A.M.A.S.E.S.
ASSOCIAZIONE PER LA MATEMATICA APPLICATA
ALLE SCIENZE ECONOMICHE E SOCIALI

ATTI
DEL VENTUNESIMO
CONVEGNO ANNUALE
A.M.A.S.E.S.

Roma, 10-13 Settembre 1997

Universita' "La Sapienza" - Roma
Universita' "Tor Vergata" - Roma
"L.U.I.S.S. Guido Carli" - Roma

Con il Contributo del Consiglio Nazionale delle Ricerche

UN SISTEMA PREVISIVO NEURALE A 2 STADI CON APPLICAZIONE A SERIE STORICHE*

PierLuigi Belcaro e Marco Corazza
Dip.to di Matematica Applicata ed Informatica — Univ.tà di Venezia

1. Introduzione

L'ipotesi classica secondo cui le variazioni dei prezzi delle attività finanziarie sono supposte indipendenti e log-normalmente distribuite, cioé

$$(1.1) \qquad r(t+\Delta t) = \ln[P(t+\Delta t)] - \ln[P(t)] \sim N(\mu\Delta t, \sigma^2\Delta t), \ \Delta t > 0$$

non è assunta in un crescente numero di lavori. In particolare, diversi studi mostrano l'inadeguatezza di una tale ipotesi, poiché le distribuzioni empiriche delle variazioni logaritmiche dei prezzi delle attività finanziarie risultano generalmente caratterizzate da molti *outlier*, non stazionarietà nel livello della varianza ed asimmetria (vedi, ad esempio: [20]) e da dipendenza di breve e lungo periodo (vedi, ad esempio: [15], [8] e [7]). A motivo di ciò, alcuni autori congetturano che tali variazioni logaritmiche dei prezzi possano essere, oltre che stocastiche, dipendenti in modo altamente non lineare (vedi, ad esempio: [19], [14] e [6]). Si noti che, in generale, sono poche le assunzioni che è possibile fare in relazione alla forma della suddetta dipendenza e, perciò, una posizione significativa nel campo delle previsioni di serie storiche finanziarie può essere occupata dalla classe dei modelli che siano congiuntamente non lineari e non parametrici. In particolare, in lavori apparsi recentemente, alcuni autori analizzano diversi mercati finanziari utilizzando modelli non lineari e non parametrici provenienti dall'ambito della *soft Artificial Intelligence* (sAI), le cosiddette Reti Neurali Artificiali (RNA) (supervisionate e *feedforward*) di tipo *Multi-Layer Perceptron* (MLP), riscontrando evidenza di predicibilità (vedi, ad esempio: [1], [2], [18], [4], [10] e [21]).

Generalmente, a causa delle esigue assunzioni a priori che è possibile stabilire in un contesto finanziario, risulta necessario considerare diverse specificazioni del modello previsionale, ciascuna delle quali sia riferita ad un particolare scenario previsivo[1]. Di conseguenza, occorre "fronteggiare" le seguenti "difficoltà:

(1.a) la gestione della (relativa) scarsità dei dati, dovuta all'impiego di molteplici scenari;

(1.b) l'ordinamento degli scenari adottati sulla base della loro capacità

---

* Sebbene le idee generali su cui si basa questo lavoro siano state sviluppate congiuntamente dai due autori, le sezioni 1, 5, 6 e 7 sono state scritte da entrambi e le sezioni 2, 3, e 4 e le elaborazioni dei dati sono state realizzate da Marco Corazza.

[1] Il senso in cui qui si intende "scenario previsivo" è definito alla sezione 2.

esplicativa dei fenomeni (finanziari) analizzati;

(1.c) l'estrazione da ciascuno degli scenari adottati dell'informazione (implicita) relativa al fenomeno analizzato, al fine di una sua successiva incorporazione in un nuovo scenario maggiormente esplicativo rispetto ai precedenti.

L'obiettivo di questo lavoro consiste nello sviluppo di un sistema previsionale neurale di tipo MLP a 2 stadi che sia in grado di affrontare con una certa efficacia le "difficoltà" sopra enumerate e, di conseguenza, che possa migliorare la "qualità" delle previsioni che si otterrebbero adottando, invece, un approccio standard ad un solo stadio. In particolare, qui analizziamo due diverse serie storiche utilizzando un tale sistema previsionale: una serie *benchmark* generata dalla mappa deterministica caotica[2] di McKey-Glass ed una serie storica finanziaria composta dai rendimenti logaritmici giornalieri del contratto *future* sul BTP.

La parte rimanente dell'articolo è strutturata nel modo seguente. Alla sezione **2.** introduciamo le definizioni e gli strumenti che utilizziamo nelle sezioni successive; alla sezione **3.** illustriamo in dettaglio l'approccio seguito nello sviluppo del sistema previsionale neurale MLP a 2 stadi; alla sezione **4.** diamo una descrizione degli insiemi di dati impiegati nelle applicazioni; alla sezione **5.** esponiamo sinteticamente lo schema valutativo adottato per la classificazione dei risultati ottenuti dall'analisi; alla sezione **6.** riportiamo i risultati prodotti dalle applicazioni realizzate ed, infine, alla sezione **7.** proponiamo alcune considerazioni conclusive.

2. Aspetti Preliminari: Definizioni e Strumenti

In questa sezione introduciamo definizioni e strumenti richiesti per lo sviluppo della nostra metodologia previsionale neurale MLP a 2 stadi.

In primo luogo, in forza del ruolo significativo assunto nell'ambito di questo lavoro, stabiliamo le due definizioni sottostanti, relative al senso in cui la terminologia "scenario previsivo" viene utilizzata nel seguito.

Definizione 2.1 *Siano* $x_i(t)$ *e* $y_j(t)$, *con* $i = 1,\dots,I$, $I \in N_0$, $j = 1,\dots,J$ *e* $J \in N_0$, $I + J$ *variabili reali all'istante* t *e siano* $x(t) = (x_i(t - N_i\Delta),\dots,x_i(t);\ i = 1,\dots,I)$ *e* $y(t) = (y_j(t+\Delta),\dots,y_j(t + M_j\Delta);\ j = 1,\dots,J)$, *con* $N_1,\dots,N_I,M_1,\dots,M_J \in N_0$ *e* $\Delta \in R_0^+$, *due vettori reali; siano* $e_l(t)$, *con* $l = 1,\dots,L$ *e* $L \in N$, L *variabili casuali all'istante* t, *le quali possono, o meno, essere indipendentemente ed identicamente distribuite e sia* $e(t) = (e_l(t);\ l = 1,\dots,L)$ *un vettore di variabili casuali; siano* p_k, *con* $k = 1,\dots,K$ *e* $K \in N_0$, K *parametri reali e sia* $p = (p_k;\ k = 1,\dots,K)$ *un vettore reale; sia* $f(\cdot,\cdot;\cdot) \subseteq R^N \times R^L \times R^K \to R^M$, *dove* $N = I + \sum_{i=1}^{I} N_i$ *e* $M = \sum_{j=1}^{J} M_j$, *una funzione tale che* $y = f(x(t), y(t); p)$.

Si definisce (t)*-scenario previsivo per il fenomeno* $y(t)$ *l'insieme*

[2] Il senso in cui qui si intende "caos deterministico" è chiarito alla sezione 4.

$FS_{y(t)} = \{x(t), e(t), p, f(\cdot;\cdot;\cdot)\}$.

Si noti che questa definizione tiene conto tanto dell'assenza di variabili casuali ($L = 0$ e, conseguentemente, $e(t) = \varnothing$) quanto della loro presenza ($L \neq 0$).

Definizione 2.2 *Siano* $y(\tau)$, *con* $\tau = t, \ldots, t + T\Delta$ *e* $T \in N_0$, $(T\Delta - t)/\Delta$ *fenomeni e sia* $y(t, t + T\Delta) = (y(\tau); \tau = t, \ldots, t + T\Delta)$ *una matrice reale.*
Si definisce $(t, t + T\Delta)$*-scenario previsivo per la matrice* $y(t, t + T\Delta)$ *di fenomeni il vettore di insiemi* $FS_{y(t,t+T\Delta)} = \{FS_{y(\tau)}; \tau = t, \ldots, t + T\Delta\}$.

Da un punto di vista qualitativo, la **Definizione 2.1** stabilisce che si possono prevedere[3] J variabili che descrivono un fenomeno, ciascuna delle quali per un prefissato numero di *step-ahead*, facendo uso di: I variabili che "spiegano" il fenomeno analizzato, ciascuna delle quali per un prefissato numero di valori ritardati; L variabili casuali che rappresentano i termini d'errore; un modello previsionale $f(\cdot;\cdot;\cdot)$ definito in precedenza e K parametri opportunamente predeterminati. In particolare, si noti che una tale definizione include tanto l'approccio univariato ($I = 1$ e $J = 1$) quanto quello multivariato ($I \neq 1$ o $J \neq 1$). Inoltre, la **Definizione 2.2** stabilisce che si possono utilizzare lo stesso modello previsionale precedentemente definito e gli stessi parametri opportunamente predeterminati allo scopo di prevedere il fenomeno considerato in istanti di tempo successivi.

Naturalmente, da un punto di vista operativo, occorre definire sia la forma funzionale adottata dal modello previsionale che, di conseguenza, la metodologia in base alla quale determinare i K parametri. In particolare, nel presente lavoro, in relazione allo sviluppo del previsore neurale a 2 stadi, ci serviamo degli MLP (supervisionati e *feedforward*) in qualità di modelli previsivi. A motivo di ciò, ci occupiamo ora della presentazione della classe di tali modelli neurali.

In generale, questi modelli presentano la medesima struttura dei grafi orientati ad archi pesati, i cui nodi siano disposti secondo un (cosiddetto) strato di input (N nodi, con $N \in N_0$), O (cosiddetti) strati nascosti, con $O \in N_0$, (rispettivamente, H_1, ..., H_O) ed un (cosiddetto) strato di output (M nodi, con $M \in N_0$). Ciascuno di questi strati risulta totalmente connesso con gli strati ad esso adiacenti, senza connessioni interne (cioè, intrastrato)[4] (si vedano, per maggiori dettagli: [12], [13] e [6]). Si noti che in un tale approccio il ruolo delle variabili $x_i(t)$, con $i = 1, \ldots, I$, e $y_j(t)$, con $j = 1, \ldots, J$, è, rispettivamente, assunto dai nodi di input e dai nodi di output. Nella parte rimanente del presente lavoro, indichiamo questi modelli neurali con la notazione $MLP(N + 1; H_1 + 1, \ldots, H_O + 1; M)$, dove "+1" indi-

---

[3] Non ha qui importanza se inefficacemente o (meglio) efficacemente.
[4] I valori assunti dai nodi di input vengono trasmessi dallo strato iniziale di input - attraverso gli strati nascosti - allo strato finale di output, senza cicli.

ca la presenza - motivata da "ragioni tecniche" - di un nodo aggiuntivo all'interno dello strato considerato. Ciascun nodo, ad eccezione dei nodi di input, è dotato di memoria locale ed è in grado di elaborare dati. Ogni nodo "non-input", infatti, risulta caratterizzato da due funzioni. La prima determina il valore d'ingresso del nodo, attraverso il calcolo di una "aggregazione" pesata dei valori confluenti nel nodo medesimo, mentre la seconda (la cosiddetta funzione di "attivazione" o di "trasferimento") determina l'output del nodo, trasformando opportunamente il valore d'ingresso (si vedano nuovamente, per maggiori dettagli: [12], [13] e[6]). La funzione di "aggregazione" e quella di "attivazione" qui adottate sono, rispettivamente, la classica sommatoria pesata e la funzione logistica.

Una volta che la forma funzionale del modello previsionale sia stata definita, è possibile scegliere una metodologia conseguente per mezzo della quale determinare i parametri (liberi) del modello stesso, vale a dire, in relazione ad un approccio neurale, i cosiddetti pesi[5]. Il valore "ottimale" di detti pesi viene determinato adottando una procedura di stima iterativa detta algoritmo di addestramento (o di apprendimento), procedura che, in questo lavoro, risulta basata sul metodo di ottimizzazione di Levenberg-Marquardt (si veda, per dettagli maggiori: [17]) anziché sul metodo classico dell'*Error Backpropagation*. Questa procedura, utilizzando i vettori di input-output $(x(p), y(p);\ p = 1,\dots,P)$, con $P \in N_0$, appartenenti ad un prefissato *data set* D, aggiorna iterativamente i pesi del modello MLP allo scopo di raggiungere il minimo (assoluto) di una funzione di costo, di solito la (ben nota) *Mean Square Error* (MSE),

$$\text{(2.1)} \qquad \text{MSE} = \frac{1}{P}\sum_{p=1}^{P}\left(\left\|y(p) - y^*(p)\right\|\right)^2 ,$$

dove $\|\cdot\|$ indica la norma euclidea, mentre $y(p)$ ed $y^*(p)$, con $p = 1,\dots,P$, sono, rispettivamente, i fenomeni osservati e quelli previsti.

Si noti che, in relazione all'algoritmo di addestramento sopra nominato, un ruolo essenziale è svolto dal cosiddetto criterio di *Stop-Learning*. In particolare, onde evitare la possibilità sfavorevole che l'algoritmo "colga" relazioni inesistenti tra il sottovettore di input $x(p)$ ed il corrispondente sottovettore di output $y(p)$, con $p = 1,\dots,P$, adottiamo qui un criterio di *Stop-Learning* basato sul cosiddetto approccio *Concurrent Descent Methodology*. Secondo tale metodologia il *data set* D viene, in primo luogo, convenientemente suddiviso in due sottoinsiemi ad intersezione vuota, denominati, rispettivamente, *training set* D_T e *validation set* D_V. Successivamente, l'algoritmo di addestramento viene fatto operare iterativamente sul sottoinsieme di *training*, fintantoché non venga raggiunto il minimo (assoluto, se possibile) sul sottoinsieme di *validation* della prefissata funzione di costo.

---

[5] Si noti che, senza perdita di generalità, nel presente lavoro i parametri di soglia, i cosiddetti *bias*, sono considerati come pesi associati ad archi (fittizi) confluenti nei nodi interessati, archi che propagano un segnale costante uguale a -1.

3. L'Approccio Neurale MLP a 2 Stadi

In questa sezione illustriamo in dettaglio il sistema previsionale neurale MLP a 2 stadi. In particolare, dapprima descriviamo questo approccio e, poi, riportiamo alcune considerazioni ad esso relative.

Durante il primo stadio prendiamo in considerazione, per lo stesso predeterminato fenomeno $y(t, t+T\Delta)$, un prefissato numero Q di $(t, t+T\Delta)$-scenari previsivi $FS_{q,y(t,t+T\Delta)}$, con $Q \in N_0$, $T \in N$ e $q = 1, \dots, Q$. Naturalmente, ognuno di tali scenari previsivi è caratterizzato da differenti variabili "esplicative" $x_{q,i}(\cdot)$, con $q = 1, \dots, Q$ e $i = 1, \dots, I$, differenti modelli previsivi neurali $MLP_q(N_q + 1; H_{q,1} + 1, \dots, H_{q,O_q} + 1; M_q)$, con $q = 1, \dots, Q$, e da differenti vettori di parametri (cioé di pesi) p_q, con $q = 1, \dots, Q$. Si noti che ognuno dei Q vettori di parametri è determinato utilizzando uno dei Q distinti *data set* $D_q = (x_q(p), y(p); p = 1, \dots, P)$, con $q = 1, \dots, Q$, ognuno dei quali viene opportunamente suddiviso in un *training set* $D_{q,T}$ ed in un *validation set* $D_{q,V}$, con $q = 1, \dots, Q$. Inizialmente, in corrispondenza di ognuna di queste Q "strutture previsive", determiniamo la corrispondente previsione del fenomeno indagato ($y_q^*(t, t+T\Delta)$, con $q = 1, \dots, Q$) e, poi, consideriamo le (differenti) previsioni di ognuno dei sottovettori di output (che sono noti) appartenenti ai corrispondenti *validation set* (cioé $y_q^*(p)$, con $q = 1, \dots, Q$ e $p = 1, \dots, P_V$). In particolare, ognuno di questi *validation set* deve essere relativo agli stessi sottovettori (cioé $y_q(p) = y(p)$, con $p = 1, \dots, P_V$).

Durante il secondo stadio consideriamo un nuovo $(t, t+T\Delta)$-scenario previsivo $FS_{II,y(t,t+T\Delta)}$ per il fenomeno $y(t, t+T\Delta)$. Ora, per quest'ultimo, le Q (differenti) previsioni precedentemente ottenute svolgono il ruolo di variabili "esplicative". In particolare, il vettore di parametri di quest'ultima "struttura previsiva" è determinato utilizzando un *data set* i cui sotto-vettori di input sono costituiti dalle Q (differenti) previsioni dei corrispondenti sottovettori di output (che sono noti) provenienti da ognuno degli scenari di primo stadio

Questa metodologia previsiva neurale MLP a 2 stadi può essere schematizzata come segue:

Remark *PRIMO STADIO.*

Passo 1 *Si considera un fenomeno* $y(t, t+T\Delta)$.

Passo 2 *Si considerano* Q *data sets* $D_q = (x_q(p), y(p); p = 1, \dots, P)$, *con* $q = 1, \dots, Q$ *e* $P < t - \Delta$, *e si suddivide opportunamente ognuno di questi nei corrispondenti training set* $D_{q,T}$ *e validation set* $D_{q,V}$.

Passo 3 Si *determinano* Q $(t, t+T\Delta)$*-scenari previsivi per mezzo di una conveniente specificazione dei corrispondenti* Q *modelli neurali*

$MLP_q(N_q + 1; H_{q,1} + 1, \dots, H_{q,O_q} + 1; M_q)$, *con* $q = 1,\dots,Q$, *e per mezzo di una conveniente determinazione dei corrispondenti* Q *vettori di pesi* p_q, *con* $q = 1,\dots,Q$.

Passo 4 *Si determinano, per ognuna di queste* Q *"strutture previsive", le corrispondenti previsioni* $y_q^*(t, t + T\Delta)$, *con* $q = 1,\dots,Q$ *del fenomeno indagato.*

Passo 5 *Si determinano, ancora per ognuna di queste* Q *"strutture previsive", le previsioni* $y_q^*(p)$, *con* $q = 1,\dots,Q$ *e* $p = 1,\dots,P$, *di ognuno dei sottovettori di output dei corrispondenti validation data sets.*

Remark *SECONDO STADIO.*

Passo 6 *Si considera un data set* $D_{II} = ((x_q(p);\ q = 1,\dots,Q), y(p);\ p = 1,\dots,P)$ *e lo si suddivide opportunamente in un training set* $D_{II,T}$ *ed in un validation set* $D_{II,V}$.

Passo 7 *Si determina un* $(t, t + T\Delta)$*-scenario previsivo di secondo stadio per mezzo di una conveniente specificazione del corrispondente modello previsivo neurale* $MLP_{II}(N_{II} + 1; H_{II,1} + 1, \dots, H_{II,O} + 1; M_{II})$ *e per mezzo di una conveniente determinazione del corrispondente vettore di pesi* p_{II}.

Passo 8 *Si determina, per mezzo di tale* $(t, t + T\Delta)$*-scenario previsivo di secondo stadio, le previsioni di II stadio* $y_{II}^*(t, t + T\Delta)$ *del fenomeno.*

In relazione a questa metodologia previsiva, si può notare quanto segue.

In generale, per garantire al modello neurale una qualche abilità effettivamente "esplicativa" (o "generalizzante"), un ruolo cruciale viene ricoperto dal corretto "dimensionamento" del modello stesso, ovvero dalla opportuna determinazione del numero di nodi che lo costituiscono. In particolare, relativamente ad un tale opportuno dimensionamento, esistono risultati, sia empirici che teorici, che forniscono dei *lower bound* per il quoziente $N_T / N \cdot H_1 \cdot \ldots \cdot N_O \cdot M$, dove N_T è la cardinalità del *training set* D_T (si veda, per maggiori dettagli: [3], i riferimenti bibliografici in esso contenuti e [13]). Per questo motivo la forma funzionale di un modello neurale effettivamente "esplicativo" è condizionato dal numero di vettori di input-output disponibili per l'addestramento. Naturalmente, sotto questo particolare punto di vista, si potrebbe verificare che analizzare il fenomeno dato per mezzo di un approccio "standard" ad 1 stadio, caratterizzato da un modello previsivo neurale $MLP(1 + \sum_{q=1}^{Q} N_q; H_1 + 1, \dots, H_O + 1; M)$, richieda per l'addestramento un numero di vettori di input-output maggiore di quello necessario per l'approccio a 2 stadi qui proposto. In particolare, è possibile provare (semplicemente) che l'approccio a 2 stadi richiede per l'addestramento un numero di vettori di "input-output" minore o uguale a quello richiesto dall'approccio "standard" ad 1 stadio se e solo se:

$$(3.1)\qquad \frac{\left(1 + \sum_{q=1}^{Q} N_q\right) H_1 + \sum_{o=1}^{O-1} (H_o + 1) H_{o+1} + (H_O + 1) M}{\max\limits_{1 \le q \le Q} \left\{ (N_q + 1) H_{q,1} + \sum_{o=1}^{O_q - 1} \left(H_{q,o} + 1\right) H_{q,o+1} + \left(H_{q,O_q} + 1\right) M \right\}} > 1 .$$

In secondo luogo, ricordando le poche assunzioni a priori che è possibile stabilire in un contesto finanziario, nel primo stadio, *ex ante*, si dovrebbero considerare tutti gli scenari previsivi "ragionevolmente" esplicativi il fenomeno analizzato ed, *ex post*, li si dovrebbe ordinare sulla base della loro effettiva abilità di spiegare, seppur implicitamente, tale fenomeno. Al fine di affrontare questa "difficoltà" proponiamo un criterio ordinante basato sul seguente indice sviluppato *ad hoc*

$$\text{(3.2)} \qquad \mathrm{CEI}_q = 100\left(1 - \frac{\mathrm{MSE}_q}{\max\limits_{1\le q\le Q}\{\mathrm{MSE}_q\}}\right) \in [0,100],\ q = 1,\dots,Q,$$

dove CEI_q, con $q = 1,\dots,Q$, è una misura della "capacità esplicativa (implicita)" del q-esimo $(t, t+T\Delta)$-scenario previsivo e MSE_q, con $q = 1,\dots,Q$, è il valore assunto dalla funzione di costo MSE associata al modello previsivo neurale del q-esimo scenario previsivo in corrispondenza dell'iterazione "ottima" di addestramento. Sulla base di questo criterio, la "capacità esplicativa (implicita)" di ognuno dei Q $(t, t+T\Delta)$-scenari previsivi considerati diminuisce od aumenta a seconda che CEI_q tenda verso 0 o verso 100. Si noti che il valore assunto da tale quantità in corrispondenza dello scenario previsivo caratterizzato dal più grande MSE è uguale a 0. Ciò, semplicemente, significa che un tale scenario previsivo mostra, tra tutti i Q utilizzati, la più bassa "capacità esplicativa (implicita)".

4. Dati e Scenari Previsivi

In questa sezione descriviamo i dati e gli scenari previsivi utilizzati nelle applicazioni. In particolare, consideriamo prima la serie *benchmark* e poi la serie storica finanziaria.

La serie *benchmark* è stata generata utilizzando la mappa non lineare deterministica di McKey-Glass:

$$\text{(4.1)} \qquad z(t+1) = \alpha z(t) + \beta\frac{z(t-\tau)}{1 + z^{\gamma}(t-\tau)}.$$

Tale mappa, per valori particolari dei suoi parametri, cioè $\alpha = -0.4$, $\beta = -5$, $\gamma = 10$ e $\tau = 5$, è caratterizzata da un comportamento cosiddetto deterministico caotico (si veda, per dettagli maggiori: [16]). In corrispondenza di questi particolari valori dei parametri, la mappa di McKey-Glass genera una serie temporale così difficilmente predicibile da essere scambiata per una serie puramente casuale. A motivo di ciò, questa mappa è un generatore di serie *benchmark* ampiamente diffuso.

Per analizzare questa prima serie temporale, consideriamo i seguenti 3 *data set* di primo stadio ($Q_B = 3$):

(4.2.1) $$D_{B;1} = \left(z(t-p), z(t+1-p);\ p = 1,\ldots,506\right)$$

(4.2.2) $$D_{B;2} = \left(z(t-3-p), z(t+1-p);\ p = 1,\ldots,506\right) \text{ e}$$

(4.2.3) $$D_{B;3} = \left(z(t-5-p), z(t+1-p);\ p = 1,\ldots,506\right)$$

($I_{B;1} = I_{B;2} = I_{B;3} = 1$, $J_B = 1$, $x_{B;1}(t) = (z(t-p);\ p = 1,\ldots,506)$, $x_{B;2}(t) = (z(t-3-p);\ p = 1,\ldots,506)$, $x_{B;3}(t) = (z(t-5-p);\ p = 1,\ldots,506)$ e $P_B = 506$), e consideriamo i seguenti tre (corrispondenti) $(t, t+15)$-scenari previsivi di primo stadio ($T_B = 15$ e $\Delta_B = 1$):

(4.3.i) $$FS_{B;i,z(t,t+15)} = FS_{B;i} = \left\{z_{B;i}(t), e_{B;i}(t), p_{B;i}, MLP_{B;i}(1+1;3+1;1)\right\}\ ,\ i = 1,2,3$$

($N_{B;1} = N_{B;2} = N_{B;3} = 1$, $H_{B;1,O_1} = H_{B;2,O_2} = H_{B;3,O_3} = 3$ e $M_B = 1$).

Si noti che soltanto il primo ed il terzo tra gli scenari previsivi sopra descritti sono caratterizzati da un sottoinsieme dell'insieme di variabili effettivamente "esplicative" (rispettivamente, l'insieme di variabili del 1° ritardo e l'insieme di variabili del 5° ritardo). Per questo motivo, ci si aspetta che soltanto, od almeno principalmente, questi ultimi scenari previsivi siano, in una qualche misura, esplicativi del fenomeno analizzato. Inoltre, si noti che, essendo conosciute a priori le proprietà analitiche del processo generatore sottostante (la mappa di McKey-Glass), la definizione della forma funzionale dei modelli previsivi neurali adottati risulta basata su di un risultato teorico di A. N. Kolmogorov (si vedano, per maggiori dettagli: [12] e [4]).

La serie storica finanziaria è composta dai rendimenti logaritmici giornalieri del contratto *future* sul BTP:

(4.4) $$r(t+1) = \ln\left[P(t+1)\right] - \ln\left[P(t)\right]$$

Per analizzare questa seconda serie storica, consideriamo i seguenti due *data set* di primo stadio ($Q = 2$):

(4.5.1) $$D_{F;1} = \left(\left(r(t+4-p),\ldots,r(t-p)\right), r(t+1-p);\ p = 1,\ldots,506\right) \text{ e}$$

(4.5.2) $$D_{F;2} = \left(\left(v(t+4-p),\ldots,v(t-p)\right), r(t+1-p);\ p = 1,\ldots,506\right)$$

dove $v(\cdot)$ è il volume[6] trattato giornalmente ($I_{F;1} = I_{F;2} = 1$, $J_F = 1$, , $M_F = 1$, $x_{F;1} = (r(t+4-p),\ldots,r(t-p))$, $x_{F;2} = (v(t+4-p),\ldots,v(t-p))$ e $P_F = 506$), e

[6] La scelta di questo secondo insieme di variabili "esplicative" è dovuta al suo utilizzo operativo in molteplici *trading rules* (si veda, ad esempio: [5]).

consideriamo i seguenti due (corrispondenti) $(t, t+15)$-scenari previsivi di primo stadio ($T_F = 15$ e $\Delta_F = 1$):

$$(4.6.\text{i}) \qquad FS_{F;i,r(t,t+15)} = FS_{F;i} = \left\{x_{F;i}(t), e(t)_{F;i}, p_{F;i}, MLP_{F;i}(5+1;3+1;1)\right\}, \; i = 1,2$$

($H_{F;1,O_1} = H_{F;2,O_2} = 3$ e $M = 1$). In particolare, in relazione alle previsioni dell'ultimo fenomeno, consideriamo il periodo di tempo che si estende dal 25 Novembre 1996 al 13 Dicembre 1996. Inoltre, si noti che la definizione della forma funzionale dei modelli previsionali neurali si basa su una struttura architetturale a strato nascosto unico, di tipo "*bottleneck*", date le buone capacità già manifestate da quest'ultima nell'affrontare molteplici problemi previsionali di natura finanziaria (si vedano, per maggiori dettagli: [18] e [4]).

5. Lo schema di Valutazione

In questa sezione descriviamo sinteticamente lo schema teorico che utilizziamo per esprimere una valutazione del grado di efficacia con cui le previsioni del fenomeno si confrontano con i corrispondenti valori osservati. Per far questo, consideriamo un insieme di indicatori provenienti tanto dalla statistica quanto dal campo "operativo" ed osserviamo il comportamento dinamico di ciascuno di essi.

Innanzitutto, consideriamo un indice statistico che fornisce una misura della "distanza" tra le previsioni ed i corrispettivi valori osservati, il *Mean Absolute Percentage Error* (MAPE), la cui formulazione analitica è la seguente:

$$(5.1) \qquad \text{MAPE}(T) = \frac{1}{T}\sum_{\tau=1}^{T}\left|\frac{y_1(t+\tau) - y_1^*(t+\tau)}{y_1(t+\tau)}\right| \in [0,+\infty), \; T = 1,\ldots,15.$$

Adottiamo una misura di errore basata sul valore assoluto (invece di una classica fondata sulla radice quadrata) a motivo della sua ampia diffusione nelle valutazioni finanziarie a carattere operativo. Inoltre, stabiliamo di esprimerla in forma percentuale per permettere la comparabilità tra valori che provengano da serie storiche diverse e, quindi, caratterizzate da scale dimensionali diverse.

E' da sottolineare che è semplice generalizzare la quantità precedente (come le successive) nel caso in cui si prevedano più variabili, ciascuna delle quali per un maggior numero di *step-ahead* (si veda, ad esempio: [4]).

In secondo luogo, consideriamo un indice statistico che fornisce una misura della correlazione lineare tra le previsioni ed i corrispondenti valori osservati, il ben noto indice R di Bravais-Pearson. In relazione a questo indice, i valori desiderabili risultano essere quelli maggiori di 0, mentre i valori ottimali quelli uguali o vicini ad 1.

In terzo luogo, consideriamo un indice statistico che fornisce una misura del grado di efficacia con cui le previsioni si presentano se paragonate a quelle che si otterrebbero corrispondentemente da un previsore di tipo *random walk*. Si tratta del ben noto indice T_r di Theil. Relativamente a questo indice, i valori desiderabili risultano essere quelli inferiori ad 1, mentre i valori ottimali sono quelli uguali o vicini allo 0.

In quarto luogo, consideriamo un indice operativo che fornisce una misura dell'"abilità" con la quale la metodologia previsionale adottata è in grado di cogliere efficacemente i punti d'inversione, l'indicatore *Correct Signum* (CS), che, analiticamente, si presenta nel modo seguente:

$$\text{(5.2)} \quad \begin{gathered} CS(T) = \frac{100}{T}\sum_{\tau=1}^{T}\beta(\tau) \in [-100,100]\% \, , \\ \beta(\tau) = \begin{cases} 1 \text{ se } & y_1(t+\tau)y_1^*(t+\tau) > 0 \\ 1 \text{ se } & y_1(t+\tau) = y_1^*(t+\tau) = 0 \\ 0 \text{ altrimenti} & \end{cases}, \quad T = 1,K,15 \, . \end{gathered}$$

Relativamente a quest'indice, i valori desiderabili risultano essere quelli maggiori di 50, mentre i valori ottimali sono quelli uguali oppure prossimi a 100. Infine, consideriamo un indice operativo che fornisce una misura della profittabilità netta (vale a dire, guadagni compensati da perdite) della metodologia previsionale utilizzata, l'*Average Relative Net Profitability* (ARNP). Quest'ultimo risulta sviluppato sulla base di una semplice strategia di *trading*, secondo la quale un'attività finanziaria va acquistata, mantenuta oppure ceduta in funzione della sua dinamica temporale di ascesa, stabilità o discesa. La formulazione analitica dell'ARNP è la seguente:

$$\text{(5.3)} \quad \begin{gathered} ARNP(T) = \frac{\sum_{\tau=1}^{T}\gamma(\tau)y_1(t+\tau)}{\sum_{\tau=1}^{T}|y_1(t+\tau)|} \in [-100,100], \; \gamma(\tau) = \begin{cases} 1 & \text{se } \beta(\tau) > 0 \\ 0 & \text{se } \beta(\tau) = 0 \\ -1 & \text{se } \beta(\tau) < 0 \end{cases}, \\ T = 1,K,15 \, . \end{gathered}$$

Relativamente a quest'ultimo indice, i valori desiderabili risultano essere quelli maggiori di 0, mentre i valori ottimali sono quelli uguali o prossimi a 100.

Va precisato che l'insieme proposto di indicatori non ha pretese di univocità né di esaustività.

6. I Risultati Sperimentali

In questa sezione riportiamo gli esiti delle sperimentazioni realizzate. In particolare, presentiamo per primi i risultati relativi al primo stadio sperimentale e, di seguito, i risultati relativi al secondo stadio.

6.1. I Risultati del 1° Stadio

Dalla lettura della **Tabella 6.1** possiamo dedurre quanto segue in relazione alla "capacità esplicativa (implicita)" degli scenari previsivi di primo stadio.

Per ciò che riguarda la serie *benchmark*, il secondo $(t, t+15)$-scenario previsivo risulta distinto dalla più bassa "capacità esplicativa (implicita)", cioè dal più elevato livello di MSE. Questo aspetto è in accordo con il fatto che soltanto il primo ed il terzo dei suoi scenari previsivi risultano caratterizzati dalla presenza delle effettive variabili "esplicative".

Tabella 6.1: CEI degli Scenari Previsivi di 1° Stadio

	$FS_{B;1}$	$FS_{B;2}$	$FS_{B;3}$		$FS_{F;1}$	$FS_{F;2}$
$CEI_{B;q}$	3.06141	0.00000	42.90821	$CEI_{F;q}$	2.75993	0.0000

Oltre a ciò, il terzo dei suddetti scenari, vale a dire quello distinto dal sottoinsieme delle variabili "esplicative" non lineari, presenta un livello di CEI più elevato rispetto al corrispondente livello del primo scenario previsivo, vale a dire quello caratterizzato dal sottoinsieme delle variabili "esplicative" lineari. Questa realtà sperimentale ben si accorda con le proposizioni teoriche della dinamica deterministico-caotica. Infatti, per opportuni valori dei loro parametri, i corrispondenti comportamenti deterministico-caotici risultano imputabili esclusivamente alle relazioni di natura non lineare.

Tabella 6.2: Comportamento dinamico dell'indice MAPE

#	$FS_{B;1}$	$FS_{B;2}$	$FS_{B;3}$	$FS_{B;II}$	#	$FS_{F;1}$	$FS_{F;2}$	$FS_{F;II}$
1	1.003	0.927	2.063	*1.142*	1	6.034	7.457	3.172
2	1.996	3.626	2.561	1.254	2	3.495	4.045	2.024
3	5.691	16.706	3.073	*4.437*	3	2.984	2.942	1.429
4	4.488	12.675	2.318	*3.355*	4	2.989	2.393	*2.635*
5	3.712	10.284	1.879	*2.701*	5	2.825	2.180	*2.604*
6	3.130	8.669	1.598	*2.258*	6	2.809	2.000	*2.779*
7	2.692	7.597	1.428	*1.943*	7	2.531	1.860	*2.490*
8	2.420	6.727	1.338	*1.721*	8	2.410	1.700	*2.289*
9	2.393	6.175	1.295	*1.540*	9	2.222	1.646	*2.064*
10	2.547	6.358	1.217	*1.474*	10	2.134	1.614	*1.982*
11	16.957	43.978	13.236	12.334	11	2.051	1.502	*2.036*
12	15.636	40.366	12.133	11.307	12	1.974	1.474	*1.954*
13	14.491	37.320	11.205	10.450	13	1.925	1.443	*1.863*
14	13.463	34.752	10.426	9.720	14	2.274	2.051	1.868
15	12.583	32.500	9.763	9.072	15	2.204	1.981	1.835

Infine, si noti che, naturalmente, la serie *benchmark* in oggetto potrebbe essere analizzata ricorrendo ad un approccio "classico" monostadio. La forma funzionale dei corrispondenti modelli previsionali neurali, essendo qui note a priori le proprietà analitiche del processo generatore sottostante, dovrebbe essere $MLP(2+1;5+1;1)$. In questo caso, il rapporto (3.1) dovrebbe essere pari a 2.1. A motivo di ciò, per assicurare al modello previsionale MLP monostadio un dimensionamento conveniente (nel senso inteso alla sezione **3.**), sarebbe richiesto un numero più elevato di vettori input-output d'addestramento che nel caso di un approccio a doppio stadio.

Per ciò che riguarda la serie storica finanziaria, il primo $(t, t+15)$-scenario previsivo presenta una "capacità esplicativa (implicita)" più elevata, vale a dire un livello di MSE più ridotto, rispetto a quanto accade per il secondo scenario. Questa evidenza sperimentale è coerente con altri risultati empirici reperibili in letteratura (si veda, ad esempio: [5]).

Tabella 6.3: Comportamento dinamico dell'indice R

#	$FS_{B;1}$	$FS_{B;2}$	$FS_{B;3}$	$FS_{B;II}$	#	$FS_{F;1}$	$FS_{F;2}$	$FS_{F;II}$
1	--	--	--	--	1	--	--	--
2	**1.000**	**1.000**	-1.000	**1.000**	2	-1.000	**1.000**	**1.000**
3	**0.998**	**0.993**	-0.847	*-0.757*	3	-0.163	**0.452**	***0.007***
4	**0.619**	**0.661**	**0.983**	**0.985**	4	-0.383	**0.635**	-0.504
5	**0.856**	**0.725**	**0.993**	**0.990**	5	-0.486	**0.618**	-0.498
6	**0.774**	**0.761**	**0.977**	**0.990**	6	-0.525	**0.637**	*-0.521*
7	**0.719**	**0.537**	**0.961**	**0.990**	7	-0.150	**0.153**	*-0.123*
8	**0.686**	**0.555**	**0.957**	**0.990**	8	-0.122	**0.152**	*-0.108*
9	**0.684**	**0.596**	**0.960**	**0.990**	9	-0.083	**0.046**	*-0.006*
10	**0.698**	**0.621**	**0.962**	**0.991**	10	-0.149	-0.073	**0.107**
11	**0.712**	**0.636**	**0.960**	**0.990**	11	-0.174	0.266	-0.326
12	**0.657**	**0.745**	**0.977**	**0.994**	12	-0.247	-0.028	*-0.086*
13	**0.802**	**0.790**	**0.987**	**0.995**	13	-0.514	-0.111	-0.051
14	**0.777**	**0.568**	**0.983**	**0.995**	14	-0.491	-0.127	-0.035
15	**0.782**	**0.568**	**0.980**	**0.995**	15	-0.499	-0.093	-0.045

Tabella 6.4: Comportamento dinamico dell'indice T_r

#	$FS_{B;1}$	$FS_{B;2}$	$FS_{B;3}$	$FS_{B;II}$	#	$FS_{F;1}$	$FS_{F;2}$	$FS_{F;II}$
1	--	--	--	--	1	--	--	--
2	1.984	4.197	2.030	**0.906**	2	**0.915**	**0.604**	***0.838***
3	2.509	6.932	1.892	1.678	3	**0.830**	**0.510**	***0.681***
4	**0.885**	**0.723**	**0.130**	***0.152***	4	**0.844**	**0.495**	**0.869**
5	1.124	1.162	**0.199**	**0.175**	5	**0.911**	**0.537**	**0.954**
6	1.058	1.194	**0.250**	**0.166**	6	**0.952**	**0.548**	1.023
7	**0.990**	1.243	**0.303**	**0.157**	7	**0.857**	**0.821**	***0.831***
8	**0.968**	1.214	**0.330**	**0.157**	8	**0.767**	**0.713**	***0.726***
9	**0.967**	1.206	**0.332**	**0.155**	9	**0.785**	**0.782**	**0.722**
10	**0.971**	1.221	**0.332**	**0.157**	10	**0.752**	**0.749**	**0.693**
11	**0.976**	1.248	**0.342**	**0.174**	11	**0.820**	**0.755**	**0.980**
12	1.032	1.027	**0.257**	**0.131**	12	1.020	**0.991**	1.113
13	1.210	1.213	**0.248**	**0.206**	13	**0.829**	**0.757**	***0.782***
14	1.136	1.382	**0.288**	**0.233**	14	**0.783**	**0.729**	***0.730***
15	1.102	1.368	**0.313**	**0.225**	15	**0.808**	**0.743**	***0.768***

Si noti che, in generale, gli scenari previsivi di primo stadio relativi alla serie *benchmark* risultano caratterizzati da un livello più elevato di CEI rispetto agli scenari previsivi di primo stadio finanziari. Quest'ultimo aspetto può essere ragionevolmente ricondotto alle maggiori "difficoltà" che i modelli neurali si trovano a fronteggiare nella previsione di serie storiche economiche reali, (probabilmente) caratterizzate da intenso rumore, piuttosto che deterministiche (seppur caotiche).

Alla **Tabella 6.2**, **6.3**, **6.4**, **6.5**, e **6.6**, riportiamo, rispettivamente, il comportamento dinamico degli indicatori MAPE, R, T_r, CS ed ARNP per entrambe le serie storiche analizzate. In particolare, tanto per gli scenari previsivi di primo stadio come per quelli di secondo stadio, utilizziamo la notazione in grassetto per indicare che il valore dell'indicatore considerato appartiene al proprio intervallo di "desiderabilità" (si noti che un intervallo siffatto non esiste per l'indice MAPE). In relazione agli scenari di secondo stadio, inoltre, utilizziamo la notazione corsiva/sottolineata per indicare che il valore considerato dell'indice è migliore o uguale al corrispondente peggiore/migliore valore degli scenari previsivi di primo stadio.

Tabella 6.5: Comportamento dinamico dell'indice CS

#	$FS_{B;1}$	$FS_{B;2}$	$FS_{B;3}$	$FS_{B;II}$	#	$FS_{F;1}$	$FS_{F;2}$	$FS_{F;II}$
1	0.000	**100.000**	0.000	*0.000*	1	0.000	0.000	0.000
2	0.000	50.000	0.000	*0.000*	2	50.000	50.000	50.000
3	0.000	33.333	33.333	33.333	3	33.333	**66.667**	**66.667**
4	25.000	50.000	50.000	50.000	4	25.000	**75.000**	***50.000***
5	40.000	**60.000**	**60.000**	**60.000**	5	20.000	**60.000**	*40.000*
6	50.000	**66.667**	**66.667**	**66.667**	6	16.667	50.000	*33.333*
7	**57.143**	**57.143**	**71.429**	**71.429**	7	28.571	42.857	42.857
8	**62.500**	**62.500**	**75.000**	**75.000**	8	25.000	50.000	50.000
9	**66.667**	**55.556**	**77.778**	**77.778**	9	33.333	44.444	**55.556**
10	**70.000**	50.000	**80.000**	**80.000**	10	30.000	40.000	50.000
11	**72.727**	45.455	**81.818**	**81.818**	11	27.273	45.455	45.455
12	**66.667**	50.000	**83.333**	**83.333**	12	25.000	41.667	41.667
13	**69.231**	**53.846**	**84.615**	**84.615**	13	23.077	38.462	46.154
14	**71.429**	50.000	**85.714**	**85.714**	14	21.429	35.714	42.857
15	**73.333**	**53.333**	**86.667**	**86.667**	15	20.000	40.000	40.000

In relazione alla serie *benchmark*, i valori degli indici alla terza colonna di ciascuna tabella sono, generalmente, peggiori dei corrispondenti valori alla seconda e quarta colonna. Una volta ancora, questo aspetto si accorda con il fatto che soltanto il primo ed il terzo dei suoi scenari previsivi risultano caratterizzati dalle variabili effettivamente "esplicative". Oltre a ciò, i valori degli indicatori posti nella seconda e quarta colonna di ciascuna tabella appartengono, nella generalità dei casi, ai corrispondenti intervalli di "desiderabilità". In particolare, i valori degli indici R, CS ed ARNP appaiono estremamente soddisfacenti.

Per ciò che riguarda la serie storica finanziaria, quest'ultima appare di nuovo più difficilmente prevedibile della serie *benchmark*. Infatti, i valori degli indici contenuti in ciascuna tabella, per quanto, in generale, appartenenti agli intervalli di desiderabilità, sembrano meno soddisfacenti di quelli ottenuti nel caso della serie *benchmark*. Si noti che ciò è, probabilmente, anche imputabile all'adozione di strutturazioni semplici per ciò che riguarda la determinazione degli scenari previsivi "ottimali" di primo stadio (non è questo l'obiettivo principale del presente lavoro).

Ancora, i valori degli indicatori collocati all'ottava colonna di ogni tabella si presentano, generalmente, un po' migliori dei corrispondenti valori posti alla settima colonna. Questo fenomeno è, in parte, in disaccordo con i risultati riportati alla

Tabella 6.1, stando alla quale il primo scenario previsivo - in luogo del secondo - dovrebbe manifestare una "capacità esplicativa (implicita)" più elevata. Nondimeno, questo medesimo aspetto risulta essere in accordo con altri risultati empirici reperibili in letteratura (si veda, ad esempio: [5]).

Tabella 6.6: Comportamento dinamico dell'indice ARNP

#	$FS_{B;1}$	$FS_{B;2}$	$FS_{B;3}$	$FS_{B;II}$	#	$FS_{F;1}$	$FS_{F;2}$	$FS_{F;II}$
1	-100.000	**100.000**	-100.000	*-100.000*	1	-100.000	-100.000	-100.000
2	-100.000	**42.971**	-100.000	*-100.000*	2	**91.156**	**91.156**	**91.156**
3	-100.000	**34.155**	-87.668	*-87.668*	3	**54.842**	**92.836**	**92.836**
4	**75.440**	**91.914**	76.954	*76.954*	4	**41.414**	**93.457**	***76.113***
5	**89.686**	**96.604**	90.322	*90.322*	5	**22.766**	**67.947**	***52.890***
6	**92.659**	**97.583**	93.112	*93.112*	6	**13.165**	**54.812**	***40.933***
7	**93.695**	**69.686**	94.084	94.084	7	**54.921**	-19.632	**69.336**
8	**94.087**	**71.568**	94.451	94.451	8	**41.046**	-8.917	**72.082**
9	**94.194**	**68.458**	94.552	94.552	9	**50.661**	-23.772	**76.636**
10	**94.235**	**67.263**	94.591	94.591	10	**32.243**	-33.091	**55.042**
11	**94.236**	**67.232**	94.592	94.592	11	**14.436**	-15.170	**34.165**
12	**54.988**	**73.853**	95.685	95.685	12	-13.006	-35.512	**1.992**
13	**65.954**	**80.223**	96.736	96.736	13	-29.042	-47.400	**20.058**
14	**70.440**	**56.477**	97.166	97.166	14	-29.721	-47.903	**18.909**
15	**72.407**	**59.373**	97.355	97.355	15	-36.615	-33.395	**7.245**

6.2. I Risultati del 2° Stadio

In generale, il secondo stadio di elaborazione neurale migliora i risultati ottenuti al termine del primo stadio per entrambe le serie analizzate. Tale miglioramento è apprezzabile dalla lettura dei livelli di CEI degli scenari previsivi di secondo stadio - infatti: $CEI_{B;II} = 50{,}08691$ e $CEI_{F;II} = 3{,}18902$ (si veda, per i livelli di primo stadio, **Tabella 6.1**) - e dai valori posti alla quinta ed ultima colonna dalla **Tabella 6.2** alla **Tabella 6.6**.

In particolare, alla **Tabella 6.7** riportiamo, per entrambe le serie analizzate, la percentuale secondo cui i valori degli indicatori al secondo stadio sono: migliori o uguali ai corrispondenti migliori valori riscontrati al primo stadio ($\uparrow$), migliori o uguali ai corrispondenti peggiori valori riscontrati al primo stadio ($\leftrightarrow$) e peggiori dei corrispondenti peggiori valori rilevati al primo stadio ($\downarrow$).

Per ciò che riguarda la serie *benchmark*, il miglioramento che il secondo stadio di previsione produce sui valori degli indicatori è ben evidente (in particolare per gli indici R, T_r e CS). Inoltre, tali valori indicatoriali risultano analoghi a quelli corrispondenti ottenuti, per la medesima serie storica, sulla base di altre tecniche di tipo sAI (si veda, ad esempio: [11]).

In relazione alla serie finanziaria, il miglioramento prodotto dal secondo stadio sui valori degli indici si presenta soddisfacente e risulta particolarmente evidente - anche dal punto di vista dell'intervallo di "desiderabilità" - per l'indice ARNP, la cui rilevanza operativa è ben manifesta. In ogni caso, soltanto in misura del 10.96% i valori indicatoriali di secondo stadio risultano essere (leggermente) peggiori dei corrispondenti valori di primo stadio.

Tabella 6.7: Percentuali di Miglioramento/Peggioramento

Indice	$FS_{B;II}$ $\uparrow$	$FS_{B;II}$ $\leftrightarrow$	$FS_{B;II}$ $\downarrow$	$FS_{F;II}$ $\uparrow$	$FS_{F;II}$ $\leftrightarrow$	$FS_{F;II}$ $\downarrow$
MAPE	40.00%	50.00%	00.00%	33.33%	66.67%	00.00%
R	85.71%	14.29%	00.00%	35.71%	42.86%	21.43%
T_r	92.86%	07.14%	00.00%	14.29%	50.00%	35.71%
CS	86.67%	13.33%	00.00%	80.00%	20.00%	00.00%
ARNP	60.00%	40.00%	00.00%	80.00%	20.00%	00.00%
Globale	**72.60%**	**27.40%**	**00.00%**	**49.31%**	**39.73%**	**10.96%**

7. Considerazioni conclusive

In quest'ultima sezione esprimiamo le seguenti considerazioni finali.

Secondo quanto i risultati presentati nella sezione **6.** lasciano emergere, il sistema di previsione neurale a doppio stadio che abbiamo qui proposto sembra in grado, attraverso un'opportuna combinazione delle previsioni associate ai diversi scenari previsivi di primo stadio, di produrre un output finale la cui qualità sia più elevata di quella dei corrispondenti output di primo stadio. Senza dubbio, prima di giungere ad un giudizio definitivo in relazione alla suddetta metodologia previsionale neurale, occorrere poter disporre dei risultati ottenuti da ulteriori analisi (tanto di natura teorica quanto di natura applicativa).

Infine, a motivo dell'evidente carattere di predicibilità individuato in relazione alla serie finanziaria, appare indispensabile lo sviluppo di un nuovo quadro teorico, all'interno del quale i modelli finanziari classici dovrebbero trovare una definizione più ampia e generale.

8. Riferimenti Bibliografici

[1] AZOFF, E. M. (1994) *Neural Network Time Series Forecasting of Financial Markets*. Wiley.

[2] BAESTAENS, D. E., VAN DEN BERGH, W. M. e WOOD, D. (1994) *Neural Network Solutions for Trading in Financial Markets*. Pitman Publishing.

[3] BAUM, E. B. e HAUSSLER, D. (1989) What Size Net Gives Valid Generalization?, in TOURETZKY, D. S. (ed.) (1989) *Advances in Neural Information Processing Systems 1*, 81-90, Morgan Kaufmann Publishers, 81-90.

[4] BELCARO, P. L., CANESTRELLI, E. e CORAZZA, M. (1996) Artificial Neural Network Forecasting Models: an Application to the Italian Stock Market, *Badania Operacyjne i Decyzje*, 3-4, 29-48.

[5] BROCK, W., LAKONISHOK, J. e LeBARON, B. (1992) Simple Technical Trading Rules and the Stochastic Properties of Stock Returns, *The J. of Finance*, XLVII, 5, 1731-1764.

[6] CAMPBELL, J. Y., LO, A. W. e MacKINLAY, A. G. (1997) *The Econometrics of Financial Data*. Princeton University Press.
[7] CORAZZA, M., MALLIARIS, A. G. e NARDELLI, C. (1996) Searching for Fractal Structure in Agricultural Futures Market, in stampa su *The J. of Futures Market*, 17 (4).
[8] DING, Z., GRANGER, C. e ENGLE, R. (1993) A Long Memory Property of Stock Returns and a New Model, *J. of Empirical Finance*, 1, 83-106.
[9] FELDMAN, K. e KINGDON, J. (1995) Neural Networks and some Applications to Finance, *Applied Mathematical Finance*, 2, 17-42.
[10] GATELY, E. (1996) *Neural Networks for Financial Forecasting*. Wiley.
[11] GIOVE, S., PELLIZZARI, P. e TEZZA, S. (1996) RBF Networks for Financial Data Analysis and Forecasting: a Fuzzy-Cluster Approach, *Badania Operacyjne i Decyzye*, 3-4, 119-130.
[12] HECHT-NIELSEN, R. (1990) *Neurocomputing*. Addison-Wesley Publishing.
[13] HERTZ, J., KROGH, A. e PALMER, R. G. (1994) *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing.
[14] HSIEH, D. A. (1991) Chaos and Nonlinear Dynamics: Application to Financial Markets, *The Journal of Finance*, 46, 5, 1839-1877.
[15] LO, A. W. (1991) Long-term Memory in Stock Market Prices, *Econometrica*, 59, 5, 1279-1313.
[16] McKEY, M. C. e GLASS, L. (1977) Oscillation and Chaos in Physiological Control Systems, *Science*, 197, 287-289.
[17] PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. e VETTERLING (1986) *Numerical Recipies. The Art of Scientific Computing*. Cambridge University Press.
[18] REFENES, A.-P. (ed.) (1995) *Neural Networks in the Capital Martkets*. Wiley.
[19] SCHEINKMAN, J. A. e LeBARON, B. (1989) Nonlinear Dynamics and Stock Returns, *The J. of Business*, 62, 3, 311-337.
[20] SIMKOWITZ, M. e BEEDLES, W. (1980) Asymmetric Stable Distributed Security Returns, *J. of the American Statistical Association*, 75, 306-312.
[21] TURBAN, E. e TRIPPI, R. R. (eds.) (1996) *Neural Networks in Finance and Investing*, Irwin.

Finito di stampare nel mese di luglio 1997
negli Stabilimenti Tipografici Carlo Colombo S.p.A.
Via Roberto Malatesta, 296
00176 Roma