LE RETI NEURALI NEL MARKETING : IL PROBLEMA DELLA SEGMENTAZIONE PER OBIETTIVI

Luca Molteni, Michele Gnecchi

Introduzione

Intendendo il marketing come *attività necessaria per pianificare e realizzare lo sviluppo, la promozione e la distribuzione di idee, beni e servizi, per creare uno scambio che soddisfi le aspirazioni degli individui e delle organizzazioni*[1] emerge come compito specifico del marketing management non sia più generalmente riferito alle attività d'affari necessarie per trasferire prodotti e servizi dal produttore al consumatore quanto piuttosto quello di monitorare, elaborare, modellare e gestire la complessità ambientale al fine di favorire modalità di integrazione fra il sistema aziendale e il sistema ambientale che consentano, in condizioni di economicità, il conseguimento degli obiettivi dell'impresa e la soddisfazione dei bisogni degli individui.

Con specifico riferimento al rapporto domanda offerta, una delle determinanti più rilevanti della complessita' ambientale, e' la **varietà**[2],

In particolare tale determinante si riferisce all'esistenza di bisogni, desideri, strategie concorrenziali e in generale ad attributi riferibili ad attori o fenomeni influenti sull'impresa, differenti gli uni dagli altri. In particolare "La maggiore sofisticatezza della domanda da un lato e l'intensificarsi del confronto concorrenziale dall'altro hanno incentivato l'impresa a ricercare una fonte per il suo vantaggio competitivo nella varietà e nella personalizzazione dell'offerta dei suoi prodotti o servizi" (Valdani, 1995).

Da un punto di vista di analisi ed approfondimento alla varietà vengono associati i processi cosiddetti di "segmentazione".

Con particolare riferimento a quest'ultimo aspetto, la segmentazione del mercato intesa come "la strategia di selezionare dei gruppi di clienti che l'impresa desidera servire adattando ai loro specifici bisogni la sua offerta di prodotti e servizi e specializzando la sua strategia di marketing" (Valdani, 1995), e' stata oggetto di numerosi studi volti ad analizzarne i diversi aspetti.

In modo particolare e' largamente condivisa l'opinione secondo la quale, per definire efficaci politiche di offerta, non sia possibile prescindere da analisi e strategie di segmentazione della domanda orientate a verificare direttamente la configurazione delle aspettative di valore dei clienti. (Valdani, 1984, 1995 ;Busacca 1994 ; Costabile 1996). Infatti le scelte di differenziazione dell'offerta conseguenti alla segmentazione hanno da un lato l'obiettivo della distinzione competitiva e dall'altro quello dell'ampliamento della varietà offerta in funzione dell'eterogeneità rilevata nelle configurazioni del valore per i clienti (Costabile, 1996).

Sebbene i primi studi in tale direzione risalgano agli anni '60 (YanKelovich, 1964), per molti anni le aziende hanno preferito orientarsi ad approcci "aprioristici" alla segmentazione, vale a dire approcci fondati su "criteri di correlazione" fra variabili descrittive e comportamento d'acquisto[3] (Valdani 1984 e 1985). I motivi di tale atteggiamento sono riconducibili da un lato alla relativa economicità di gestione di tali approcci, dall'altro alle caratteristiche di mercato relativamente omogeneo e caratterizzato da livelli concorrenziali non esasperati.

Dall'altro canto "la tendenza verso il progressivo inasprimento della pressione concorrenziale" riconducibile, fra l'altro alla diffusivita' delle nuove tecnologie, alla flessibilità dei nuovi processi produttivi e allo sviluppo internazionale delle imprese, ha evidenziato sempre con maggior enfasi come un simile approccio alla segmentazione costituiscono sempre più un limite per un impresa che e' chiamata non più a gestire le conseguenze di eventi che si sono già manifestati, bensì di eventi che non si sono ancora evidenziati, un impresa per ciò "proattiva" [4].

Da qui un interesse sempre maggiore per un differente approccio alla segmentazione, cosiddetto "a posteriori", vale a dire fondato in prevalenza sui "criteri della strumentalità del prodotto" e consistenti in processi di segmentazione basati sulle differenze ***osservate*** nella domanda, con riferimento alle modalità d'uso del prodotto, alle percezioni e agli atteggiamenti, ai benefici attesi e al comportamento d'acquisto (Valdani, 1995). In un ottica di analisi del valore, tale approccio permette di superare il principale limite dell'approccio "aprioristico" alla segmentazione legato all'impossibilità di una "misurazione empirica" delle configurazioni del valore che guidano il comportamento di scelta dei clienti. Infatti i segmenti identificati con riferimento per esempio a criteri demografici quali l'età o la provenienza geografica ben difficilmente possono offrire informazioni relativamente alla catena dei valori del cliente cosi da permettere una differenziazione dell'offerta volta all'incremento del valore stesso per il cliente.

La segmentazione "value-based" a posteriori non e' esente da limiti. Innanzitutto a differenza dell'approccio a priori e' caratterizzata da un non indifferente impegno economico relativo alla necessita' di svolgere delle approfondite e particolareggiate indagini sul campo. Soprattutto e' necessario sottolineare come una volta identificati dei gruppi di clienti omogenei in termini di valore percepito non sempre e' possibile individuare delle differenze significative

in termini di profili socio-demografici, vale a dire potrebbe non essere possibile identificare in maniera precisa i target cui veicolare efficaci politiche di marketing mix (comunicazione, promozione e/o distribuzione)[5].

Alla luce di questa premessa, volta a sottolineare come la segmentazione possa rappresentare un valido strumento di analisi capace di affinare le capacita' conoscitive dell'azienda con riferimento all'ambiente in cui opera permettendole quindi, fra l'altro, di anticipare e reagire ai cambiamenti che influenzano il mercato dei suoi prodotti, scopo di questo lavoro vuole essere quello di esporre alcune considerazioni in relazione alle diverse metodologie statistiche e/o matematiche che possono essere impiegate per ottenere una segmentazione della domanda.

All'interno del ricco patrimonio metodologico proposto relativamente ai diversi algoritmi di segmentazione e' possibile individuare due approcci : la segmentazione per ***omogeneità*** e la segmentazione per ***obiettivi***.

Obiettivo delle metodologie proprie del primo approccio (quali per esempio la cluster analysis) e' quello di raggruppare gli elementi del campione sulla base della loro similarita' riferita ad un particolare insieme di variabili (bisogni, attitudini, motivazioni d'uso occasioni d'uso, stili di vita o benefici ricercati) in gruppi caratterizzati da un'elevata variabilità esterna e una elevata omogeneità interna.

Con il secondo approccio si vuole suddividere la popolazione considerata in sub-popolazioni utilizzando una variabile dipendente conosciuta a priori : per questo motivo si parla anche di "tecniche di classificazione" ossia quei processi attraverso i quali si cerca di ridurre la complessita' di specifici fattori ambientali (ad es. mercati, concorrenti, consumatori,...) attraverso il loro raggruppamento in insiemi aventi caratteristiche simili secondo specifici modelli.

Lo scopo di tali modelli e' duplice ; da un lato identificare le variabili che "spiegano" in modo più significativo un determinato comportamento cosi' da fornire al management utili informazioni per l'ottimizzazione delle scelte relative al marketing mix ; dall'altro e' possibile date n osservazioni, distinte in due o più gruppi, stabilire una regola di decisione per assegnare nuove unita ai gruppi, ciò presuppone l'identificazione di una funzione di classificazione che separa in modo ottimale i gruppi, per ridurre il rischio di future classificazioni erronee.

Indipendentemente dal settore in cui opera un azienda (largo consumo, industriale o servizi) le applicazioni di tali metodologie sono molteplici.

Considerando i clienti attuali, l'applicazione delle tecniche di segmentazione per *obiettivi* può essere finalizzata: a) allo studio del grado di conoscenza di un particolare prodotto o azienda; b) alla misurazione della customer satisfaction (elemento chiave per la fidelizzazione della clientela); c) allo studio di diversi comportamenti d'acquisto relativamente sia alle

motivazioni di scelta sia alle occasioni d'uso di un particolare prodotto/servizio; d) alla caratterizzazione dei clienti fedeli; e) all'analisi delle preferenze per diverse forme distributive; f) alla tipizzazione di particolari comportamenti d'acquisto; g) all'analisi degli effetti di diversi mix di prodotto/servizio.

Considerando i clienti potenziali sono da segnalare gli studi finalizzati: a) all'individuazione di nuovi clienti; b) alla selezione dei diversi mix di prodotto/servizi ottimali per fronteggiare la concorrenza; c) allo studio degli effetti dei diversi mix comunicazionali; d) all'analisi della rete di vendita con riferimento sia alla definizione degli obiettivi sia alla valutazione delle performance.

Obiettivi dell'analisi.

Obiettivo del lavoro è presentare una rassegna delle principali tecniche statistiche multivariate utili per elaborare *processi di classificazione*, confrontando nel contempo la capacità delle tecniche stesse di generare risultati utili in termini di validità esplicativa e di ricadute operative.

Di fronte al proliferare di pacchetti software applicativi che propongono tecniche dedicate all'analisi di segmentazione, si vogliono fornire alcuni suggerimenti utili per scelte corrette ed efficaci a tutti coloro che a vario titolo (ricercatori, analisti e manager) affrontano un tema di tale rilevanza.

In particolare si valuterà il ruolo delle reti neurali, uno strumento relativamente recente che, se impiegato congiuntamente alle metodologie statistiche tradizionali, sembra fornire dei risultati particolarmente interessanti.

I database considerati costituiscono un sottoinsieme delle informazioni raccolte in due indagini distinte, relative rispettivamente al lancio di *una carta di credito* (utilizzabile per pagare servizi specifici) e all'analisi della *soddisfazione della clientela* di vari istituti finanziari. Nel primo caso, si è cercato di mettere in relazione l'interesse verso alcune caratteristiche della carta di credito e la propensione al suo impiego; nel secondo caso, si è voluto comprendere in che modo numerosi aspetti della relazione e della struttura dei servizi influenzassero il livello di soddisfazione, misurato come disponibilità a consigliare ad altri il proprio istituto.

Ai fini dell'applicabilità di tutte le tecniche prescelte, si sono considerate unicamente variabili esplicative quantitative, misurate a livello di intervallo (valutazioni di importanza / soddisfazione su caratteristiche dei prodotti / servizi analizzati). Le variabili dipendenti sono invece dicotomiche. Inoltre, in entrambe le analisi, l'80% delle osservazioni è stato utilizzato per stimare i parametri dei diversi modelli e il 20% come test fuori campione.

Tecniche statistiche e software applicativo utilizzate.

Tra le tecniche statistiche multivariate utili per la discriminazione, si sono in particolare considerate l'analisi *discriminante lineare*, la *regressione logistica*, il *metodo CHAID* e le *reti neurali* ; per quanto alcune di queste tecniche risultino ampiamente conosciute, per esigenze di maggior chiarezza di seguito vengono presentati alcuni sintetici cenni metodologici.

Analisi Discriminante Lineare.

L'analisi discriminante consente di studiare ed esaminare le differenze esistenti fra *g* gruppi (definiti a priori, con *g* ε 2) di *n* osservazioni, in relazione a diverse variabili esplicative; grazie alla tecnica, le informazioni sulle medesime variabili possono anche essere in un secondo momento usate per classificare opportunamente nuove osservazioni, di cui non si conosca il gruppo di appartenenza.

Nell'analisi discriminante lineare, le variabili indipendenti sono combinate in una o più equazioni lineari dette *funzioni discriminanti :*

$$d_{ij} = \alpha_{0i} + \alpha_{1i}x_{1j} + \alpha_{2i}x_{2j} + \ldots + \alpha_{pi}x_{pj} \qquad i = 1,2,\ldots,\min(p, g-1);\ j = 1,2,\ldots,n$$

Come evidenziato dalla precedente scrittura, il numero massimo di funzioni discriminanti individuabili è pari al minimo tra il numero di variabili esplicative e il numero di gruppi individuati a priori meno 1. Le analisi condotte nel presente studio riguardano variabili dipendenti dicotomiche e pertanto prevedono la possibilità di calcolare un'unica funzione discriminante ; si noti che i coefficienti delle funzioni discriminanti vengono calcolati in modo che tali funzioni massimizzino la differenza complessiva tra i gruppi in relazione alle variabili indipendenti.

I valori d_{ij}, detti *discriminant scores*, possono essere ricavati per ciascuna delle *j* osservazioni analizzate e, nel caso in cui le variabili indipendenti risultino significative, consentono di prevedere correttamente l'appartenenza ad uno dei gruppi. Tale significativita' viene valutata attraverso un opportuno test F e il relativo *p-value*.

L'ottimalità dell'analisi discriminante lineare è garantita quando quattro ipotesi fondamentali sono soddisfatte. In primo luogo, per ogni variabile esplicativa, le osservazioni devono essere tra di loro indipendenti ; inoltre le variabili esplicative debbono seguire per ciascun gruppo una distribuzione normale multivariata (e ciò implica che ciascuna variabile debba essere normale). In terzo luogo, le varianze delle variabili esplicative per ciascun gruppo dovrebbero risultare nella popolazione uguali o molto simili ; infine, per ciascuna coppia di variabili, le covarianze nella popolazione dovrebbero per i diversi gruppi risultare uguali o molto simili.

Nelle applicazioni pratiche, spesso queste condizioni non risultano soddisfatte ; nonostante ciò, i risultati dell'analisi discriminante sono ritenute in molti casi sufficientemente attendibili per fornire valutazioni qualitative sulla relazione indagata.

Analisi Logit (regressione logistica).

L'analisi logit fornisce ancora un modello lineare facilmente interpretabile per studiare l'andamento di una variabile dipendente qualitativa ; più specificamente, le variabili esplicative sono combinate in una funzione lineare utilizzata per stimare il logaritmo del rapporto tra la probabilità che un evento (appartenenza a un gruppo) accada e la probabilità che lo stesso non accada. Tale logaritmo è noto come "*log odds*" o anche "*logit*" e può essere espresso nel modo seguente :

$$\log \frac{\pi_j}{1-\pi_j} = \log(O_j) = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \beta_p x_{pj} \qquad j = 1,2,\ldots,n$$

Nel nostro caso, i "*log odds*" stimati consentono di prevedere l'appartenenza di un'osservazione ad un determinato gruppo, ancora definito a priori. Più specificamente, per ogni osservazione viene stimata la probabilità di appartenere ad una delle due categorie definita dalla variabile dipendente, dato il livello delle variabile esplicative ; l'osservazione viene quindi collocata nel gruppo con la più elevata probabilità stimata.

Quanto più rilevanti e significative sono le variabili esplicative utilizzate, tanto più precisa è la riassegnazione operata dall'equazione logit sulla base delle variabili esplicative stesse; la verifica di tale significatività avviene utilizzando (nel caso di variabili indipendenti categoriche) un opportuno test χ^2 e il relativo *p-value*.

Il vantaggio rispetto all'analisi discriminante sta nella minore gravosità delle ipotesi sottostanti. Unica condizione per l'applicabilità è infatti che per ogni variabile esplicativa le osservazioni siano indipendenti, mentre non risultano necessarie le ipotesi di normalità e di uniformità delle matrici di varianza e covarianza nei gruppi.

Analisi CHAID.

La tecnica *CHAID* viene tipicamente utilizzata nelle analisi di segmentazione, quando l'obiettivo finale è di ripartire una popolazione (campione) in segmenti molto differenziati rispetto a un criterio di riferimento.

Più specificamente, il metodo *CHAID* si basa su una procedura iterativa. Nella prima fase, la popolazione viene divisa in due o più gruppi sulla base della variabile esplicativa più

significativamente legata alla variabile dipendente prescelta ; ciascuno di questi gruppi ottenuti nella prima fase viene ulteriormente diviso in sottogruppi più piccoli sulla base delle rimanenti variabili esplicative. Il processo di divisione termina quando non possono più essere individuati predittori statisticamente significativi ; i sottogruppi finali (segmenti) vengono quindi rappresentati in un diagramma ad albero.

I segmenti individuati sono mutuamente esclusivi ed esaustivi. In altre parole, i segmenti non si sovrappongono e ciascuna unità della popolazione è contenuta in un unico segmento. Dal momento che ogni segmento è definito da combinazioni delle variabili esplicative, è possibile riclassificare ogni osservazione (azienda) nel segmento appropriato semplicemente conoscendo tali variabili ; per questo motivo il metodo *CHAID* viene talvolta utilizzato in alternativa all'analisi discriminante e al metodo logit.

In particolare nel presente lavoro si è utilizzata una versione dell'algoritmo *CHAID* modificata rispetto a quella proposta da Kass nel 1980.

Le differenze più marcate si riferiscono al metodo impiegato per calcolare il livello di significatività per ciascuna variabile esplicativa. Il fattore di aggiustamento di Bonferroni adottato (necessario per evitare le distorsioni da test ripetuti) risulta meno conservativo rispetto a quello utilizzato nell'algoritmo originario proposto da Kass.

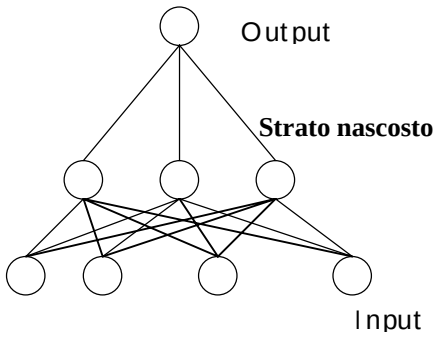
Reti Neurali.

Le reti neurali sono state originariamente create con l'intento di studiare il comportamento del sistema nervoso. Come è noto, tale sistema è costituito da un numero elevatissimo di cellule nervose (*neuroni*), collegati da connessioni dette *sinapsi*. Ogni neurone è in grado di eccitarsi e di trasmettere uno stimolo elettrico in risposta alle combinazioni di stimoli che riceve dagli altri neuroni. La trasmissione degli stimoli dipende dalla natura delle sinapsi e tale natura si modifica durante la crescita per effetto dell'*apprendimento*.

In ambito economico, le reti neurali hanno cominciato ad essere impiegate in alternativa ai più semplici e comuni modelli ipotetico - deduttivi, per superare i vincoli posti da tali modelli alla ricerca delle reali dipendenze. Le reti neurali sono infatti in grado di realizzare funzioni molto complesse a più valori in più variabili apprendendo da sole come riprodurre i comportamenti osservati sperimentalmente ; particolari tecniche di addestramento consentono inoltre di trovare le strutture più semplici per approssimare gli andamenti studiati.

Uno degli algoritmi più utilizzati è quello di *retropropagazione dell'errore*. Esso si basa sul confronto iterativo tra la prestazione fornita da una rete, in risposta ad una combinazione di stimoli in ingresso, con quella desiderata e nella conseguente modifica delle interconnessioni, fino ad ottenere la prestazione voluta.

L'impiego di una rete neurale per la segmentazione a priori prevede, nel caso più semplice, il ricorso ad una rete strutturata secondo tre livelli: un livello di unità di input, che riunisce i neuroni riferiti a ciascuna variabile di input, uno o più livelli nascosti e un livello di unità di output cui è associata la o le variabile dipendente/i.



Il principale limite all'impiego di tale metodologia nell'analisi dei dati di marketing è legato alla difficoltà di "spiegare" il comportamento di una rete. Infatti una volta ottenuta una struttura che sia in grado non solo di "fittare" al meglio l'insieme dei dati forniti, ma anche di generalizzare interpretando delle nuove osservazioni, resta sempre il problema di poter valutare il ruolo delle diverse variabili indipendenti. A differenza delle tecniche statistiche non è possibile, se non in via approssimata, attribuire un peso alle diverse variabili esplicative. Per ovviare almeno in parte a questa limitazione e per poter comparare i risultati ottenuti con questa tecnica con quelli propri delle tecniche tradizionali, nel presente lavoro si è utilizzato il metodo di interpretazione proposto per una rete con un unico strato nascosto da Yoon, Swales e Margavio [1993]. In particolare il legame esistente tra ciascuna variabile di input X_i e di output O_j è misurato come:

$$\mathrm{RS}_{ji} = \frac{\sum_{k=1}^{n} W_{ki} U_{jk}}{\sum_{i=1}^{m} ABS\left(\sum_{k=1}^{n} W_{ki} U_{jk}\right)}$$

RS_{ji} è il peso relativo tra la *i*-esima (tra *m*) unità di input e la *j*-esima unità di output, W_{ki} è il peso del legame tra la *k*-esima (tra *n*) unità nascosta e la *i*-esima unità di input, e U_{jk} è il peso della "sinapsi" che collega la *j*-sima unità di output alla *k*-esima unità nascosta.

Software applicativo utilizzato.

Non essendo ancora al momento disponibile la versione definitiva di un pacchetto statistico in cui risultino integrate tutte le procedure citate, si è fatto ricorso a software applicativi diversi (PC IBM compatibili):

SPSS per Windows 6.1 (analisi discriminante lineare e regressione logistica)

Knowledge Seeker 3.0 (CHAID)

WinNN 0.96 (reti neurali - algoritmo di back-propagation)

I risultati dell'analisi.

Primo esempio : la proposta di una nuova carta di credito.

Il primo database utilizzato è emerso dall'indagine condotta su 350 aziende con personale che effettua viaggi all'estero per motivi di lavoro, intervistando il dirigente responsabile per la scelta degli strumenti finanziari (alternativamente il direttore amministrativo, il direttore del personale, il responsabile dell'ufficio viaggi, il titolare). Più specificamente, si è cercato di mettere in relazione l'interesse verso alcune caratteristiche della nuova carta di credito proposta (*tav. 1*) e la propensione alla sua adozione (*tav. 2*). L'inclusione nell'analisi del livello di soddisfazione globale è dovuto alla necessità di verificare l'influenza del gradimento per il mix di caratteristiche proposte sulla propensione ad adottare la carta.

Tav. 1 - Variabili esplicative: gradimento per caratteristiche carta
1) diffusione carta
2) semplicità uso
3) modalità pagamento
4) completezza documentazione fatture
5) costo carta "base"
6) sconti fedeltà
7) agevolazioni/ premi
8) limite spesa
9) soddisfazione globale

Come precedentemente evidenziato, circa il 20% delle osservazioni è stato utilizzato per una verifica della capacità esplicativa al di fuori del campione utilizzato per la stima dei parametri, evidenziando quindi la robustezza dei modelli proposti. Si noti inoltre che il campione risulta quasi perfettamente diviso tra favorevoli e non favorevoli all'adozione della carta proposta.

Nel caso delle reti neurali, si sono incluse nell'analisi tutte le variabili esplicative e sono stati verificati modelli a tre strati con differente numero di unità nello strato nascosto ; il trade off tra complessità della rete e capacità previsiva ha fatto cadere la scelta su una rete con quattro unità nello strato nascosto.

Nel caso delle rimanenti tecniche si sono utilizzate procedure di selezione delle variabili, adattando i livelli di significatività sulla base della capacità previsiva in campione e fuori campione.

Tav. 2 - Variabile dipendente:
propensione a suggerire uso carta (1=sì / 2=no)

% favorevoli nel campione:	54.6%
% non favorevoli nel campione:	45.4%
n. oss. utilizzate per stima:	281
n. oss. utilizzate per controllo:	69
n. totale osservazioni:	350

In *tav. 3* sono riportati tassi di riclassificazione corretta in campione e fuori campione. Come si può facilmente notare, il modello basato sulle reti neurali mostra una performance chiaramente superiore alle rimanenti tecniche relativamente alle previsioni "interne" al campione ; tale vantaggio viene tuttavia praticamente annullato considerando la capacità previsiva fuori campione, che risulta equivalente a quella dell'analisi discriminante lineare e molto vicino a quello degli altri due metodi. Si noti inoltre come le performance con le osservazioni di controllo risultino in questo caso superiori rispetto a quelle ottenute con le osservazioni utilizzate per la stima dei modelli ; nel momento in cui risulterà più facile il dialogo tra i differenti pacchetti e/o tutte le procedure verranno integrate in un unico applicativo, sarà evidentemente opportuno condurre analisi di sensitività del risultato a modifiche nella partizione del campione totale.

Tav. 3 - Sintesi dei risultati:
% casi correttamente riclassificati

Metodo	*campione*	*controllo*
Analisi discriminante lineare	77.2%	82.6%
Regressione logistica	75.4%	81.1%
CHAID	76.9%	81.1%
Reti Neurali	86.8%	82.6%

Nella *tav. 4* ritroviamo invece esplicitato l'ordine di significatività delle variabili esplicative utilizzate. I risultati hanno necessariamente un'interpretazione differente per le diverse tecniche.

Più specificamente, nel caso dell'analisi discriminante l'ordine di significatività riguarda rispettivamente il test F relativo alle procedura di selezione delle variabili ; le variabili effettivamente incluse nel modello (con un limite inferiore di significatività pari al 90%) sono state evidenziate con un asterisco, mentre per le rimanenti si è indicato l'ordine della correlazione con la funzione discriminante individuata.

Nel caso dell'analisi logit, l'ordine di significatività delle variabili incluse è legato alla sequenza con cui sono state inserite nel modello (ancora con un limite inferiore del 90%), mentre per le rimanenti variabili si è presentato l'ordine di significatività legato alla potenziale inclusione nel modello nella fase finale della procedura di selezione. Si noti che sia nel caso dell'analisi discriminante che dell'analisi logit la considerazione dei modelli completi (con tutte le variabili esplicative incluse) non porta a miglioramenti nella capacità di riclassificazione.

Nel caso del metodo CHAID, l'ordine presentato per le variabili inserite nel modello riguarda la sequenza nell'inclusione (gestita manualmente per migliorare la performance del modello), mentre per le variabili escluse si è presentato l'ordine di significatività della relazione con la variabile dipendente.

Per quanto riguarda infine la misura di impatto delle variabili nel caso delle reti neurali, si è utilizzato il già citato metodo di interpretazione proposto da Yoon, Swales e Margavio, includendo tutte le variabili esplicative nell'analisi.

Tav. 4 - Sintesi dei risultati: ordine significatività variabili

	a.d.l.m.	*r.l.*	*CHAID*	*reti*
diffusione carta	8	7	1(*)	6
semplicità uso	6	6	7	3
modalità pagamento	9	5	8	8
completezza documentazione fatture	5	8	9	5
costo carta "base"	4	4	6	4
sconti fedeltà	2 (*)	2 (*)	3	7
agevolazioni/ premi	7	9	4	9
limite spesa	3 (*)	3	5	2
soddisfazione globale	1 (*)	1 (*)	2(*)	1

Come si può facilmente osservare, il *livello di soddisfazione globale* risulta estremamente rilevante per tutti i modelli di analisi, mentre *limite spesa* e *sconti fedeltà* in tre casi su quattro. Eccezioni notevoli sono *diffusione carta* e *semplicità uso*, rilevanti rispettivamente solo nel caso di CHAID e delle reti. In estrema sintesi, esiste pertanto coerenza tra i risultati, ma non si può parlare di totale sovrapponibilità.

La *tav. 5* presenta con maggiore dettaglio i dati evidenziati nella *tav. 3*. Si osservi in particolare come la performance dell'analisi discriminante lineare e dell'analisi CHAID sia soprattutto legata alla capacità di individuare correttamente una categoria di intervistati (i favorevoli all'adozione della carta), mostrando invece tassi di errore rilevanti per la rimanente categoria. Più omogeneo è invece il comportamento della regressione logistica e delle reti neurali.

Tav. 5 -Dettaglio risultati: % casi correttamente riclassificati per segmento

Metodo	*campione*		*controllo*	
	favorev.	non favor.	favorev.	non favor.
Analisi discriminante lineare	84.0%	69.5%	95.0%	64.3%
Regressione logistica	77.3%	73.3%	87.8%	71.4%
CHAID	88.0%	64.0%	92.7%	64.3%
Reti Neurali	92.0%	80.9%	87.8%	75.0%

Secondo esempio: il grado di "fidelizzazione" della clientela di un istituto finanziario.

Il secondo database impiegato è relativo ad un'indagine su un campione di circa 280 clienti di un istituto finanziario ai quali era stato chiesto, fra l'altro, un giudizio "overall" sull'istituto stesso ed una valutazione dell'importanza da loro attribuita ad alcuni elementi che caratterizzano il processo di scelta. Più in dettaglio si voleva mettere in relazione l'interesse verso alcune caratteristiche del servizio ed il giudizio complessivo sull'istituto (*tav. 6*) con la propensione a consigliarlo a parenti e amici (*tav. 7*).

Tav. 6 - Variabili esplicative: importanza attributi
1) costo del servizio
2) tempistica invio avviso di scadenza
3) servizi accessori
4) tempo disbrigo pratiche
5) tempo necessario per la liquidazione
6) adeguatezza del rimborso
7) aspetto esteriore degli uffici
8) assistenza collaborazione
9) orari di apertura
10) chiarezza dei contenuti del contratto
11) conoscenza del personale
12) facilità di lettura del contratto
13) fiducia nell'operatore
14) garanzia del nome della società
15) solidità economica della società
16) vicinanza / comodità degli uffici
17) adeguamento del valore
18) giudizio complessivo

Anche in questo caso circa il 20% delle osservazioni è stato dedicato alla verifica della capacità esplicativa esterna al campione utilizzato per la stima dei parametri, evidenziando quindi la robustezza dei modelli proposti.

Tav. 7 - Variabile dipendente: propensione a consigliare la propria società (1=sì / 2=no)	
% propensi a consigliare nel campione	53.7%
% non propensi a consigliare nel campione	46.3%
n. oss. utilizzate per stima:	229
n. oss. utilizzate per controllo:	54
n. totale osservazioni:	283

In *tav. 8* sono riportati tassi di riclassificazione corretta in campione e fuori campione. Come si può notare, il modello basato sulle reti neurali (con tre unità nello strato nascosto) mostra ancora una volta, anche se in misura inferiore rispetto all'esempio precedente, una performance superiore alle rimanenti tecniche relativamente alle previsioni "interne" al campione ; tale vantaggio viene annullato considerando la capacità previsiva fuori campione, che risulta inferiore rispetto a quella ottenuta con le altre tre metodologie, con una leggera dominanza di regressione logistica e CHAID.

Tav. 8 - Sintesi dei risultati: % casi correttamente riclassificati		
Metodo	*campione*	*controllo*
Analisi discriminante lineare	67.7%	55.6%
Regressione logistica	66.4%	57.4%
CHAID	66.1%	56.4%
Reti Neurali	69.9%	53.7%

Nella *tav.9* ritroviamo invece esplicitato l'ordine di significatività delle variabili esplicative utilizzate ; le modalità di interpretazione dell'impatto sono le medesime descritte precedentemente per il primo esempio applicativo. I livelli di significatività utilizzati per le procedure di selezione delle variabili, pur essendo più bassi rispetto alla precedente verifica applicativa (livello intorno al 70%), sono stati scelti in modo da massimizzare la capacità di riclassificazione di analisi discriminante, regressione logistica e CHAID in campione e fuori campione.

Entrando nello specifico, il *giudizio complessivo* risulta estremamente rilevante per i modelli "tradizionali" di analisi, mentre *il tempo necessario per la liquidazione, la fiducia*

nell'operatore e la chiarezza dei contenuti del contratto emergono nel modello ottenuto con le reti neurali.

In sintesi, appaiono ancora una volta sostanzialmente sovrapponibili i risultati relativi all'analisi discriminante lineare, alla regressione logistica e al metodo CHAID, mentre si differenziano chiaramente quelli ricavati con le reti neurali.

Tav. 9 - Sintesi dei risultati: ordine significatività variabili

	a.d.l.m.	*r.l.*	*CHAID*	*reti*
1) costo del servizio	5(*)	6(*)	4	12
2) tempistica invio avviso di scadenza	14	14	9	13
3) servizi accessori	18	10(*)	11	18
4) tempo disbrigo pratiche	2(*)	2(*)	2(*)	11
5) tempo necessario per la liquidazione	4(*)	4(*)	6	1
6) adeguatezza del rimborso	8	9(*)	5	17
7) aspetto esteriore degli uffici	3(*)	3(*)	3(*)	15
8) assistenza collaborazione	17	13	12	10
9) orari di apertura	15	11	15	16
10) chiarezza dei contenuti del contratto	10	16	13	3
11) conoscenza del personale	13	12	7	7
12) facilità di lettura del contratto	9	8(*)	10	14
13) fiducia nell'operatore	6(*)	5(*)	17	2
14) garanzia del nome della società	7(*)	7(*)	18	6
15) solidità economica della società	12	17	8	5
16) vicinanza / comodità degli uffici	11	18	14	8
17) adeguamento del valore	16	15	16	9
18) giudizio complessivo	1(*)	1(*)	1(*)	4

La *tav. 10* presenta con maggiore dettaglio i dati evidenziati nella *tav.8*.

L'analisi CHAID è caratterizzata da una performance legata soprattutto alla capacità di individuare correttamente una categoria di intervistati (i propensi a consigliare la propria società), mostrando invece tassi di errore rilevanti per la rimanente categoria. Omogeneo risulta invece il comportamento delle restanti metodologie.

Tav.10 -Dettaglio risultati: % casi correttamente riclassificati per segmento

Metodo	*campione*		*controllo*	
	propensi	non prop.	propensi	non prop.
Analisi discriminante lineare	70.5%	64.5%	60.0%	50.0%
Regressione logistica	75.4%	56.0%	66.7%	45.8%
CHAID	77.0%	53.3%	70.0%	41.7%
Reti Neurali	80.3%	57.9%	60.0%	45.8%

Conclusioni.

In entrambe le analisi condotte, i tassi di corretta riclassificazione nel campione risultano più elevati con l'impiego delle reti neurali. Tuttavia, questa situazione di dominanza non si manifesta per la riclassificazione fuori campione: il metodo di regressione logistica e l'analisi discriminante lineare (nonostante il mancato superamento dei test di uguaglianza delle matrici di varianza e covarianza) sembrano infatti rivelarsi un poco più efficaci nella generalizzazione. La performance della tecnica CHAID risulta valida globalmente, ma si evidenziano alcune difficoltà a evidenziare con precisione specifici segmenti di clientela.

Infine, relativamente all'impatto delle variabili esplicative, le tecniche tradizionali hanno fornito indicazioni pressoché equivalenti, mentre qualche differenza significativa sembra emergere con il metodo delle reti neurali.

Bibliografia:

Biggs D., Barry de Ville & ED Suen, "A method of choosing multiway partitions for classification and decision trees" Journal of Applied Statistics, vol.18 No. 1, 1991 pp. 49-62.

Busacca B., Le risorse di fiducia dell'impresa", UTET, Torino, 1994

Chilanti M., "I Modelli di Previsione delle Insolvenze mediante Reti Neurali o Analisi Discriminante: Quale Confronto?" - AF - n. 15 - 1994

Costabile M. "Misurare il valore per il cliente" UTET Torino1996

DeMaris, A., "Logit Modeling : Practical Applications", Sage University Paper Series on Quantitative Applications

Fabbri G., Orsini R. , "Reti Neurali per le Scienze Economiche", Muzzio Editore, 1993.

Kotler, P., "Marketing management - analysis, planning, implementation and control, 7 th edition, Prentice-Hall International 1991

Kass, G.V. "An exploratory technique for investigating large quantities of categorical data", Applied Statistics, 29,2, pp.119-127.

Magidson J., "Improved Statistical Techniques for Response Modeling" - Journal of Direct Marketing - vol. 2 n. 4 - Autunno 1988

Magidson J., "The Chaid Approach to Segmentation Modeling" in *Advanced Methods of Marketing Research* - a cura di R.P. Bagozzi - Basil Blackwell - 1994

Molteni L., "L'analisi multivariata nelle ricerche di marketing" - EGEA 1993.

Vacca' S., Imprese e sistema industriale in una fase di rapida trasformazione tecnologica", Giuffre', Milano 1986

Valdani E., "Le decisioni di marketing nell'impresa industriale, Giuffre', Milano 1984

Valdani E., "Marketing Strategico - Un'impresa proattiva per sviluppare capacita' market driving e valore" ETASLIBRI 1995.

Vicari S., Troilo G., "Errore e apprendimento nelle politiche di marketing" in "Scritti in onore di Carlo Fabrizi" AAVV, CEDAM Padova 1995

Vicari S., " Strategie competitive nei mercati senza confine" Economia & Management, vol. 1 EGEA 1988

Vicari S., " Nuove dimensioni della concorrenza - strategie nei mercati senza confini. EGEA 1989

Yoon, Swales e Margavio, "A Comparison of Discriminant Analysis versus Artificial Neural Networks", J. Opl. Res. Soc. 44, n. 1, pp. 51-60 - 1993.

Note:

[1] Tale definizione e' stata sviluppata da un gruppo di studiosi all'interno dell'American Marketing Association nel 1985.

[2] L'altra determinate e' la **variabilità.**

[3] Per un maggiore approfondimento dei diversi criteri di segmentazione "a priori" : criteri geografici, criteri socio-demografici, criteri psicografici, criteri comportamentali, si rimanda fra gli altri a Costabile 1996.

[4] Su tali aspetti per un approfondimento si rimanda, fra gli altri, a Vacca 1986, Vicari 1989, Valdani 1991, Busacca 1994 e Valdani 1995.

[5] Al riguardo e' stato osservato, comunque, che "e' lecito domandarsi se la mancanza di differenze sotto il profilo descrittivo non renda, altresì inefficaci i convenzionali approcci aprioristici alla segmentazione.....Con la significativa differenza che, nel caso di una segmentazione "value based", a conclusione del processo il management dell'impresa ha comunque sviluppato la consapevolezza delle configurazioni d'offerta che generano maggiore valore per la domanda" (Costabile, 1996). Inoltre la definizione delle azioni di marketing dovrebbe comunque essere indirizzata secondo quanto emerso dalla segmentazione "value-based" relativamente all'omogeneità per valori, pervenendo alla descrizione dei segmenti in un secondo tempo con il sostegno delle diverse tecnologie interattive oggi giorno sempre più' presenti (numeri verdi, sistemi distributivi on demamd, internet,...) (Vicari e Troilo, 1995).