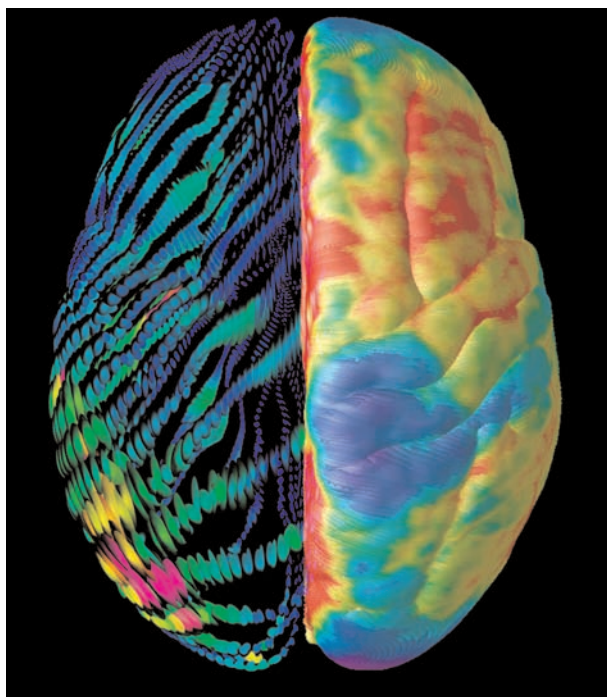


Marco Mozzoni

ALZHEIMER

**Come diagnosticarlo
precocemente
con le reti neurali artificiali**



Semeion Centro Ricerche

FrancoAngeli

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio “Informatemi” per ricevere via e.mail le segnalazioni delle novità.

Marco Mozzoni

ALZHEIMER

**Come diagnosticarlo
precocemente
con le reti neurali artificiali**

FrancoAngeli

Volume pubblicato con il contributo della Bracco S.p.A.



In copertina: “Mappatura della variabilità cerebrale nelle popolazioni umane. Immagine di risonanza magnetica generata utilizzando algoritmi avanzati”. Gentilmente concessa da Paul Thompson, Andrew Lee, Kiralee Hayashi, Agatha Lee, Arthur Toga, *Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, CA, USA.*

Copyright © 2010 by FrancoAngeli s.r.l., Milano, Italy.

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni qui sotto previste. All'Utente è concessa una licenza d'uso dell'opera secondo quanto così specificato:

1. L'Utente è autorizzato a memorizzare l'opera sul proprio pc o altro supporto sempre di propria pertinenza attraverso l'operazione di download. Non è consentito conservare alcuna copia dell'opera (o parti di essa) su network dove potrebbe essere utilizzata da più computer contemporaneamente;
 2. L'Utente è autorizzato a fare uso esclusivamente a scopo personale (di studio e di ricerca) e non commerciale di detta copia digitale dell'opera. Non è autorizzato ad effettuare stampe dell'opera (o di parti di essa).
- Sono esclusi utilizzi direttamente o indirettamente commerciali dell'opera (o di parti di essa);
3. L'Utente non è autorizzato a trasmettere a terzi (con qualsiasi mezzo incluso fax ed e-mail) la riproduzione digitale o cartacea dell'opera (o parte di essa);
 4. è vietata la modificazione, la traduzione, l'adattamento totale o parziale dell'opera e/o il loro utilizzo per l'inclusione in miscellanee, raccolte, o comunque opere derivate.

I may, from the foregoing, state an anatomical fact which, in my opinion, is of major importance in judging the physiology of the nervous system: that is, the existence of a diffuse network...

Camillo Golgi

Nobel Lecture – December 11, 1906

Ringraziamenti

Si ringraziano in particolare:

Prof.ssa Gabriella Bottini
Università degli Studi di Pavia

Prof. Massimo Buscema
Direttore, Centro di Ricerca Semeion, Roma

Dr. Enzo Grossi
Direttore Medico, Dipartimento Farma Italia, Bracco SpA

Indice

Introduzione	pag.	9
1. Intelligenza artificiale e ricerca biomedica	»	11
2. Reti neurali artificiali: che cosa sono e a che cosa servono	»	15
3. Demenza: la diagnosi clinica attuale	»	24
4. Alzheimer: biomarcatori candidati e modelli computazionali	»	31
5. È possibile “automatizzare” la diagnosi dell’Alzheimer?	»	54
6. Il Nun Study scandagliato con le reti	»	59
7. Maschi <i>versus</i> femmine: per le reti è un Alzheimer differente	»	65
8. Diagnosi precoce e prognosi dell’Alzheimer con le reti neurali	»	68
9. Elettroencefalografia e reti nell’Alzheimer	»	77
10. Le nuove reti per il <i>data mining</i> nell’Alzheimer	»	83
11. Patologie neurodegenerative: un problema di rete?	»	102
12. Il campione e l’individuo	»	106
Bibliografia	»	109

Introduzione

Sulla base di quanto consensualmente espresso dalla comunità scientifica internazionale, in questo lavoro verrà sottolineata l'utilità di un approccio integrato alle patologie neurodegenerative, sempre più lette da ricercatori e clinici di tutto il mondo in termini di disfunzione cerebrale a livello di "network".

Passeremo tematicamente in rassegna i più recenti studi sull'applicabilità e sulle potenzialità delle reti neurali artificiali nel contesto della diagnosi e della prognosi dell'Alzheimer, con particolare attenzione agli importanti risultati ottenuti negli anni dalla ricerca di matrice italiana.

Faremo una panoramica sull'intelligenza artificiale nell'ambito medico-biologico, illustreremo le caratteristiche e le applicazioni consolidate e in via di sviluppo delle reti neurali artificiali, forniremo un quadro di insieme dell'attuale processo valutativo che porta alla diagnosi clinica di demenza, descriveremo puntualmente lo stato dell'arte della ricerca sui biomarcatori candidati e sui modelli computazionali dell'Alzheimer.

In una progressione contestualizzata e storicizzata nei limiti concessi dall'impostazione "per nodi" della materia, mostreremo le evidenze di maggiore validità dell'approccio non-lineare rispetto alle tecniche di analisi dei dati della statistica tradizionale; illustreremo infine i nuovissimi algoritmi per il *data mining* semantico e le loro prime applicazioni nel contesto dell'Alzheimer.

La dissertazione è compendata da diagrammi di flusso, tabelle e grafici adattati dagli studi a nostro giudizio più significativi.

1. Intelligenza artificiale e ricerca biomedica

Il termine Artificial Intelligence (AI) è stato coniato dal matematico John McCarthy per l'ormai storico Dartmouth summer research project del 1956 "allo scopo di non lasciare dubbi sull'obiettivo della nuova disciplina: simulare l'intelligenza umana"¹.

L'intelligenza artificiale (IA) ha progredito mirabilmente negli anni e l'interesse per la sua applicazione nell'ambito della ricerca medica e biologica è andato viepiù crescendo, ma solo di recente si sono potuti ottenere modelli sufficientemente dettagliati del funzionamento cerebrale capaci di influenzarne i progetti: in precedenza infatti, in assenza di strumenti in grado di esplorare il cervello con sufficiente risoluzione, gli "ingegneri dell'IA" hanno sviluppato tecniche proprie, non esattamente basate sulla retroingegnerizzazione dell'intelligenza naturale (Boden, 2006; Kurzweil, 2008).

Sin dal suo nascere i ricercatori hanno intuito le potenzialità dell'applicazione dei metodi dell'IA alle scienze della vita e nel 1965 hanno dato vita presso l'Università di Stanford al progetto Dendral (Dendritic algorithm), con l'obiettivo di sviluppare uno strumento di ausilio alla chimica organica per l'identificazione di molecole sconosciute attraverso l'analisi del loro spettro di massa e tenendo in considerazione tutte le conoscenze di riferimento del settore a quel momento acquisite: primo esempio di "automatizzazione dei processi decisionali umani", vi hanno lavorato gomito a gomito informatici, chimici, genetisti, filosofi della scienza (Lindsay, 1980). Negli anni Settanta l'intelligenza artificiale è stata sostanzialmente identificata con un metodo specifico, i cosiddetti "sistemi esperti", perché gran parte dei ri-

¹ Dal comunicato stampa del Dartmouth college, USA, del 13 marzo 2006 di annuncio del congresso di celebrazione dei 50 anni dallo storico meeting del 1956; Dartmouth college office of public affairs, Press release 3/13/2006.

cercatori erano impegnati nello sviluppo di regole logiche capaci di simulare i processi decisionali di esperti umani. Il procedimento consisteva nell'intervistare esperti di diversi settori, come tecnici e medici, per "codificare le loro regole di decisione". In questa direzione sono stati messi a punto diversi sistemi di diagnosi medica (Perry, Pillar, 1990) quali: Mycin, progettato sempre alla Stanford University per diagnosticare e identificare la cura delle malattie infettive, la cui abilità diagnostica è stata messa alla prova in studi di valutazione risultando paragonabile a quella di medici esperti (Shortliffe *et al.*, 1974); CASNET (Causal Associational Networks), sistema a rete semantica che formulava raccomandazioni per il trattamento sulla base di tre livelli descrittivi: osservazione, stato patofisiologico, categoria della malattia; Internist-1, sistema computerizzato di diagnosi in medicina interna generale (Miller, Pople, Myers, 1982).

Nel 1980 la neo costituita American Association for Artificial Intelligence (AAAI) istituisce un gruppo dedicato alle applicazioni mediche dell'IA (AAAI-M). Nel 1986 viene fondata nel vecchio continente l'European society on artificial intelligence in medicine, per promuovere la ricerca e le applicazioni dell'IA nel comparto medico e organizzare biennialmente conferenze di aggiornamento in materia, che prendono il nome di *Artificial Intelligence in Medicine Europe* (AIME). Al congresso AIME del 1991 Edward "Ted" Shortliffe, medico e matematico fra gli sviluppatori di Mycin a Stanford i cui realistici giudizi verranno tenuti a riferimento da molti come una sorta di cartina di tornasole sullo stato effettivo del settore, ritiene opportuno parlare ancora di "Adolescence of AI in medicine" (Shortliffe, 1993).

Con il procedere della ricerca è diventato sempre più chiaro che i processi decisionali umani non si fondano esclusivamente su regole logiche, ma su tipi di evidenza "più morbidi" e che gli spunti decisionali sono influenzati anche dalla combinazione di evidenze derivanti dalle esperienze precedenti. All'inizio degli anni Novanta i sistemi esperti hanno così incorporato "incertezza" e "probabilità". Nell'ultimo decennio la logica bayesiana – delineata in vero nel Settecento dal matematico Thomas Bayes per stabilire la probabilità condizionale di eventi futuri in base a occorrenze simili del passato – ha dato fondamento matematico alla combinazione nelle c.d. "reti di credenze" o reti bayesiane di milioni di regole probabilistiche. I sistemi basati su questo metodo raccolgono costantemente dati dall'esperienza, imparando continuamente e migliorando progressivamente i loro processi decisionali. Come sottolinea oggi Nichelatti "il ragionamento bayesiano è naturalmente connesso all'attività medica in quanto estremamente simile a un processo diagnostico" (Nichelatti *et al.*, 2008). Ancora un matematico, Andrei Markov, ha gettato alla fine dell'Ottocento il seme per

lo sviluppo in anni a noi più recenti di modelli di valutazione delle probabilità di accadimento di determinate sequenze di eventi: i “modelli markoviani” sono stati utilizzati in numerose applicazioni e ricerche biomediche (Rabiner, 1989). Gli algoritmi ricorsivi rappresentano invece un metodo che consente – in sintesi – di ridurre un problema a un insieme di sottoproblemi più semplici da risolvere: una procedura ricorsiva chiama se stessa ciclicamente, ha una o più condizioni di terminazione e a ogni chiamata si avvicina alla condizione di terminazione; questo algoritmo risulta particolarmente utile per eseguire compiti ripetitivi su un set di input variabili (Jiang, Cooper, 2007). Un altro paradigma di auto-organizzazione è rappresentato dagli “algoritmi genetici”, altrimenti detti “evolutivi” proprio perché emulanti l’evoluzione, con tanto di riproduzioni e mutazioni geniche. Il loro funzionamento è basato sulla definizione iniziale di un elenco di parametri quale codice genetico dell’algoritmo e sulla generazione casuale di migliaia di altri codici, considerati organismi soluzione simulati; successivamente, questi organismi vengono messi alla prova in ambienti simulati, dove i migliori sopravvivono e possono riprodursi trasmettendo ai discendenti i propri codici, con mutazioni genetiche casuali consentite. Il risultato è la creazione di generazioni che si evolvono. Utilizzando gli algoritmi evolutivi possiamo attendere l’emergenza di una soluzione da un processo iterativo basato su competizione e velocissimi miglioramenti evolutivi. Gli algoritmi evolutivi – che fanno parte della c.d. teoria del caos o della complessità – vengono utilizzati per catturare schemi altrimenti non identificabili all’interno di dati caotici e si prestano bene per affrontare problemi in cui l’eccessiva numerosità delle variabili in campo non consente il calcolo di soluzioni analitiche precise (Cinquemani *et al.*, 2008).

La biologia, la chimica e la medicina si trovano oggi a fronteggiare nuove sfide legate alla travolgente quantità dei dati da dover prendere in esame e alla necessità di darvi una rapida interpretazione: gli approcci computazionali quali gli algoritmi evolutivi, i sistemi *fuzzy* e le reti neurali artificiali vengono usati con frequenza crescente per fare fronte a questo problema, alla luce del rumore, della non-linearità e delle dinamiche temporali caratteristiche di questi dati. Questi metodi possono essere usati per sviluppare modelli robusti di elaborazione, da impiegarsi singolarmente o insieme agli approcci statistici tradizionali. Ciò è particolarmente vero per il *data mining*, in cui la modellizzazione è una componente chiave della comprensione scientifica (Fogel, 2008).

In tale direzione, una delle branche dell’IA più interessanti e promettenti è quella che si occupa delle reti neurali artificiali (ANN, da Artificial Neural Networks), algoritmi matematici generati da macchine di calcolo capaci di

apprendere dai dati e di “catturare la conoscenza” in essi contenuta. Le reti neurali artificiali istruite possono rappresentare veri e propri modelli digitali del cervello biologico, capaci di rilevare in estesi set di dati complesse relazioni non-lineari fra le variabili (dipendenti e indipendenti), di difficile identificazione da parte dell’essere umano e della statistica tradizionale. Ai nostri giorni si registra nei diversi settori della medicina un interesse e un utilizzo crescente delle ANN, in cardiologia, in radiologia, nella farmacoepidemiologia, nel *data mining* medico e nei processi diagnostici (Patel J., Goyal, 2007).

Nel luglio 2007 si è tenuto ad Amsterdam l’undicesimo congresso AI-ME: a 15 anni dalla celebre relazione di Ted Shortliffe² sull’“adolescenza” del settore, si parla finalmente della “Maturity of artificial intelligence in medicine”³. In particolare viene sottolineato che “proprio per il fatto che i metodi dell’intelligenza artificiale sono sempre più integrati nelle applicazioni mediche, l’IA risulta sempre meno visibile rispetto a quanto lo era nel suo periodo eroico... ed è proprio questo l’indicatore del suo successo” (Patel V.L. *et al.*, 2008; Bellazzi, Abu-Hanna, 2009).

Nel settembre 2008 a Stoccolma prende il via il primo congresso internazionale di “neuroinformatica”. Organizzato dall’International Neuroinformatics Coordinating Facility (INCF) – ente nato da una costola dell’OECD Global science forum al fine di promuovere cooperazione interdisciplinare nello sviluppo di database e infrastrutture computazionali per le neuroscienze – *Neuroinformatics 2008* è la prima vetrina mondiale dei progressi conseguiti da questo innovativo campo di ricerca, che tratta di atlanti digitali del cervello, neuroimmagine, elettrofisiologia, genetica, genomica, ingegneria neuromorfica, interfacce uomo-macchina, neuro-ontologia, neuroinformatica clinica, analisi bayesiana, modelli stocastici di segnalazione neuronale, mappe cognitive distribuite, dinamiche dei network corticali, metodi matematici per la modellizzazione di fenomeni su larga scala, causalità non-lineare, architetture semantiche di integrazione dei dati, modelli computazionali delle funzioni cellulari astrogliali nell’Alzheimer, piattaforme di condivisione su web di dati neurofisiologici e neuroscientifici quale per esempio il Neuroscience Information Framework (NIF)⁴.

² Il 1° luglio 2009 Ted Shortliffe, oggi professore di medicina e di informatica biomedica all’Arizona State University, viene nominato presidente dell’American Medical Informatics Association (AMIA) di Bethesda, USA (ref. AMIA press release, September 3, 2008).

³ Patel V.L. *et al.* (2008), “Maturity of artificial intelligence in medicine”, position paper, *Artif. intell. med.*

⁴ Gli atti di Neuroinformatics 2008 sono pubblicati su *Frontiers in neuroinformatics*.

*2. Reti neurali artificiali: che cosa sono e a che cosa servono**

Come spiega il professor Massimo Buscema del Centro di ricerche Semeion di Roma (Buscema, 1998a; 1998b; 1998c; 1998d; 1998e; Buscema, Gallucci, Grossi, 2006; Buscema *et al.*, 2004; 2005; Buscema, Capriotti *et al.*, 2007; Buscema, Grossi, 2008), le reti neurali artificiali sono “meccanismi di elaborazione di dati che auto-producono le proprie regole sulla base di una esplorazione iterativa dell’evidenza empirica disponibile. In altre parole, il ricercatore non è obbligato fin dall’inizio a specificare una precisa ipotesi – e quindi una classe di possibili regole che connettono i dati di input a quelli di output – per la quale cercare un riscontro fattuale, ma deve soltanto valutare l’algoritmo più adatto per la determinazione della regola appropriata. Ciò rende le ANN particolarmente utili quando si fronteggia un problema per il quale si possiede una rilevante massa di dati, senza avere una precisa idea dei processi che li hanno generati. In questo senso, le ANN si differenziano tanto dai sistemi esperti, in grado di lavorare su basi di dati ristrette in condizioni di scarsa comprensione del problema ma caratterizzati da una struttura molto rigida, quanto dai modelli fisici, che hanno bisogno di una precisa specificazione teorica dalla cui correttezza dipende interamente la qualità delle loro elaborazioni. Rispetto ad altre tecniche di analisi, dunque, le ANN sono utili quando si dispone di molti dati, ma non di una buona ipotesi esplicativa per organizzarli”.

Le ANN sono adattive: scoprono dinamicamente le regole “sfumate” che connettono vari insiemi di dati. Ciò significa che le regole che esse selezionano dipendono criticamente dall’insieme di dati che viene loro presentato. Man mano che nuovi dati vengono aggiunti, le ANN aggiusteranno

* Si ringrazia il Centro di ricerche Semeion di Roma per le informazioni tecniche rese disponibili per la stesura del presente capitolo.

le loro regole di conseguenza, integrando i vecchi dati con i nuovi, senza bisogno di alcun intervento esterno. L'aggiornamento continuo gestito da una ANN consente, quindi, di produrre una "banca dati dinamica", le cui regole saranno automaticamente raffinate dalle ANN in diretto rapporto con l'evoluzione temporale del problema reale che l'ANN deve gestire. Se l'evoluzione del problema richiede una drastica ridefinizione delle regole, l'ANN procederà pertanto a una radicale ristrutturazione del proprio modello di interdipendenza che lega le diverse variabili.

ANN: che cosa sono, in sintesi

- Sono algoritmi matematici (implementati con software) capaci di apprendere dai dati e di catturare la conoscenza in essi contenuta (Patel J.L., Goyal, 2007).
- Sono meccanismi di elaborazione di dati che auto-producono le proprie regole sulla base di una esplorazione iterativa dell'evidenza empirica disponibile (Buscema, Gallucci, Grossi, 2006; Buscema *et al.*, 2004; 2005; Buscema, Capriotti *et al.*, 2007; Buscema, Grossi, 2008).

Le ANN generalizzano, quindi, predicono e riconoscono: quando una ANN è stata addestrata con opportuni dati a determinare le regole sfumate che meglio descrivono un certo fenomeno, essa è anche in grado di effettuare delle generalizzazioni corrette, rispondendo appropriatamente anche a dati che non ha mai avuto occasione di elaborare, purché a loro volta rappresentativi del fenomeno in esame.

Le ANN possono essere usate, in generale, per affrontare due tipi di problemi complessi. Il primo di questi problemi è il seguente: date N variabili, quale è il miglior modo di connetterle sensatamente, ovvero, in linguaggio tecnico, qual è la loro matrice di connessione migliore? Un problema di questo tipo si definisce "endogeno", in quanto le variabili note sono di per sé sufficienti per trovare una buona soluzione. Il secondo tipo di problemi che una ANN può affrontare è esprimibile nel modo seguente: date N variabili, su cui è facile avere dati, e M variabili, diverse dalle prime e sulle quali è difficile e/o costoso e/o possibile avere i dati, valutare i valori delle M variabili, tramite i dati delle N variabili. Un problema di questo tipo lo definiamo "esogeno". Questo tipo di problemi viene chiamato di previsione quando le M variabili sono variabili che seguono nel tempo le N variabili; viene chiamato di riconoscimento e/o di discriminazione e/o di estrazione dei tratti fondamentali quando le M variabili dipendono da qualche tipologia statica e/o dinamica. Concettualmente si tratta dello stesso tipo di fenomeno: trami-

te i valori di alcune variabili note, determinare i valori di altre variabili non note. Perché problemi di questo tipo siano affrontabili correttamente con ANN è necessario seguire un preciso protocollo di validazione. È necessario avere un buon campione di casi, per ciascuno dei quali le N variabili (note) e le M (da prevedere) siano entrambe note e attendibili. Questo campione di dati completi sarà necessario per addestrare l'ANN e per valutarne le prestazioni previsionali. Il protocollo di validazione prevede che una parte dei casi del campione completo (campione di taratura) siano usati per addestrare l'ANN (*training set*), mentre la parte restante dei casi sia utilizzata per valutare la capacità previsionale dell'ANN (*testing set* o *validation set*). In questo modo si ha la possibilità di controllare, prima dell'uso reale, l'affidabilità dell'ANN nel trattare il problema in oggetto.

ANN: vantaggi applicativi

- *Non-linearità* → migliore adesione ai dati.
- *Insensibilità al rumore* → previsioni accurate in presenza di incertezza nei dati ed errori di misurazione.
- *Elevato parallelismo* → velocità di elaborazione e tolleranza a “guasti hardware”.
- *Apprendimento e adattamento* → il sistema è in grado di “aggiornare” (update) la sua struttura interna in risposta a cambiamenti ambientali (input esterni).
- *Generalizzazione* → applicazione del modello risultante a dati nuovi (non appresi).

Il numero di problemi affrontabili tramite ANN, trattandosi di una metodologia trasversale, è teoricamente illimitato (fig. 1). Buscema sottolinea che varrebbe la pena di affrontare un problema con le reti neurali nei seguenti casi:

- il problema è complesso, cioè la funzione che si intende computare non è riducibile a una funzione lineare;
- si hanno molti dati ma nessun modello fisico-matematico è disponibile per approssimare in modo utile la funzione che li connette;
- le variabili in gioco sono molte e di natura diversa (metriche, categoriche ecc.);
- si hanno molti dati, l'attendibilità di alcuni dei quali è incerta;
- altre metodologie non hanno fornito risultati soddisfacenti;
- è utile una soluzione anche approssimata del problema e non solo una soluzione certa o la migliore in assoluto;

- una soluzione accettabile del problema produce grandi risparmi in termini umani e/o economici;
- si ha una notevole serie di casi che mostrano il comportamento “strano” del fenomeno che rappresenta il problema (solo per le ANN eteroassociate);
- “si ha veramente voglia di risolvere il problema”.

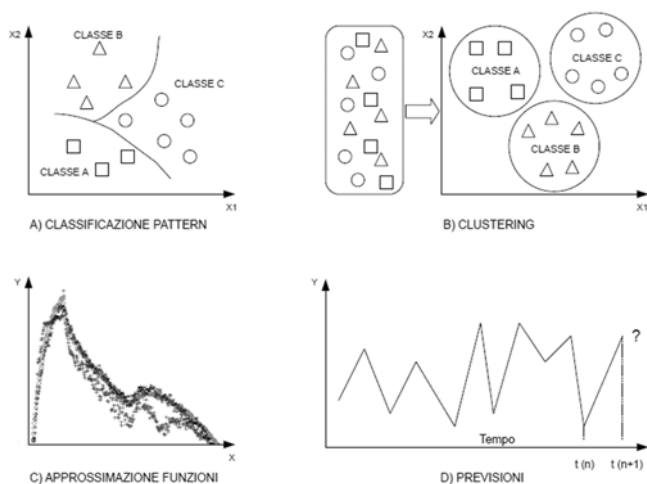
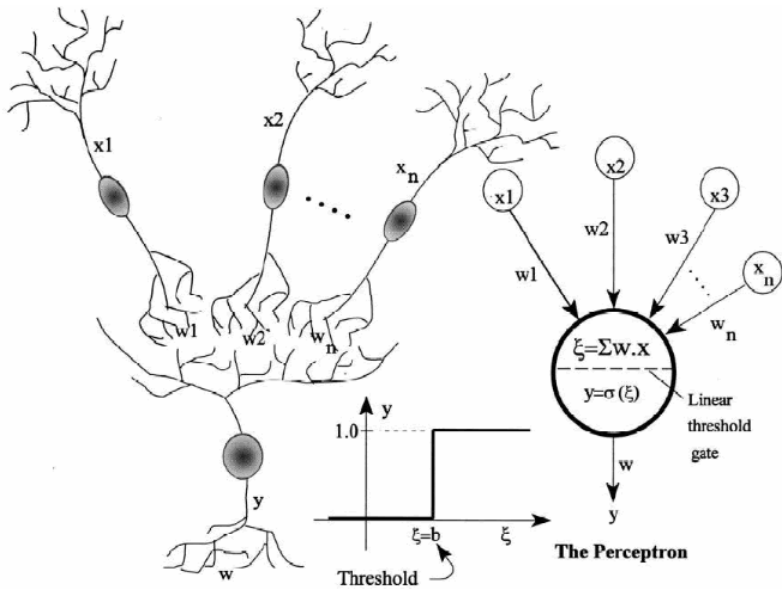


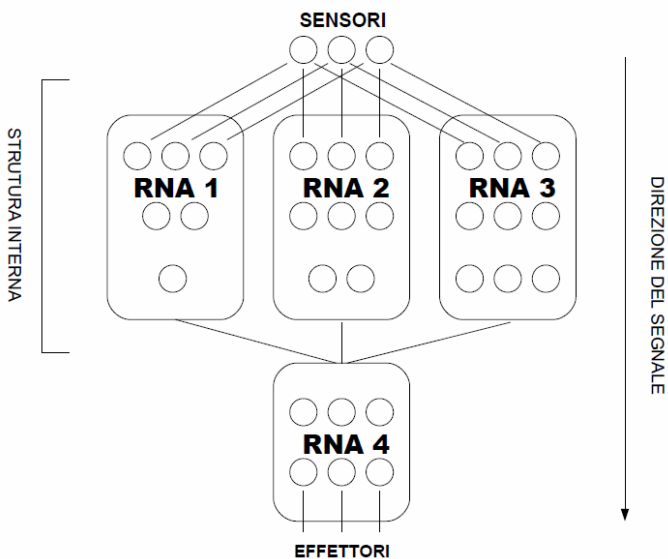
Fig. 1 – Utilizzo delle reti neurali artificiali per (a) classificazioni, (b) clustering, (c) approssimazione di funzioni, (d) previsioni

Dal punto di vista teorico, non è utile assumere che ogni specifica base di dati ha una famiglia di ANN che approssima in modo ottimo la sua funzione. È più corretto ritenere che ogni base di dati possa essere letta da famiglie diverse di ANN, ciascuna delle quali fornisce una modellizzazione matematica diversa dalla stessa base di dati. ANN diverse rispondono a domande diverse. Di fronte a una stessa base di dati, ANN diverse possono cooperare, in sequenza e in parallelo, per modellizzare, in modi più opportuni le une rispetto alle altre, lo stesso problema. Ciò permette di pensare a una teoria degli organismi artificiali, nella quale ANN diverse siano organizzate per simulare processi complessi (vedi figg. 2 e 3). Una tale possibilità di “ibridazione” tra ANN apre la porta anche a schemi di ibridazione con altre metodiche di computazione quali i sistemi esperti, i sistemi *fuzzy*, gli algoritmi genetici. In questa direzione i ricercatori si sono già mossi, realizzando diverse applicazioni, accademiche e industriali.



Fonte: Rosenblatt (1962).

Fig. 2 – Il perceptrone di Rosenblatt



Fonte: Mozzoni (2008).

Fig. 3 – Un organismo artificiale, rete di reti