risposte. Questo può voler dire che molte persone potrebbero essere interessate, conoscendone i vantaggi, a comprare on-line. Purtroppo, però, spesso i vantaggi sono offuscati dalle difficoltà e dai rischi. Riferendoci anche a quanto detto da *marketingroutes.com*, oltre alle difficoltà sopraindicate, è opportuno aggiungere anche il problema della banda larga che scarseggia, la cattiva usabilità dei siti internet, e la carenza di mentalità imprenditoriale nel mondo delle vendite on-line.

La quarta sezione del questionario riguarda i dati anagrafici degli intervistati.

Tabella 6.13 – Distribuzione per sesso degli intervistati ( % di intervistati) :

	<i>Maschi</i> 62%	<i>Femmine</i> 38%
--	----------------------	-----------------------

Tabella 6.14 – Distribuzione per età degli intervistati ( % di intervistati) :

	<i>15-18</i>	<i>19-23</i>	<i>24-28</i>	<i>29-35</i>	<i>36-45</i>	<i>46-60</i>	<i>&gt;60</i>
	8%	4%	25%	38%	17%	8%	0%

Tabella 6.15 - Residenza degli intervistati( % di intervistati) :

	<i>centro</i> 25%	<i>periferia</i> 29%	<i>provincia</i> 46%
--	----------------------	-------------------------	-------------------------

Tabella 6.16 – Livello d’istruzione degli intervistati ( % di intervistati) :

	<i>Media inferiore</i>	<i>Diploma di maturità</i>	<i>Diploma di Laurea</i>	<i>abilitazione e specializzazione</i>	<i>Dottorato di ricerca</i>
	17%	33%	33%	8%	8%

Tabella 6.17 – Professione degli intervistati ( % di intervistati):

	<i>studente</i>	<i>Impiegato pubblico</i>	<i>Dirigente</i>	<i>In cerca di occupazione</i>	<i>Docenti</i>	<i>Medici</i>	<i>altro</i>
	17%	17%	13%	15%	8%	7%	33%

Le tabelle da 6.13 a 6.17) presentano i risultati riguardanti le principali caratteristiche anagrafiche degli intervistati, il loro livello di istruzione e la professione; da una prima analisi si nota la varietà del campione intervistato. La metà degli intervistati sono in possesso di laurea e tra questi il 17% possiede anche titoli d’istruzione post-laurea. Inoltre non bisogna farsi trarre in inganno da quel 17% di intervistati che dichiarano di possedere soltanto il titolo di scuola media, poiché una cospicua parte di essi sta frequentando un istituto d’istruzione superiore.

## 6.5 Alcuni approfondimenti sulla metodologia di supporto alle decisioni

Prima di scegliere la metodologia che minimizza i gap negativi fra performance richieste e performance effettive, per poi integrarla nell’applicazione realizzata, sono state effettuate alcune verifiche, utilizzando altre metodologie decisionali, con diverse variazioni nelle tecniche di selezione delle opzioni. Le tabelle da 6.18 a 6.22 evidenziano le opzioni di cellulari selezionate applicando altre metodologie decisionali.

Nella tabella 6.18 sono evidenziati i telefonini selezionati in ordine decrescente dal punto di vista della differenza semplice fra l’indicatore sintetico standardizzato di performance ottenuto in base alla media ponderata dei valori degli indicatori di performance dei diversi attributi e l’indicatore standardizzato di prezzo. Nella quarta riga della tabella è riportato l’indicatore standardizzato di performance, nella sesta riga l’indicatore standardizzato di prezzo, nella settima riga la differenza semplice fra l’indicatore standardizzato di performance e l’indicatore standardizzato di prezzo. Nelle righe successive sono riportati i valori degli indicatori di performance dei telefonini in termini dei diversi attributi considerati.

Tabella 6.18 – Telefoni cellulari selezionati in ordine decrescente dal punto di vista della differenza semplice fra l'indicatore sintetico standardizzato di performance ottenuto in base alla media ponderata dei valori degli indicatori di performance dei diversi attributi e l'indicatore standardizzato di prezzo.

	Marchio	Motorola	Lg	Lg	Momo design	Lg					
	Modello	E770	U310	U300	MD1	L600V					
	Performance	61.7%	70.5%	69.7%	65.1%	69.6%					
	Perform_standardiz.	0.814676	1.24607	1.20685	0.98135	1.20195					
	Prezzo	49 €	99 €	99 €	99 €	169 €					
	Prezzo standardiz.	-1.25729	-0.739642	-0.739642	-0.739642	-0.0149341					
	Differenza Qualità/Prezzo	2.07197	1.98571	1.94649	1.72099	1.21688					
★★★★☆	Presenza_MMS	✓	✓	✓	✓	✓					
★★★★★	Videomessaggio	✓	✓	✓	✗	✓					
★★★★★	Internet	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS
★★★★☆	E-mail	✓	✓	✓	✓	✓					
★★★★☆	Colori_Display	80%	262144	80%	262000	80%	262000	80%	262144	80%	262144
★★★★☆	Risoluzione_Display	77%	38720	77%	38720	77%	38720	77%	38720	77%	38720
★★★★☆	Connettività	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth
★★★★★	Risoluz_fotocamera	16%	307200	66%	1310720	66%	1310720	66%	1310720	66%	1310720
★★★★★	Zoom_digitale	42%	4X	14%	2X	14%	2X	28%	2X	14%	2X
★★★★★	Flash	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
★★★★☆	Videocamera	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
★★★★☆	Lettore_mp3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
★★★★☆	Pda_Organizer	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
★★★★☆	Radio_Am/Fm	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
★★★★☆	Vivavoce	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
★★★★★	Videochiamata	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
★★★★☆	Dimensioni in mm3	30%	107226	50%	84020	32%	104489	64%	70993	48%	86316
★★★★☆	Peso in gr.	41%	108	60%	89	53%	93	46%	100	50%	95
★★★★★	Autonomia_stand_by	44%	200ore	48%	220ore	48%	220ore	0%	0ore	62%	280ore
★★★★★	Autonomia_conversaz.	33%	3ore	20%	2ore	20%	2ore	0%	0ore	38%	4ore
★★★★★	Memoria Totale	53%	16	46%	8	46%	8	61%	25	38%	7
★★★★★	Memoria espandibile	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
★★★★☆	Design a conchiglia	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Apertura a scorrimento	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗

Importanza attribuita dai consumatori a ciascuna caratteristica:

★★★★☆ nulla, ★★★★★ bassa, ★★★★★ Media, ★★★★★ elevata, ★★★★★ Molto elevata.

Disponibilità dell'attributo nel telefonino considerato: ✓ Presente, ✗ Assente

Nella tabella 6.19 sono riportati i primi sei telefoni cellulari selezionati in ordine decrescente dal punto di vista della differenza semplice fra l'indicatore sintetico standardizzato di performance ottenuto in base alla media ponderata dei valori dei logaritmi degli indicatori di performance dei diversi attributi e l'indicatore standardizzato di prezzo. Nella quarta riga della

tabella è riportato l'indicatore standardizzato di performance, nella sesta riga l'indicatore standardizzato di prezzo, nella settima riga la differenza semplice fra l'indicatore standardizzato di performance e l'indicatore standardizzato di prezzo. Nelle righe successive sono riportati i valori degli indicatori di performance dei telefonini in termini dei diversi attributi considerati.

Tabella 6.19 – Telefoni cellulari in ordine decrescente dal punto di vista della differenza semplice fra l'indicatore sintetico standardizzato di performance ottenuto in base alla media ponderata dei valori dei logaritmi degli indicatori di performance dei diversi attributi e l'indicatore standardizzato di prezzo.

	Marchio	Motorola		Lg		Lg		Momo design		Lg		Nokia	
	Modello	E770		U310		U300		MD1		L600V		6070	
	Performance	72%		79.2%		78.8%		73.2%		78.8%		60.4%	
	Perform_standardiz.	0.833251		1.18263		1.16322		0.89148		1.16322		0.270364	
	Prezzo	49 €		99 €		99 €		99 €		169 €		89 €	
	Prezzo_standardiz.	-1.25729		-0.739642		-0.739642		-0.739642		-0.0149341		-0.843171	
	Differenza Qualità/Prezzo	2.09054		1.92227		1.90286		1.63112		1.17815		1.11354	
★★★★☆	Presenza_MMS	✓		✓		✓		✓		✓		✓	
★★★★☆	Videomessaggio	✓		✓		✓		✗		✓		✗	
★★★★☆	Internet	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS	50%	EDGE
★★★★☆	E-mail	✓		✓		✓		✓		✓		✓	
★★★★☆	Colori_Display	80%	262144	80%	262000	80%	262000	80%	262144	80%	262144	60%	65536
★★★★☆	Risoluzione_Display	77%	38720	77%	38720	77%	38720	77%	38720	77%	38720	55%	20480
★★★★☆	Connettività	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	50%	IrDA
★★★★☆	Risoluz_fotocamera	16%	307200	66%	1310720	66%	1310720	66%	1310720	66%	1310720	16%	307200
★★★★☆	Zoom_digitale	42%	4X	14%	2X	14%	2X	28%	2X	14%	2X	28%	2X
★★★★☆	Flash	✗		✗		✗		✗		✗		✗	
★★★★☆	Videocamera	✓		✓		✓		✓		✓		✓	
★★★★☆	Lettore_mp3	✓		✓		✓		✓		✓		✓	
★★★★☆	Pda_Organizer	✗		✗		✗		✗		✗		✗	
★★★★☆	Radio Am/Fm	✗		✗		✗		✗		✗		✓	
★★★★☆	Vivavoce	✓		✓		✓		✓		✓		✓	
★★★★☆	Videochiamata	✓		✓		✓		✓		✓		✗	
★★★★☆	Dimensioni in mm3	30%	107226	50%	84020	32%	104489	64%	70993	48%	86316	47%	86847
★★★★☆	Peso in gr.	41%	108	60%	89	53%	93	46%	100	50%	95	62%	88
★★★★☆	Autonomia_stand_by	44%	200ore	48%	220ore	48%	220ore	0%	0ore	62%	280ore	66%	300ore
★★★★☆	Autonomia_conversaz.	33%	3ore	20%	2ore	20%	2ore	0%	0ore	38%	4ore	29%	3ore
★★★★☆	Memoria Totale	53%	16	46%	8	46%	8	61%	25	38%	7	30%	3
★★★★☆	Memoria espandibile	✓		✓		✓		✓		✓		✓	
★★★★☆	Design a conchiglia	✗		✓		✓		✓		✓		✗	

Nella tabella 6.20 sono riportati i primi sei telefoni cellulari in ordine decrescente dal punto di vista della differenza semplice fra l'indicatore sintetico standardizzato di performance ottenuto in base alla minimizzazione della somma degli scarti negativi dei valori degli indicatori di performance dei diversi attributi rispetto ai valori indicati come soddisfacenti e l'indicatore standardizzato di prezzo.

Nella quarta riga della tabella è riportato l'indicatore standardizzato di performance, nella sesta riga l'indicatore standardizzato di prezzo, nella settima riga la differenza semplice fra l'indicatore standardizzato di performance e l'indicatore standardizzato di prezzo. Nelle righe

successive sono riportati i valori degli indicatori di performance dei telefonini in termini dei diversi attributi considerati.

Tabella 6.20 – Telefoni cellulari in ordine decrescente dal punto di vista della differenza semplice fra l'indicatore sintetico standardizzato di performance ottenuto in base alla minimizzazione della somma degli scarti negativi dei valori degli indicatori di performance dei diversi attributi rispetto ai valori indicati come soddisfacenti e l'indicatore standardizzato di prezzo.

	Marchio	Motorola	Lg	Lg	Momo design	Lg	Nokia						
	Modello	E770	U310	U300	MD1	L600V	6070						
	Performance	72%	79.2%	78.8%	73.2%	78.8%	60.4%						
	Perform_standardiz.	0.740349	0.982003	0.952533	0.522271	1.05862	-0.102493						
	Prezzo	49 €	99 €	99 €	99 €	169 €	89 €						
	Prezzo_standardiz.	-1.25729	-0.739642	-0.739642	-0.739642	-0.0149341	-0.843171						
	Differenza Qualità/Prezzo	2.09054	1.92227	1.90286	1.63112	1.17815	1.11354						
★★★★☆	Presenza_MMS	✓	✓	✓	✓	✓	✓						
★★★★☆	Videomessaggio	✓	✓	✓	✗	✓	✗						
★★★★☆	Internet	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS	50%	EDGE
★★★★☆	E-mail	✓	✓	✓	✓	✓	✓						
★★★★☆	Colori_Display	80%	262144	80%	262000	80%	262000	80%	262144	80%	262144	60%	65536
★★★★☆	Risoluzione_Display	77%	38720	77%	38720	77%	38720	77%	38720	77%	38720	55%	20480
★★★★☆	Connettività	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth	50%	IrDA
★★★★☆	Risoluz_fotocamera	16%	307200	66%	1310720	66%	1310720	66%	1310720	66%	1310720	16%	307200
★★★★☆	Zoom_digitale	42%	4X	14%	2X	14%	2X	28%	2X	14%	2X	28%	2X
★★★★☆	Flash	✗	✗	✗	✗	✗	✗						
★★★★☆	Videocamera	✓	✓	✓	✓	✓	✓						
★★★★☆	Lettore_mp3	✓	✓	✓	✓	✓	✓						
★★★★☆	Pda_Organizer	✗	✗	✗	✗	✗	✗						
★★★★☆	Radio Am/Fm	✗	✗	✗	✗	✗	✗						
★★★★☆	Vivavoce	✓	✓	✓	✓	✓	✓						
★★★★☆	Videochiamata	✓	✓	✓	✓	✓	✓						
★★★★☆	Dimensioni in mm3	30%	107226	50%	84020	32%	104489	64%	70993	48%	86316	47%	86847
★★★★☆	Peso in gr.	41%	108	60%	89	53%	93	46%	100	50%	95	62%	88
★★★★☆	Autonomia_stand_by	44%	200ore	48%	220ore	48%	220ore	0%	0ore	62%	280ore	66%	300ore
★★★★☆	Autonomia_conversaz.	33%	3ore	20%	2ore	20%	2ore	0%	0ore	38%	4ore	29%	3ore
★★★★☆	Memoria Totale	53%	16	46%	8	46%	8	61%	25	38%	7	30%	3
★★★★☆	Memoria espandibile	✓	✓	✓	✓	✓	✓						
★★★★☆	Design a conchiglia	✗	✓	✓	✓	✓	✓						

Le tabelle 6.18, 6.19, 6.20 evidenziano risultati simili; cellulari selezionati secondo metodi differenti hanno un prezzo molto conveniente, ma non rispecchiano del tutto le aspettative dell'utente. La selezione del Motorola E770 sarà dipesa molto dal bassissimo prezzo, solo 49 Euro, ma tenendo conto delle caratteristiche che maggiormente interessano l'utente considerato, notiamo che questo cellulare dispone di una fotocamera con scarsissima risoluzione, scarso zoom digitale, non ha il flash, autonomia della batteria appena sufficiente e non presenta design a conchiglia. Stessa cosa per il cellulare Lg U310 che non ha il flash, ha uno scarsissimo zoom digitale e bassa autonomia della batteria.

Il metodo utilizzato per la selezione dei risultati di Tabella 6.20, rimane comunque uno dei metodi migliori, perché riesce a limitare le compensazioni tra performance positive e

performance negative, con il pericolo che il risultato venga inficiato dall'influenza di valori outliers.

La tabella 6.21 mostra i telefonini selezionati in ordine decrescente dal punto di vista dell'indicatore standardizzato di performance ottenuto in base alla minimizzazione della somma degli scarti negativi dei valori degli indicatori di performance dei diversi attributi rispetto ai valori indicati come soddisfacenti. Nella quarta riga della tabella è riportato l'indicatore standardizzato di performance, nella sesta riga l'indicatore standardizzato di prezzo, nella settima riga la differenza semplice fra l'indicatore standardizzato di performance e l'indicatore standardizzato di prezzo. Nelle righe successive sono riportati i valori degli indicatori di performance dei telefonini in termini dei diversi attributi considerati.

Tabella 6.21 - Telefonini in ordine decrescente dal punto di vista dell'indicatore standardizzato di performance ottenuto in base alla minimizzazione della somma degli scarti negativi dei valori degli indicatori di performance dei diversi attributi rispetto ai valori indicati come soddisfacenti (*performance standardizzata*).

Marchio	NOKIA		Nokia		Nokia		Nokia		Motorola		Sony Ericsson		Nokia	
Modello	N71		N73		N70		6288		V3XX		Z610i		N70	
Performance	83.2%		81%		78.6%		76.8%		75.6%		75.5%		75.4%	
Perform_standardiz.	1.94272		1.81306		1.6716		1.56551		1.49478		1.48889		1.48299	
Prezzo	359 €		379 €		269 €		279 €		299 €		249 €		279 €	
Prezzo standardiz.	1.95213		2.15919		1.02036		1.12389		1.33095		0.813303		1.12389	
Differenza Qualità/Prezzo	-0.00940416		-0.346131		0.651239		0.441617		0.16383		0.675584		0.359101	
★★★★★ Presenza_MMS	✓		✗		✓		✓		✓		✓		✓	
★★★★★ Videomessaggio	✓		✓		✓		✓		✓		✓		✓	
★★★★★ Internet	75%	UMTS	75%	UMTS	75%	UMTS	75%	UMTS	100%	HSDPA	75%	UMTS	75%	UMTS
★★★★★ E-mail	✓		✓		✓		✓		✓		✓		✓	
★★★★★ Colori Display	80%	262144	80%	262000	80%	262000	80%	262000	80%	262000	80%	262144	80%	262000
★★★★★ Risoluzione Display	88%	76800	88%	76800	77%	36608	88%	76800	88%	76800	77%	38720	77%	36608
★★★★★ Connettività	100%	IrDA+Bluetooth	100%	IrDA+Bluetooth	100%	Bluetooth	100%	IrDA+Bluetooth	100%	Bluetooth	100%	Bluetooth	100%	Bluetooth
★★★★★ Risoluz_fotocamera	83%	1920000	100%	3145728	83%	1920000	83%	1920000	66%	1310720	83%	1920000	83%	1920000
★★★★★ Zoom_digitale	100%	20X	100%	20X	100%	20X	85%	8X	85%	8X	14%	2X	100%	20X
★★★★★ Flash	✓		✓		✓		✓		✗		✗		✓	
★★★★★ Videocamera	✓		✓		✓		✓		✓		✓		✗	
★★★★★ Lettore_mp3	✓		✓		✓		✓		✓		✓		✓	
★★★★★ Pda_Organizer	✓		✓		✓		✗		✗		✗		✓	
★★★★★ Radio_Am/Fm	✓		✓		✓		✓		✗		✗		✓	
★★★★★ Vivavoce	✓		✓		✓		✓		✓		✓		✓	
★★★★★ Videochiamata	✓		✓		✓		✓		✓		✓		✓	
★★★★★ Dimensioni in mm3	9%	130246	34%	102410	13%	125707	39%	96503	54%	80549	43%	92120	13%	125707
★★★★★ Peso in gr.	19%	139	35%	116	28%	126	36%	115	46%	100	39%	110	28%	126
★★★★★ Autonomia_stand_by	48%	216ore	77%	350ore	53%	240ore	55%	250ore	73%	330ore	88%	400ore	53%	240ore
★★★★★ Autonomia_conversaz.	38%	4ore	38%	4ore	29%	3ore	38%	4ore	30%	3ore	64%	6ore	29%	3ore
★★★★★ Memoria Totale	84%	128	69%	42	61%	22	46%	10	69%	50	100%	1000	61%	22
★★★★★ Memoria espandibile	✓		✓		✓		✓		✓		✓		✓	
★★★★★ Design a conchiglia	✓		✗		✗		✗		✓		✓		✗	
★★★★★ Apertura a scorrimento	✗		✗		✗		✓		✗		✗		✗	

La tabella 6.21 mostra come il cellulare Nokia N71, con un valore negativo della differenza *Qualità - Prezzo* (-0,346), non sia un prodotto conveniente pur avendo una performance totale delle sue caratteristiche dell'83,2% . Il Nokia N70 è il primo cellulare che presenta un valore positivo della differenza *Qualità - Prezzo*, e presenta anche un ottimo indice di performance del 78,6 per cento; alcuni consumatori potrebbero rimanere comunque insoddisfatti della scelta se il prezzo pari a 269 € fosse molto più alto di quanto avrebbero voluto spendere; potrebbe quindi essere utile definire da parte del consumatore un eventuale tetto massimo di spesa.

La tabella 6.22 deriva da una metodologia simile a quella utilizzata per la tabella 6.21, ma presenta anche un tetto massimo di spesa.

Tabella 6.22 – Telefoni cellulari in ordine decrescente dal punto di vista dell'indicatore standardizzato di performance ottenuto in base alla minimizzazione della somma degli scarti negativi dei valori degli indicatori di performance dei diversi attributi rispetto ai valori indicati come soddisfacenti (*performance standardizzata*), con un tetto massimo di spesa di 200 euro.

Marchio	MOTOROLA	Nokia	LG	Samsung	Lg	Lg	Lg	Lg
Modello	V3x	6234	U900	SGH Z230	L600V	U310	U310 Duetto	U300
Performance	75.1%	71.4%	69%	68.9%	68.2%	66.9%	66.9%	66.4%
Perform_standardiz.	1.46531	1.24723	1.10578	1.09988	1.05862	0.982003	0.982003	0.952533
Prezzo	199 €	199 €	199 €	199 €	169 €	99 €	199 €	99 €
Prezzo standardiz.	0.295655	0.295655	0.295655	0.295655	-0.0149341	-0.739642	0.295655	-0.739642
Differenza Qualità/Prezzo	1.16966	0.951578	0.810122	0.804228	1.07356	1.72164	0.686348	1.69217
★★★★ Presenza_MMS	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Videomessaggio	✗	✓	✓	✓	✓	✓	✓	✓
★★★★ Internet	75% UMTS	25% GPRS	75% UMTS					
★★★★ E-mail	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Colori Display	80% 262144	80% 262144	80% 262144	80% 262144	80% 262144	80% 262000	80% 262000	80% 262000
★★★★ Risoluzione Display	88% 76800	88% 76800	88% 76800	77% 38720	77% 38720	77% 38720	77% 38720	77% 38720
★★★★ Connettività	100% Bluetooth	100% IrDA+Bluetooth	100% Bluetooth					
★★★★ Risoluz_fotocamera	83% 1920000	83% 1920000	50% 1228800	66% 1310720	66% 1310720	66% 1310720	66% 1310720	66% 1310720
★★★★ Zoom_digitale	85% 8X	85% 8X	14% 2X					
★★★★ Flash	✓	✗	✓	✗	✗	✗	✗	✗
★★★★ Videocamera	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Lettore_mp3	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Pda_Organizer	✗	✗	✗	✗	✗	✗	✗	✗
★★★★ Radio_Am/Fm	✗	✓	✗	✗	✗	✗	✗	✗
★★★★ Vivavoce	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Videochiamata	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Dimensioni in mm3	33% 102841	45% 89424	28% 108933	39% 96660	48% 86316	50% 84020	50% 84020	32% 104489
★★★★ Peso in gr.	29% 125	39% 110	39% 110	58% 90	50% 95	60% 89	60% 89	53% 93
★★★★ Autonomia_stand_by	55% 250ore	75% 340ore	28% 130ore	55% 250ore	62% 280ore	48% 220ore	48% 220ore	48% 220ore
★★★★ Autonomia_conversaz.	48% 5ore	38% 4ore	38% 4ore	38% 4ore	38% 4ore	20% 2ore	20% 2ore	20% 2ore
★★★★ Memoria Totale	76% 64	76% 64	69% 51	61% 28	38% 7	46% 8	46% 8	46% 8
★★★★ Memoria espandibile	✓	✓	✓	✓	✓	✓	✓	✓
★★★★ Design a conchiglia	✓	✗	✗	✓	✓	✓	✓	✓
★★★★ Apertura a scorrimento	✗	✗	✓	✗	✗	✗	✗	✗

La tabella 6.22 mostra come, stabilendo un tetto massimo di spesa di 200 euro, si è riusciti a individuare un prodotto conveniente, il Motorola V3X, soddisfacendo quasi del tutto le esigenze dell'utente; un limite di questo cellulare è tuttavia l'impossibilità di inviare videomessaggi; ma anche spendendo di più non si riuscirebbe comunque a soddisfare tutte le esigenze di ciascun consumatore, essendo il telefono cellulare un prodotto non ancora personalizzabile.

## 6.6 Conclusioni

Come evidenziato nei capitoli precedenti, l'analisi dei processi decisionali dei consumatori è stata oggetto negli ultimi decenni dell'attenzione di molti studiosi di economia, di psicologia cognitiva, e d'informatica.

Le decisioni di acquisto sono particolarmente complesse nel caso di prodotti caratterizzati da molti attributi, come per esempio i telefoni cellulari. Il fondamento teorico dell'analisi delle decisioni nel caso di prodotti caratterizzati da molti attributi è la teoria di Lancaster, secondo cui il valore complessivo di un bene per un consumatore deriva dal valore delle sue diverse caratteristiche. Le due principali applicazioni empiriche di questa teoria sono la metodologia dei

prezzi edonici, che cerca di stimare il valore delle diverse caratteristiche di un prodotto mediante un'analisi econometrica in cui il prezzo delle diverse versioni di un prodotto è la variabile dipendente e i valori delle diverse caratteristiche sono variabili esplicative, e la *conjoint analysis*, una metodologia di analisi proposta nell'ambito della psicologia matematica, e ampiamente utilizzata nelle ricerche quantitative di marketing quando non sono disponibili stime affidabili dei prezzi di mercato per le diverse versioni del prodotto.

In teoria, un procedimento razionale di valutazione delle diverse opzioni richiederebbe la conoscenza della relazione fra l'intensità in cui i diversi attributi sono presenti in ciascuna opzione e il valore per il consumatore della opzione stessa; ottenuta in questo modo una stima monetaria del valore per il consumatore delle diverse opzioni, si potrebbero confrontare questi valori con i rispettivi prezzi, e decidere di acquistare l'opzione a cui corrisponde il maggior surplus netto del consumatore; ma in concreto gran parte delle decisioni dei consumatori sono adottate attraverso delle procedure molto semplificate ("*euristiche*") che tengono conto a volte soltanto di una piccola parte delle informazioni necessarie per una decisione ottimale.

L'applicazione presentata in questo capitolo aiuta il consumatore a selezionare le opzioni più adatte a soddisfare le sue esigenze, tenendo presenti gran parte degli attributi che definiscono un prodotto, in questo caso un telefono cellulare. Ciò è possibile attraverso una metodologia di selezione dei prodotti che minimizza i gap negativi tra performance minima richiesta e performance effettiva per i diversi attributi. L'applicazione qui proposta presenta anche una nuova interfaccia grafica (Figura 6.3) che consente all'utente online di evidenziare gli attributi per lui più importanti, e quindi definire le sue preferenze con molta semplicità (interagendo con il mouse), evitando la confusione possibile quando gli attributi rilevanti sono molti. La pagina di output (Figura 6.6) presenta in ordine decrescente di preferenza le dieci opzioni selezionate dall'applicazione.

La prima versione dell'applicazione è stata testata presso il reparto di telefonia di un centro commerciale locale, suscitando molta curiosità, ed ottenendo valutazioni abbastanza soddisfacenti; il 41% dei consumatori interessati ad acquistare un cellulare, acquistarono uno dei prodotti filtrati dal programma. Alcuni portali di *shopping comparison* hanno espresso interesse per l'applicazione online presentata in questo capitolo<sup>68</sup>.

Per confermare la validità della metodologia decisionale adottata dall'applicazione qui proposta, nell'ultima parte di questo capitolo sono stati presentati i risultati ottenibili applicando altre metodologie decisionali per selezionare i telefoni cellulari che meglio possano corrispondere alle esigenze e aspirazioni dei consumatori; i risultati del confronto appaiono incoraggianti per l'applicazione qui proposta.

---

<sup>68</sup> In particolare, il Dr. Andrea Boscaro, country manager di Pangora Italia, azienda leader in servizi di ricerca e confronto di prodotti utilizzati dai migliori comparatori europei, ha espresso il suo apprezzamento con la seguente e-mail: "*Complimenti. Davvero una applicazione molto carina e ben fatta. Buon lavoro.*" Pangora, azienda leader in Europa nell'ambito dei servizi di ricerca e confronto di prodotti su internet, offre ai portali soluzioni di shopping completi e agli shop un'ampia rete di distribuzione per i loro prodotti. Pangora è presente in Germania, Francia, Austria, Gran Bretagna e Italia, con 70 portali partner e 2.600 negozi che utilizzano il suo know-how tecnologico per un totale di 14 milioni di offerte ogni giorno. Pangora Italia può contare nel nostro paese su accordi con portali quali Quotidiano.net, Lycos, Excite, Tuttogratis e realtà dello shopping comparison come BuyCentral, Costameno, Msn Shopping, Kataweb Shopping, Repubblica Shopping e Techeye (Mytech).

## Conclusioni

Come detto nell'introduzione, il principale obiettivo di questa tesi di dottorato è stato quello di analizzare il comportamento dei consumatori (*consumer behaviour*) durante la fase di scelta di un prodotto da acquistare. A questo scopo si è proceduto analizzando i processi decisionali in condizioni di incertezza e i diversi sistemi di supporto alle decisioni, tecnologici e non, oggi a disposizione di aziende e privati, che si basano su metodologie di "filtraggio" delle informazioni. E' stato, quindi, proposto un sistema di supporto alle decisioni di acquisto basato sul principio della minimizzazione degli scarti negativi fra i livelli di performance delle diverse caratteristiche di un prodotto e i livelli minimi di performance richiesti; lo stesso è stato sperimentato per l'acquisto di un telefono cellulare mediante un sito creato ad hoc "*SceltaGiusta.it*" e consultabile all'indirizzo <http://www.sceltagiusta.it>.

Pertanto, nel primo capitolo sono state esaminate le principali teorie economiche e psicologiche inerenti i processi decisionali, sottolineando come l'interazione tra diversi approcci disciplinari possa consentire di migliorare gli esiti dei processi decisionali. Le ricerche effettuate evidenziano che il problema di una eccessiva varietà di scelte ha cominciato ad apparire sempre più rilevante a partire dai primi anni duemila, soprattutto in connessione con lo sviluppo degli acquisti su Internet. Mentre nei negozi tradizionali esistono dei vincoli fisici che limitano il numero delle varietà di ciascun bene che possono essere mostrate ai potenziali acquirenti, nei siti di vendite online non esistono praticamente limiti, e possono essere rese disponibili quindi migliaia di varietà per ciascun bene. Diversi studi hanno evidenziato che una possibile reazione dei consumatori al sovraccarico informativo è il differimento della decisione di acquisto, per il timore di non riuscire ad assumere la decisione migliore; i consumatori, tuttavia, trovano difficile rinunciare a considerare una parte rilevante delle informazioni disponibili. Le principali "anomalie" cognitive, e quindi decisionali, legate al funzionamento della memoria umana sono, secondo McFadden le seguenti: l'intensità delle emozioni diminuisce nella memoria con il passar del tempo, le informazioni più recenti e significative sono quelle più presenti in memoria in ciascun istante, la ricostruzione in memoria di informazioni passate è influenzata dal contesto e dagli stimoli attuali, le coincidenze sono memorizzate in modo più forte e persistente, la nozione soggettiva del tempo tende a comprimere nella memoria i periodi lontani nel tempo. Secondo Misuraca, Fasolo e Cardaci, le difficoltà nelle decisioni di acquisto possono dipendere da fattori interni e da fattori esterni, rispetto ai soggetti che devono prendere la decisione. I fattori esterni, oltre al numero di opzioni, sono il numero di attributi che caratterizzano le diverse opzioni, il loro livello di dettaglio, le interrelazioni fra gli attributi. I fattori interni sono costituiti dallo stile decisionale di ciascuna persona, e dagli indicatori dell'importanza che per ciascuna persona hanno i diversi attributi.

La Dell, azienda leader nella produzione di computer, ha cercato di ovviare *all'information overload*, conquistando una posizione di leadership con un modello di produzione dei computer in base alle specificazioni richieste da ciascun cliente; i produttori di automobili, elettrodomestici e altri beni complessi stanno cercando di imitare, in parte, il modello di produzione e vendita sperimentato con successo da Dell, consistente essenzialmente in una sorta di "personalizzazione di massa" dei prodotti. Una nuova rivoluzione tecnologica che consente ai consumatori di prendere decisioni migliori in condizioni di incertezza è rappresentata dall'uso degli "agenti intelligenti", in grado di rendere automatiche le ricerche di informazioni, e di apprendere da modifiche dell'ambiente circostante.

Nel secondo capitolo della tesi si parla d'intelligenza artificiale, settore in pieno sviluppo, soprattutto in ambienti online dove tutte le informazioni sono a portata di mouse. In particolar modo, alcune ricerche presentate in questo capitolo evidenziano che anche nelle vendite online gli agenti intelligenti sono sempre più utili per aiutare l'utente/consumatore ad effettuare scelte di acquisto in grado di corrispondere meglio alle esigenze ed aspirazioni di ciascuno. In generale, l'interazione con sistemi automatici di elaborazione delle informazioni coinvolge oggi tramite

Internet una parte significativa della popolazione in tanti paesi; in questo ambito, agenti intelligenti artificiali riescono a muoversi e interagire, anche con esseri umani. Molte ricerche, svolte sia nelle università che in imprese, pubbliche e private, nel settore della *new economy* e del commercio elettronico, hanno evidenziato la grande rilevanza di questo fenomeno. La ricerca teorica ha cercato di dare all'intelligenza artificiale un apparato formale di riferimento che possa permettere la progettazione di sistemi intelligenti artificiali sempre più sofisticati; i principali campi di indagine sono lo studio di logiche specifiche per formalizzare e analizzare il ragionamento automatico degli agenti intelligenti, riuscire a percepire e rappresentare il mondo esterno, decidere in modo razionale attraverso l'utilizzo di tecniche decisionali in situazioni in cui le informazioni sono incomplete o incerte e quindi apprendere in modo automatico servendosi anche di sistemi che si ispirano al funzionamento del cervello umano, come le reti neurali. Tutto ciò stimola l'integrazione di tecniche diverse in architetture di agenti intelligenti, in grado di svolgere compiti specifici sia nel mondo fisico (robot) che nel mondo virtuale di internet (*softbot*). Con riferimento a internet, ad esempio, un tema di grande attualità è quello dei sistemi multi-agenti, con le problematiche connesse alla loro coordinazione e alla condivisione di conoscenza. In questo capitolo, pertanto, viene evidenziato come può essere utile per un agente intelligente fare riferimento anche alla teoria delle decisioni analizzata nel primo capitolo (in particolare paragrafo 1.2), per dimostrare come un agente è razionale se sceglie l'azione che porta alla più alta utilità attesa, in base al principio della massimizzazione dell'utilità attesa. Le decisioni, spesso, riguardano la scelta fra opzioni caratterizzate da molti attributi; infine vengono presentate alcune tecniche in grado di valutare le informazioni utili per indirizzare le scelte tra più opzioni multi-attributo.

Nel terzo capitolo vengono descritte le principali metodologie decisionali adottate dagli agenti intelligenti capaci di filtrare le informazioni online, in particolar modo per gli acquisti online; si è discusso, in particolare, di *information retrieval*, vale a dire, recupero delle informazioni rilevanti, sottolineando i grandi benefici apportati dai motori di ricerca per il *World Wide Web*, dove un utente può inserire un quesito (*query*), e ottenere una lista di pagine rilevanti. Le principali metodologie di filtraggio delle informazioni, sono le metodologie cognitive (*content based*), le metodologie "sociali" o "collaborative", oppure una sintesi dei due approcci. In sintesi, le metodologie cognitive selezionano le opzioni sulla base dei contenuti, per esempio la presenza o meno in un testo di alcune parole, oppure la presenza o meno in un telefonino di particolari caratteristiche, ma con lo svantaggio di non poter estendere le valutazioni a prodotti significativamente diversi da quelli valutati dall'utente, con il rischio di perdere importanti opportunità. Alcuni sistemi *content based* selezionano le opzioni, massimizzando l'utilità attesa; ad esempio i sistemi compensatori, adottati da alcuni *price comparison* statunitensi, che tendono a compensare, per ogni opzione di prodotto, le performance negative con le performance positive. Le metodologie di filtraggio collaborative si basano, invece, sulle relazioni fra le valutazioni di persone diverse; l'applicazione più interessante è rappresentata dai filtri collaborativi, che cercano di aiutare le persone a decidere sulla base delle opinioni di altri, nell'ipotesi che le persone che hanno evidenziato preferenze analoghe in passato concorderanno anche in futuro. Le metodologie di filtraggio collaborative presentano tuttavia alcuni svantaggi: non possono essere applicate né a nuovi utenti né a nuove opzioni; quando si presenta un nuovo utente non si sa con quali vecchi utenti sono correlate le sue valutazioni, e quindi non esistono elementi per consigliargli delle opzioni; quando invece ci sono nuove opzioni, non essendo state ancora valutate da nessuno, non si sa a chi consigliarle. Per una efficace implementazione di un sistema collaborativo di raccomandazione è necessario quindi che sia stata raggiunta una massa critica di valutazioni; il sistema funziona bene soltanto se c'è un gruppo sociale di persone che collaborano, e nel quale sono molte le persone con sistemi di preferenze simili. I sistemi di raccomandazione collaborativi, al momento, sono comunque i sistemi più utilizzati online, soprattutto per la valutazione di prodotti o servizi, per la loro semplicità e perché con questi sistemi non è necessario considerare il contenuto, vale a dire le caratteristiche delle opzioni, per

potere effettuare delle raccomandazioni. Comunque, pur tenendo conto dei diversi vantaggi offerti dall'utilizzo di sistemi di *collaborative filtering*, nel terzo capitolo sono stati riportati i risultati di alcuni studi che mostrano come l'utilizzo di un agente *content based filtering* può risultare molto più efficace e affidabile nella selezione di prodotti che corrispondano meglio alle preferenze di ciascun consumatore. Ciò perchè un sistema di raccomandazione *content based*, avvalendosi dei valori di performance degli attributi di ogni opzione di prodotto, effettua delle comparazioni di prodotto con dati dimostrabili e non rilevati dalle indicazioni di un gruppo qualsiasi di opinionisti. Secondo Häubl e Trifts, i sistemi *content-based filtering* consentono ai consumatori di prendere decisioni migliori con un minore sforzo cognitivo, allentando il *trade-off* tra grado di accuratezza delle decisioni e sforzo sostenuto [Häubl, Trifts, 2000].

Nel quarto capitolo si è ritenuto opportuno presentare pregi e difetti dei siti e-commerce italiani e i principali aspetti evolutivi dei comparatori di prezzo online, prima di effettuare delle indagini esplorative negli ambienti online dedicati agli acquisti osservando come si comportano i consumatori in questi ambienti (capitolo 5). Quando un utente abbandona la sessione di acquisto di un sito e-commerce senza comprare, i motivi possono essere tanti: l'umore dell'utente, l'aspetto grafico e comunicativo del sito, il prezzo troppo alto del prodotto, tempi troppo lunghi e/o procedimenti complessi per acquistare un prodotto, processi di pagamento non molto trasparenti, spese di spedizione troppo elevate. I risultati delle ricerche analizzate nel quarto capitolo hanno messo in evidenza che per stimolare le vendite online può essere utile applicare dei prezzi in linea con il mercato, visualizzare e specificare dall'inizio i prezzi per le spedizioni, predisporre la possibilità di conservare il proprio carrello per un periodo di almeno 15-30 giorni, effettuare delle promozioni sui prodotti (per esempio la consegna gratuita), avere un catalogo di prodotti ben fornito e aggiornato, limitare il numero di passaggi per confermare l'ordine, garantire l'assistenza in tempo reale, la disponibilità di un numero verde, la possibilità di parlare con il responsabile vendite per chiarire eventuali dubbi. Ma la nuova frontiera del commercio elettronico è l'evoluzione dei siti di comparazione delle offerte dei prodotti venduti online, che permettono di facilitare il processo decisionale dei consumatori. Molti di questi siti, al momento, utilizzano semplici matrici di comparazione attributi-opzioni di prodotto, oppure sistemi di *collaborative filtering*. I sistemi che consentono scelte più razionali, perchè più vicine alle preferenze del consumatore/utente, sono i sistemi di raccomandazione *content based filtering*, ma sono ancora relativamente pochi, i siti di *price comparison* che li adoperano; quelli disponibili sono in genere predisposti in America, mentre sono quasi del tutto assenti in Italia. Nel quarto capitolo viene presentata un'analisi del complesso sistema delle reputazioni online, nel quale rientrano ovviamente i sistemi *collaborative filtering*. Utilizzando un qualsiasi motore di ricerca, ci si può fare rapidamente un'idea sulla reputazione acquisita da un prodotto o servizio, un venditore o produttore o anche un professionista. Le valutazioni online vengono fatte nelle aste elettroniche, nei social networks, in forum di discussione, nei sistemi recensori; in qualsiasi ambiente, la reputazione serve a creare fiducia nel venditore (o nel produttore o nel professionista) al quale si paga il bene (o a cui si assegna l'esecuzione del progetto). Esistono però, diversi studi che sollevano forti dubbi sul funzionamento effettivo del sistema reputazionale di eBay (il cosiddetto "*feedback forum*"), oppure sul sistema adottato da Amazon che si basa principalmente sui *ranks*, vale a dire i punteggi ottenuti da ogni revisore (*reviewer*). Si stanno diffondendo online sempre più siti che si offrono come "guardiani della reputazione" (*reputation guardian*), che offrono i propri servizi a tutti coloro che sono interessati ad avere una corretta reputazione online, o anche a nascondersi dai motori di ricerca. Quindi siti, forum, blog, motori di ricerca, social network, video online, etc., possono risultare utili per valutare un prodotto, servizio, venditore o professionista, ma è necessario essere consapevoli, di quanto sia possibile trovarsi, ora più di prima, davanti a degli inganni.

Nel quinto capitolo sono illustrati i risultati di alcune indagini realizzate per poter osservare più in profondità le difficoltà riscontrate dai consumatori negli ambienti online dedicati agli acquisti. Prendendo come prodotto di indagine il telefono cellulare, sono stati

considerati attributi, opzioni, ed offerte disponibili online. In particolare, è stata effettuata un'analisi delle opinioni rilasciate dagli utenti (consumatori e/o esperti) su uno dei principali *price comparison* italiani. La matrice di correlazione tra le variabili relative alle caratteristiche dei cellulari venduti online su Kelkoo.it (Tabella 5.2) evidenzia che la qualità del display è correlata positivamente con la risoluzione della fotocamera (scatti di qualità, hanno bisogno di display di qualità), e che la connettività senza fili è correlata positivamente con la presenza di sistema operativo Windows, mail, connessione a Internet, memoria espandibile, mp3, peso, prezzo, risoluzione del display e della fotocamera, *touch screen* e videocamera. I risultati della regressione multivariata (Tabella 5.3) evidenziano quali caratteristiche presentano delle relazioni statisticamente significative con il prezzo; in particolare, sono significativamente diverse da zero, le stime dei coefficienti delle variabili autonomia di conversazione, memoria espandibile, peso, risoluzione della fotocamera, zoom, autonomia in standby, dimensioni, flash, e connessione a internet; è significativamente positivo il segno dei coefficienti delle variabili: autonomia in standby, connessione a internet, peso, risoluzione fotocamera, invece è significativamente negativo il segno dei coefficienti delle variabili autonomia in conversazione, dimensione, flash, memoria espandibile. Il peso è una caratteristica negativa per l'acquirente, ma risulta direttamente proporzionale al prezzo; ciò deriva probabilmente dal fatto che i telefonini cosiddetti *smartphone* perché ultratecnologici, caratterizzati da prezzi molto alti, utilizzano molta energia e quindi devono essere dotati di grandi batterie che fanno aumentare di molto il peso complessivo del telefonino. In modo analogo si può probabilmente spiegare il coefficiente significativamente negativo per la variabile autonomia di conversazione: diversi telefonini di elevate prestazioni (in particolare per il display) e prezzo elevato consumano molta energia, e hanno quindi una bassa autonomia di conversazione. Analizzando in modo approfondito il mercato della telefonia mobile, basandoci in una prima indagine soltanto sui cellulari venduti a luglio del 2008 sul sito Kelkoo.it e in una seconda indagine sull'intera popolazione di cellulari venduti nell'ultimo decennio, i risultati ottenuti mediante la *cluster analysis* (Tabelle 5.8 e 5.9) appaiono alquanto concordanti. In entrambe le analisi sono stati ottenuti dei *cluster* con cellulari raggruppati per vecchio o nuovo modello, per tecnologia di base, media o alta, oppure per dimensione, quindi cellulari ultrasottili, di dimensioni ridotte, o *smartphone* di grandi dimensioni. Sostanziali differenze sono dovute soprattutto alla velocità di evoluzione tecnologica delle caratteristiche che definiscono un cellulare. Tra i telefoni cellulari venduti a luglio 2008 una percentuale elevata dispone di display *touch screen*, Gps, tastiera qwerty, tecnologia Wi-Fi e connessione a Internet superveloce tramite Hsdpa; considerando invece l'intera popolazione di telefoni cellulari usciti sul mercato fra il 1998 e il 2008, i telefonini che hanno queste caratteristiche rappresentano una percentuale molto bassa del totale, mentre moltissimi dispongono soltanto di funzione MMS, connettività a infrarossi, e velocità di connessione a Internet Gprs o al massimo Edge. Tutto ciò si vede meglio osservando le Figure 5.18 e 5.19, che mostrano la penetrazione sul mercato delle principali caratteristiche di un telefono cellulare. Per esempio, le caratteristiche Gps, videochiamata, quantità di colori del display, memoria totale e *touch screen* hanno avuto un forte sviluppo negli ultimi 3 anni. In particolare, la funzione *touch screen*, quasi sconosciuta fino agli inizi del 2007, è ormai disponibile in quasi tutti i cellulari con ampi display, mentre per le caratteristiche connettività senza fili e risoluzione della fotocamera la crescita è stata meno rapida. Altre caratteristiche quali ad esempio mms, mp3, mail e videocamera, hanno avuto una forte crescita dal 2002 al 2005, e poi si sono stabilizzate, anche perché ormai si trovano in quasi tutti i cellulari. Una terza indagine è stata svolta con l'obiettivo di valutare il comportamento degli opinionisti online; alcuni dei risultati ottenuti sono evidenziati nella tabella 5.12, che mostra come i 15 cellulari che hanno avuto un maggior numero di valutazioni su Kelkoo.it hanno destato maggiore curiosità subito dopo l'uscita sul mercato, ottenendo punteggi elevati per quanto riguarda lo stile del telefonino, una qualità che si apprezza prima dell'acquisto. A volte accade però che l'apparenza inganna, e quindi diversi consumatori che, attratti dal design innovativo, si

aspettavano anche performance superiori, spesso sono rimasti delusi; invece i punteggi rilasciati per le caratteristiche: autonomia della batteria, facilità d'uso, e qualità in generale non superano di molto la sufficienza; lo stesso vale per il rapporto qualità/prezzo, rendendo quindi, alquanto dubbia la convenienza dell'acquisto. Ciò si deduce anche dall'albero decisionale raffigurato in Figura 5.27; prendendo come variabile di stima la valutazione totale del prodotto, risulta che il 74% degli utenti che hanno valutato massima la qualità del prodotto (indice di qualità = 100), considerano importantissima la qualità della batteria; si può quindi ipotizzare che quando si acquista un cellulare ovviamente si tiene conto molto dello stile del prodotto, anche perché è quello che più colpisce l'utente, soprattutto quando si acquista in un negozio tradizionale; però si rimane soddisfatti dell'acquisto fatto, e quindi si consiglia di acquistare lo stesso telefonino, soltanto quando si valuta alta la qualità del prodotto, che dipende molto anche dall'autonomia della batteria.

Le ricerche effettuate nei settori della psicologia cognitiva, dell'intelligenza artificiale e delle metodologie di filtraggio delle informazioni (capitoli 1, 2 e 3), l'analisi degli ambienti online, italiani ed esteri, dedicati agli acquisti (capitolo 4), e le indagini empiriche illustrate nel capitolo 5, hanno consentito di acquisire maggiore consapevolezza dei limiti dei sistemi adottati per facilitare gli acquisti online. Nel sesto capitolo della tesi è stato presentato un nuovo sistema di supporto decisionale *content based filtering* (paragrafo 3.3). Questo sistema si basa su una metodologia decisionale e su una interfaccia utente che consentono di orientare il processo decisionale verso scelte in grado di corrispondere meglio alle esigenze e aspirazioni dei consumatori, in particolare nel caso di prodotti caratterizzati da molti attributi, come per esempio i telefoni cellulari, per i quali il processo decisionale risulta molto complesso. Il fondamento teorico dell'analisi delle decisioni nel caso di prodotti caratterizzati da molti attributi è la teoria di Lancaster, secondo cui il valore complessivo di un bene per un consumatore deriva dal valore delle sue diverse caratteristiche. In teoria, un procedimento razionale di valutazione delle diverse opzioni richiederebbe la conoscenza della relazione fra l'intensità in cui i diversi attributi sono presenti in ciascuna opzione e il valore per il consumatore della opzione stessa; ottenuta in questo modo una stima monetaria del valore per il consumatore delle diverse opzioni, si potrebbero confrontare questi valori con i rispettivi prezzi, e decidere di acquistare l'opzione a cui corrisponde il maggior surplus netto del consumatore; ma in concreto gran parte delle decisioni dei consumatori sono adottate attraverso delle procedure molto semplificate ("euristiche") che tengono conto in genere soltanto di una piccola parte delle informazioni necessarie per una decisione ottimale. L'applicazione presentata nel sesto capitolo consente di evitare eccessive semplificazioni, aiutando ciascun consumatore a selezionare le opzioni che meglio corrispondono alle sue esigenze e aspirazioni, tenendo presente la gran parte degli attributi di un telefono cellulare. Il sistema di supporto alle decisioni proposto nel capitolo sesto adotta una metodologia di filtraggio diversa rispetto a quelle utilizzate dai principali sistemi di raccomandazione *content based filtering*, ovvero i sistemi di aiuto alle decisioni di "compensatori", così denominati perché per ogni opzione performance negative per alcuni attributi possono essere compensate da performance positive per altri attributi (paragrafo 3.5.1); la metodologia decisionale di un sistema compensatorio si basa sulla teoria della massimizzazione dell'utilità attesa, ma si corre il rischio di selezionare opzioni sbagliate quando gli attributi sono inversamente correlati (paragrafo 6.2). Invece, la metodologia decisionale del sistema qui proposto, in sintesi, tende a minimizzare, per ogni attributo di ogni opzione di prodotto, il gap tra performance richiesta dal consumatore e performance effettiva; e quindi presentare le opzioni di prodotto con i più bassi valori complessivi di gap negativi. La nuova applicazione presenta anche una nuova interfaccia grafica (Figura 6.3), che consente all'utente online di considerare eventualmente soltanto alcuni degli attributi per lui più importanti, e quindi definire le sue preferenze con molta semplicità (interagendo con il mouse), evitando possibili confusioni, soprattutto quando il numero di attributi è particolarmente elevato. La pagina di output (Figura 6.6) presenta, in ordine di preferenza decrescente, le dieci opzioni selezionate

dall'applicazione. La prima versione dell'applicazione è stata testata presso il reparto di telefonia di un centro commerciale della provincia di Cosenza, suscitando molta curiosità, ed ottenendo risultati abbastanza soddisfacenti (il 41% dei consumatori interessati ad acquistare un cellulare ha effettivamente acquistato uno dei prodotti filtrati dal programma). Alcuni esponenti di portali di *shopping comparison* hanno manifestato interesse per l'applicazione proposta in questa tesi<sup>69</sup>.

Per verificare la validità della metodologia decisionale adottata nell'applicazione qui proposta, nell'ultima parte del sesto capitolo sono presentati i risultati ottenuti applicando diverse metodologie decisionali per selezionare i telefoni cellulari che meglio possano corrispondere alle esigenze e aspirazioni dei consumatori. I risultanti del confronto appaiono alquanto positivi per l'applicazione proposta nel sesto capitolo di questa tesi.

---

<sup>69</sup> In particolare, Andrea Boscaro di Pangora Italia, azienda leader in servizi di ricerca e confronto dei prodotti utilizzati dai migliori comparatori europei, ha espresso per e-mail la seguente valutazione: "Complimenti. Davvero una applicazione molto carina e ben fatta. Buon lavoro". Pangora è una delle aziende leader in Europa nell'ambito dei servizi di ricerca e confronto di prodotti su internet; essa offre ai portali soluzioni di shopping completi e agli shop un'ampia rete di distribuzione per i loro prodotti. Pangora è presente in Germania, Francia, Austria, Gran Bretagna e Italia, con 70 portali partner e 2.600 negozi che utilizzano il suo know-how tecnologico per un totale di 14 milioni di offerte ogni giorno. Andrea Boscaro è il country manager di Pangora Italia, che ha stipulato accordi con portali quali Quotidiano.net, Lycos, Excite, Tuttogratis; e realtà importanti dello *shopping comparison* come BuyCentral, Costameno, Msn Shopping, Kataweb Shopping, Repubblica Shopping e Techeye (Mytech).

## Riferimenti Bibliografici

- Alba J. W., J. Linch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, S. Wood (1997), Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplace, *Journal of Marketing*, 61, pp. 38-53
- Ancarani F. (2003), Le opportunità per il pricing in rete, *economia & management*, n. 6, pp. 55-70.
- Anderson C. (2007), *La coda lunga*, Codice edizioni, Torino; edizione originaria: *The Long Tail: Why the Future of Business Is Selling Less of More*, Chris Anderson, 2006.
- Anstine, J. (2000), "Consumers' Willingness to Pay for Recycled Content in Plastic Kitchen Garbage Bags: A Hedonic Price Approach", *Applied Economics Letters* 7, 35-39.
- Antinucci F. (2008), La rete e l'abilitazione cognitiva, *Sistemi Intelligenti*, n. 2, pp. 347-350.
- Antonelli E. (2007), Il benessere soggettivo nella prospettiva psicosociale: una rassegna, *Giornale italiano di psicologia*, n. 1, pp.57-116.
- Arcuri E. (2008), Scelte di consumo e valore, *Sistemi Intelligenti*, n. 1, pp. 79-92.
- Ariely D., Linch J. G., Aparicio M. (2004), Learning by collaborative and individual-based recommendation agent, *Journal of Consumer Psychology*, 14, pp. 81-94.
- Auh, S., Johnson, M. D. (2005). Compatibility effects in evaluations of satisfaction and loyalty. *Journal of Economic Psychology* 26, 35-57.
- Baer, D. M., Wolf, M. M., Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91-97.
- Bagozzi, R. P. (2000). The poverty of economic explanations of consumption and an action theory alternative. *Managerial and Decision Economics*, 21, 95-109.
- Balabanovi'c M., Shoham Y. (1997), Fab: Content-Based, Collaborative Recommendation, *Communications of the ACM*, Vol. 40 (3), pp 66-72, 1997.
- Baldassarre G., Marocco D., Midolli M. (2006), La via italiana alla vita artificiale, *Sistemi Intelligenti*, n. 1, pp. 3-6.
- Ballotta G., Gallo S., Legrenzi P. (2007), L'economia comportamentale: analisi del caso Unicredit, *Sistemi Intelligenti*, n. 3, pp. 359-380.
- Barberis, N., Shleifer, A., Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49, 307-343.
- Bellifemine F., Caire G., Greenwood D. (2004), *Developing multi-agent systems with JADE*, Wiley, Chichester, UK.
- Benartzi, S., Thaler, R. H. (1995) . Myopic Loss Aversion and the Equity Premium Puzzle *The Quarterly Journal of Economics*, 110, n. 1.
- Berger J. O., (1985), *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, second edition.
- Berlyne D.E. (1960). *Conflict, arousal, and curiosity*. Boston, MA: McGraw-Hill.
- Berners-Lee T., Hendler J., Lassilla O, (2001), "The Semantic Web", in *Scientific American*, May, pp. 29-37.
- Berners-Lee T., Hall W., Shaldbot N. (2006), "The Semantic Web Revisited", in *IEEE Intelligent Systems*, vol. 21, n. 3, pp. 96-101.
- Berry, S., Levinsohn, J., Pakes, A. (1995), "Automobile Prices in Market Equilibrium", *Econometrica* 63, 841-890.
- Berry, S., Levinsohn, J., Pakes, A. (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market", *Journal of Political Economy* 112, 68-105.

- Bettman J.R., Johnson E.J., Luce M.F., Payne J.W. (1993). Correlation, conflict, and choice. *Journal of Experimental Psychology: Learning, Memory and cognition*, 19 (4), 931-951
- Biggiere L., Laise D. (2004), Comportamento organizzativo e aiuto alle decisioni multicriteriali, *Sistemi Intelligenti*, n. 3, pp. 435-455.
- Bilotta E., (1999). Dalla Scienza Cognitiva agli Agenti Intelligenti, *Memoria*, pp. 79-81, 86-103, 108-112, 141-147, 173-203.
- Bilotta E., Pantano P. (2006), Strade nella complessità: percorsi fra arte e scienza, *Sistemi Intelligenti*, n. 1, pp. 37-46.
- Blaise Pascal, *Pensieri*, Mondadori, Milano, 1976
- Bonnetain, P. (2003), "A Hedonic Price Model for Islands", *Journal of Urban Economics* 54, 368-377.
- Bown N. J., Daniel Read, Barbara Summers. (2003) The lure of choice. *Journal of Behavioral Decision Making* 16:4, 297-308
- Bown N. J., 2007, The relevance of judgment and decision making research for marketing: Introduction to the special issue, *Marketing Theory*, Vol. 7, No. 1, pp. 5-11
- Breese J., Heckerman D., Kadie C. (1998), "Empirical Analysis of Predictive Algorithms for collaborative Filtering", Proceedings of the fourteenth Conference on Uncertainty in Artificial Intelligence, Madison, Wisconsin.
- Buchanan, L., O'Connell A. (2006), A brief history of Decision Making, *Harvard Business Review*, gennaio, pp. 32-51.
- Burke R. (2002), "Hybrid recommender systems: Survey and experiments", *User Modeling and User-Adapted Interaction*, Vol. 12, n. 4, pp. 331-370.
- Burman E. (2002), *Internet*, nuovo Leviatano, Etas, Milano.
- Busemeyer, J. R., Townsend J. T. (1993). Decision Field Theory: A dynamic cognition approach to decision making. *Psychological Review*, 100, pp. 432-459.
- Busemeyer J. R., Barman R., Mehta S., Chaturvedi A. (2007), Context effects and models of preferential choice: implications for consumer behavior *Marketing Theory*, Vol. 7, No. 1, pp. 39-58.
- Camerer C. F., Lowenstein G., Rabin M. (a cura di) (2004), *Advances in Behavioral Economics*, Princeton University Press.
- Caligiore D., Tria M., Parisi D. (2006). Some adaptive advantages of the ability to make predictions. In Nolfi S. et al., *Proceedings of The Ninth International Conference on the Simulation of Adaptive Behavior (SAB'06) 25 - 29 September 2006, CNR, Rome, Italy* (pp. 17-28). Berlin, Germany: Springer Verlag.
- Camerer C., Loewenstein G., Prelec D. (2004), Neuroeconomia, ovvero come le neuroscienze possono dare nuova forma all'economia, *Sistemi Intelligenti*, n. 3, pp. 337-418.
- Camerer, C. F., Loewenstein, G., Rabin, R. (eds.) (2003) *Advances in Behavioral Economics*, Coval.
- Cappellin A. (2005), Avatar e simulazioni, *Sistemi Intelligenti*, n. 1, pp. 45-58..
- Cardaci (a cura di) (2001), *Ciber-psicologia. Esplorazioni cognitive di Internet*, Carocci, Roma.
- Chanel, O., Gérard-Varet, A., Ginsburgh, V. (1996), "The Relevance of Hedonic Price Indices: The Case of Paintings", *Journal of Cultural Economics* 20, 1-24.
- Chernev A., (1997), The effect of common features on brand choice: Moderating role of attribute importance, *Journal of Consumer Research*, 23, pp.304-311.
- Chiang, K.P., and R.R. Dholakia, (2003), Factors Driving Consumer Intention to Shop Online: An Empirical Investigation, *Journal of Consumer Psychology*, Vol. 13, No. 1 -2, pp. 177 -183.
- Christensen, G. L., & Olson, J. C. (2002). Mapping consumer' mental models with ZMET. *Psychology and Marketing*, 19, 477-502.
- Church K., Patil R. (1982), Coping with syntactic ambiguity or how to put the block in the box on the table. *Comput. Linguist.* 8, 3-4 (Jul. 1982), pp. 139-149.
- Consumer Reports (2008), 2008 Buying guide, Consumers Union, Yonkers, NY.

- Coombs C. H. (1964) *A Theory of Data*, New York: John Wiley and Sons.
- Coombs C. H. Dawes R. M., Tversky A. (1970), *Mathematical Psychology An Elementary Introduction*. Englewood Cliffs, N. J.: Prentice-Hall.
- Court A. (1939), "Hedonic Price Indexes with Automotive Examples", in: General Motors Corporation (Ed.), *The Dynamics of Automobile Demand*, 99-117.
- Coval J. D. & Shumway, T. (2005). Do Behavioral Biases Affect Prices? *The Journal of Finance*, 60, 1 – 34.
- Daniel, K., Hirshleifer, D., Subrahmanyam, A. (1998). Investor Psychology and Security Markets: Under- and Overreactions. *The Journal of Finance*, 6, 1839 – 1885.
- De Bondt, W., Palm, F., Wolff, C. (2004). Introduction to the special issue on behavioral finance. *Journal of Empirical Finance*, 11, 423-427.
- Degeratu A., Rangaswamy A., Wu J. (2000), Consumer choice behaviour in online and traditional supermarkets: the effect of brand name, price, and other search attributes, *International Journal of Research in Marketing*, 17, pp.55-78.
- Del Missier F., Ranyard R., Pietroni D. (2008), Interazioni virtuose tra la psicologia delle decisione ed economia comportamentale: il caso dell'inflazione percepita, *Sistemi Intelligenti*, n. 1, pp. 39-56.
- Dewenter, R., Haucap, J., Luther, R., Rötzel, P. (2006). Hedonic Prices in the German Market for Mobile Phones. *Telecommunication Policy*, 30, forthcoming.
- Dewenter R., Haucap J., Luther R., Rötzel P. (2007) Hedonic prices in the German market for mobile phones, *Telecommunication Policy*, 31, 4-13.
- Diamond P., Vairtiainen H. (a cura di) (2007), *Behavioral Economics and its applications*, Princeton University Press.
- DiClemente, D. F., Hantula, D. A. (2003). Applied behavioral economics and consumer choice, *Journal of Economic Psychology*, 24, 589–60.
- Di Nocera F., Terenzi M., Ferlazzo F. (2007), Misurare l'informazione Handling: un contributo sperimentale sulla sensibilità della Scala "maneggevolezza" di Us.E.1.0, *Giornale italiano di psicologia*, n. 1, pp.223-236.
- Dipayan, B. (2004). Economics of information in the Web economy: Towards a new theory? *Journal of Business Research*, 57, 724-733.
- Doyle J. (1982), Analysis of feedback systems with structured uncertainty, IEE Proceedings, V129, Part D, No.6.
- Evans, D. A. (2006). Subject Perceptions of Confidence and Predictive Validity in Financial Information Cues. *The Journal of Behavioral Finance*, 7, n. 1.
- Fasolo B., Motta M., Misuraca R. (2005), Il processo decisionale *online*: rassegna di studi empirici e confronto tra siti internet per l'aiuto alle decisioni negli Stati Uniti e in Italia, *Giornale Italiano di Psicologia*, XXXII, 4, pp. 693-710.
- Fasolo B., McClelland G. H., Todd P. M. (2007) Escaping the tyranny of choice: when fewer attributes make choice easier, *Marketing Theory*, 7 (I), 13-26.
- Fasolo B., McClelland G. H., Todd P. M. (2007), Escaping the tyranny of choice: when fewer attributes make choice easier, *Marketing Theory*, n. 7 (1), pp. 13-26.
- Fasolo B., Misuraca R., Cardaci, M (2007), *I Processi Decisionali: Sfide, Paradossi e Supporti*. Il Mulino, Bologna.
- Favaro C., Burigana I. (2007), Ipotesi di "attrattiva della scelta". Una verifica mediante questionario, *Giornale italiano di psicologia*, n. 1, pp. 237-246.
- Feigenbaum L., Herman I., Hongsermeier T., Neumann E., Stephens S. (2007), "The Semantic Web in Action", in *Scientific American*, December, pp. 90-97.
- Fenton-O'Creevy, M. Nicholson N., Soane E., Willman P. (1999), Trading on illusions. Paper to Academy of Management Conference, Chicago.
- Filbeck, G., Hatfield, P. Horvath, P. (2005). Risk Aversion and Personality Type. *The Journal of Behavioral Finance*, vol. 6, n. 4.

- Fischhoff B. (1991). Value elicitation: Is there anything in there? *American Psychologist*, 46, 835-847.
- Forsythe R., Nelson F., Neumann G. R., Wright J. (1992), Anatomy of an Experimental Political Stock Market, *American Economic Review*, vol. 82, no. 5, pp. 1142–61; D. Shaw and B. Charles Roxburgh, “Hidden Flaws in Strategy,” *McKinsey Quarterly*, 2003, N. 2.
- Foxall, G. R. (1987). Radical behaviorism and consumer research theoretical promise and empirical problems. *International Journal of Research in Marketing*, 4, 111–129.
- Foxall, G. R. (1994). Consumer choice as an evolutionary process: An operant interpretation of adopted behavior. *Advances in Consumer Research*, 21, 312–317.
- Frey B. S., Stutzer A. (a cura di) (2007), *Economics and Psychology*, CESifo Seminar Series, The MIT Press, Cambridge, MA.
- Fundenberg D. (2006), Advancing Beyond Advances in Behavioral economics, *Journal of Economic Literature*, settembre, pp. 694-711.
- Gabor A., Granger C. W. J. (1966), Prices as an Indicator of Quality: Report on an Inquiry. *Economica*, 33, 43-70.
- Gamble A. (2006). Euro illusion or the reverse? Effects of currency and income on evaluations of prices of consumer products. *Journal of Economic Psychology*, 27, 531-542.
- Gardener, D. W. (1971), Is There a Generalized Price-Quality Relationship? *Journal of Marketing Research*, 8, 241-3.
- Gigerenzer G., Goldstein D. G. (1996), Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* 103: pp. 650-669.
- Giles, David E. (1982), “The Interpretation of Dummy Variables in Semi-logarithmic Equations”, *Economics Letters* 10, 77-79.
- Goedhart M, Koller T., Wessels D. (2005), Do fundamentals - or emotions - drive the stock market?, *McKinsey quarterly*, edizione speciale su value and performance.
- Goldberg D., Nichols D., Oki B., Terry D. (1992), “Using collaborative filtering to weave an information tapestry“, *Communications of the ACM*, vol. 35, n.12, pp. 61-70.
- Good N., Schafer B. J., Konstan J. A., Borchers A., Sarwar B., Herlocker J., Riedl J. (1999), “Combining collaborative filtering with personal agents for better recommendations”, *American Association for Artificial Intelligence, Proceedings of the sixteenth national conference on Artificial intelligence*, pp. 439-446.
- Green P. E., Carmone F. J., (1971), *Consumer Evaluation of Discount Cards: A Conceptual Model and Experimental Test*. Working Paper, University of Pennsylvania, February.
- Green P. E., Vithala R. (1971), Conjoint Measurement for Quantifying Judgmental Data. *Journal of Marketing Research*, 8, 355-63.
- Green P. E., Srinivasan V. (1978) Conjoint analysis in consumer behaviour: issues and outlook, *Journal of Consumer Research*, 5, 103-23.
- Griffin, A., & Hauser, J. R. (1993). The voice of the customer. *Marketing Science*, 12, 1–27.
- Griliches, Zvi (1961), “Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change”, in: *The Price Statistics of the Federal Government*, NBER: New York, 173-196.
- Griliches, Zvi (1994), “Hedonic Prices Indexes for Personal Computers: Intertemporal and Interspatial Comparisons”, *Economics Letters* 44, 353-357.
- Guido G., Peluso A. M., Rizzato F. (2005), Il mercato on line del comparison shopping: come classificare gli infomediari, *Economia & Management*, 2, pp. 55-76.
- Guo F. Y., Holyoak K. J. (2002), Understanding similarity in choice behavior: A connectionist model. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society* (pp. 393–398). Mahwah, NJ: Erlbaum.
- Ha, H.-Y., Perks, H. (2005). Effects of consumer perceptions of brand experience on the web: brand familiarity, satisfaction and brand trust. *Journal of Consumer Behaviour*, 4, 438-452.

- Hammond J. S., Keeney R. L., Raiffa H. (2006), The hidden traps in Decision Making, *Harvard Business Review*, gennaio, pp. 118-130.
- Hammond J. S., Keeney R. L., Raiffa H. (1999), *Smart Choices*, a practical guide to making better life decisions, Broadway Books, New York.
- Handgraaf M. J. J., Van Raaij W. F., (2005), Fear and loathing no more: The emergence of collaboration between economist and psychologist, *Journal of Economic Psychology*, 26, pp. 387-391.
- Hantula, D. A., DeNicolis Bragger, J., Rajala, A. (2001). Slips and falls in stores and malls: Implications for community-based prevention. *Journal of Prevention and Intervention in the Community*, 22, 67–80.
- Harvard Business Review (2001), *Harvard Business Review on Decision Making*, Harvard Business School Press.
- Häubl G., Trifts V., (2000), Consumer decision making in online shopping environments: The effects of interactive decision aids, *Marketing Science*, 19, pp. 4-21.
- Hauser, J. R., & Clausing, D. (1988). The house of quality. *Harvard Business Review* (May–June), 63–73.
- Hayes, S. C., & Cone, J. D. (1981). Reduction of residential consumption of electricity through simple monthly feedback. *Journal of Applied Behavior Analysis*, 14, 81–88.
- Haykin S., (1994), *Neural Networks: a comprehensive foundation*, Upper Saddle River, New Jersey: Prentice Hall.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243–266.
- Hertwig R., Todd P. M. (2003), More is not always better: The benefits of cognitive limits. In L. Macchi and D. Hardman, editors, *Psychological perspectives on reasoning, judgment and decision making* (pp. 213–231). Chichester, England: Wiley.
- Hilton D J. (2001) The Psychology of Financial Decision-Making: Applications to Trading, Dealing, and Investment Analysis, *The Journal of Psychology and Financial Markets*, 2, 37-53.
- Hirshleifer, D. (2001) Investor Psychology and Asset Pricing. *The Journal of Finance*, 56, 1533 - 1597.
- Hirshleifer D., Subrahmanyam A., Titman S. (2006). Feedback and the success of irrational investors. *Journal of Financial Economics*, 81, 311-338.
- Hoch S. J., Kunreuther H. C., Gunther R. E. (a cura di) (2001), *Wharton on Making Decisions*, Wiley, Hoboken, N.J.
- Hogarth, R. M., Karelaia, N. (2006). "Take-the-best" And other simple strategies: Why and when they work "Well" With binary cues. *Theory and Decision*, 61, 205-249.
- Horn T. J., Lovallo D. P., Viguerie S. P. (2006), Learning to let go: Making better exit decisions, *McKinsey quarterly*, maggio.
- Hostlera, R. E., Yoona, V. Y., Guimaraesb, T. (2005). Assessing the impact of internet agent on end users' performance. *Decision Support Systems*, 41, 313– 323.
- Howard J. A., Sheth J. N. (1969), *The Theory of Buyer Behavior*, New York: John Wiley and Sons.
- Iyengar S. S., Lepper M. R. (2000), "When choice is demotivating: can one desire too much of a good thing?", in *Journal of Personality and Social Psychology*, n°79.
- Hunsberger L., Grosz B. J., (2000), A Combinatorial Auction for Collaborative Planning. In *Proceedings of the Fourth International Conference on Multi-Agent Systems*, Boston, Massachusetts, IEEE Computer Society Press, pp. 151-158.
- Jacoby J., Olson J. C., Rafael A. P. (1970), Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality. *Purdue Papers in Consumer Psychology*, Paper No. 111.
- Jacoby J, Speller D. E., Kohn C. A. (1974), Brand Choice Behaviour as a Function of Information Load, *Journal of Marketing Research*, vol. 11, no. 1, pp.63-69.

- Jacoby J., Speller D. D. , Berning C. A. (1974), Brand choice behavior as a function of information load: replication and extension. *Journal of Consumer Research*, 1 , 33-42.
- Jeffrey, H. J. (2006). Expanding the Range of Behavioral Factors in Economic Simulations. *The Journal of Behavioral Finance*, 7, n. 2.
- Johnson R. M. (1974) Trade-off analysis of consumer values, *Journal of Marketing Research*, 11, 121-217.
- Kahneman D., Tversky A. (1974) Judgment under uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.
- Kahneman D., Tversky A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.
- Kahneman D., Tversky A. (1981a), The framing of decisions and the psychology of choice. *Science*, 211, 453-458.
- Kahneman D., Slovic P., Tversky A. (1982). Judgment under uncertainty: Heuristics and Biases. New York: Cambridge University Press.
- Kahneman D. e Tversky A. (1984), "Choice, values and frames", *American Psychologist* Vol. 39, pp. 341-350.
- Kahneman D., Tversky A. (1992), "Advance in Prospect Theory: cumulative representation of uncertainty", *Journal of Risk and Uncertainty* Vol. 5, pp. 297-323.
- Kahneman D., Tversky A. (a cura di) (2000), Choices, values and frames, Cambridge university Press.
- Keeney R. L., Raiffa H. (1976), Decisions with Multiple Objecties: Preference and Value Tradeoffs,, Wiley, New York.
- Kim B. M., Li Q., Park C. S., Kim S. G., Kim J. Y. (2006), A new approach for combining content-based and collaborative filters, *Journal of Intelligent Information Systems*, 27, pp. 79-91.
- Kim J. K., Cho Y. H., Kim W. J., Kim R. J., Suh J. H. (2002), A personalized recommendation procedure for internet shopping support, *Electronic Commerce research and applications*, n. 1, pp. 301-313.
- Kivetz R., Simonson I. (2000), The effects of incomplete information on consumer choice, *Journal of Marketing Research*, 37, pp. 427-448.
- Koehler D. J., Harvey N. (a cura di) (2007), Blackwell Handbook of Judgement & Decision Making, Blackwell Publishing.
- Kohli R., Jedidi K. (2007), Representation and Inference of Lexicographic Preference Models and their Variants, *Marketing Science*, maggio-giugno, pp. 380-399.
- Kohonen T. (1995), Self-Organizing Maps, prima edizione, Springer-Verlag, New York (seconda edizione 1997, terza edizione 2001).
- Konstan J. A., Miller B. N., Maltz D., Herlocker J. L., Gordon L. R., Riedl J. (1997), "GroupLens: Applying collaborative filtering to usenet news", *Communications of the ACM*, vol. 40, n.3, pp. 77-87.
- Konstan J. A., Riedl J., Schafer J. B. (1999) "Recommender systems in e-commerce", *Proceedings of the First ACM Conference on Electronic Commerce*, pp. 158-166.
- Kothari, S.P. , Lewellen, J., Warner, J. B. (2006). Stock returns, aggregate earnings surprises, and behavioral finance, *Journal of Financial Economics* , 79, 537-568.
- Kruskal, Joseph B. Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data. *Journal of the Royal Statistical Society, Series B*, 1965, 27, 251-63.
- Kucuk S. U., Krishnamurthy S. (2007), "An analysis of consumer power on the Internet", *Technovation*, 27, pp. 46-56.
- Lancaster K. J. (1966) New approach to consumer theory, *Journal of Political Economy*, 74, 132-157.
- Lancaster, K. (1971), *Consumer Demand*, Columbia University Press: New York.
- Leavitt, Harold J. A Note on Some Experimental Findings About The Meaning of Price. *Journal of Business*, 1954, 28, 205-10.

- Lee B-C., Ang L., Dubelaar C., (2005), Lemons on the Web: A signalling approach to the problem of trust in Internet commerce, *Journal of Economic Psychology*, 26, pp. 607-623.
- Lee W. (2004). Applying domain knowledge and social information to product analysis and recommendations: an agent-based decision support system. *Expert Systems*, 21, n. 3.
- Lefkoff-Hagius, R., & Mason, C. H. (1993). Characteristic, beneficial, and image attributes in consumer judgments of similarity and preference. *Journal of Consumer Research*, 20, 100–110.
- Lemire D., Maclachlan A, Slope One Predictors for Online Rating-Based Collaborative Filtering, In *SIAM Data Mining (SDM'05)*, Newport Beach, California, April 21-23, 2005.
- Levin M., Levin P., Heath, (2003), Product category dependent consumer preferences for online and offline shopping features and their influence on multi-channel retail alliances, *Journal of Electronic Commerce Research*, Vol. 4, N° 3, pp. 85-93.
- Lichtenstein S., Slovic P. (a cura di) (2006), *The Construction of Preferences*, Cambridge University Press.
- Li Q., Kim B. M. (2003), “Clustering approach for Hybrid recommender System”, Proceedings of the IEEE/WIC Conference on Web Intelligence, pp. 33-38.
- Lindsley, O. R. (1962). A behavioral measure of television viewing. *Journal of Advertising Research*, 2, pp. 2–12.
- Lindstrom M., Peppers Don, Rogers M. (2001), *ClicKs, Bricks & Brands*, Etas, Milano.
- Lokken, S. L., Cross, G. W., Halbert, L. K., Lindsey, G., Derby, C., Stanford, C. (2003). Comparing online and non-online shoppers. *International Journal of Consumer Studies*, 27, 126–133.
- Lopes L. L., (1994), Psychology and economics: Perspectives on risk, cooperation, and the marketplace, *Annual Review of Psychology*, 45, pp.197-227.
- Luce R. D. (1964), Simultaneous ConJoint Measurement: A New Type of Fundamental Measurement, *Journal of Mathematical Psychology*, 1, 1-27.
- Luhn H. P. (1958), “A business intelligence system”, *IBM Journal of Research and Development*, vol. 2, n. 4, pp. 314-319.
- Luthje, C. (2004). Characteristics of innovating users in a consumer goods field: An empirical study of sport-related product consumers. *Technovation*, 24, 683–695
- Lynch P.D., R. J. Kent, and S. S. Srinivasan, (2001), The Global Internet Shopper: Evidence from Shopping Tasks in Twelve Countries,” *Journal of Advertising Research*, Vol. 41, 3, pp.15-23.
- Malone T. W. , Grant K. R., Turbak F. A., Brobst S. A., Cohen M. D. (1987), “Intelligent information sharing systems”, *Communications of the ACM*, vol. 30, n.5, pp. 390-402.
- Mandel N., Johnson E.J. (2002). When web pages influence choice. *Journal of Consumer Research*, 29, 235-245
- Marconi D. (2006), “ L’intelligenza artificiale ha 50 anni “ (Introduzione), *Sistemi Intelligenti*, n. 3, pp. 339-344.
- Marmo S. (2007), L'uso della reputazione nelle applicazioni internet. Un antico artefatto per un nuovo contesto sociale, *Sistemi Intelligenti*, n. 1, pp. 91-112.
- Marmo S. (2006), L'uso della reputazione nelle applicazioni internet: prudenza o cortesia? L'approccio socio-cognitivo. In Greco, A., Penco, C., Sandini, G., Zaccaria, R. (Ed.), *Scienze Cognitive e Robotica* (pp. 130-135). Genova.
- Maurer, R., Pitzer, M., Sebastian, S., (2004), “Hedonic Prices Indices for the Paris Housing Market”, *Allgemeines Statistisches Archiv* 88, 303-326.
- McConnell, J. D. (1968a), An Experimental Examination of the Price-Quality Relationship. *Journal of Business*, 40, 439-44.
- McConnell, J. D. (1968b), The Price-Quality Relationship in an Experimental Setting. *Journal of Marketing Research*, 5, 300-3.
- March J. G. (1994), *Prendere decisioni*, il Mulino, Bologna.

- Markov A. A. (1971), Extension of the limit theorems of probability theory to a sum of variables connected in a chain". reprinted in Appendix B of: R. Howard. *Dynamic Probabilistic Systems, volume 1: Markov Chains*. John Wiley and Sons.
- Martelli S., Gaglio S. (2004), Immagini della emergente società in rete, Franco Angeli, Milano.
- Misuraca R., Fasolo B., Cardaci M. (2007), I Processi decisionali, Paradossi, sfide, supporti, il Mulino, Bologna.
- McFadden D., (2005). "Razionalità per economisti?", in: *Critica della ragione economica*, a cura di M. Motterlini, M.P. Palmarini, Il Saggiatore, pp. 25-76.
- McFadden D. (2006), Free Markets and Fettered Consumers, *American Economic Review*, marzo, pp. 5-24.
- Miller G. A. (1956), "The magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information" , *Psychological Review*, n. 63, pp. 81-97).
- Misuraca R. (2001), Impatto psicologico dell'e-commerce, in Cardaci (2001), pp. 67-89.
- Montgomery H., Decision-Making (in Swedish). In L.-G. Lundh, H. Montgomery, and Y. Wærn, editors, *Cognitive Psychology*, 6, pp. 171–188.
- Motterlini, M., Guala F. (2005), *Economia Cognitiva e Sperimentale*, Università Bocconi Editore, Milano.
- Motterlini, M., Palmarini M.P. (2005), *Critica della ragione economica*, il Saggiatore, Milano
- Nerlove, Marc (1995), "Hedonic Price Functions and the Measurement of Preferences: The Case of Swedish Wine Consumers", *European Economic Review* 39, 1697-1716.
- Neumann J. v., Morgenstern O., (1944), *Theory of Games and Economic Behavior*, Princeton University Press, New Jersey.
- Nilsson N. J. (2002), *Intelligenza artificiale*, edizione italiana a cura di S. Gaglio, Apogeo, Milano; edizione originale: *Artificial Intelligence: A new Synthesis*, Morgan Kaufman Publishers, Inc., 1998.
- Oard D. W. (1997), "The state of the art in text filtering", *User Modeling and User-Adapted Interaction*, Vol. 7, n. 3, pp. 141-178.
- Oberlechner, T., & Hocking, S. (2004). Information sources, news, and rumours in financial markets: Insights into the foreign exchange market. *Journal of Economic Psychology*, 25, 407-424.
- O'Curry Fogel, S., Berry, T. (2006). The Disposition Effect and Individual Investor Decisions: The Roles of Regret and Counterfactual Alternatives. *The Journal of Behavioral Finance*, 7, n.2.
- Oja M., Kaski S., Kohonen T. (2003), "Bibliography of self-organizing map (som) papers: 1998 2001 addendum, *Neural Computing Surveys*, n. 3, pp. 1-156.
- Olson, J. C., & Reynolds, T. J. (1983). Understanding consumers' cognitive structures: Implications for advertising strategy. In L. Percy & A. Woodside (Eds.), *Advertising and consumer psychology* (pp. 77–90). Lexington, MA: Lexington.
- Parisi D. (2001), *Simulazioni, la realtà rifatta nel computer*, il Mulino, Bologna.
- Peng, L. & Xion, W. (2006). Investor attention, overconfidence and category learning, *Journal of Financial Economics*, 80, 563-602.
- Percival-Straunik L. (2001), *E-Commerce*, The Economist Books.
- Peterson, Robert A. The Price-Perceived Quality Relationship: Experimental Evidence. *Journal of Marketing Research*, 1970, 7, 528-8.
- Pietroni D., Pieti L., Rumiati R. (2005), Progettare siti web "cognitivamente compatibili": un contributo sperimentale, *Sistemi Intelligenti*, n. 2, pp. 283-302.
- Prather, L. J. & Middleton, K. L. (2006). Timing and selectivity of mutual fund managers: An empirical test of the behavioral decision-making theory, *Journal of Empirical Finance*, 13, 249-273.
- Rabin, M. (1998). Psychology and Economics, *Journal of Economic Literature*, 36 11-46.

- Raiffa H., (1968) *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*, Addison-Wesley, Reading, MA.
- Rajala, A. K., & Hantula, D. A. (2000). Towards a behavioral ecology of consumption: Delay-reduction effects on foraging in a simulated Internet mall. *Managerial and Decision Economics*, 21, 145–158.
- Rao V. R. (1972), Saliency of Price in the Perception of Product Quality- A Multidimensional Measurement Approach. In Niel H. Bordon, Jr. (ed.), *Proceedings of the 1971 Fall Conference of American Marketing Association*, Chicago, Illinois (1972).
- Remondino M. (2005), Paradigmi e tipologie di agenti software per lo studio dei sistemi sociali complessi, *Sistemi Intelligenti*, n. 1, pp. 25-44.
- Reynolds, T. J., & Gutman, J. (1988). Laddering theory, method, analysis, and interpretation. *Journal of Advertising Research* (February/March), 11–31.
- Resnick P., Iacovou N., Suchak M., Bergstorm P., Riedl J. (1994), “GroupLens: An open architecture for collaborative filtering of netnews”, *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, pp. 175-186.
- Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in internet transactions: empirical analysis of eBay's reputation system. In M. R. Baye (ed.), *The Economics of the Internet and E-Commerce*. Vol. 11 of *Advances in Applied Microeconomics*. Amsterdam, Elsevier Science.
- Roe R., Busemeyer, J. R., Townsend J. T. (2001) Multi-Alternative Decision Field Theory: A Dynamic Connectionist Model of Decision-Making. *Psychological Review*, 108, 370-392.
- Romesburg H. C. (2004), *Cluster analysis for researchers*, Lulu Press, North Carolina.
- Roskam E. E. (1968), *Metric Analysis of Ordinal Data in Psychology*. Van Voorshoten.
- Ross Don (2005), *Economic Theory and Cognitive Science, Microexplanation*, The Mit Press, Cambridge, MA.
- Roxburgh C. (2003), Hidden Flaws in Strategy, *McKinsey quarterly*, maggio.
- Roxburgh C. (2006), The human factor in strategic decisions, *McKinsey quarterly*, febbraio.
- Rumiati R., Bonini N. (2001), *Psicologia della decisione*, il Mulino, Bologna.
- Rumiati R., Rubaltelli E., Mistri M. (2008), Cognizione, emozione e neuroeconomia, *Sistemi Intelligenti*, n. 1, pp. 7-22.
- Russel S., Norvig P. (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Education Italia, Milano; edizione originale: *Artificial Intelligence: A Modern Approach*, seconda edizione, Prentice Hall, Pearson Education, New Jersey, 2003.
- Russo V., (2004), Processi di costruzione del significato: il sistema cognitivo, *Psicologia del consumatore: consumi e costruzione del significato*, a cura di G. Siri, pp. 3-63
- Russo V., Ceresa S. (2006), Le ricerche esplorative, *La Ricerca sul Consumatore: domande di marketing e prodotti classici*, a cura di G. Siri, pp 20-22, 40-42, 84-99, 121-122, 105-130, 131-167, 184-202.
- Samuel A. L. (1967), Some studies in machine learning using the game of checkers. ii - recent progress. *IBM Journal of Research and Development*, 11(6), pp. 601-617.
- Samuel A. L. (1959) Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):211-229.
- Scitovsky T. (1945), Some Consequences of the Habit of Judging Quality by Price. *Review of Economic Studies*, 12, 100-05.
- Schwartz B. (2004), *The Paradox of Choice: Why More is Less*, New York, Ecco.
- Scott W. G., Sebastiani R., (2001), *Dizionario di Marketing*, a cura di Il Sole 24 Ore.
- Shafir E., Simonson, I., Tversky A. (1993), Reason based Choice. *Special Issue: Reasoning and Decision Making, Cognition*, 49 (1-2), 11-36.
- Shanteau J., Weiss D. J., Thomas R., Pounds, J., (2000), Performance-based assessment of expertise: How can you tell if someone is an expert? *European Journal of Operations Research*.

- Shardanand U., Maes P. (1995), "Social information filtering: algorithms for automating «word of mouth»", in Proceedings of the SIGCHI Conference on Human factors in computing system, pp. 210-217, ACM Press/Addison-Wesley, New York.
- Shefrin, H. (2002). *Beyond Greed and Fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press.
- Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26, 278–292.
- Shleifer, A. (1999). *Inefficient Markets: An Introduction to Behavioral Finance*, Oxford University Press.
- Shleifer, A. & Vishny, R. (1997). The Limits of Arbitrage. *Journal of Finance*, 52, 35-55.
- Simmons, C. G., Bickart, B. A., & Lynch, J. G. (1993). Capturing and creating public opinion in survey research. *Journal of Consumer Research*, 20, 316–329.
- Simon H. A. (1947). *Administrative Behaviour*, New York, NY: Macmillan.
- Simon H. A. (1972), *Human Problem Solving*, Englewood Cliffs, Prentice Hall
- Simon H. A. (1981), *The Influence of Automation on Society*, pp. 69-74.
- Simon H. A. (1985), *Causalità, razionalità, organizzazione*, Il Mulino, Bologna.
- Simonson I., Tversky A., (1992), Choice in Context; Tradeoff contrast and extremeness aversion, *Journal of Marketing Research*, 29 (3), 281-295.
- Simonson I., Carmon Z., O'Curry S. (1994), "Experimental evidence on the negative effect of product features and sales promotions on brand choice" ", *Marketing Science*, Vol. 13 No.1, pp.23-40.
- Siri G.(2001), *La Psiche del consumo*, Franco Angeli, Milano.
- Siri G. (a cura di) (2006), *La ricerca sul consumatore*, McGraw-Hill, Milano.
- Slovic, P. (1995). The construction of preference. *American Psychologist*, 50, 364-371.
- Svenson O., Process description of decision making. *Organizational Behaviour and Human Performance* 23 (1979), pp. 86–112.
- Smith, C. L. & Hantula, D. A. (2003). Pricing effects on foraging in a simulated Internet shopping mall. *Journal of Economic Psychology*, 24, .
- Smith R. E., & Swinyard W. R. (1983). Attitude-behavior consistency: The impact of product trial versus advertising. *Journal of Marketing Research*, 20, 257–267.
- Snelders, D., Schoormans, J. P.L. (2004). An exploratory study of the relation between concrete and abstract product attributes. *Journal of Economic Psychology*, 25 803–820.
- Soman D. (2004), The effect of time delay on multi-attribute choice, *Journal of Economic Psychology*, pp. 153-175.
- Stafford, J. E., Ben M. (1969), The Price-Quality Relationship: An Extension. *Journal of Marketing Research*, 6, 456-8.
- Stigler G. (1961), "The Economics of Information," *Journal of Political Economy*, June. (JSTOR)
- Fehr E., Schmidt K.M., (1999), *A theory of fairness, competition and cooperation*, Quarterly Journal of Economics, n.144, pp.817-868.
- Stracca, L. (2004). Behavioral finance and asset prices: Where do we stand? *Journal of Economic Psychology*, 25, 373-405.
- Tabossi P. (1988), *Intelligenza naturale e intelligenza artificiale*, il Mulino, Bologna.
- Tapscott D., Williams A. D. (2007), *Wikinomics. La collaborazione di massa che sta cambiando il mondo*, Rizzoli – Etas, Milano, edizione originale: *Wikinomics. How Mass Collaboration Changes Everything*, Portfolio, Penguin Group.
- Tedeschi M. (2008), Alcune idee sulle preferenze del consumatore, *sistemi Intelligenti*, n. 1, pp. 23-38.
- Tesauro G. (1992), Practical Issues in Temporal Difference Learning, *Machine Learning* 8:257-277.
- Thaler R. (1985) Mental accounting and human choice, *Marketing Science*, 4, 199-214.

- Thaler R. H., Cass R. S. (2003), Libertarian Paternalism, *American Economic Review (Papers and Proceedings)* 93(2), pp. 175-179.
- Tyszka T., Przybyszewski K. (2006). Cognitive and emotional factors affecting currency perception. *Journal of Economic Psychology*, 27, 518-530.
- Tull, S. S., Boring R. A., Gonsoir M. H. (1964), A Note on the Relationship of Price and Imputed Quality. *Journal of Business*, 38, 186-91.
- Turing A. (1936), "On computable numbers, with an application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society*, 2nd series, 42, 230-265.
- Turing A. (1950), "Computing machinery and intelligence", *Mind*, 59, 433-460.
- Tversky A. (1967a), A General Theory of Polynomial Conjoint Measurement. *Journal of Mathematical Psychology*, 1967a, 4, 1-20.
- Tversky A., (1967b), Additivity, Utility and Subjective Probability. *Journal of Mathematical Psychology*, 4, 175-201.
- Tversky A. (1972). Elimination by aspects: a theory of choice. *Psychological Review*, 79, 281-299.
- Tversky A. (2004), Preference, belief, and Similarity, selected writings, a cura di Shafir E., The MIT Press, Cambridge, MA..
- Tversky, A., Sattah S. (1979) Preference Trees, *Psychological Review* 86, 542-573.
- Tversky A., Shafir E. (1992), The disjunction effect in choice under uncertainty. *Psychological Science*, 3, 5, 305-309.
- Tversky A., Shafir E., Simonson I., (1993), Reason-based choice, *Cognition*, 49, 2, 11-36.
- Van der Sar, N. L. (2004). Behavioral finance: How matters stand. *Journal of Economic Psychology*, 25, 425-444.
- Van Raaij, W.F., Schoonderbeek, W.M. (1993), Meaning structure of brand names and extensions. In: Van Raaij et al. (Eds.), *European advances in consumer research 1*. Provo, UT: Association for Consumer Research, pp. 479-484.
- Varian H. R., Farrel J., Shapiro C. (2005). "Introduzione all'economia dell'informazione", *Etas*
- Vernon D. (2007), Una panoramica sui sistemi cognitive artificiali, *Sistemi Intelligenti*, n. 2, pp. 259-312.
- Waugh, Frederick V. (1928), "Quality Factors Influencing Vegetable Prices", *Journal of Farm Economics* 10, 185-196.
- Wellman M. P., Greenwald A., Stone P. (2007), *Autonomous Bidding Agents, Strategies and lessons from the Trading Agent Competition*, The MIT Press, Cambridge, MA.
- Wilks, Y. (1975), An Intelligent Analyzer and Understander of English, *Comm. ACM* 18, 5, 264-274.
- Wind J., MahaJan V. (2001), *Digital Marketing*, Etas, Milano.
- Wooldridge M. (2002), *An introduction to Multiagent Systems*, Wiley, Chichester, UK.
- Yee M., Dahan E., Hauser J. R., Orlin J. (2007), Gredoid-Based Noncompensatory Inference, *Marketing Science*, luglio-agosto, pp. 532-549.
- Zeithaml V. A. (1988), Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, 52, 2-22.
- Zweig G., Russell S. (1998), Speech recognition with dynamic Bayesian networks. In *Proceedings of the Fifteenth National/Tenth Conference on Artificial intelligence/innovative Applications of Artificial intelligence* (Madison, Wisconsin, United States). American Association for Artificial Intelligence, Menlo Park, CA, pp. 173-180.