

SOCIOLOGIA E RICERCA SOCIALE

Rivista fondata da Gianni Statera

anno XXIV, n. 70, 2003 nuova serie

Comitato scientifico:

A. Abruzzese (Roma I), S.S. Acquaviva (Padova), G.D. Amendola (Bari), A. Baratta (Saarbrücken), F. Barbano (Torino), G. Barbero (Roma I), G. Bechelloni (Firenze), L. Benadusi (Roma I), E. Campelli (Roma I), A. Cavazzani (Cosenza), M. Corsale (Roma I), R. De Giorgi (Lecce), A. de Lillo (Milano), P. De Nardis (Roma I), M. Fedele (Roma I), G. Genaro (Catania), F. Leonardi (Catania), M. Livolsi (Milano Iulm), P. Marconi (Roma I), C. Marletti (Torino), A. Marradi (Firenze), G. Marselli (Napoli), F. Martinelli (Roma I), U. Melotti (Roma I), M. Morcellini (Roma I), M. Negrotti (Urbino), G. Pagliano (Roma III), A. Palazzo (Pisa), G.P. Prandstraller (Bologna), M. Santuccio (Padova), G. Sertorio (Torino), G. Sivini (Cosenza), M. Toscano (Pisa).

Direttore: Enzo Campelli

Redazione:

M.S. Agnoli, L. Ciampi, A. Fasanella, A. Marinelli, C. Lombardo, S. Mauceri e S. Tusini (segreteria di redazione), S. Vergati.

Direzione e redazione: Via Savoia, 80 - 00198 Roma

Amministrazione e distribuzione: v.le Monza, 106 - 20127 Milano - tel. (02)2837141 - Casella postale 17175 - 20100 Milano

Abbonamento 2003: Italia € 46,50, Estero € 73,00; da versare sul ccp 17562208 intestato a FrancoAngeli s.r.l., Milano

Per i soci Ais - Associazione italiana di sociologia in regola con le quote 2002: € 44,00

SOCIOLOGIA E RICERCA SOCIALE USUFRUISCE DI UN CONTRIBUTO CNR

Testata iscritta nel registro della stampa ex art. 11 L. 416/81, al n. 2048 del 12.1.1987 (PCM, Div. X Editoria)

Autorizzazione del Tribunale di Milano n. 762 del 3.12.1999 - Direttore responsabile: Alda D'Eusanio - Quadrimestrale - Sped. in abbonamento postale - 45% - art. 2, comma 26/b, Legge 662/96 - Filiale di Milano - Copyright 2002 © by FrancoAngeli s.r.l. - Stampa: Tipomozza, Milano

I quadrimestre 2003 - finito di stampare nel settembre 2003

Sommario

Saggi

- M.A. Fabiano, *La sociologia di Alfredo Niceforo. Le trasformazioni del positivismo italiano* pag. 5

Strumenti e ricerche

- C. Corposanto, G. Manzo, *Disuguaglianze educative e loro trasformazioni nel tempo: problemi tecnici e metodologici di un'analisi compiuta mediante reti neurali artificiali* » 43
- S. Calandi, *Il campionamento: analisi del concetto di rappresentatività* » 70
- A.N. Feldman, *L'uso dell'Analisi Statistica dei Dati Testuali in una ricerca sistematica* » 96
- G. Franchi, *«Vergini si diventa e non si nasce»: l'infibulazione, tra differenza culturale e violazione dei diritti umani* » 112

Note, rassegne e dibattiti

- P. Cavara, *Destra e sinistra: sulla riproposizione critica delle categorie politiche moderne* » 149
- F. Pezzulli, *Una nuova emigrazione?* » 160

Strumenti e ricerche

*Disuguaglianze educative e loro trasformazioni nel tempo: problemi tecnici e metodologici di un'analisi compiuta mediante reti neurali artificiali**

di Cleto Corposanto** e Gianluca Manzo***

1. Oggetto e contesto problematico dell'analisi

Le reti neurali artificiali possono essere proficuamente utilizzate per studiare l'evoluzione temporale delle disuguaglianze di opportunità educative? Qual è il grado di affidabilità e la portata esplicativa dei risultati cui esse conducono? Quali i problemi tecnici che le caratterizzano se applicate a questo specifico aspetto del più vasto oggetto della stratificazione sociale?

Il presente articolo si pone l'obiettivo di rispondere a questo tipo di interrogativi osservando il comportamento di uno specifico modello neurale artificiale - il «perceptrone multistrato o multilivello» - all'opera su alcuni dati disponibili per il nostro paese.

Un tale interesse conoscitivo s'inserisce in un dibattito tecnico-metodologico che ha attraversato la sociologia della stratificazione sociale da oltre un ventennio: lo studio dei cambiamenti temporali delle disuguaglianze educative, infatti, ha costantemente posto specifici e complessi problemi di analisi statistica.

Gli anni '60 e, soprattutto, '70 furono monopolizzati dall'utilizzo della regressione lineare nella forma articolata dei modelli path-analitici. Nel quadro della *status attainment research*, a seguito dello studio capitale di Blau e Dun-

* Il presente articolo rappresenta il risultato di una riflessione comunemente condotta dai due autori che ha preso spunto dalle ricerche da tempo portate avanti dal primo sui temi della classificazione nelle scienze sociali - si veda, in particolare, Corposanto, 2001 - e dal lavoro di tesi (*Disuguaglianza sociale, opportunità educative e reti neurali artificiali. Teoria e metodo a fondamento di nuovi modelli di analisi dei dati*) discusso dal secondo nel settembre 2001 presso la Facoltà di sociologia dell'Università di Trento. Sebbene quanto segue derivi dal nostro sforzo comune, per motivi organizzativi i paragrafi 1, 2, 5 e l'appendice statistica si devono a Gianluca Manzo mentre i paragrafi 3 e 4 a Cleto Corposanto.

** Professore associato presso il Dipartimento di scienze umane e sociali della Facoltà di sociologia dell'Università di Trento; cleto.corposanto@soc.unitn.it.

*** Dottorando in Epistemologie e metodologia delle scienze sociali presso le Università Paris IV-Sorbonne e Trento; mgianluca@wanadoo.fr.

can (1967), analizzare l'evoluzione temporale della struttura della stratificazione sociale rinviava a comparare i parametri di un modello di regressione stimato separatamente per gruppi di età. Alcuni studi oramai classici - Duncan (1967), Treiman (1970), Duncan, Featherman e Duncan (1972), Jencks e collaboratori (1972), Featherman e Hauser (1976), Halsey (1977) - forniscono un chiaro esempio di quest'impostazione tecnica.

Il bilancio che col trascorrere degli anni e delle ricerche si poté trarre sull'andamento temporale delle asimmetrie educative non fu particolarmente incoraggiante: alla fine degli anni '70, il solo elemento sul quale si sarebbe potuto ragionevolmente convenire consisteva nel riconoscere il carattere estremamente contraddittorio dei risultati empirici.

Si deve a Robert Mare il merito di avere spiegato formalmente la natura di queste contraddizioni. Lo studioso americano, in due articoli successivi pubblicati agli inizi degli anni '80 - «Social background and school continuation» (1980); «Change and stability in educational stratification» (1981) -, mostra che i modelli di regressione lineare sono uno strumento statisticamente inadeguato per studiare l'andamento temporale delle disuguaglianze educative. Due dimensioni analitiche fondamentali devono essere distinte a questo proposito: da un lato, la distribuzione dei titoli di studio e la loro struttura in termini di dispersione, dall'altro, l'allocatione dei singoli in questa struttura ed i principi o i meccanismi che la regolano (1981, pp. 73-4). Mare mostra matematicamente che i parametri dei modelli di regressione confondono sistematicamente queste due fonti di cambiamento sottostanti l'associazione tra origine sociale ed esito educativo (*ibid.*, pp. 75-76, 78). L'instabilità dei risultati deriva dunque dall'incapacità tecnica della regressione di scorporare dal legame tra le due variabili gli effetti derivanti dalla loro struttura o dispersione. Lo studioso americano dimostra formalmente che il principale merito della regressione logistica è esattamente quello di controllare gli effetti delle distribuzioni marginali e di restituire il valore dell'associazione tra origine e titolo di studio al netto di questi effetti (1980, p. 297; 1981, p. 74).

Come osservato da Blossfeld e Shavit (1993, p. 4), il contributo di Robert Mare ha costituito una specie di rivoluzione concettuale e metodologica nello studio degli andamenti temporali delle disuguaglianze di opportunità educative.

Attualmente, infatti, la maggior parte delle ricerche empiriche condotte su questo fenomeno si avvalgono di una strumentazione statistica basata sulla regressione logistica e sulla modellistica log-lineare (Benedusi, 1993; Blossfeld, 1993; Blossfeld e Shavit, 1993; Cobalti, 1990; Cobalti, 1992; Cobalti e Schizzerotto, 1993; Cobalti e Schizzerotto, 1994; Garnier e Raffalovich, 1984; Jonsson, 1993; Jonsson e Mills, 1993a e 1993b; Jonsson e Erikson, 2000; Kerckhoff e Trott, 1993; Muller e Karle, 1993; Schizzerotto, 1997; Shavit e Westerbeek, 1997; Vallet e Thélot, 2000).

Ciononostante è ugualmente evidente che le acquisizioni tecnico-metodologiche derivanti dai lavori di Mare non sono state ancora in grado di convincere l'intera comunità scientifica tanto da omogeneizzarla attorno ad una pratica di ricerca standardizzata.

Ne sono dimostrazione, in primo luogo, importanti ricerche sugli andamen-

ti temporali della stratificazione sociale e educativa - per es., Vrooman e Dronkers (1986); Treiman e Yip (1989) - che continuano ad operare attraverso versioni più o meno articolate e complesse della regressione lineare multipla. Qualche autore ha più generalmente osservato che vi sono alcuni segni di una ripresa di vigore degli studi condotti in termini di «path models of stratification» (Grusky, 2001, p. 14449).

Particolarmente significativi, inoltre, ci sembrano alcuni elementi della sociologia francese della stratificazione. Essa, senza dubbio profondamente influenzata dalle analisi realizzate da Raymond Boudon, in particolare, nel lavoro oramai classico *L'inégalité des chances* (1973), sembra integrare con difficoltà la svolta realizzatasi nella comunità scientifica internazionale a seguito dei lavori di Mare.

Prova ne è il dibattito tecnico-metodologico - Barbut (1984 e 1985); Combesse (1984 e 1985); Florens (1984); Gremy (1984); Merlié (1985); Prévot (1985) - svoltosi sulle pagine della *Revue française de sociologie* negli anni 1984 e 1985. I numerosi articoli che, partendo dai problemi tecnici sollevati dal lavoro di Boudon citato poco sopra, cercarono di fare chiarezza sulla possibilità di studiare i cambiamenti temporali delle asimmetrie educative attraverso misure basate su semplici proporzioni, infatti, sembrano completamente ignorare che gli articoli di Robert Mare, in particolare quello del 1981¹, avevano già dato una risposta definitiva alla questione. Inoltre, è interessante notare che la più importante ricerca comparativa organizzata a livello internazionale sul fenomeno delle disuguaglianze educative e su suoi cambiamenti nel tempo (Blossfeld e Shavit, 1993), la quale si dà esplicitamente come quadro tecnico-metodologico comune quello ispirato dai lavori di Mare, non contempla la Francia. Infine, vi sono giovani ricercatori (per es., Bulle, 2002) che dedicano parte delle loro energie intellettuali a denunciare le insufficienze per gli studi di stratificazione della misura di associazione - gli *odds-ratios* - su cui si fondano matematicamente tanto la regressione logistica che la modellistica log-lineare².

La complessità tecnico-metodologica della questione è dunque evidente.

Dal nostro punto di vista, non si tratta tuttavia di rimettere in discussione le acquisizioni derivanti dai lavori di Robert Mare: è, infatti, nostra convinzione che quella parte della tecnica statistica fondata sugli *odds-ratios* sia la scelta adeguata allorché si voglia indagare l'evoluzione temporale delle distanze relative tra i gruppi sociali negli esiti scolastici³.

Il nostro obiettivo dunque non è quello di cercare di stabilire l'adeguatezza

1. Si tenga presente che in questo articolo Mare propone una critica esplicita e formale all'impostazione tecnica di Boudon (1981, pp. 74-6), mostrando come essa sia del tutto assimilabile ad un particolare tipo di modello di regressione lineare (modello lineare di probabilità). In questo modo, Mare spiega efficacemente talune contraddizioni innegabili esistenti nei risultati empirici del sociologo francese.

2. Vi sono ovviamente alcune eccezioni a quello che potremmo definire il «caso francese»: i lavori recenti di Vallet (1999 e 2000) ne rappresentano l'esempio più evidente.

3. Teniamo a precisare che questo non significa minimamente che non vi siano alcuni aspetti delle proposizioni dello studioso americano suscettibili di essere migliorate: da questo punto di vista, ci sentiamo molto vicini ad alcuni recenti lavori come quello di Breen e Jonsson, 2000.

dei modelli neurali per lo studio di questo specifico aspetto della stratificazione sociale al fine di sostituirli all'analisi log-lineare: più semplicemente, cercheremo di vedere se essi sono ugualmente utilizzabili. Ciò implicherà principalmente tentare di capire se essi sono idonei ad indagare il legame tra origine e titolo di studio al netto degli effetti che la struttura delle variabili indipendenti e dipendenti può avere sull'associazione in quanto tale.

2. Reti neurali artificiali e perceptrone multilivello⁴

Le reti neurali artificiali possono essere definite come una famiglia di modelli computazionali non lineari la cui struttura matematica è nata dall'idea di riprodurre la configurazione ed il funzionamento del sistema nervoso umano. Tre elementi essenziali sono presenti in ogni modello neurale: un certo numero di «nodi o neuronodi»; un certo numero di «connessioni o pesi»; un processo iterativo di modificazione del valore di quest'ultimi, detto «processo di apprendimento». Da un punto di vista matematico, i «nodi» non sono nient'altro che una funzione composta: i segnali di ingresso sono sintetizzati in essi sotto forma additiva e ponderata; a questa somma pesata, essi applicano una funzione matematica, detta «funzione di attivazione», che produce un valore responsabile dell'attivazione o della non attivazione del nodo. Le «connessioni» sono concretamente semplici parametri numerici che possono assumere valori positivi o negativi: la conoscenza di una rete - l'informazione che essa estrae dai dati che analizza - è incorporata nella struttura valoriale dei pesi. Il «processo di addestramento» è un algoritmo iterativo che modifica progressivamente tale struttura: l'obiettivo ultimo di questo processo è quello di dare alla rete la capacità di generalizzare a dati sconosciuti le informazioni estratte dai dati esaminati.

Secondo la configurazione assunta dai nodi, la struttura delle connessioni e, soprattutto, il tipo d'algoritmo d'apprendimento implementato, si possono avere innumerevoli topologie neurali. In particolare, con riferimento a quest'ultimo elemento, si è soliti operare una generale distinzione tra «reti ad ap-

4. Per una discussione a diversi livelli di approfondimento e di specificità delle reti neurali artificiali, si possono utilmente consultare i seguenti lavori: Bellacicco, 1997; Bellacicco e Lauro, 1997; Breda, 1999; Breda e Buscema, 1999; Buscema, 1999a, voll. 1-2 e 1999b; Carrella, 1995; Cheng e Titterington, 1994; Corposanto, 2001; Davino, Gerghi e Vistocco, 1997; Davino, Mola e Siciliano, 1997; Davino e Vistocco, 1997; Ferrari, 1996; Floreano, 1996; Guizzardi, 1997; Maarit e Hakkinen, 2000; Massimi, 1999; Matera, 1999; Paik, 2000a e 2000b; Parisi, 1989 e 1999; Patarnello, 1992; Pizzuti, 1999; Ripley, 1993; Rumelhart *et al.*, 1986; Schiavo e Lisi, 1997; Schumacher, Rossener e Vach, 1996; Stone, 1986; Vach *et al.*, 1996; Vicino, 1999; Zeng, 1999; Woelfel, 1993. Rispetto all'estensione della letteratura tecnica sulla modellistica neurale, gli studi volti ad applicarle in campo sociologico sono estremamente rari: nel più generale ambito delle scienze sociali, eccettuato un promettente filone in campo economico-finanziario, le applicazioni più frequenti sembrano essere quelle di natura politologia (Zeng, 1999 e bibliografia ivi citata: nota 3, p. 521; Woelfel, 1993). Per quanto riguarda l'argomento specifico affrontato in questo articolo, a nostra conoscenza, si tratta del primo tentativo in tal senso; più in generale, siamo riusciti a trovare solo due articoli volti ad utilizzare le reti neurali artificiali per studiare le disegualianze sociali (Meraviglia, 1996 e 1999; in senso molto più largo, Carbone e Piras, 1999).

prendimento supervisionato» e «reti ad apprendimento non supervisionato o per auto-organizzazione». Con il primo termine si indicano quei modelli in cui la risposta corretta viene fornita esternamente alla rete ed il processo di mutamento dei valori sinaptici si realizza tramite un progressivo confronto tra questa risposta corretta e la risposta fornita dalla rete con l'obiettivo ultimo di minimizzare una qualche misura di errore. I «modelli auto-organizzati», invece, si caratterizzano per la totale assenza di risposte target fornite esternamente e realizzano l'apprendimento unicamente in conformità a qualche regola di plasticità sinaptica osservando autonomamente le regolarità presenti nei dati.

Lo specifico tipo di rete che utilizzeremo nel presente articolo fa parte della prima classe di modelli, quelli ad apprendimento supervisionato: segnatamente, si tratta della topologia neurale conosciuta sotto il nome di «perceptrone multi-strato o multi-livello».

Morfologicamente, essa si caratterizza per la presenza di tre classi funzionali di nodi: uno strato di nodi di *inputs*, i quali ricevono l'informazione dall'esterno e la trasmettono alla rete; uno strato di unità di *outputs*, il quale fornisce il risultato della rete; uno (o più) strato di nodi nascosti - detti «unità hidden» - che riceve segnali dalle unità d'ingresso e li trasmette a quelle d'uscita. Questo strato intermedio è particolarmente importante: è in esso, infatti, che viene immagazzinata la conoscenza della rete.

Per quanto riguarda la struttura delle connessioni, essa è di tipo monodirezionato: ogni unità dello strato precedente è connessa con tutte le unità dello strato successivo senza che vi siano né legami tra unità del medesimo strato né connessioni tra strati non immediatamente successivi. L'informazione passa dunque dagli strati di *input* a quelli di *output* senza alcuna forma di circolarità o di ricorrenza: in questo senso, si parla spesso di «rete feed-forward».

Trattandosi dell'algoritmo d'apprendimento, il perceptrone multistrato utilizza una procedura - messa a punto verso la meta degli anni '80 (Rumelhart *et al.*, 1986, pp. 324-7, 352-61) - nota come «regola delta generalizzata» o «algoritmo di back-propagation». Essa consiste nella minimizzazione di una funzione di errore che calcola lo scarto tra la risposta data dalla rete e quella corretta. I pesi vengono aggiornati proporzionalmente alle derivate parziali di tale funzione secondo un processo «a ritroso» che muove cioè dagli strati inferiori a quelli superiori passando per gli strati intermedi. Questo meccanismo di propagazione all'indietro dell'errore costituisce uno degli elementi di maggiore forza del perceptrone multi-strato: fino alla sua apparizione, infatti, i modelli neurali non erano in grado di mutare il valore delle proprie connessioni se non per uno strato.

Sono principalmente due gli elementi matematici e computazionali che rendono questa topologia neurale estremamente attraente tanto in generale quanto più in particolare per l'analisi empirica che abbiamo intenzione di realizzare nel presente lavoro.

In primo luogo, si è dimostrato che un perceptrone multistrato con almeno uno strato di unità *hidden* possiede la proprietà di un «approssimatore universale» ovvero sia d'un operatore computazionale in grado di stimare qualunque tipo di funzione (Cheng *et al.*, 1994, p.20; Rumelhart *et al.*, 1986, p. 319; Vach *et al.*, 1996, p. 684; Paik, 2000a, pp. 430-1; Zeng, 1999, p. 501; Woelfel, 1993). In particolare, se le unità nascoste contengono una funzione di attiva-

zione non lineare, questo modello neurale si è rivelato estremamente sensibile alla scoperta di relazioni non lineari tra le variabili in input (Buscema, 1999a, p. 61; Floreano, 1996, p. 79; Paik, 2000a, pp. 430, 432; Woelfel, 1993). Alcuni autori al fine di sottolineare questa caratteristica parlano delle reti multistrato come «modelli non-lineari input-output» o, più semplicemente, «modelli o metodi non lineari» (Maarit e Hakkinen, 2000, p. 242; Paik, 2000b, p. 5; Zeng, 1999, pp. 500-1, 512).

Secondariamente, il perceptrone multilivello a *back-propagation* intrattiene una relazione di analogia particolarmente forte ed evidente con una parte importante della statistica tradizionale. Se, in generale, l'avviso secondo cui le reti neurali artificiali restano dei metodi statistici multivariati sembra diffondersi progressivamente (Maarit e Hakkinen, 2000, p. 242), le reti multistrato sono sempre più insistentemente associate e paragonate ai modelli di regressione - lineare o non, secondo il tipo di funzione di attivazione presente nei neuroni (Cheng *et al.*, 1994, p. 16; Stone, 1986, pp. 453 e ss.; Ripley, 1993, p. 59; Woelfel, 1993, p. 67). Più in particolare, poi, si sostiene che, laddove la funzione di attivazione delle unità nascoste assume una forma logistica, il perceptrone multistrato è completamente equivalente ad un modello di regressione logistica (Davino, Mola e Siciliano, 1997, p. 25; Paik, 2000a, p. 430 e 2000b, p. 5; Schumacher *et al.*, 1996; Zeng, 1999).

Si tratta di un elemento particolarmente apprezzabile per i nostri fini, nella misura in cui la maggior parte delle analisi contemporanee condotte sugli andamenti temporali delle disuguaglianze educative si basano su questa tecnica di analisi o sulla modellistica log-lineare. L'equivalenza tra reti multistrato ed analisi logistica, quindi, sembrerebbe attribuire alle prime le stesse «virtù statistiche» delle seconde quanto al problema - richiamato nel paragrafo introduttivo - degli effetti delle distribuzioni marginali sulla consistenza delle relazioni tra le variabili.

Queste potenzialità del perceptrone multistrato a *back-propagation* non devono tuttavia far dimenticare alcuni aspetti problematici che lo caratterizzano e che pongono specifiche difficoltà di natura operativa: nel quadro degli obiettivi del presente articolo, ci limitiamo ad esplicitarne due.

Primariamente, si tratta della definizione dell'architettura della rete: nella misura in cui i nodi di *input* e di *output* sono automaticamente definiti dalla struttura rispettivamente delle variabili indipendenti e dipendente, il problema si riduce alla costruzione di uno strato adeguato di nodi hidden. A questo riguardo, si incontra ciò che è stato talvolta definito il «dilemma della BP» (Buscema, 1999b, pp. 189-90): da un lato, un numero ridotto di unite intermedie impedisce alla rete di cogliere in profondità le relazioni esistenti nei dati, dall'altro, uno strato di nodi nascosti troppo articolato potrebbe appiattire la rete sui dati di addestramento riducendone la capacità di generalizzazione - problema detto di «overfitting» (Floreano, 1996, p. 92). Rispetto all'assenza in letteratura di una risposta definitiva e teoricamente fondata⁵, nella nostra ana-

5. Sebbene in generale si è portati a considerare il processo di selezione dell'architettura ottimale della rete come un processo per tentativi ed errori (Schiavo e Lisi, 1997, p. 113;

lisi adotteremo un'ottica esplorativa ed incrementale che ci porterà a costruire per ogni relazione indagata una pluralità di architetture neurali caratterizzate da una complessità crescente tanto per numero di strati intermedi che per numero di nodi all'interno di questi ultimi.

In seconda istanza, vi è una difficoltà strettamente legata a questo primo aspetto: di fronte ad una pluralità di strutture multistrato, come selezionare quella migliore? Altrimenti detto, quali criteri utilizzare per valutare la bontà della performance dei nostri modelli?

In ordine di importanza decrescente, adotteremo sistematicamente e congiuntamente i parametri decisionali seguenti. Prima di tutto si osserverà il comportamento classificatorio della rete nei termini della «correttezza della classificazione» che essa è in grado di fornire ovverosia della capacità di assegnare i casi ai gruppi cui questi appartengono realmente (Buscema, 1999a, vol. 1, pp. 82-3). Con il termine «classification rate» si indicherà una semplice misura che calcola il numero di casi correttamente classificati sul totale dei casi analizzati (Meraviglia, 1999, p. 206)⁶. Secondariamente, compareremo costantemente il classification rate ottenuto sui dati sconosciuti con quello esibito dalla rete sul set di addestramento: nel caso in cui sia presente un problema di «overfitting», quest'ultimo dovrebbe essere più elevato del secondo. A parità di comportamento classificatorio, pertanto, privilegeremo sempre la rete caratterizzata dal classification rate più basso in fase di addestramento. In terzo luogo, prenderemo in considerazione una misura di errore - Rmse, «root mean square error» (calcolata come la radice quadrata dell'errore quadratico medio) - volta a fornire un'informazione sulla distanza media intercorrente tra l'output della rete e quello campionario⁷. Infine, un ultimo criterio valutativo consisterà nel grado di complessità del modello: come orientamento generale, per ragioni di parsimonia e d'interpretabilità, a parità di comportamento classificatorio, si sceglierà sempre la rete caratterizzata dalla struttura più semplice.

Armata di questi criteri strettamente operativi, non rimane che osservare come le nostre reti multistrato a *back-propagation* con funzione di attivazione logistica nei nodi intermedi si comportano allorché calate nel contesto dei dati empirici.

Zeng, 1999, p. 522, nota 13), bisogna riconoscere che sono disponibili in letteratura un certo numero di indicazioni operative (Carrella, 1995, p. 208; Ferrari, 1996, p. 51; Floreano, 1996, pp. 93-4, 96; Ripley, 1993, p. 49; Zeng, 1999, p. 522, nota 13): noi utilizzeremo tali accorgimenti a mo' di guida generale. Si osservi, d'altro canto, che questa fase dell'analisi neurale è ragionevolmente assimilabile a ciò che nella statistica tradizionale - si pensi all'analisi log-lineare - è la fase di «selezione del modello» (Cheng *et al.*, 1994, pp. 4-5; Schumacher *et al.*, 1996, p. 670).

6. Nella misura in cui essa sarà calcolata sui casi che la rete non ha analizzato in fase di addestramento, questa misura rappresenta un'indicazione della sua capacità generalizzatrice. Per questa ragione, la si può ritenere una buona misura della performance della rete: «(...) un buon modello è quello che generalizza bene su dati nuovi», sostiene Zeng (1999, p. 508); «la previsione diventa il test principale, conclusivo nel rifiutare o meno una determinata specificazione», nota Guizzardi (1999, p. 54).

7. Consci del carattere approssimativo e problematico di questa misura (Buscema, 1999a, vol. 1, pp. 86-7), tuttavia, la utilizzeremo nel proseguo dell'analisi come criterio valutativo secondario e sussidiario.

3. Dati e variabili

Il contesto spazio-temporale dell'analisi empirica oggetto del presente articolo è rappresentato dal nostro paese durante i primi ottant'anni del ventesimo secolo.

I dati su cui si baseranno le elaborazioni seguenti sono quelli raccolti nel 1985 dall'indagine di mobilità italiana condotta da Barbagli, Capecchi, Cobalti, de Lillo e Schizzerotto, contenenti informazioni su 5016 soggetti di entrambi i generi (occupati e non) e di età compresa tra i 18 ed i 65 anni⁸. Questo campione, nazionalmente rappresentativo, è stato costruito mediante un'articolata procedura di campionamento a stadi⁹.

Noi utilizzeremo un sotto-campione di 4107 casi. Rispetto al campione originario, sono stati eliminati tutti coloro che al momento dell'intervista erano studenti o che avevano meno di 24 anni, il che permette di lavorare con soggetti che hanno tutti terminato la propria carriera educativa (la stessa operazione è realizzata da Cobalti e Schizzerotto, 1993). A causa delle specifiche caratteristiche tecniche della modellistica neurale, tale campione sarà sistematicamente articolato in tre insiemi di dati rigorosamente distinti: il *training set* (insieme di dati sul quale la rete compie il proprio addestramento) la cui numerosità ammonta sempre all'80% del campione originario; il *testing set* (insieme di dati per i quali la rete non conosce la risposta classificatoria e sui quali viene testata la capacità predittiva e generalizzativa della rete medesima), la cui numerosità ammonta sempre al 10% del campione originario; il *validation set* (insieme di dati sul quale viene monitorato il comportamento della rete mentre il processo di addestramento è ancora in corso¹⁰), la cui numerosità ammonta sempre al 10% del campione originario. Nell'assegnazione dei casi ai tre gruppi, si seguirà sistematicamente una procedura rando-

8. A parziale integrazione di questi dati e per un'esigenza specifica verranno utilizzati alcuni dati sul contesto svedese raccolti nel 1974: il campione che ci interesserà ammonta a 5828 casi. Tali dati sono stati tratti dal più ricco ed articolato *dataset* costituito nell'ambito del progetto Casmin diretto da Muller e Goldthorpe, progetto che ha dato origine ad importanti pubblicazioni come quelle di Erikson e Goldthorpe, 1992 e di Muller e Karle, 1993.

9. Siamo consci di esporci all'obiezione di utilizzare dei dati «obsoleti». Vi è ciononostante una buona ragione che giustifica questa scelta: l'obiettivo stesso della nostra analisi - valutazione della bontà di uno strumento di analisi dei dati eccentrico rispetto alla statistica tradizionale per lo studio di un fenomeno altamente complesso come l'andamento temporale delle disuguaglianze educative - ci ha consigliato di utilizzare dei dati la cui qualità fosse indubbia e sui quali fossero già state fatte numerose analisi. Il primo elemento consente di eliminare quante più possibili fonti di variazioni esterne che potrebbero influenzare il comportamento della modellistica neurale; il secondo fornisce un insieme coerente di risultati cui confrontare quelli ottenibili mediante le reti. Si tenga presente, in ogni caso, che alcune analisi condotte sui dati più recenti raccolti nel 1997 nell'ambito della prima ondata dell'«Indagine longitudinale sulle famiglie italiane» condotta dall'Università di Trento, forniscono un'immagine di sostanziale stabilità piuttosto che di cambiamento della struttura delle disuguaglianze sociali del nostro paese (Pisati e Schizzerotto, 1999a e 1999b).

10. Questo *dataset* è utilizzato dai quei software che utilizzano come procedura per limitare il problema dell'*overfitting* quella di interrompere la fase di addestramento allorché l'errore sul *validation set* incomincia ad aumentare.

mica automaticamente generata dal *software* così come la lettura dei casi operata dalla rete assumerà anch'essa una sequenza casuale¹¹.

Per quanto riguarda la misura delle variabili, la posizione sociale d'origine sarà operativizzata tanto in termini di classe sociale che di raggruppamenti distributivi. Dal primo punto di vista, si adopererà lo schema di classe ad otto posizioni - «imprenditori», «liberi professionisti», «dirigenti», «classe media impiegatizia», «piccola borghesia urbana», «piccola borghesia agricola», «classe operaia urbana», «classe operaia agricola» - lungamente discusso da Schizzerotto (1988 e 1993, pp. 63-9) e testato empiricamente in modo approfondito da Cobalti e Schizzerotto (1994). Quanto all'aspetto distributivo della stratificazione, utilizzeremo la scala di desiderabilità sociale delle occupazioni - articolata in un continuum unilineare che prende valori da 9.97 a 90.20 - costruita per il nostro paese da De Lillo e Schizzerotto (1985).

La posizione educativa dei rispondenti, invece, sarà operativizzata come il più alto titolo di studio formalmente conseguito, secondo le modalità «senza titolo o licenza elementare», «licenza di scuola media inferiore», «licenza di scuola media superiore», «laurea». Adopereremo tale schema di misurazione anche per il livello d'istruzione del padre e della madre del rispondente allorché quest'informazione sarà inserita tra le variabili indipendenti al fine di operativizzare le risorse culturali della famiglia di origine. A tal fine si utilizzerà spesso anche una misura dell'istruzione dei genitori in termini di numero di anni di scuola massimamente ottenuti, secondo una scala di valori che va da un minimo di 1 anno ad un massimo di 19.

Un'ultima considerazione prima di passare all'analisi empirica: nelle nostre intenzioni, questo schema di operativizzazione delle variabili si vorrebbe espressivo di un certo numero di convinzioni teoriche che, per limiti di spazio, ci limitiamo ad elencare.

In primo luogo, l'idea che la stratificazione sociale debba essere concettualizzata tanto nei suoi aspetti gerarchico-relazionali quanto nella sua articolazione distributiva: da qui la considerazione congiunta di uno schema di classe e di una scala di prestigio. Secondariamente, l'idea che la posizione sociale di

11. Il *software* di elaborazione dati utilizzato in quest'analisi è «Neural connection 2.0», prodotto e commercializzato dalla Spss Inc. (Spss Inc., 1997). Questo programma di implementazione delle reti neurali, parallelamente alla procedura del *validation set* per ridurre il problema dell'*overfitting*, consente anche di adottare un'ulteriore strategia nota come «save best net», in base alla quale ad ogni ciclo di addestramento viene utilizzata la matrice delle connessioni che ha fornito i risultati migliori. Tale accorgimento sarà adottato nella nostra analisi, seguendo una linea operativa diffusa in letteratura (Carboni e Piras, 1999; Meraviglia, 1999). Si consideri inoltre che: in merito al tempo di addestramento, tutte le reti saranno addestrate per 400 cicli; relativamente all'inizializzazione dei pesi, si seguirà pressoché sempre la seguente procedura: estrazione casuale dei valori da una distribuzione uniforme multivariata all'interno di un *range* pari a $-0.1/+0.1$; quanto ai valori del tasso di apprendimento e del *momentum* essi assumeranno, per tutte le reti, valori diversi a seconda degli *step* in cui si articola il processo di addestramento: il tasso di apprendimento assume una sequenza di valori pari a 0.9/0.6/0.4/0.2 mentre il *momentum* assumerà valori pari a 0.1/0.3/0.5/0.6. Questo tipo di combinazione risponde all'esigenza di costruire un processo di apprendimento progressivamente più sofisticato che eviti, al tempo stesso, i minimi locali della funzione di errore.

origine debba essere costantemente concettualizzata anche in relazione alle risorse di natura simbolica e culturale: da qui la presenza sistematica nei modelli che si costruiranno del livello educativo parentale. Da ultimo, l'idea secondo cui anche l'istruzione si compone tanto di un aspetto discreto, gerarchico e relazionale - aspetto credenzialistico del titolo di studio - che di una dimensione continua, lineare e distributiva - istruzione come investimento «capitalistico»: da qui l'inserimento di una variabile ordinale e di una variabile metrica per operativizzare la posizione educativa dei singoli.

4. L'analisi empirica mediante reti neurali artificiali

Qual è l'evidenza empirica attualmente disponibile per il nostro paese quanto alle evoluzioni temporali delle disuguaglianze di opportunità educative?

In estrema sintesi, considerando unicamente quegli studi che partendo dalle acquisizioni tecnico-metodologiche degli inizi degli anni ottanta si avvalgono dell'analisi logistica e log-lineare, si ha a che fare principalmente con due classi di risultati.

Da un lato, un certo numero di ricerche (Cobalti, 1990; Cobalti e Schizzerotto, 1993 e 1994; Schizzerotto, 1990) evidenziano una situazione di sostanziale stabilità delle asimmetrie educative tranne una riduzione localizzata alle classi agricole. Questo risultato sarebbe d'altro canto completamente coerente con quello messo in evidenza a livello internazionale dalla più importante ricerca comparativa realizzata agli inizi degli anni '90 (Blossfeld e Shavit, 1993, fatta eccezione per la Svezia e l'Olanda). Dall'altro lato, tuttavia, alcune ricerche - per ora ancora minoritarie - hanno individuato una tendenza perequativa consistente e generalizzata all'opera ai livelli inferiori del nostro sistema scolastico (Benedusi, 1993; Shavit e Westerbeek, 1997). Il che sarebbe consistente con un certo numero di lavori empirici condotti separatamente su diverse realtà nazionali - Svezia, Olanda, Inghilterra, Francia - volti a rivedere ed ad approfondire i risultati emersi dalla ricerca comparativa diretta da Shavit e Blossfeld (per es., Jonsson, 1993; Jonsson e Mills, 1993b, p. 421; Jonsson e Erikson, 2000, p. 370; Blossfeld e Shavit, 1993, pp. 16-17; Jonsson e Mills, 1993b; Vallet e Thélot, 2000, pp. 3, 19, 21, 22, 26-29, 31).

In che modo, dunque, le reti neurali artificiali si inseriscono in questo quadro di ricerca? Per rispondere a questo interrogativo, abbiamo costruito le seguenti tre coorti di età, seguendo la ripartizione adottata da alcune analisi condotte sugli stessi dati (Cobalti, 1990, p. 202; Cobalti e Schizzerotto, 1993, p. 161): nati tra il 1920 ed il 1933; tra il 1934 ed il 1947; tra il 1948-1961. Secondariamente, si è proceduto a stimare una serie di reti multistrato all'interno di ogni coorte: il comportamento classificatorio delle reti dovrebbe restare stabile tra le coorti nell'ipotesi di un'invarianza temporale degli effetti dell'origine sul destino educativo dei singoli; mutare nel senso di un peggioramento della percentuale di corrette classificazioni nel caso di un indebolimento del legame tra origine e titolo di studio.

Le tab. 1 riporta il miglior modello neurale stimato all'interno di ogni coor-

te¹²: l'ipotesi di una stabilità generalizzata delle disuguaglianze educative non sembra corroborata dalla reti neurali artificiali.

Tab. 1 - Var. ind.: classe sociale paterna e materna, prestigio occupazionale paterno, titolo studio paterno e materno, anni scuola paterni e materni; Var. dip.: titolo studio rispondente

Coorte	Struttura rete	Funz. Attiv.	% training	% testing	Rmse
20-33 (N: 1072)	27*2*4	sigmoide	75,73	77,57	0,278
34-47 (N: 1368)	27*10*8*4	sigmoide	58,80	59,52	0,363
48-61 (N: 608)	27*6*4	sigmoide	47,21	51,82	0,393
<i>Modello su intero campione</i>					
(N: 3695)	27*7*4	sigmoide	56,39	57,84	0,363

Il comportamento delle reti muta sistematicamente in connessione con il gruppo di età considerato, evidenziando un effetto d'interazione in senso equitativo tra origine e coorte. Nella coorte più anziana la rete permette di classificare correttamente il destino educativo del 77,57% - con un incremento dunque di circa venti punti percentuali rispetto all'intero campione non strutturato per età. Questo valore scende poi a 59,52% nella coorte intermedia per abbassarsi ancora fino a 51,82% tra i soggetti più giovani: una differenza di circa 26 punti percentuali che, unitamente all'assenza di qualunque problema di *overfitting*, non sembra lasciare dubbi sul fatto che secondo le reti il potere esplicativo delle variabili di origine perde di consistenza passando da una coorte all'altra.

Nella misura in cui le ricerche realizzate nel nostro paese che sostengono una stabilità delle asimmetrie educative, hanno ciononostante riscontrato una attenuazione dello svantaggio educativo tradizionalmente proprio delle classi agricole (Cobalti, 1990, Cobalti e Schizzerotto, 1993, pp. 171, 175, 178 e 1994, pp. 173-4), si potrebbe ipotizzare che il peggioramento della capacità predittiva mostrato dalle reti sia dovuto al solo miglioramento delle chances educative di queste classi: eliminandole dall'analisi, dunque, i modelli dovrebbero presentare analoghe percentuali di corrette classificazioni tra le coorti.

La tab. 2 riporta i risultati dei modelli stimati in assenza della piccola borghesia agricola e della classe operaia agricola: sebbene il comportamento classificatorio delle reti tenda ad essere più omogeneo rispetto a quello ottenuto sull'intero spettro delle classi sociali, non sembra scomparso un sistematico effetto d'interazione tra origine e coorte. La percentuale di corrette classifica-

12. Per limiti di spazio abbiamo deciso di riportare per ogni analisi unicamente il modello selezionato e non l'intera serie di architetture costruite: quest'ultima è ovviamente disponibile presso gli autori. Un'opzione puramente stilistica, invece, è alla base della scelta di presentare tutte le tabelle nell'*appendice statistica*.

zioni si riduce da 57,38% nella coorte più anziana a 48,60% in quella più giovane, passando per un valore di circa 63% nella coorte intermedia, valore che sembra ben attribuibile a qualche fattore distorsivo presente nei dati complice la ridotta numerosità campionaria sui cui si lavora (il leggero grado di *overfitting* presente nei modelli può ricondursi alla stessa fonte). Non vi è dubbio dunque che il minor potere classificatorio delle reti così come un campo di variazione più limitato del *classification rate* indicano chiaramente che le classi agricole sono state interessate da un consistente processo perequativo e che esso corrisponde ad una larga parte delle trasformazioni della struttura dei vantaggi e degli svantaggi educativi intercorrenti tra le classi sociali. Ciononostante, nella misura in cui un sistematico peggioramento del comportamento classificatorio delle reti persiste, non sembra del tutto ragionevole imputare tale tendenza perequativa interamente alla situazione di queste classi.

Tab. 2 - Var. ind.: classe sociale paterna e materna, prestigio occupazionale paterno, titolo studio paterno e materno, anni scuola paterni e materni: Var. dip.: titolo studio rispondente (Analisi senza piccola borghesia agricola e classe operaia agricola)

Coorte	Struttura rete	Funz. Attiv.	% training	% testing	Rmse
20-33 (N: 608)	27*1*4	sigmoide	65,63	57,38	0,353
34-47 (N: 840)	27*9*4	sigmoide	52,30	63,10	0,353
48-61 (N: 1069)	27*12*4	sigmoide	49,12	48,60	0,395

Al fine di incrementare il grado di comparabilità tra la nostra analisi e quelle condotte sugli stessi dati agli inizi degli anni '90, abbiamo pensato di misurare la posizione sociale d'origine nel loro stesso modo ovvero sia attraverso la sola posizione di classe familiare costruita secondo il «principio di dominanza» (Cobalti, 1990; Cobalti e Schizzerotto, 1993 e 1994).

Tab. 3 - Var. ind.: classe sociale familiare: Var. dip.: titolo studio rispondente

Coorte	Struttura rete	Funz. Attiv.	% training	% testing	Rmse
20-33 (N:1072)	8*15*4	sigmoide	72,81	75,70	0,300
34-47 (N: 1255)	8*1*4	sigmoide	57,17	57,94	0,368
48-61 (N:1368)	8*6*4	sigmoide	41,86	43,07	0,410

Le tabb. 3 e 4 mostrano i risultati dei modelli stimati rispettivamente con e senza le classi agricole: come si vede chiaramente, anche in questo caso il comportamento classificatorio delle reti indica nettamente una perdita di rilevanza dell'origine sociale nel processo d'ottenimento di un titolo di studio. Senza ripeterci nel commento delle tabelle, ci preme tuttavia porre l'accento su una differenza rispetto ai risultati evidenziati nella fase precedente dell'analisi: il valore costantemente più basso della percentuale di corrette classificazioni esibito dai sei modelli neurali scelti. Si tratta di un risultato importante

perché indica che una misurazione quanto più adeguata dell'origine sociale necessita di uno spettro di indicatori più ampio della semplice classe sociale familiare.

Tab. 4 - Var. ind.: classe sociale familiare; Var. dip.: titolo studio rispondente (Analisi senza piccola borghesia agricola e classe operaia agricola)

Coorte	Struttura rete	Funz. Attiv.	% training	% testing	Rmse
20-33 (N: 607)	8*15*4	sigmoide	58,55	50,82	0,375
34-47 (N: 829)	8*3*4	sigmoide	48,86	45,78	0,400
48-61 (N: 1060)	8*6*4	sigmoide	43,04	38,68	0,410

Rispetto al quadro dell'evidenza empirica delineato poco sopra, dunque, la modellistica neurale nella versione qui adottata sembra confermare l'ipotesi minoritaria di un declino delle disuguaglianze di opportunità educative nel nostro paese piuttosto che il risultato maggiormente consolidato di una persistenza e resistenza al cambiamento del complesso di vantaggi e svantaggi educativi costruito attorno all'origine sociale.

Trattandosi, come detto, di un risultato ancora minoritario, ottenuto per di più con una tecnica di analisi pressoché inutilizzata per lo studio di questo fenomeno, ci sembra ragionevole approfondire l'analisi in due direzioni.

In primo luogo, abbiamo pensato di testare la bontà delle reti analizzando un contesto - quello svedese - per il quale vi è completo accordo in letteratura sull'esistenza di una forte e consistente tendenza perequativa in materia di disuguaglianze educative (Blossfeld e Shavit, 1993, pp. 15-7; Jonsson, 1993; Jonsson e Erikson, 2000, p. 370; Jonsson e Mills, 1993b, p. 421; Shavit e Westerbeek, 1997, p. 91; Goldthorpe, 1996, pp. 498-9; Goldthorpe e Breen, 1997, pp. 182-3). Nel caso in cui le reti fornissero un risultato in linea con quelli ottenuti tramite l'analisi logistica e log-lineare, disporremmo di una conferma esterna, per così dire, della bontà e dell'affidabilità di questo strumento di analisi dei dati.

Secondariamente, potrebbe essere utile osservare se le reti sono in grado di confermare alcune proposizioni derivabili dagli studi, in particolare quello di Shavit e Westerbeek (1997), che hanno sostenuto un indebolimento nel nostro paese del legame tra origine e destino educativo: nel caso di un risultato positivo, si potrebbe parlare di una «conferma dall'interno» dei risultati ottenuti mediante reti neurali artificiali.

Quanto al primo aspetto, quindi, basandosi sui dati lungamente analizzati all'interno del progetto Casmin (Erikson e Goldthorpe, 1992; Muller e Karle, 1993), si è provveduto in primo luogo a ricodificare tanto lo schema di classe quanto la struttura dei titoli di studio tradizionalmente impiegati per questi dati, così da rendere entrambe le variabili analoghe a quelle fin qui utilizzate nelle nostre analisi. In un secondo tempo, abbiamo costruito tre serie di reti volte a stimare la relazione tra classe sociale paterna e titolo di studio del rispondente nelle tre differenti coorti - nati tra il 1910 ed 1924; tra il 1925 ed il 1939; tra

il 1940 ed 1954 - in cui si è ripartito il campione utilizzato, 5828 soggetti (maschi e femmine).

La tab. 5 riporta i risultati ottenuti con i migliori modelli stimati: il comportamento delle reti sembra muoversi in modo assolutamente coerente con l'evidenza empirica accumulata in sede internazionale con altre tecniche di analisi.

Tab. 5 - Var. ind.: classe sociale paterna*; Var. dip.: titolo studio rispondente (analisi dei dati svedesi, progetto Casmin)

Coorte	Struttura rete	Funz. Attiv.	% training	% testing	Rmse
10-24 (N: 1633)	6*12*4	sigmoide	74,79	79,14	0,260
25-39 (N: 1838)	6*6*4	sigmoide	61,74	58,15	0,340
40-54 (N: 2359)	6*15*4	sigmoide	44,17	43,22	0,358
<i>Modello su intero campione</i>					
(N: 5828)	6*1*4	Sigmoide	58,70	58,49	0,343

* Forma binaria. Schema di classe di Erikson e Goldthorpe (1992, pp. 38,39) ricodificato in schema di classe di Cobalti e Schizzerotto (1994, p. 44).

Tra i soggetti più anziani, le reti classificano correttamente il 79,14% dei casi analizzati in fase di testing, evidenziando un effetto d'interazione tra origine e coorte chiaramente indicato dall'incremento del classification rate - 58,49% - proprio del modello stimato sull'intero campione non strutturato per coorte. Nella seconda coorte, poi, la capacità predittiva della rete declina in modo consistente: solo il 58,15% dei casi processati in fase di test è classificato correttamente. Un decremento di oltre 21 punti percentuali che sembrerebbe confermare la riduzione del peso della classe sociale paterna muovendosi tra i primi due gruppi di età. Tendenza perequativa, d'altro canto, che continua ulteriormente prendendo in considerazione i soggetti appartenenti alla coorte più giovane: la percentuale di corrette classificazioni in fase di generalizzazione scende in questo caso a 43,22%. Si tratta di una percentuale di corrette classificazioni di circa quindici punti percentuali inferiore a quella riscontrata nella coorte precedente, il che mostra come tra i soggetti più giovani il peso dei condizionamenti esercitati dall'origine sociale sul destino educativo individuale tende a declinare rispetto ai soggetti più anziani.

Complessivamente considerati, dunque, i risultati derivati dall'applicazione della modellistica neurale ai dati svedesi, confermano l'evoluzione temporale diagnosticata in sede internazionale nel senso di un declino delle disuguaglianze di opportunità educative in Svezia.

Passando ora al secondo «approfondimento», se i nostri risultati colgono una tendenza equitativa effettivamente all'opera nel nostro paese, il comportamento delle reti dovrebbe muoversi in due ulteriori direzioni deducibili, in particolare, dallo studio di Shavit e Westerbeek (1997).

In primo luogo, nella misura in cui questi ultimi sostengono che il fattore

responsabile della perequazione delle *chances* educative verificatosi a livello di conseguimento della licenza media inferiore ed all'atto dell'iscrizione alla secondaria superiore consiste nel livello culturale della famiglia di origine (*ibid.*, pp. 101, 103, 107), l'analisi mediante reti neurali dovrebbe essere in grado di mettere in luce un indebolimento del peso di questa risorsa.

Al tal fine, si sono costruite tre serie di reti volte a stimare la relazione tra il livello di istruzione della famiglia ed il titolo di studio del rispondente all'interno delle tre differenti coorti in cui si è diviso il campione. Affinché l'ipotesi possa dirsi verificata, si dovrebbe riscontrare un peggioramento del comportamento classificatorio delle reti passando dai modelli stimati tra i soggetti più anziani a quelli relativi al gruppo di età più giovane.

La tabella 6 riporta i risultati di quest'operazione: la nostra aspettativa sembra largamente confermata.

Tab. 6 - Var. ind.: titolo studio paterno e materno, anni scuola paterni e materni*. Var. dip.: titolo studio rispondente

Coorte	Struttura rete	Funz. Attiv.	% training	% testing	Rmse
20-33 (N:1072)	10*6*4	sigmoide	75,38	78,50	0,278
34-47 (N:1255)	10*4*4	sigmoide	58,88	59,52	0,360
48-61 (N:1368)	10*12*4	sigmoide	44,90	53,28	0,400
<i>Modello su intero campione</i>					
(N: 3695)	10*13*4	sigmoide	54,73	52,55	0,373

* Forma binaria. Schema di classe di Erikson e Goldthorpe (1992, pp. 38-9) ricodificato in schema di classe di Cobalti e Schizzerotto (1994, p. 44).

Si vede chiaramente, infatti, che la capacità predittiva dei modelli declina attraverso le coorti: la percentuale di corrette classificazioni passa da 78,50% tra i soggetti più anziani a 59,52% nella coorte intermedia per scendere a 53,28% tra i rispondenti più giovani¹³. Osservando il comportamento classificatorio della rete sull'intero campione, si riconosce chiaramente un effetto d'interazione tra coorte e risorse culturali della famiglia di origine: sebbene l'ottimo comportamento delle reti mostri la rilevanza di queste risorse nel processo di costruzione sociale dei destini educativi individuali, è indubbio che il comportamento peggiori indicando un'attenuazione nel tempo di tali effetti sperequativi.

13. Si noti che la riduzione del *classification rate* è più pronunciata passando dalla prima alla seconda coorte (circa 19 punti percentuali in meno) che passando dalla seconda alla terza (oltre 6 punti percentuali in meno). Questo dato sembra consistente con un risultato di dettaglio emerso dall'analisi di Shavit e Westerbeek: in merito al conseguimento della licenza media, infatti, i due studiosi riscontrano che gli effetti declinanti del livello culturale della famiglia d'origine, sono più estesi per i soggetti nati durante gli anni '40 che per quelli nati negli anni '50 (1997, p. 101).

In seconda istanza, nella misura in cui i dati riportati da Shavit e Westerbeek mostrano che la riduzione del nesso tra origine sociale ed esito educativo verificatosi, secondo le loro analisi, a livello di conseguimento della licenza media inferiore ed in connessione con l'iscrizione alla secondaria superiore è più esteso in corrispondenza del primo punto di transizione che del secondo (1997, pp. 101, 104), anche le reti dovrebbero essere in grado di evidenziare questa differenza d'intensità.

Al fine d'indagare questo aspetto, si è proceduto a stimare una serie di reti multistrato volte a indagare gli effetti della posizione sociale di origine ai vari punti del nostro sistema scolastico¹⁴, aggiungendo la coorte all'insieme di variabili indipendenti dirette ad operativizzare l'origine sociale dei rispondenti.

Se la riduzione degli effetti dell'origine fossero effettivamente più forti in relazione al conseguimento della licenza media inferiore che all'atto dell'iscrizione alla secondaria superiore, il peso della variabile indicatore della coorte dovrebbe essere più importante alla prima transizione rispetto alla seconda. Nel linguaggio delle reti, ciò significa che, inserendo la coorte nei due modelli, il miglioramento del comportamento rispetto ai modelli senza coorte dovrebbe essere più consistente in corrispondenza del primo punto di transizione che del secondo. Ci si potrebbe aspettare, inoltre, che l'inserimento della coorte non comporti alcun incremento nell'adattamento del modello in riferimento al conseguimento del diploma di laurea, giacché a questo livello non sembra essersi verificato alcun processo equitativo.

La tab. 7 riporta il comportamento della modellistica neurale: le aspettative appena esplicitate sembrano largamente confermate dai dati.

Tab. 7 - *Var. ind.*: classe sociale paterna, prestigio occupazionale paterno, titolo studio paterno e materno, anni scuola paterni e materni, coorte¹; *Var. dip.*: conseguimento (almeno) secondaria inferiore [prima trans.]; (almeno) secondaria superiore [seconda trans.]; laurea [terza trans.]

<i>Transizione</i>	<i>Struttura rete</i>	<i>Funz. Attiv.</i>	<i>% training</i>	<i>% testing</i>	<i>Rmse</i>
Prima (N: 3695)	20*3*2	sigmoide	77,06	79,46	0,380
Seconda (N: 1965)	20*3*2	sigmoide	67,37	67,51	0,440
Terza (N: 541)	20*1*2	sigmoide	65,66	66,67	0,480
<i>Modello senza coorte</i>					
	19*9*2	sigmoide	74,76	76,49	0,410

A livello della prima transizione, le reti permettono di classificare corret-

14. In questi modelli, la variabile dipendente assume la forma di una variabile dicotomica strutturata nel modo seguente: «non possesso della licenza secondaria inferiore-possesso della licenza secondaria inferiore (almeno)»; «non possesso della licenza secondaria superiore-possesso della licenza secondaria superiore (almeno)»; «non possesso della laurea-possesso della laurea». Per questo tipo di analisi, seguiamo la categorizzazione normalmente utilizzate per studio delle transizioni scolastiche nel nostro paese (Cobalti, 1990, p. 203; Cobalti e Schizzerotto, 1993, p. 162).

tamente in fase di generalizzazione il 79,46% dei casi: rispetto al modello privo della variabile «coorte», si osserva un incremento della capacità predittiva della rete di circa tre punti percentuali (2,97). La rete, dunque, riconosce la pertinenza dell'informazione relativa al periodo di nascita al fine di distinguere in modo più adeguato il destino educativo dei singoli. Tale rilevanza si riduce in corrispondenza della seconda transizione: in questo caso la rete classifica correttamente il 67,51% dei casi, con un incremento rispetto al modello senza coorte di solo mezzo punto percentuale, incremento di circa sei volte inferiore rispetto a quello registrato con riferimento alla prima transizione.

Si tratta di un comportamento classificatorio perfettamente in linea con l'ipotesi formulata in base ai risultati ottenuti da Shavit e Westerbeek.

Alla terza transizione, d'altro canto, le reti esibiscono un comportamento che, per quanto non contraddica le nostre ipotesi, necessita di essere interpretato più attentamente. Come spiegare, infatti, che il modello migliore classifica correttamente il 66,67% in fase di generalizzazione, evidenziando un certo decremento - 1,85 punti percentuali - del potere predittivo delle reti rispetto al modello senza «coorte»? Al di là della possibile azione di variazione stocastiche presenti nei dati, ci sembra ragionevole ipotizzare la spiegazione seguente. Tra i soggetti più giovani diminuiscono le chances di conseguire la laurea per la semplice ragione che è mutata la composizione della popolazione iscritta alle secondarie superiori: è stato evidenziato, infatti, come nelle coorti più giovani sia aumentata la quota di soggetti che frequenta gli istituti tecnici e professionali, istituti dai quali tipicamente sono minori i flussi di passaggio all'università (Cobalti, 1990, pp. 201, 207-8; Cobalti e Schizzerotto, 1993, pp. 158, 172, 174).

Chiarita in questo modo l'apparente eccentricità del risultato riscontrato alla terza transizione, nel complesso anche questa seconda direzione di analisi sembrerebbe confermare che l'iniziale risultato fornito dalla modellistica neurale in merito ad un processo perequativo delle chances educative nel contesto italiano ha una certa fondatezza ed è del tutto coerente con quello riscontrato da Shavit e Westerbeek nell'analisi del 1997.

In estrema sintesi, pertanto, la modellistica neurale ha delineato il seguente scenario quanto al problema dell'evoluzione temporale delle disuguaglianze educative nel nostro paese.

In primo luogo essa ha colto l'esistenza di una tendenza verso il declino degli effetti che l'origine sociale esercita sul destino educativo dei singoli (tabb. 1 e 3): le reti neurali, dunque, hanno permesso di acquisire ulteriori informazioni a favore di un risultato minoritario che solo qualche ricerca ha messo in luce nel nostro paese (Benadusi, 1993; Shavit e Westerbeek, 1997). Secondariamente, tale processo perequativo assume particolare rilevanza ai livelli inferiori del sistema educativo (tab. 7); in esso un ruolo particolarmente rilevante è stato giocato dalla riduzione del complesso di svantaggi educativi che ha tradizionalmente caratterizzato la condizione delle classi agricole (tabb. 2 e 4). In terzo luogo, le reti neurali hanno mostrato altresì che le tendenze perequative registratesi, originano tanto da una riduzione degli effetti delle risorse economiche (tabb. 1 e 3) - per come colte dalle variabili occupazionali e di classe - quanto da quelli connessi al livello culturale della famiglia di origine

(tab. 6). Quest'ultimo risultato, d'altro canto, sembra coerente con quello riscontrato nel contesto svedese, nella misura in cui gli studiosi che se ne sono occupati riconducono l'attenuazione del nesso tra origine ed esito educativo sia alla riduzione del peso del capitale culturale disponibile in famiglia che al miglioramento delle condizioni di vita: quest'ultimo fattore assumerebbe particolare rilevanza per le classi inferiori (per es., Jonsson, 1993, pp. 125-6). Infine, considerando l'interazione tra origine e genere, la modellistica neurale mostra che il declino del peso della prima si è verificato per entrambi i generi ma in tempi diversi: la riduzione dei condizionamenti esercitati dalla posizione sociale di appartenenza si è registrata per le donne prevalentemente nella coorte più giovane (tab. 8).

Tab. 8 - *Andamento temporale del peso dell'origine per genere (Prospetto riassuntivo)*

Coorte	Maschi (training)	Maschi (test)	Femmine (training)	Femmine (test)
20-33	94,96	94,00	80,21	77,59
34-47	55,75	42,19	64,36	73,77
48-61	42,16	49,25	49,64	55,71

Senza dubbio, dunque, un utilizzo guidato dalla teoria, misto a svariati spunti di natura esplorativa, ha permesso alla modellistica neurale di inserirsi a pieno titolo in un dibattito complesso quale quello inerente alle dinamiche temporali della disuguaglianza sociale, ed in particolare, educativa.

Ciò detto, siamo veramente in diritto di affermare che questo strumento di analisi possiede la stessa affidabilità mostrata dall'analisi logistica e log-lineare nel corso di circa quindici anni di ricerca empirica?

5. Una discussione critica: carattere problematico dei risultati

Al fine di rispondere a questo interrogativo, è necessario prendere in considerazione un aspetto della modellistica neurale che è rimasto fin qui latente ed implicito.

Come evidenziato da taluni autori (per es., Ferrari, 1996, p. 107; più implicitamente, Paik, 2000a, pp. 447-8), le reti neurali supervisionate ovvero quelle che necessitano di una variabile di target per portare a termine l'addestramento, presentano alcuni problemi in merito all'affidabilità dei risultati laddove la distribuzione delle variabili di target possiede una configurazione fortemente asimmetrica.

Il problema origina dal fatto che se la rete in fase di addestramento processa casi caratterizzati per la maggior parte dal medesimo valore sulla variabile di target (che rappresenta la nostra variabile dipendente), allorché in fase di generalizzazione si trova a dover classificare casi che presentano un valore differente tenderà a rispondere secondo le modalità più frequentemente viste durante l'addestramento. Nella misura in cui, poi, anche in sede di generaliz-

zazione la maggior parte dei casi è caratterizzata dal medesimo valore, lo stesso presente in fase di *training*, è fortemente plausibile ipotizzare che la rete classifichi correttamente una consistente quota di casi per la sola corrispondenza tra casi visti nei due momenti della stima dei modelli. Per dirla diversamente, allorché tanto in fase di *training* quanto in fase di *testing* la rete vede solo casi caratterizzati dallo stesso valore sulla variabile di *target*, presenterà un'alta percentuale di corrette classificazioni per il semplice fatto che non sussistono fattori per i quali incorrere in errore. In assenza di variabilità tanto in fase di addestramento quanto in fase di *testing*, dunque, la rete apprende male ma classifica bene un alto numero di casi.

Questo tipo di problema, ci sembra evidente, è direttamente chiamato in causa dallo studio delle evoluzioni temporali delle disuguaglianze di opportunità educative.

Nelle coorti più anziane, infatti, la distribuzione dei titoli di studio assume una configurazione marcatamente asimmetrica nella misura in cui solo una quota molto ridotta della popolazione ha accesso alle istituzioni educative: la maggior parte dei soggetti, dunque, sarà caratterizzata dall'assenza di un titolo di studio o, nel migliore dei casi, dal possesso di un titolo di studio inferiore.

Da ciò discende che nell'inserimento nella rete della variabile *target*, la maggior parte dei soggetti presenteranno il valore numerico utilizzato in fase di operativizzazione per indicare la categoria di coloro che non posseggono alcun titolo di studio o sono in possesso di un titolo di studio inferiore; la medesima situazione conseguentemente si registra in fase di generalizzazione. Date queste caratteristiche della distribuzione della variabile dipendente, pertanto, scarsissime saranno le possibilità che la rete sbagli in un numero di casi superiore a quello dei soggetti che presentano le caratteristiche atipiche.

Va da sé che tale situazione muta nella misura in cui la distribuzione della variabile dipendente diventa maggiormente simmetrica.

Ma ciò, nel caso dello studio dell'andamento temporale delle asimmetrie educative, è esattamente quello che si verifica spostandosi verso le coorti più giovani a seguito del processo di espansione educativa. Una quota maggior di soggetti accede ai sistemi scolastici e si amplia la gamma di titoli di studio che è possibile conseguire soprattutto a livello secondario, tanto inferiore che superiore. In questa situazione, dunque, la rete viene addestrata su un insieme di casi caratterizzato da maggiore variabilità ed in fase di generalizzazione si trova a lavorare su soggetti che presentano una medesima variabilità. Da ciò discende che si ampliano le possibilità che la rete commetta degli errori: tali errori, tuttavia, potrebbero derivare prevalentemente dalla mutata distribuzione della variabile di *target*.

Il corollario che si trae da questo ragionamento è il seguente: allorché si valuta il declino degli effetti dell'origine sull'esito educativo dei singoli osservando la capacità classificatoria della rete, si potrebbe attribuire un declino di quest'ultima all'attenuazione del nesso tra origine ed istruzione mentre tale declino emergerebbe prevalentemente a seguito della maggiore simmetria assunta dalla distribuzione della variabile dipendente.

Una dimostrazione concreta di tale «illusione ottica» la si può ottenere, per esempio, comparando la trasformazione della distribuzione di frequenza della

variabile di target (titolo di studio) attraverso le coorti ed il rispettivo comportamento classificatorio della rete.

La tab. 9 riporta tale comparazione: è evidente che esiste una significativa corrispondenza tra il mutamento in senso simmetrico della distribuzione della variabile dipendente (*target*) ed il peggioramento della capacità predittiva della rete.

Tab. 9 - Distribuzione variabile di target e comportamento classificatorio della rete

Titolo di studio	%	Corrette class. Test (Classe soc. fam.: tab. 3)	Corrette class. Test (Risorse culturali: tab. 6)
<i>Coorte 20-33 (N: 1072)</i>			
Senza tit. o licenza elementare	71.9	75.70	78.50
Licenza media inferiore	15.1		
Diploma media superiore	9.8		
Laurea	3.2		
<i>Coorte 34-47 (N: 1255)</i>			
Senza tit. o licenza elementare	53.8	57.94	59.52
Licenza media inferiore	25.9		
Diploma media superiore	14.0		
Laurea	6.3		
<i>Coorte 48-61 (N: 1368)</i>			
Senza tit. o licenza elementare	20.8	43.07	53.28
Licenza media inferiore	37.7		
Diploma media superiore	31.4		
Laurea	10.1		

Ripetendo la stessa operazione per le evoluzioni temporali delle asimmetrie educative per gli uomini e le donne separatamente considerati, si osserva la medesima corrispondenza.

I dati riportati in tab. 10, infatti, evidenziano chiaramente che la percentuale di corrette classificazioni esibita dalle reti peggiora per gli uomini e per le donne in tempi diversi perché la distribuzione dei titoli di studio assume una configurazione maggiormente simmetrica in coorti differenti per gli uni e per le altre.

Se, dunque, il problema è stabilire l'utilità delle reti neurali artificiali per lo studio dei cambiamenti temporali cui sono potenzialmente soggette le disuguaglianze di opportunità educative, non vi è alcun dubbio che la questione tecnica appena discussa concorre fortemente ad indirizzare la valutazione nel senso di una risposta negativa. Da questo punto di vista, la modellistica neurale, almeno nella specifica versione da noi utilizzata, si troverebbe di fronte a

difficoltà analoghe, sebbene per ragioni diverse, a quelle cui deve far fronte la regressione lineare allorché applicata a questo preciso oggetto di ricerca. I parametri di quest'ultima - come mostrato da Robert Mare negli articoli discussi nel paragrafo introduttivo - confondono sistematicamente gli effetti delle distribuzioni marginali ed il valore dell'associazione netta tra origine e titolo di studio. Il comportamento classificatorio delle reti multistrato è profondamente influenzato dalla distribuzione di frequenza della variabile-*target* sulla quale si realizza il processo di addestramento: anche in questo caso, pertanto, l'indebolimento dei legami tra origine ed esito educativo originerebbe da un mero artefatto tecnico-matematico piuttosto che da una effettiva tendenza presente nei dati.

Tab. 10 - Distribuzione variabile di *target* e comportamento della rete nei sottocampioni per genere

Titolo di studio	Uomini	Donne	Corrette class. Test Uomini (Origine sociale: tab.9)	Corrette class. Test Donne (Origine sociale: tab.9)
<i>Coorte 20-33 (Nm=497, Nd=575)</i>				
Senza tit. o licenza elementare	67,2	76,0	94,00	77,59
Licenza media inferiore	17,5	13,0		
Diploma media superiore	11,3	8,5		
Laurea	4,0	2,4		
<i>Coorte 34-47 (Nm=641, Nd=614)</i>				
Senza tit. o licenza elementare	46,6	61,2	42,19	73,77
Licenza media inferiore	28,5	23,1		
Diploma media superiore	16,1	11,9		
Laurea	8,7	3,7		
<i>Coorte 34-47 (Nm=670, Nd=698)</i>				
Senza tit. o licenza elementare	18,4	23,1	49,25	55,71
Licenza media inferiore	39,7	35,8		
Diploma media superiore	31,5	31,4		
Laurea	10,4	9,7		

Da ciò deriva che la validità dei risultati empirici della nostra analisi è profondamente compromessa: le reti neurali multistrato, a causa della loro particolare struttura matematico-computazionale, non paiono particolarmente adatte per studiare questo specifico aspetto della stratificazione sociale.

Ciononostante, vi sono in particolare alcuni fattori che si devono necessariamente considerare prima di destituire definitivamente di fondamento la nostra analisi.

In primo luogo, osservando i dati riportati in tab. 9, si vede chiaramente come a parità dell'andamento della distribuzione di frequenza della variabile dipendente la rete fornisca risultati differenti in relazione alle diverse variabili in base alle quali essa viene addestrata. Ciò indica che la rete è in grado di cogliere il peso differente delle diverse variabili e riesce a stimare per ognuna di essa tipi di relazione tra loro eterogenei: se la significatività delle variabili, infatti, non intervenisse nel processo di stima sottostante al comportamento classificatorio della rete in fase di generalizzazione, si dovrebbero osservare analoghe sequenze predittive attraverso le tre coorti in relazione alle differenti variabili. Il che è piuttosto evidente anche allorché si osservano attentamente i risultati comparati in tab. 10: nella prima coorte, per esempio, pur in presenza di una distribuzione della variabile titolo di studio tra i due generi non marcatamente differente in termini di grado di asimmetria, si evidenzia una differenza nel comportamento classificatorio della rete di oltre sedici punti percentuali. D'altro canto, tra gli uomini della seconda e terza coorte, a fronte di una distribuzione che diventa visibilmente più simmetrica si registra un incremento della percentuale di corrette classificazioni piuttosto che un declino come ci si sarebbe aspettati. Altrimenti detto, sebbene la struttura della variabile-target influenza largamente il comportamento classificatorio della rete in fase di generalizzazione, è ugualmente chiaro che quest'ultimo non è il semplice prodotto automatico del processo di addestramento.

Secondariamente, le reti neurali artificiali sono state in grado di mettere in evidenza un declino delle diseguaglianze educative sui dati svedesi ovvero sia con riferimento ad un contesto empirico per il quale vi è completo accordo in letteratura sul fatto che esso sia stato effettivamente interessato da un processo perequativo delle chances educative. Inoltre, le reti hanno permesso di confermare un certo numero d'ipotesi di dettaglio legate all'affermazione più generale concernente un possibile indebolimento delle asimmetrie educative nel nostro paese.

Infine, non si deve dimenticare che la letteratura matematico-statistica, peraltro ancora scarsa, sulle specifiche relazioni tra modelli di regressione logistica e reti multistrato con funzione di attivazione sigmoidea, sostiene l'equivalenza tra queste due tecniche di analisi dei dati (Davino, Mola e Siciliano, 1997; Paik, 2000a e 2000b; Schumacher *et al.*, 1996; Zeng, 1999): essa dichiara, dunque, per quanto implicitamente, che la modellistica neurale in questa versione è in grado di tenere sotto controllo le variazioni che si registrano nella distribuzione marginale delle variabili indipendenti e dipendenti.

Alla luce di tale complesso di considerazioni, pertanto, ci sembra ragionevole proporre una valutazione «sfumata» sulla bontà dei modelli neurali, nella specifica versione qui adottata, allorché essi si applichino allo studio degli andamenti temporali delle diseguaglianze di opportunità educative.

Vi sono certamente buone ragioni per affermare che, laddove le reti vengano utilizzate per studiare fenomeni caratterizzati da un sistematico mutamento delle distribuzioni marginali, esse non siano in grado di fornire informazioni chiare su quale parte del proprio comportamento classificatorio sia riconducibile ai mutamenti in parola e quale invece possa essere attribuita alle trasformazioni che hanno interessato l'associazione tra le variabili in gioco al netto

di tali mutamenti.

D'altro canto allo stato attuale del dibattito statistico-matematico sulla modellistica neurale artificiale - soprattutto con riferimento ad un suo utilizzo in uno specifico contesto di ricerca sociologica quale lo studio della disegualianza sociale - non sembra ragionevolmente fondato rifiutare categoricamente i risultati ottenuti dalla nostra analisi. Il fatto che ogni aspetto di tali risultati possa essere inserito all'interno di precise ipotesi teoriche e di evidenze empiriche disponibili, pare un elemento di non secondaria importanza al fine di mantenerli piuttosto che di rigettarli.

In ultima analisi, dunque, l'atteggiamento maggiormente fondato sembrerebbe quello di considerare congiuntamente tanto i risultati empirici ottenuti quanto le considerazioni critiche proposte poco sopra. Entrambi gli aspetti dovrebbero assumere carattere ipotetico e vorrebbero essere volti ad aprire piuttosto che a chiudere definitivamente alcune strade di ricerca: l'unica conclusione azzardata che si potrebbe trarre, a nostro parere, consiste nel presentare come definitivamente dimostrati l'uno o l'altro dei due aspetti posti in luce.

Riferimenti bibliografici

- M. Barbut, 1984, «Note sur quelques indicateurs globaux de l'inégalité», *Revue française de sociologie*, XXV, 4, 1984, pp. 609-23.
- M. Barbut, 1985, «Commentaire sur la note de J-C. Combessie», *Revue française de sociologie*, XXVI, 1985, pp. 659-60.
- A. Bellacicco, 1997, *Su un modello di previsione basato su rete neurale*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- A. Bellacicco, A. Lauro, 1997, *Prefazione*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- L. Benadusi, 1993, «Diseguaglianze educative: un assaggio sulle variazioni nel tempo e nello spazio», *Sociologia e ricerca sociale*, 42.
- P.M. Blau, O.D. Duncan, 1967, *The American Occupational Structure*, New York, Wiley.
- H.-P. Blossfeld, 1993, *Changes in Educational Opportunity in the Federal Republic of Germany. A Longitudinal Study of Cohorts born between 1916 and 1965*, in Blossfeld e Shavit (eds.), 1993.
- H.-P. Blossfeld, Y. Shavit, 1993, *Persisting Barriers. Changes in Educational Opportunities in Thirteen Countries*, in Blossfeld e Shavit (eds.), 1993.
- H.-P. Blossfeld, Y. Shavit (eds.), *Persistent Inequality: Changes Educational Attainment in Thirteen Countries*, Boulder, Westview press.
- R. Boudon, 1973, *L'inégalité des chances. La mobilité dans les sociétés industrielles*, Paris, Colin.
- M. Breda, 1999, *Self-organization Map*, in Buscema e Semeion group, 1999a.
- M. Breda, M. Buscema, 1999, *Reti neurali ricorrenti*, in Buscema e Semeion group, 1999a.
- R. Breen, J.O. Jonsson, 2000, «Analyzing educational careers: a multinomial transition model», *American sociological review*, 65, pp. 754-72.
- N. Bulle, 2002, «L'évolution comparée des inégalités. Une approche explicative», comunicazione presentata al Convegno «L'explication de l'action sociale. Problèmes méthodo-

- logiques et prospectives de recherche», Università di Trento, 29-30 apr. 2002.
- M. Buscema, Semeion group, 1999a, *Reti neurali artificiali e sistemi sociali complessi*, Milano, Angeli.
- M. Buscema, 1999b, *Back Propagation Networks*, in Buscema e Semeion group, 1999a.
- V. Carbone, G. Piras, 1999, *Orientamento e dispersione scolastica: le reti neurali artificiali come supporto dell'azione formativa*, in Buscema e Semeion group, 1999a.
- G. Carrella, 1995, *L'officina neurale. Viaggio tra la teoria e la pratica delle reti neurali*, Milano, Angeli.
- B. Cheng, D.M. Titterington, 1994, «Neural networks: a review from a statistical perspective», *Statistical science*, 9, 1, pp. 2-54.
- A. Cobalti, 1990, «Schooling inequality in Italy: trends over time», *European sociological review*, 3, pp. 199-214.
- A. Cobalti, 1992, «Origine sociale e livello di istruzione: un modello», *Polis*, 6, 1, pp. 117-45.
- A. Cobalti, 1993, *Sistemi scolastici, mobilità, diseguaglianze*, in L. Gallino (a c. di), 1993, *Diseguaglianze ed equità in Europa*, Bari, Laterza.
- A. Cobalti, A. Schizzerotto, 1993, *Inequality of educational opportunity in Italy*, in Blossfeld e Shavit (eds.), 1993.
- A. Cobalti, A. Schizzerotto, 1994, *La mobilità sociale in Italia. L'influenza dei fattori di diseguaglianza sul destino educativo, professionale e sociale dei singoli nel nostro paese*, Bologna, il Mulino.
- J.-C. Combessie, 1984, «L'évolution comparée des inégalités: problèmes statistiques», *Revue française de sociologie*, XXV, pp. 233-54.
- J.-C. Combessie, 1985, «Paradoxe des fonctions de concentration», *Revue française de sociologie*, XXVI, pp. 653-8.
- C. Corposanto, 2001, *La classificazione in sociologia. Reti neurali, discriminant e cluster analysis*, Milano, Angeli.
- C. Davino, M. Gerghi, D. Vistocco, 1997, *Sulla stabilità delle regole di classificazione nelle reti neurali*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- C. Davino, F. Mola, R. Siciliano, 1997, *Modelli statistici e reti neurali: il caso del modello a bilanci latenti*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- C. Davino, D. Vistocco, 1997, *Reti neuronali e statistica: bibliografia, software e siti internet*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- A. De Lillo, A. Schizzerotto, 1985, *La valutazione sociale delle occupazioni. Una scala di stratificazione occupazionale per l'Italia contemporanea*, Bologna, il Mulino.
- B. Duncan, 1967, «Education and social background», *American journal of sociology*, pp. 363-72.
- O.D. Duncan, D.L. Featherman, B. Duncan, 1972, *Socioeconomic Background and Achievement*, New York, Seminar press.
- R. Erikson, J. Goldthorpe, 1992, *The Constant Flux. A Study of Class Mobility in Industrial Societies*, Oxford, Clarendon Press.
- L.D. Featherman, M.R. Hauser, 1976, «Equality of schooling: trends and prospect», *Sociology of Education*, 49, pp. 99-120.
- A. Ferrari, 1996, *Aspetti applicativi delle reti neurali artificiali*, Milano, Angeli.

- D. Floreano, 1996, *Manuale sulle reti neurali*, Bologna, il Mulino.
- J.-P. Florens, 1984, «Inégalité et dépendance statistique», *Revue française de sociologie*, XXV, pp. 255-63.
- A.M. Garnier, E.L. Raffalovich, 1984, «The evolution of equality of educational opportunities in France», in *Sociology of education*, 57, pp. 1-11.
- J.H. Goldthorpe, 1996, «Class analysis and the reorientation of class theory: the case of persisting differentials in education attainment», *British journal of sociology*, 47, 3, 1996, pp. 481-505.
- J.H. Goldthorpe, R. Breen, 1997, *Explaining Educational Differentials: Towards a Formal Rational Action Theory*, in J. Goldthorpe, 1997, *On sociology, Numbers, Narratives, and the Integration of Research and Theory*, Oxford University Press, Oxford.
- J.-P. Grémy, 1984, «Sur les différences entre pourcentage et leur interprétation», *Revue française de sociologie*, XXV, pp. 396-420.
- J. Grusky, 2001, «Social stratification», *International Encyclopedia of the Social and Behavioral Sciences*, Elsevier, vol. 21, pp. 14443-52.
- A. Guizzardi, 1997, *Metodi di previsione della domanda e reti neurali*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- H.A. Halsey, 1977, *Towards Meritocracy? The Case of Britain*, in A.H. Halsey e J. Karabel (a c. di), 1977, *Power and Ideology in Education*, Oxford, Oxford University Press.
- C. Jencks, M. Smith, H. Acland, M.J. Bane, D. Cohen, H. Gintis, B. Heyns, S. Michelson, 1972, *Inequality. A Reassessment of the Effect of Family and Schooling in America*, New York, Basic Books; tr. it., *Ineguaglianza. Una rivalutazione del ruolo della scuola e della famiglia*, Roma, Armando, 1983.
- O.J. Jonsson, 1993, *Persisting Inequalities in Sweden*, in Blossfeld e Shavit (eds.), 1993.
- J.O. Jonsson, R. Erikson, 2000, «Understanding educational inequality: the swedish experience», *L'année sociologique*, 50, 2, pp. 345-82.
- J.O. Jonsson, C. Mills, 1993a, «Social class and educational attainment in historical perspective: a Swedish-English comparison. Part 1», *British journal of sociology*, 44, 2, pp. 213-347.
- J.O. Jonsson, C. Mills, 1993b, «Social class and educational attainment in historical perspective: a Swedish-English comparison. Part 2», *British journal of sociology*, 44, 2, pp. 403-428.
- C.A. Kerckhoff, M.J. Trott, 1993, *Educational Attainment in a Changing Educational System. The Case of England and Wales*, in Blossfeld e Shavit (eds.), 1993.
- P. Maarit, H. Hakkinen, 2000, «Neural network used analyze multiple perspectives concerning computer-based learning environments», *Quality and quantity*, 34, pp. 237-58.
- R. Mare, 1980, «Social background and school continuation decisions», *Journal of the american statistical association*, 75.
- R. Mare, 1981, «Change and stability in educational stratification», *American sociological review*, 46.
- G. Massimi, 1999, *Le reti di Hopfield*, in Buscema e Semeion group, 1999a.
- F. Matera, 1999, *Fuzzy Artmap Classification Network*, in Buscema e Semeion group, 1999a.
- C. Meraviglia, 1996, «Models of rappresentation of social mobility and inequality systems. A neural approach», *Quality and quantity*, 30, pp. 231-52.
- C. Meraviglia, 1999, *Questioni di mobilità maschile e femminile: tecniche a confronto*, in Buscema e Semeion group, 1999a.

- D. Merllié, 1985, «Analyse de l'interaction entre variables. Problème statistique ou sociologique», *Revue française de sociologie*, XXVI, pp. 629-52.
- W. Muller, W. Karle, 1993, «Social selection in educational systems in Europe», *European sociological review*, 9.
- H. Paik, 2000a, «Comments on neural networks», *Sociological methods and research*, 28, 4, pp. 425-453.
- H. Paik, 2000b, «Television viewing and high school mathematics achievement: a neural network analysis», *Quality and quantity*, 34, pp. 1-15.
- D. Parisi, 1989, *Intervista sulle reti neurali*, Bologna, il Mulino.
- D. Parisi, 1999, *Mente: i nuovi modelli della vita artificiale*, Bologna, il Mulino.
- S. Patarnello, 1992, *Le reti neuronali. Semplificare la complessità con l'aiuto dell'informatica*, Milano, Angeli.
- M. Pisati, A. Schizzerotto, 1999a, «Changes in the italian regime of social fluidity: higher barriers between manual and non manual classes», paper non pubblicato.
- M. Pisati, A. Schizzerotto, 1999b, «Pochi promossi, nessun bocciato. La mobilità di carriera in Italia in prospettiva comparata e longitudinale», in *Stato e Mercato*, 56, pp. 249-79.
- S. Pizzuti, 1999, *I sistemi neuro fuzzy*, in Buscema e Semeion group, 1999a.
- J. Prévot, 1985, «A propos d'indices et de comparaisons de proportions», *Revue française de sociologie*, XXVI, pp. 601-28.
- D.B. Ripley, 1993, *Statistical Aspects of Neural Networks*, in O.E. Borndorff-Nielsen, J.L. Jensen, W.S. Kendall (eds.), *Networks and Chaos. Statistical and Probabilistic Aspects*, London, Chapman and Hall.
- D.E. Rumelhart, E.G. Hinton, R.J. Williams, 1986, *Learning Internal Representation by Error Propagation*, in D.E. Rumelhart, J.L. McClelland (eds.), 1986, *Parallel distributed processing. Explorations in the microstructure of cognition*, Cambridge, Mit Press.
- R.A. Schiavo, F. Lisi, 1997, *Reti neurali e modelli caotici per la previsione di serie storiche finanziarie: un'applicazione ai tassi di cambio*, in A. Bellacicco e N.C. Lauro (a c. di), 1997, *Reti neurali e statistica*, Milano, Angeli.
- A. Schizzerotto, 1988, *Il concetto di classe sociale: rilevanza e limiti*, in Schizzerotto (a c. di), 1988, *Classi sociali e società contemporanea*, Milano, Angeli.
- A. Schizzerotto, 1990, «Stabilità e mutamento nelle diseguaglianze educative collegate alla classe di origine, al genere e all'appartenenza territoriale», *Scuola democratica*.
- A. Schizzerotto, 1993, *Problemi concettuali e metodologici nell'analisi delle classi sociali* in M. Palumbo (a c. di), *Classi, diseguaglianze e povertà. Problemi di analisi*, Milano, Angeli.
- A. Schizzerotto, 1997, «Perché in Italia ci sono pochi diplomati e pochi laureati? Vincoli strutturali e decisioni razionali degli attori come cause della contenuta espansione della scolarità superiore», *Polis*, 11, 3, pp. 345-65.
- M. Schumacher, R. Rossner, W. Vach, 1996, «Neural networks and logistic regression: part I», *Computational statistics and data analysis*, 21, 6, pp. 661-68.
- Y. Shavit, K. Westerbeek, 1997, «Istruzione e stratificazione in Italia: riforme, espansione educativa e uguaglianza delle opportunità», *Polis*, 11, 1, pp. 91-109.
- Spss Inc., 1997, *Neural Connection 2.0. User's Guide*, Spss Inc. and recognition systems.
- O.G. Stone, 1986, *An Analysis of the Rule and the Learning of Statistical Association*, in D.E. Rumelhart, J. McClelland (eds.), 1986, *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, Cambridge, Mit Press.
- J.D. Treiman, K.-B. Yip, 1989, *Educational and Occupational Attainment in 21 countries*,

- in L.M. Kohn (ed.), 1989, *Cross-national Research in Sociology*, Newbury park, Sage publications.
- J.D. Treiman, 1970, *Industrialization and Social Stratification*, in E.O. Laumann (ed.), *Social Stratification: Research and Theory for the 1970s*, Indianapolis, Bobbs-Merrill.
- W. Vach, R. Rossner, M. Schumacher, 1996, «Neural networks and logistic regression: part II», *Computational statistics and data analysis*, 21, 6, pp. 683-701.
- L.-A. Vallet, 1999, «Quarante années de mobilité sociale en France. L'évolution de la fluidité sociale à la lumière de modèles récents», *Revue française de sociologie*, 40, 1, pp. 5-64.
- L.-A. Vallet, C. Thélot, 2000, «La réduction des inégalités sociales devant l'école depuis le début du siècle», *Economie et société*, 4.
- F. Vicino, 1999, *Reti neurali artificiali e statistica*, in Buscema e Semeion group, 1999a.
- J.C. Vrooman, J. Droonkers, 1986, «Changing educational attainment: some evidence from the Netherlands», *Sociology of education*, 1986, pp. 69-78.
- L. Zeng, 1999, «Prediction and Classification with neural network models», *Sociological methods and research*, 27, 4, pp. 499-524.
- J. Woelfel, 1993, «Artificial neural networks in policy research: a current assessment», *Journal of communication*, 43, 1.

Pervenuto in redazione nel marzo 2003