# UN MODELLO NEURALE PER L'ADEGUAMENTO DEI CAMPIONI ESTIMATIVI "SCARSI" DEFINITO PER LA SELEZIONE DEI CANONI DI LOCAZIONE DA BANCHE-DATI

Gianluigi De Mare\*

# INTRODUZIONE¹ ED OBIETTIVI

Fin dai primi anni '90, la necessità di avviare una riforma strutturale del meccanismo di revisione degli estimi, finalizzata alla individuazione di parametri che fossero effettivamente ancorati al mercato reale<sup>2</sup>, ha innescato un confronto dialettico sugli elementi da porre a riferimento per la definizione dell'imponibile catastale. Il dibattito, svoltosi in ambito giuridico e legislativo, ha portato in seguito a riconoscere l'esigenza di affidare un ruolo di rilievo ai valori immobiliari<sup>3</sup>. Questa scelta è stata motivata dal carattere poco trasparente del mercato delle locazioni, fortemente

<sup>\*</sup> Ingegnere, Dottore di Ricerca in Metodi di Valutazione.

Le premesse di carattere legislativo, qui sinteticamente richiamate, sono in dettaglio riportate nel lavoro di N. Morano e G. De Mare, *Il regime attuale delle locazioni per la revisione degli estimi urbani*, pubblicato in questo stesso volume degli atti della ricerca. A tale lavoro si rimanda per ogni approfondimento anche in merito alla fase di rilevazione e selezione dei dati.

<sup>&</sup>lt;sup>2</sup> Vedi A. De Santis, Processo di integrazione catasto conservatorie: un importante passo verso la probatorietà, XXVIII incontro Ce.S.E.T., Roma 19/10/98.

<sup>&</sup>lt;sup>3</sup> L'utilizzo dei valori per la revisione è previsto già del DPR 1142/49, ma al loro impiego si riserva una condizione di eccezionalità connessa con l'eventuale assenza o inefficienza del mercato delle locazioni.

influenzato dal regime vincolistico all'epoca in vigore<sup>4</sup>. Sui valori immobiliari furono determinate le tariffe di estimo urbano approvate e promulgate con Decreti Ministeriali 20/01/90 e 27/09/91, poi annullati dal Tar del Lazio con sentenza n.1184 del 1992.

Si è passati dopo ad un'apparente equiparazione dei due procedimenti di stima (art. 5 del DPR 138/98): quello basato sui canoni annui ordinariamente ritraibili e quello basato sui valori di mercato degli immobili.

Le recenti norme che regolano il sistema delle locazioni<sup>5</sup>, hanno capovolto lo scenario che aveva acceso il dibattito.

Con le leggi intervenute negli ultimi 5 anni, il mercato delle locazioni ha acquisito i connotati di una fonte di dati utili per la revisione degli estimi urbani, molto più consistente e trasparente rispetto a quella del mercato delle compravendite.

Per il 2000, da 14.980 atti registrati presso l'Agenzia delle Entrate di Salerno, 1.160 sono risultati atti di locazioni residenziali stipulati esclusivamente nell'area urbana del capoluogo. Gli atti classificabili nella tipologia del "canale libero" di cui alla legge 431 del 1998 e nel filone dei "patti in deroga" di cui alla legge 359 del 1992 sono risultati 450. Sono queste le due tipologie contrattuali, tra le tante attualmente esistenti, che rispondono ai vincoli della normativa vigente sulla revisione degli estimi urbani.

Con il presente lavoro è stato definito un modello econometrico che utilizza le reti neurali. La funzione del modello è quella di provvedere, una volta addestrata la rete sui canoni reali di

<sup>4</sup> Vedi P.Carrer in: Sulla fiscalità immobiliare un po' di chiarezza non fa mai male, Osservatorio NOMISMA sul Mercato Immobiliare, Il rapporto quadrimestrale, luglio 1991; Precisazioni di natura tecnica ed economico estimativa afferenti le nuove tariffe d'estimo, Rivista del Catasto n.2/92; I valori di mercato dei nuovi estimi non capovolgono i criteri di imposta, Il sole 24 ore, n.22 del 16/05/92.

<sup>&</sup>lt;sup>5</sup> Vedi N.Morano e G.De Mare, op.cit.

mercato, alla selezione - dai 450 contratti di locazione residenziale - di quelli con canone di ammontare prossimo ai livelli di mercato. Il campione finale viene costituito con i dati rilevati dal mercato e con quelli selezionati dal modello. Assume così consistenza adeguata per l'applicazione di procedimenti statistico-estimativi.

Caratteristiche operative del modello sono la rapidità di esecuzione, la facilità di automatizzazione, l'oggettività e la semplicità di impiego ai fini di un agevole trattamento di masse di dati.

### I DATI DISPONIBILI SUL TERRITORIO

Per la definizione e la sperimentazione del modello lo studio è stato circoscritto alla microzona catastale n.5 del territorio del comune di Salerno, per la quale gli atti di locazione residenziale enucleati dai 1.160 riscontrati per l'intero centro urbano, sono risultati in tutto 397.

Si tratta di un'area piuttosto estesa che confina con il centro attraverso il quartiere del "Torrione", longitudinalmente limitata ad ovest dal mare e ad est dalla sede della ferrovia, e che si estende a sud fino alla zona di "Parco Arbostella". Nella fascia più centrale (Via Baratta, S.Mobilio, Guercio, Torrione) vi sono edifici multipiano di buona qualità edilizia. Spostandosi verso sud, rimangono ben curate le arterie principali di scorrimento (Via Posidonia, Trento, Lungomare Marconi, Colombo) mentre più modesta è la qualità edilizia nelle aree interne, per divenire in talune spiccatamente popolare. In Via Pisacane e in Via Volontari della Libertà la qualità dell'edificato è di medio livello, di contro diffusa è la presenza di verde ed ampi sono gli spazi di rispetto tra i fabbricati.

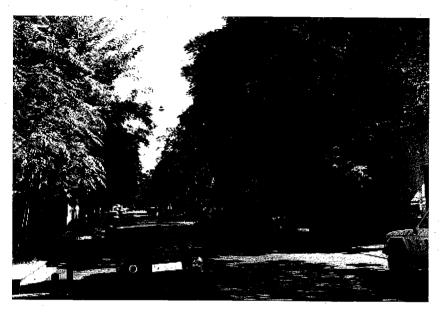
Molto scarsa è in generale, per l'intera microzona, la presenza di posti auto di pertinenza degli appartamenti. Tutti i palazzi che lo necessitano sono dotati di ascensore. Carente la disponibilità di spazi verdi, mentre i trasporti pubblici sono garantiti su gomma e con buona efficienza prevalentemente lungo le principali direttrici di traffico (circuito del lungomare con traiettoria interna lungo Via Trieste e Posidonia). Risultano piuttosto diffusi gli esercizi commerciali, ben concentrate le scuole ed i pochi impianti sportivi. Particolare cura è posta dai condomini e dagli inquilini alla manutenzione degli spazi comuni degli edifici.

Esempi della qualità edilizia di porzioni del territorio di indagine vengono riportate nelle fotografie seguenti.



Figura 1 – La centrale via S. Mobilio.

Figura 2 - Via Pisacane, alberature diffuse.



Per l'analisi statistico-estimativa è stato costituito un campione di studio formato da 155 contratti di locazione stipulati nel 2000. Sono contratti che rientrano nel "canale libero" della 431/98 o nella normativa dei "patti in deroga" (ex lege 359/92). Di essi sono stati rilevati la data di stipula, la durata, le condizioni di adeguamento al costo della vita e le forme di aggiornamento, l'importo del canone, l'ammontare della cauzione, le caratteristiche dei contraenti, l'ubicazione dell'immobile (via, numero civico, piano, interno), la consistenza (in vani), la presenza di accessori e di mobilio. Sull'intero campione è stata poi condotta un'indagine ricognitiva per la determinazione delle caratteristiche previste dal DPR 138/98 all'art.8 comma 7, riguardanti il "fattore edilizio".

E' stato in tal modo costituito un database alfanumerico digitale e fotografico.

Sempre per le unità immobiliari residenziali sono state quindi verificate le condizioni del mercato delle locazioni nella zona di indagine mediante ricorso ai dati delle agenzie appartenenti alla federazione italiana degli agenti immobiliari professionali (FIAIP). Sono stati così raccolti i dati di mercato relativi a 20 contratti stipulati nel corso del 2000, per i quali pure è stato costituito un database.

#### IL MODELLO

Il modello econometrico qui proposto è fondato sull'utilizzo di reti neurali, specificamente sull'impiego di un modello di rete noto come *Multi-layer perceptron* ovvero percettrone multistrato. Partendo dai due insiemi di dati sopra descritti, quello composto dai 155 contratti registrati, rilevati cioè presso l'Agenzia delle Entrate, e l'altro costruito sui dati forniti per 20 contratti dalle agenzie immobiliari FIAIP, vengono individuati dal modello, tra i contratti di locazione registrati, quelli di essi che riportano un canone "veritiero", ossia un ammontare che con "buona" probabilità è quello realmente pattuito.

La scelta dell'algoritmo neurale è connessa con l'obiettivo di pervenire a risultati affidabili ma, peraltro, generabili da operatori privi di expertise spinte in campo statistico. Le reti neurali, tra l'altro, ben si adattano alla natura del problema per la cui soluzione a poco serve la interpretazione e la visualizzazione delle forme funzionali del fenomeno. Le reti neurali sono infatti come scatole nere in grado di sviluppare funzioni interpretative, lineari e non, molto complesse, le cui forme rimangono però occulte all'operatore, il quale deve regolarsi in base alla bontà logico-statistica dei risultati ed, ovviamente, ai parametri di efficienza delle funzioni<sup>6</sup>.

<sup>&</sup>lt;sup>6</sup> Vedi S.Zani, Analisi dei dati statistici, Vol.II, Giuffrè, Milano 2000, pg. 397.

Nella fattispecie, un altro vantaggio derivante dall'uso di tale strumento consegue dalla possibilità di poter addestrare la rete attraverso un set di Training (cosa invece impossibile in altre tecniche di classificazione quale ad esempio la *Cluster analisys*). I neuroni artificiali apprendono i principi logico-matematici sui quali dovranno valutare i casi che verranno ad essi successivamente sottoposti (Testing set) per la enucleazione di quelli congruenti con il tipo definito, oggetto della ricerca.

## IL DATASET DI ADDESTRAMENTO (TRAINING)

Per la fase di apprendimento, come si è detto, debbono essere selezionati due set di dati. Il primo è costituito dall'insieme dei dati delle locazioni effettivamente rilevate sul mercato, insieme M, mentre il secondo insieme R<sub>i</sub>, archivia un sottoinsieme di dati tratti dai contratti ufficialmente registrati presso l'Agenzia delle Entrate. Per ognuno dei casi dei due set di dati vanno portate in conto le seguenti variabili:

- ammontare del canone di locazione (milioni/mese);
- consistenza dell'unità residenziale (vani);
- livello di piano;
- livello delle finiture esterne dell'edificio in cui l'unità è ubicata;
- livello delle finiture interne dell'edificio in cui l'unità è ubicata;
- ambito urbano o zona di ubicazione dell'immobile.

Dalla consistenza dell'unità, misurata in vani, vanno esclusi i servizi e la cucina, oltreché i disimpegni e l'ingresso, quando non costituiscano vani indipendenti.

Per il livello di piano è utilizzata una scala di punti per la quale

all'*iesimo* piano corrisponde un punteggio pari ad *i*. Unica eccezione riguarda le unità al pian terreno, per le quali si adopera il punteggio 0,5 e non 0, uguale a quello utilizzato per i piani rialzati.

Per la valutazione delle *finiture esterne* ed *interne* dell'edificio è utilizzata una scheda che considera le principali variabili connotative della qualità. La scala di valutazione adoperata è la seguente: ottimo punti 5, normale punti 3, scadente punti 1.

# Scheda di rilevazione delle caratteristiche edilizie

Finiture esterne prev	<u>raienti                                   </u>		
_	Ottimo	Stato d'uso Normale	Scadente
Pittura		Ш	
Intonaco			
Piastrelle			
Altro			·
<u> </u>			
Finiture interne prev	alenti		·
		Stato d'uso	
	Ottimo	Normale	Scadente
1) Tipo prevalente di pavimentazione di atrio e scale			
Marmo o pietra naturale			
simili			
Materiali plastici e simili			
2) Tipo prevalente di finitura delle pareti di atrio e sc	ale		
Pittura			
Intonaco			
Piastrelle			
Plastica		[	

Per condizioni di manutenzione intermedie vanno impiegati i valori 2 e 4. Per le *finiture interne* è definito un solo indicatore sintetico, somma dei valori raggiunti dalle parti costituenti (pavimentazioni e pareti). L'indicatore va decurtato del 15% per pavimentazioni e finiture delle pareti realizzate in materiale plastico o simile.

All'ambito urbano o zona di ubicazione dell'immobile è attribuito il seguente punteggio: I per le aree più popolari e degradate, 2 per le zone in condizioni ordinarie caratterizzate da edilizia di livello medio (Via Baratta, Mobilio, Guercio, Torrione, Trento, Posidonia), 3 per tutta la fascia costiera, in ragione dell'incremento dovuto alla panoramicità. Al risultato del computo è aggiunto un maggior valore pari a 0,5 quando l'unità abitativa è ubicata ad un livello di piano dal quale gode di particolare esposizione e panoramicità.

Poiché l'addestramento ha la funzione di regolare e modificare la struttura della rete neurale per la successiva analisi e classificazione del campione di Testing, diventa necessario che i dati di Training contengano informazioni corrette sulle differenze tra i due dataset M e R<sub>i</sub>. Perché la rete interpreti tale differenza in funzione dell'obiettivo è necessario rapportare le caratteristiche immobiliari individuate e misurate, che concorrono alla formazione del canone, con il canone stesso. Pertanto le variabili di input non sono quelle dei dataset iniziali ma sono ottenute dividendo il canone per le variabili originarie. Si hanno quindi:

- canone/vano (milioni/mese x vano);
- canone/livello di piano (milioni/mese x piano);
- canone/finiture esterne (milioni/mese x indicatore);
- canone/finiture interne (milioni/mese x indicatore);
- canone/zona (milioni/mese x indicatore).

# LE FASI DEL MODELLO

Per rendere la procedura immune da errori che potessero scaturire da meccanismi di selezione soggettivi si è definito un approccio di tipo probabilistico-iterativo. La fase di Training viene condotta sull'insieme M, costituito dai dati mercantili relativi a 20 contratti di locazione, e sull'insieme R<sub>i</sub>, che per analogia di consistenza è anch'esso formato dai dati concernenti 20 casi. In questo caso però relativi a contratti registrati presso l'Agenzia delle Entrate. Il dataset M rimane invariato anche in seguito. In effetti la difficoltà da superare è proprio la scarsità dei dati di mercato. I 20 casi registrati, invece, variano ad ogni iterazione della procedura, in quanto scelti di volta in volta in modo casuale tra i 155 costituenti il campione di partenza. Per ogni iterazione, la rete neurale, una volta addestrata, è in grado di distinguere, tra i rimanenti 135 casi, quelli che presentano canoni corrispondenti ai livelli espressi dal mercato, da quelli i cui canoni sono affetti da probabili "decurtazioni strategiche".

Al fine di ottenere una sufficiente stabilità dei risultati, fortemente influenzati dalle caratteristiche del Training set, la procedura è implementata per un numero n (minimo n = 10) di iterazioni. I risultati che vengono confermati dall'80% delle iterazioni sono considerati affidabili.

### APPLICAZIONE E RISULTATI

Una rete neurale, dunque, impara dal Training set. Se ad essa viene attribuita una struttura molto complessa, per tipo di funzioni di trasferimento prescelte e/o per numero di layer e di neuroni adottato, la sua capacità di imparare dal Training set si amplifica. Di contro, con l'aumentare della complessità della struttura, aumentano le probabilità di incorrere nel problema dell'overfitting,

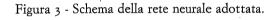
<sup>&</sup>lt;sup>7</sup> Vedi S.Zani, op.cit., pg. 394.

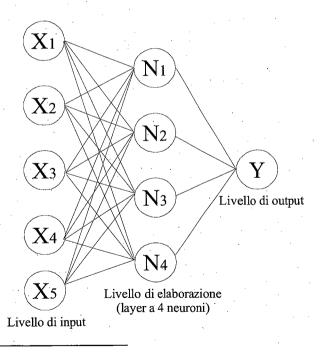
in pratica si riduce la flessibilità del modello ovvero la sua capacità di generalizzare l'analisi. Ciò, a causa dell'irrigidimento dei criteri selettivi utilizzati sui dati di Testing.

Al giusto equilibrio della struttura concorrono l'esperienza del progettista della rete, il tipo di problema da affrontare e l'impiego di regole empiriche<sup>8</sup>.

Nel caso in esame è stata adottata una rete con un unico layer, 4 neuroni ed una funzione di trasferimento lineare (vedi figura 1). Gli esperimenti condotti con modelli più complessi si sono rivelati inefficaci.

La tabella 1, riportata a pag. 420, dà conto dei risultati della elaborazione.





<sup>&</sup>lt;sup>8</sup> Vedi S.Cammarata, Reti neuronali, Etaslibri, Milano 1997, pg. 77.

Nella colonna 2, con la cifra 1 sono indicati i dati registrati per i quali, in base alla conoscenza del locale mercato delle locazioni, si sospetta che vi siano omissioni nell'ammontare dichiarato; con la cifra 2 sono invece indicati i dati ritenuti "veritieri".

Seguono le colonne, dalla 3 alla 12, che contengono i risultati ottenuti con i 10 differenti set random di Training (R<sub>i</sub>). Per ogni colonna, i dati in grassetto sottolineato sono quelli appartenenti al set di Training; gli altri sono invece quelli scaturiti dalla fase di Testing, per cui l'1 indica che l'informazione sul canone non è"veritiera", in quanto distante dai livelli espressi dal mercato, ed il 2 invece che il dato è classificato come "veritiero".

Le colonne 3, 4 e 5 riguardano i set di Training ottenuti effettivamente con la procedura di selezione "Random" prevista dal software utilizzato. Le seguenti sei colonne, dalla 6 alla 11, si riferiscono a campioni definiti con procedura sequenziale (che rientra pur sempre nel modello random), prescelta per facilità di implementazione. La colonna 12 riguarda un insieme R<sub>i</sub> selezionato volontariamente e costituito da dati di Training desunti da contratti di locazione nei quali appariva evidente l'assenza di "trasparenza".

Quando la rete si addestra restituisce tra i dati di output un'analisi del Training set, nella quale viene espresso un parere sui dati utilizzati. Nel caso in esame, i risultati di ogni operazione comprendono un "responso di appartenenza" nel quale la rete suggerisce quali dati è preferibile eliminare dal set M per trasferirli in R e viceversa. Tali informazioni sono presenti in Tab.1. In ogni colonna, i dati di Training da trasferire da R a M sono segnati con la cifra 2 tra quelli riportati in rosso. I dati che l'output dell'elaborazione suggerisce di trasferire da M a R sono quelli indicati con t nella parte B riguardante i Dati di mercato

Nella parte B della tabella sono riportati inoltre i coefficienti riassuntivi. Alla riga A si legge il numero di casi analizzati dalla rete che hanno rispettato le ipotesi formulate nella colonna 2; la relativa percentuale è riportata alla riga D (ad esempio per il primo campione i valori sono rispettivamente 116/135 e 86%). Alla riga B è riportato il rapporto tra numero di casi "veritieri" individuati dalla rete e numero di casi sospettati "veritieri". La relativa espressione percentuale è riportata alla riga E. Alla riga C è presentato il numero totale di dati "veritieri" individuati dalla rete. Infine vengono segnalati, nelle righe "inf R" e "inf M", il totale di casi dei Training set che dovrebbero migrare da R a M e viceversa.

La colonna 13 riporta i risultati ottenuti complicando la rete con l'utilizzo di una funzione di trasferimento di tipo non lineare (sigmoide). Si evidenzia l'immediato abbattimento dei valori di "inf R" e "inf M", a dimostrazione dell'ottimo adattamento ai dati di Training, di contro si registra la più bassa percentuale di casi che rispettano le previsioni stimate nella colonna 2, appena il 74%.

Nelle colonne 14, 15 e 16 viene infine segnalata (con la cifra 1) l'appartenenza di determinati casi al sottogruppo dei dati "veritieri" rispettivamente nel 100% delle iterazioni del procedimento, nel 90% e nell'80%.

Come si può notare al termine della parte A il numero totale di casi "drenati" dall'insieme dei dati registrati, in quanto ritenuti "veritieri" dal modello, assume rispettivamente ammontare pari a 13, 20 e 30. Questi casi, aggiunti a quelli del campione M, incrementano rispettivamente del 65%, del 100% e del 150%, la numerosità dei dati affidabili sui quali eseguire indagini statistico-estimative.

Tabella 1 - Quadro dei risultati (parte A).

	Colonne														
<u> </u>	Tall	<u> </u>	4	F	à	7			r	144	40	40		145	40
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
		Γr.	Tr.	Tr.	Tr.4 Seq.	Tr.5 Seq.	Tr.6 Seq.	Tr.7 Sea.	Tr.8 Seq.	Tr.9 Seq.	Tr. 10		100	90	80
n.	11	1	2	.3	1-	21-	41-	61-	81-	101-	Selez.	Sigm.de	%	%	%
		.R	R	. R	20	40	60	80	100	120	00.02.		~	"	
					Dat	i rileva	ti pres	so L'A	genzia	delle E	ntrate		•		
1	1	1	1	1	1	1	1	1	1	1	<u>1</u>	1 .	. 0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
-3	1	2	1	2	1	2 .	_ 1	1	2	1	2	1 .	0	0	0
4	1	1	1	<u>1</u>	1	-1	1.	1	1	1	1	1	0	0	-0
5	1	2	1	· 2	· <u>2</u>	2	2	2	1	.2	2	2	0	0	1
6	1	1	1	1	1	_1	1	1	1	1	1	1	0	0	0
7	1	2	2	2	2	_1	2	2	1	2	2	2	0	0	1
8	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
9	1	1	1	1	. 1	1	1	1	1	1.	1	1 .	0	0	0
10	2	1	2	2	1	2	2.	2	2	2	. 2	2	0	0	1
11	1	1	1	1	1	1	1	1	2	1	11	1	0	0	0
12	2	1	2_	. 2	1	1	. 2	2	1	2	2	1	0	0	0
13	1	1	11	1	1	1	1	1	1	1	1	1	0	0	0
14	1	1	1	1	1	1	, 1	1	1	1_	. 1	11	0	0	0
15	1	1	1	1_	1	1	1	1	1	1_	1	. 1	0	0	0
16	1	1	1	1	1	1	_ 1	1	1	_ I	1	1	0	0	0
17	1	1	1	1	1	1	1	1	1	1	1	1 .	0	0	0
18	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
19	2	2	.2	2	2	2	2	2	1	2	2	2	0	1_	1
20	1	1	1	1	1	1	1	<u>l</u>	1	1	1	1	0	0	0
21	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
22	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
23	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	. 0	0	0
25	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
26	2	2	1	2	2	2	2	1	2	1	2	2	0	0	0
27	1	1	1	1	i	1	1	1	1	1	1	1	0	0	0
28	1	2	1	2	2	2	2	2	2	2	2	2	0	1	1
29	1	1	1	1	1	1	1	1	1	1	1 2	1	0	0	0
30	2	1	2 1	2	1	1	2	2	1	2	2	1	0	.0	0
31 32	1	-		1	1	1	1	1	1	1	2 .	2	0	0	0
33	1	2	2 ·	2	2	1 2	1	1	1	2	1	1 2	0	0	0
34		2	1	2	1	1	2	2	1		2	2	. 0	1_0	1
35	1	1	2	1	1	1	2	1	1	2	1	2	. 0	0_	0
36	1	1	. 1	2	1	<u>1</u> 1	1	2.	1	1	$\frac{1}{2}$	1	0	0_	.0
37	1	1	1	1	2	- <u>1</u> -	2	1	1	1	1	2	0	0	0
38	1	1	1	1	1	_ <u>-</u>	1	1.	1	· 1	1	1	0	0	0
39	1	1	1	1	1	<u>1</u>	2	1	.1	1	1	1	0	0	-
40	1	1	1	1	1	1	1	1	1	1	1	1	0.	0	0
	++	1	1	. 1	1		1	1	1		1	l	U	L 0	-
	$\perp \perp$									L		L		L	لـــــا

	Colonne														
1	2	3	4	5	6	7 .	8	9	10	11	12	13	14	15	16
	Ť		Tr.		Tr.4	Tr.5	Tr.6	Tr.7	Tr.8	Tr.9		<del>                                     </del>		<del>  .</del> .	
n.		Tr.	1 r.	Tr. 3	Seq.	Seq.	Seq.	Seq.	Seq.	Seq.	Tr. 10	Sigm.de	100	90	80
".		R	Ŕ	· R	1-	21-	41-	61-	81-	101-	Selez.	j Sigiri.de	%	%	%
41	1	1	1	-	_20_	40	60_	80	100	120	<del></del>				
41	1	1	1	1	1	1	1	1	1	1	1	1	0_	0	0
43	1	<del> </del>	1	1	1	1	1	1	1	1	1	1	0_	0	0
44	1	2	2	2	2	2	1	1	2	1	1	1 2	0	0	0
45	2	2	1	2	2	2	1	2	2	2	$\frac{1}{2}$	2	0	0	0
46	2	$\frac{2}{2}$	2	2	2	2	1	1	2	1	2	2	0	0	0
47	$\frac{\bar{2}}{2}$	$\frac{2}{2}$	1	2	2	2	2	2	2	2	2	2	0	1	1
48	1	$\frac{\tilde{1}}{1}$	2	1	1	ī	$\frac{z}{1}$	1	1	1	1	1	0	0	0
49	2	2	1	2	2	, 2	1	2	2	2	2	2	0	0	1
50	1	1	1	2	1	1	1	1	2	1	2	2 .	0	0	0
51	1	1	1	1	1	I	1	1	1	1	$-\frac{2}{1}$	1	0	0	0
52	1	1	Ī	1	1	1	1	1	1	1	1 .	2	0	0	0
53	1	1	1	1	1	1	2	2	1	1	2	2	.0	0	0
54	1	1	1	2	1	ī	2	2	1	1	2	2	0	0	0
55	Ī	1	1	1	1	1	1	1	1	1	1	1	0	0	0
56	2	2	1	1	2	2 .	1	1	2	1	2	1	0	0	0
57	1	1	1	1	1	1	1	1	1	1	1.	1	0	0	0
_58	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
59	2	2	2	2	2	· 2	1	1	2	1	2	2	0	0	0
60	2	2	.1	1	2	2	1	1	2	1	2	1	0	0	0
_61	2	2	1	<u>2</u>	2	_2	1	1	· 2	1	2	2	0	0	0
_62	1	_1_	1.	. 1	1	1	1_	1	1	1	1	1	0	0	0
63	1	1	1	1	1	_1	11	1	1	_1	1	1	0 .	0	0
_64	1	1	2	1	1	1	1_	1	1	_1	2	1	. 0	0	0
_65	2	2	2	2	2	_2	2_	2	2	2	2	2	1	1	1
_66	2	_2_	1	2	2	2	1	1	2	_1_	2	<u>· 1</u>	0	0	0
67	1	1	1	1	1	1	1_	1	1	_1	1	11	0	0	0
68	1	1	2	1	2	2	.2	1	1	2	2	1	0	0	0
69	1	2	2		_2_	2	2	1	2	2	2	11	0	0	1
70	1	1	1	1	2	1	1	1	2	1	2	1	0	0	0
71 72	1 2	2	1	1	-1	1	1	1	1	1	1	1	0	0	0
$\frac{12}{73}$	1	1	<u>2</u>	1	2	2	2	2	2	.2	2	. 2	1	1	1
74	1	$\frac{1}{1}$	1	1	1	1	1	1	1	1	1		0	0	0
$\frac{74}{75}$	1	1	1	- <u>1</u> -	1	1 1	1 1	_	1	1	1	1	0	0	0
76	2	2	2	$\frac{1}{2}$	2	2	$\frac{1}{2}$	1 2	1 2	1	1	1	0	0	0
77	2	2	1	2	2	2	$\frac{2}{2}$	2	1	2 2	2	2	1	-1	1
78	1	1	1	1	$\frac{2}{1}$	1	1.	1	1	1	2	2 .	0	0	1
79	2	2	1	2	2	2	1	1	2	2	12	1 2	0	0	0
80	1	1	1	1	$\frac{2}{1}$	1	$\frac{1}{1}$	1	1	1	$-\frac{2}{1}$	1	0	0	0
81	1	1	1	1	$\frac{1}{1}$	1	1	1	1	1	$-\frac{1}{1}$	1	0	0	0
82	1	1	1	1	1	1	1	1	1	1	- <u>1</u>	1	0	0	0
83	1	1	1	1	1	$\frac{1}{1}$	1	1	1	1	1	1	0	0	0
84	1	2	1	$\frac{1}{1}$	1	1	1	1	1	1	1	$\frac{1}{2}$	0	0	0
	П									-			<del>- ` -</del>	-	Ů.

	Colonne														
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
-	F		_	T-	Тг.4	Tr.5	Tr.6	Tr.7	Tr.8	Tr.9					
n.		Tr. 1	Tr.	Tr	Seq.	Seq.	Seq.	Seq.	Seq.	Seq.	Tr. 10	Sigm.de	100	90	80
11.		Ŕ	R	Ř.	1- 20	21- 40	41- 60	61- 80	81- 100	101- 120	Selez.	,	- %	%	%
85	2	2	2		20	2	2	2	2	2	2	2	1	1	1
86	1	1.	1	1	1	1	1	1	1	1	1	1	0	0	0
87	2	2	2	2	2	2	2	2	1	2	. 2	2	0	1	Ĭ
88	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1.
89	1	1	1	1	1	• 1	1	1	1	1	1	1	0	0	0
90	1	1	2	2	2	i	2	2	1	2	2	2	0	0	0
91	1	- 1	1	1	1	1	1	1	1	1	1	1	0_	0	0
92	1	1	1	1	1	1	1	I	1	1	1	2	0	0	0
93	1	1	ı	1	1	1	2	1	1	1	2	2	0	0	0
94	1	1	1	1	1	1	1	1	1	1	1	-1	0	0	0
95	1	1	1	1	1	1	1	_1	1	1	1	1	0	0	0
96	1	1	1	2	1	1	1	2	1	2	2	2	0	0_	0
97	2	1	. 1	1	1	1	1	1	1_	1	1	1	0	0	0
98	1	1	1.	1	1	_11	1	1	1	1	1	1	0	0	0
99	1	1	1	1	_1	• 1	1_	1	1	1	1	1	0	0	0
100	1	1	1	1	_1	1	1_	1	1	1	1	1	0_	0	0
101	2	2	2	2	2	2	2_	1	2	1	2	2	0	0	1
102	2	2	2	1	2	2_	2.	1	2	1	2	2	0	0	0
103	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
104	1	1	1	1	1	1	1	1	1 .	1	1	1	0	0	0
105	1	1	1	1	1	1	1	1	2	1	2	1	0	0	0
106	2	1	1	1_	1	1	1	1	1	1 1	1	1	0	0	0
107	1	1	2	1	1	1	1	2	1	1	2	1	0	.0	0
108	2	-	2	1	2	2	2	1	2	2	2	2	10	0	1
110	$\frac{ 2 }{1}$	1	1	1	1	1	1	1	1	1	1	1	0	0	0
111	2		1	2	2	2	1	1	2	2	2	1	0	0	0
112	2		2	2	2	2	1	2	2	2	$\frac{2}{2}$	2	0	1	1
113	1	-	1	1	1	1	1	1	1	1	1	1	0	0	0
114	1	_	1	1	1	1	1	1	1	1	1	1	0	0	0
115	2	-	2	2	2	2	. 2	2	1.	2	2	2	.0	1	1
116	1	_	1	1	1	1	1	1	1	1	1 .	2	0	0	0
117	1	1	2	2	1	1	1	. 2	1	1	2	1	0	0	0
118	1	1	1	-1	1	1	1	1	1	1	1	1	0	0	0
119	1	2	2	2	2	2	.2	2	2	2	2	2	1	- 1	. 1
120	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
121	1	1	1	1	1	1	1.	1	1	1	1	1	0	0	0
122	1	-	1	1	2	1	2	1	. 1	1	1	11	0	0	0.
123	1	-	1	1	1	1	1.	1 '	1	1	1.	. 1	0	0	0
124	1	_	1	1	1	1	1	1	1	1	2	2	0	0	0
125	2	+-	2	2	2	1.	2	2	1	2	2	2	0	0	0
126	2	_	1	2	. 2	1	2	2	1	2	2	2	0	0	0
127	1	-	1	1	1	1	1	1	1	1	1	1	0	0	0
128	1	1	1	1	1	1	1	1	1	1	11	1	10	0	0
		1					Ι.								
			<u> </u>	l	<u> </u>	Ь_			<u> </u>		L		⊥		Ь

	Colonne														
1	2	3	4	5	6	7	8	9 .	10	11	. 12	13	14	15	16
n.		Tr. 1 R	Tr. 2 R	Tr. 3 R	Tr.4 Seq. 1- 20	Tr.5 Seq. 21- 40	Tr.6 Seq. 41- 60	Tr.7 Seq. 61- 80	Tr.8 Seq. 81- 100	Tr.9 Seq. 101- 120	Tr. 10 Selez.	Sigm.de	100 %	90 %	80 %
129	1	1	2	2	. 2	1	2	2.	1	2	2	2	0	0	0
130	1	1	- 1	. 1	1	1	1	1	1_	1	1	1	0	0	0
131	1	1	2	2	. 2	1	2	2	1	2	2	2	0_	0	0_
132	1	1	1	1	1	1	1	1	1	1.	1	1	0_	0	0
133	1	2	2	1	2	2	2	1 .	2	1	_2	2	0	0	0.
134	1	1	1	1	1	1 .	1	_1	1	1	<u>· 1</u>	1	0	0	0
135	1	2	1_	2	l	1	1	2	2	1_1_	2	2	0	0	0
136	1	1	1	1 -	1	1	1	1	1	1	1	1	0	0	0
137	1	1	1	1	2	1	1_	1	2	1	2	2	0	0	0
138	1	2	1	2	1	2	1	· 1	2	1	2	1	0	0	0
139	1	1	1	1_	1	1	1_	1.	1	1	1	1	0	0	0
140	1	1	1	1	1	1	1	1	1_	1	1	1 .	0_	0	O
141	1	1	1	1.	-1	1	. 1	1	1_	1	1	1.	0_	0	0
142	2	2	2	2	2	2	2	. 2	. 2	2	. 2	2 .	11	1	1
143	1	1	1.	1.	2_	1	1	1	1	1	1	1	0 -	0	0
144	2	2	1 -	2	2	2_	1	1	2	1	2	1	0	0	0
145	1	1	1	1	1	1_	1_	1 .	1	1	1	1	0	0	0
146	1	1	1_	. 1	1	1	1_	1	i	1_	1	1	0	0	0_
147	1	1	1	1	1	1	1_	1	1	1_	. 1	1	0	0	0
148	1	1	2	1	1	1	1_	1	2	1	2	1	0	0	0
149	2	2	2	2 ·	2	2	2	2	2_	2.	2	2	1	1	1
150	2	2	2	2	· 2	2	2	2 .	2.	2	2	_2	1	1	1
151	1	-	2	2.	2	2	2	2	2	2	2	2	1	1	1
152	2	_	2	2	2	2	2	2	2	2	2	2	1	1	1
153	2	2	1	2	2	2	1	1	2	2	2	1	10	.0.	0
154	2	2	1	_2	2	2.	2	1	2	2	2	2	0	0	1
155	2	2	2	2	2	2	2	2	2	2	2	2	1_	1	1
	L		L -			<u> </u>	L	L	<u> </u>	ļ		ļ	L	<u> </u>	<u> </u>
L	上		<u> </u>					L	<u> </u>	L		tot	13	20	30

Tabella 1 - Quadro dei risultati (parte B).

Colonne  6	
Training Training B 9 10 Sigmoide Seq. Seq. Selez. Selez. t t t t t t t t t t t t t t t t t t t	
Training Training Sigmoide Seq. Seq. Seq. Selez. Selez. t t t t t t t t t t t t t t t t t t t	
Training Training Seq. Seq. Seq. Seq. Seq. Seq. t t t t t t t t t t t t t t t t t t t	
Training Training 8 9 Seq. 81-100 101-120 t t t t t t t t t t t t t t t t t t t	
100ne	
10   9   10   9   10   10   9   10   9   10   10	-
Training 7 Seq. 61-80 i mercato	<b>-</b>
Seq. 41-60 Dati c	
7 Training 5 Seq. 21-40	t
6 8 4 4 8 Seq. 1-20 1-20	+
Training 3 Random t	
Training Training Training 1 2 3 3 Random Random Random t t t t t t t t t t t t t t t t t t t	-
Training 1 t t	
8	
1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	15 15

							_			_	_				
-		9	80%				-						•		
lonne		15	%08 %06							·			,		
		14	100%												
		13	Sigmoide						00/135	29/39	51	74%	74%	က	0
		12	Training 10	Selez.					103/135	40/42	70	%9/	95%	0	0
		11	Training 9	101-120					110/135 114/135	25/34	37	84%	73%	8	22
		10	Training 8 Sed	81-100				<b>,</b>		29/38	45	81%	%92	4	5
	Colonne	6	Training 7	61-80			٠		105/135	21/35	41	%82	%09	9	5
	ပိ	8	raining 6	35q. 41-60		+			120/135   108/135   105/135	25/35	45	%08	%12	7	က
		7	Training 5	21-40				+		33/38	42	%68	87%	4	3
		9	Training 4	2eq.					115/135	35/39	51	%98	%06	.3	1
		5	Training 3	Random		t			111/135	33/39	51	82%	85%	3	က
		4	Training 2.	Random		ţ			106/135	22/36	37	%62	61%	9	9
		8	Training 7	Random				+	116/135	32/37	46	86%	%98	5	3
		7							⋖	m	ပ	Ω	Ш	፮ଝ	ijΣ
		-			17	18	19	20							
		_	1							_			_		

#### CONCLUSIONI

La procedura di selezione descritta permette un rilevante incremento (ben oltre il 100%) della consistenza del campione di analisi costruito in base ad informazioni di mercato. In realtà, poiché l'applicazione sviluppata nel presente lavoro è finalizzata alla definizione di un campione di studio - relativo ad unità immobiliari in locazione residenziale - ai fini della revisione delle tariffe d'estimo, i risultati perseguiti risultano notevolmente più rilevanti (incremento ben oltre il 200%), dovendosi svolgere la ricerca con riferimento non ad un solo anno ma ad un'intera epoca censuaria. Dacché se con l'analisi condotta sui contratti di locazione registrati nel corso del 2000 si sono selezionati 30 dati "trasparenti", è plausibile che se ne individuino almeno altri 30 per il 2001.

Il modello si distingue inoltre per la semplicità di impiego, la portata dei risultati e la facilità di automazione. Sfruttando il serbatoio di dati fornito dagli uffici dell'Agenzia delle Entrate ed un breve e rapido lavoro di ricognizione sul territorio per l'acquisizione di informazioni sui caratteri del mercato delle locazioni, il modello consente di incrementare un dataset essenziale composto da pochi dati di mercato, portandolo ad assumere le dimensioni di una robusta banca dati. A partire da questo archivio diventa poi agevole sviluppare ulteriori analisi o applicare procedimenti statistico-estimativi per la revisione delle tariffe ed il classamento degli immobili.

Nell'esempio, partendo da un dataset di 20 dati di mercato si perviene ad un campione finale di 80 dati, che si dimostra sufficiente per svolgere analisi statistiche su una popolazione di 11.000 elementi (quante sono le unità immobiliari a destinazione residenziale che compongono la microzona n.5), in applicazione

della formula9:

$$n = \frac{n_0 * N}{n_0 + (N - 1)} \cup 83$$

dove

N = consistenza della popolazione;n = consistenza del campione;

$$n_0 = \frac{Z^2 * \sigma^2}{e^2};$$

avendo assunto

Z = 1,96 pari ad un livello di confidenza del 95%; e = £.20.000, è l'errore di campionamento ammesso;  $\sigma$  = £.93.000<sup>10</sup>, è lo scarto quadratico medio della popolazione.

### **BIBLIOGRAFIA**

Berenson - Levine, "Statistica per le scienze economiche", Zanichelli, Bologna 1993.

G.Fabbri, R.Orsini, "Reti neurali," Franco Muzzio Editore, Padova 1993.

S.Cammarata, "Reti neuronali", Etas Libri, Milano 1997.

S.Zani, "Analisi dei dati statistici", Vol. II, Giuffrè, Milano 2000.

N.Morano, B.Manganelli, "Comparative performance of structural equations system and neural networks for real estate appraisal", New logics for the new economy", VIII SIGEF CONGRESS, Edizioni Scientifiche Italiane, Roma 2001.

<sup>&</sup>lt;sup>9</sup> Vedi Berenson-Levine, *Statistica per le scienze economiche*, Zanichelli, Bologna 1993, pgg. 296-305.

 $<sup>^{10}</sup>$   $\sigma$  è calcolato sui 50 canoni/mese x vano di mercato ottenuti nell'applicazione proposta, i quali hanno media pari a £.380.000, su cui l'e assunto rappresenta appena il 5%.