

# Intelligenza artificiale nella gestione dei processi siderurgici

Valentina Colla, Marco Vannucci, Gianluca Nastasi, Silvia Cateni

Le tecniche di intelligenza artificiale hanno trovato un vasto impiego in ogni campo industriale e per una moltitudine di applicazioni. Le ragioni di tale successo stanno sia nella potenza di questi algoritmi che nel progresso tecnologico degli strumenti tramite i quali essi vengono implementati. Inoltre, il progresso dell'automazione nella gestione dei processi produttivi ha reso disponibile e fruibile una enorme mole di dati, i quali contengono spesso buona parte del know-how di un'azienda. Le tecniche di intelligenza artificiale possono consentire di sfruttare appieno il contenuto informativo di questi dati, estraendone la "conoscenza vera" del processo, grazie alla loro flessibilità, alla loro capacità di rappresentare problemi altamente complessi e non completamente conosciuti e ad una gran quantità di algoritmi di apprendimento efficienti ed adatti ad ogni applicazione. Nel settore siderurgico, l'applicazione di tali tecniche è sempre più frequente e proficua e si prevedono incoraggianti sviluppi per il futuro.

## Keyword

*Intelligenza artificiale, controllo di processo, processi siderurgici, automazione industriale.*

Le tecniche di intelligenza artificiale, quali le reti neurali, la logica fuzzy ed i sistemi di controllo che ne derivano, nonché gli algoritmi genetici, hanno trovato un vasto impiego in ogni campo industriale e per una moltitudine di applicazioni. Le ragioni di tale successo stanno, da un lato, nella potenza di questi algoritmi, ma sicuramente anche nel progresso tecnologico degli strumenti tramite i quali essi vengono implementati (in primo luogo i processori).

Nella gestione dei processi produttivi, il progresso della sensoristica e dell'automazione nonché degli strumenti per la gestione dati ha reso disponibile e fruibile una enorme mole di dati, i quali contengono spesso buona parte del know-how di un'azienda. Tuttavia il contenuto informativo di questi dati non sempre viene sfruttato appieno, ossia i dati vengono immagazzinati per avere un archivio storico relativamente ai processi esaminati, ma non sempre si procede ad una analisi che permetta l'estrazione di informazione fruibile in termini di conoscenza del processo produttivo ed utilizzabile per migliorare ed ottimizzare la gestione del processo stesso. L'estrazione di conoscenza dai dati grezzi (in inglese Data Mining) è una materia vasta e complessa che richiede l'applicazione anche congiunta di tecniche di analisi derivanti dalla statistica tradizionale e di tecniche di intelligenza artificiale.

Tali tecniche si dimostrano particolarmente adatte per la loro flessibilità, la capacità di rappresentare problemi altamente complessi e non completamente conosciuti e soprattutto l'enorme bagaglio di algoritmi di apprendimento che sfruttano i dati disponibili al fine di ottimizzare, ad esempio, i parametri interni di un modello (affinché esso rappresenti nella maniera più fedele possibile la realtà) oppure un insieme di regole di controllo espresse in forma linguistica (al fine di migliorare la gestione del processo controllato, secondo criteri di bontà precedentemente definiti). Tali algoritmi di adattamento consentono ai sistemi sviluppati di adattarsi continuamente ed automaticamente alla evoluzione dei cicli produttivi.

È tuttavia importante concepire l'impiego degli strumenti dell'intelligenza artificiale non come alternativo ma come complementare alla conoscenza ed alla esperienza pratica che il personale che lavora in azienda a tutti i livelli (dal manager di processo all'operaio specializzato) matura sul campo. Sono disponibili tecniche e metodologie per integrare tale conoscenza all'interno dei sistemi che vengono sviluppati e, comunque, chi si fa carico dello sviluppo di tali sistemi trae un enorme vantaggio dalla informazione che gli viene fornita da chi opera direttamente su impianti e/o gestisce i cicli produttivi.

Nel paragrafo che segue sono introdotte le principali tecniche di intelligenza artificiale (reti neurali, logica fuzzy, algoritmi genetici). Il terzo paragrafo è dedicato ad una serie di applicazioni delle tecniche di intelligenza artificiale relative ai processi siderurgici. Infine viene presentata una discussione conclusiva, volta a mettere in luce vantaggi e problematiche relative all'impiego delle tecniche di intelligenza artificiale e

V. Colla, M. Vannucci, G. Nastasi, S. Cateni - Scuola Superiore Sant'Anna, Laboratorio PERCRO, Polo Sant'Anna Valdera, Pontedera, (PI), Italy

trarre conclusioni su quali siano i requisiti base per l'impiego pratico di tali tecniche in un contesto industriale.

## Principali tecniche di intelligenza artificiale

È naturalmente difficile riassumere una materia così vasta in poche pagine: il tentativo che viene fatto nel presente contesto è quello di fornire per ciascuna tecnica il principio ispiratore, le caratteristiche fondamentali ed il meccanismo base di funzionamento, lasciando i dettagli e gli approfondimenti alla volontà ed alla curiosità del lettore.

### Reti Neurali

Le reti neurali artificiali (Artificial Neural Networks – ANN) sono un modello di calcolo ispirato al sistema nervoso centrale degli esseri viventi sia nella struttura che nel funzionamento. In particolare una rete neurale è costituita da un insieme di elementi di calcolo fortemente interconnessi detti neuroni la cui cooperazione permette complesse elaborazioni dei dati [1].

Le ANN per la loro flessibilità d'impiego sono state utilizzate in molteplici ambiti e per vari compiti come la regressione, la classificazione, il clustering e il filtraggio di dati: hanno ottime capacità di generalizzazione e adattamento e tollerano il rumore presente nei dati. Il punto di forza dei modelli basati sull'uso di una ANN consiste nella capacità di scoprire le relazioni esistenti fra i dati (eventualmente fra gli input e gli output) basandosi sull'osservazione di un insieme di dati utilizzato nella fase di addestramento. In questa fase infatti, tramite processi che simulano i meccanismi di apprendimento degli esseri viventi, la rete modifica i propri parametri interni al fine di modellare le relazioni esistenti fra i pattern forniti per l'addestramento. Esistono varie categorie di ANN: la classificazione più comune le divide in supervisionate e non supervisionate.

Alle reti supervisionate nella fase di addestramento viene presentato un insieme di dati di input corredato dei corrispondenti output desiderati: i parametri interni della rete vengono modificati da un opportuno algoritmo al fine di riprodurre la relazione fra input e output. La più comune tipologia di rete supervisionata è la Feed-Forward Neural Network (FFNN). In tale tipo di rete, mostrata nella figura 1(a), i neuroni sono disposti in strati successivi in cui i neuroni di strati adiacenti sono totalmente connessi fra loro e a ciascuna di queste connessioni è associato un peso. Durante il calcolo il primo strato (detto di ingresso) riceve gli input dall'esterno, passa il proprio risultato al successivo strato (detto hidden) e così ancora fino all'uscita del segnale tramite lo strato di output. Il segnale si propaga nei vari strati attraverso i singoli neuroni come illustrato nella figura 1(b) secondo l'equazione (1) in cui  $out_1$  e  $out_2$  rappresentano le uscite dei neuroni  $N_1$  e  $N_2$  rispettivamente,  $b_3$  è un parametro detto polarizzazione (o bias) del neurone  $N_3$  e  $f$  è la funzione di attivazione di  $N_3$ .

$$out_3 = f(w_1 out_1 + w_2 out_2 - b_3) \quad (1)$$

I parametri di una ANN di tipo FFNN che vengono modificati durante il processo di addestramento sono costituiti dai pesi associati alle singole connessioni e dai bias di ciascun neurone mentre resta fissa la funzione di trasferimento. Tipiche funzioni di trasferimento comunemente utilizzate sono la funzione sigmoide, la lineare o quelle a scalino.

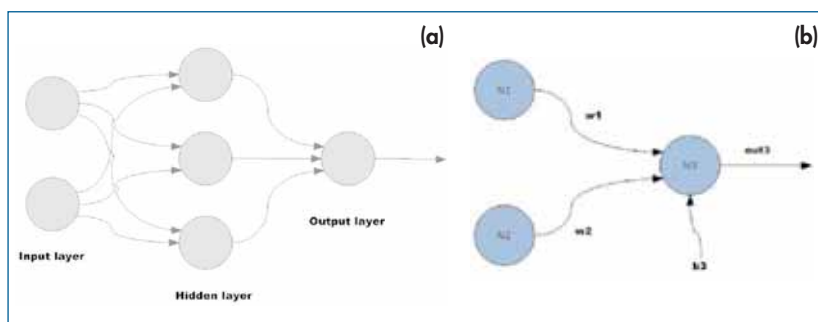


Figura 1 - Un esempio di rete Feed-Forward a 2 strati (a) e un dettaglio del calcolo relativo ad un singolo neurone (b)

Esiste una vasta gamma di algoritmi di training ma in generale il loro funzionamento si basa su una esplorazione della superficie di errore tramite il metodo della discesa del gradiente al fine del raggiungimento del minimo di tale superficie in funzione dei parametri della rete. Il più celebre fra questi algoritmi è la backpropagation [2] che modifica i parametri della rete in maniera proporzionale alla discrepanza fra la sua uscita e l'output desiderato e propaga all'indietro tale errore modificando in maniera appropriata i pesi di tutte le connessioni.

Le reti non supervisionate sono addestrate utilizzando una serie di input pattern cui non corrisponde una uscita desiderata: sono solitamente utilizzate per compiti quali la classificazione, il filtraggio e il clustering. La più conosciuta fra tali reti è la Self Organizing Map (SOM) [3]. Essa ha la proprietà di suddividere i pattern di input in più categorie sulla base delle somiglianze fra gli stessi creando un insieme di classi (la mappa, corrispondente ai neuroni della rete) che fungono da rappresentanti per gruppi di input, in modo che input simili siano messi in corrispondenza con la medesima classe. Inoltre tali classi sono create nello spazio di input in modo da riflettere distribuzione e topologia dell'insieme dei pattern utilizzati per l'addestramento della rete.

### Logica e sistemi fuzzy

La teoria degli insiemi fuzzy, elaborata da L.A. Zadeh [4] a metà degli anni '60, introduce un metodo per la risoluzione problemi tramite il cosiddetto ragionamento approssimato. Nel ragionamento approssimato un elemento appartiene ad un insieme con un certo grado di appartenenza, estendendo così la logica tradizionale (crisp) in cui un elemento appartiene o meno ad un insieme in modo "totale". Un generico insieme fuzzy  $A$  è definito come  $A = \{(x, \mu_A(x)) \mid x \in X\}$  dove  $X$  è l'universo di definizione (universe of discourse) e  $\mu_A(x)$ , con

$\mu_A: X \rightarrow [0, 1]$  è detta funzione d'appartenenza (membership function). Un insieme fuzzy è dunque composto da coppie di valori  $(x, \mu_A(x))$ :  $\mu_A(x)=0$  denota l'estraneità dell'elemento  $x$  all'insieme  $A$ , viceversa per  $\mu_A(x)=1$ . Tipiche funzioni d'appartenenza sono a triangolo, a trapezio, la gaussiana, la sigmoide e la curva di Bell. La logica fuzzy rappresenta un'estensione della logica booleana (logica bivalente) che consente di rappresentare anche le sfumature di verità. Per manipolare le variabili fuzzy occorre estendere gli operatori logici AND, OR e NOT (ovvero intersezione, unione e negazione).

Nella prima metà degli anni '70 E. Mamdani [5] estende la teoria degli insiemi fuzzy, consentendo la costruzione di sistemi di controllo attraverso la definizione di regole fuzzy della forma "if  $x$  is  $A$  then  $y$  is  $B$ ", dove  $A$  e  $B$  sono valori linguistici definiti per mezzo di insiemi fuzzy e  $x$  è una variabile crisp definita sull'universo del discorso  $X$ . La parte *if* delle regole fuzzy (detta antecedente o premessa) e la parte *then* (chiamata conseguenza o conclusione) possono essere costituite da più parti unite da operatori logici. Nella logica fuzzy una proposizione condizionale come quella sopra descritta è vera con un certo grado nell'intervallo  $[0, 1]$ , mentre nella logica classica una proposizione è semplicemente vera o falsa. La conclusione viene influenzata dalla premessa tramite la funzione di implicazione  $I: [0, 1] \times [0, 1] \rightarrow [0, 1]$  la quale, date due preposizioni fuzzy  $P$  e  $Q$ , definisce il valore di verità della proposizione condizionale "if  $P$  then  $Q$ ". Data una regola generica e complessa come:

$$\text{if } x \text{ is } A \text{ AND } y \text{ is } B \text{ then } z \text{ is } C \quad (2)$$

l'algoritmo da eseguire per la sua interpretazione è il seguente:

- *Fuzzificazione degli ingressi*: vengono calcolati i valori di tutte le parti della premessa utilizzando le rispettive funzioni d'appartenenza e ottenendo così valori compresi nell'intervallo  $[0, 1]$ .
- *Applicazione degli operatori logici*: se la premessa è costituita da più variabili, si applicano gli operatori logici, ottenendo un singolo valore compreso fra 0 e 1.
- *Applicazione della funzione implicazione*: il valore ottenuto al passo precedente viene utilizzato per ottenere l'insieme fuzzy di uscita, rappresentato da una funzione d'appartenenza opportunamente modificata dalla funzione di implicazione.

Tuttavia per ottenere un sistema utile, occorre definire un certo numero di regole, ognuna delle quali darà in uscita un insieme fuzzy modificato dalla funzione di implicazione. Per aggregare questi risultati vanno eseguiti i seguenti passi, in successione ai tre descritti precedentemente:

- *Aggregazione delle conclusioni*: gli insiemi fuzzy ottenuti dalle singole regole vengono combinati ottenendo un singolo insieme per ogni variabile d'uscita. L'aggregazione può essere definita in diversi modi: alcuni dei metodi più usati sono max (massimo), probor (or probabilistico) e sum (somma).
- *Defuzzificazione*: ricava un singolo valore reale a partire dall'insieme fuzzy aggregato. Esistono diversi paradigmi, fra cui il più usato è quello del centroide.

La figura 2 riporta uno schema tipico di un Sistema di Inferenza Fuzzy (FIS).

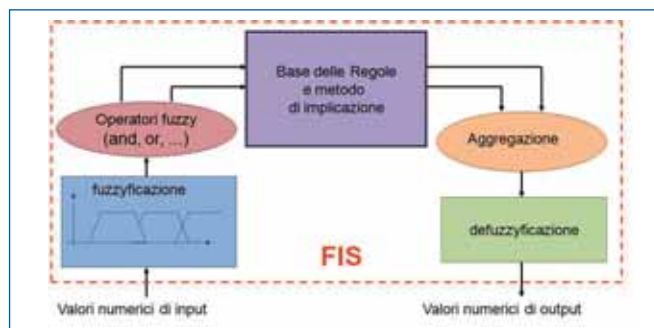


Figura 2 - Schema di un Sistema di Inferenza Fuzzy

La logica fuzzy e i FIS sono largamente impiegati nei sistemi di supporto alle decisioni, come descritto in [6] e in [7]. I principali vantaggi di questo approccio sono:

- *flessibilità*: le regole che compongono il sistema fuzzy sono facilmente modificabili;
- *scalabilità*: è possibile aggiungere o rimuovere variabili e regole senza incidere sulla restante parte del sistema;
- *semplicità*: l'impiego della logica fuzzy, e quindi delle variabili linguistiche, rende possibile, da una parte, la comprensione del sistema anche da non esperti nel settore. D'altra parte, questo genere di sistemi risulta particolarmente adatto ad implementare sistemi di controllo o valutazione in cui la base di conoscenza proviene da esperti e personale tecnico che ha scarsa familiarità con il formalismo matematico ma una notevole conoscenza del problema in oggetto.

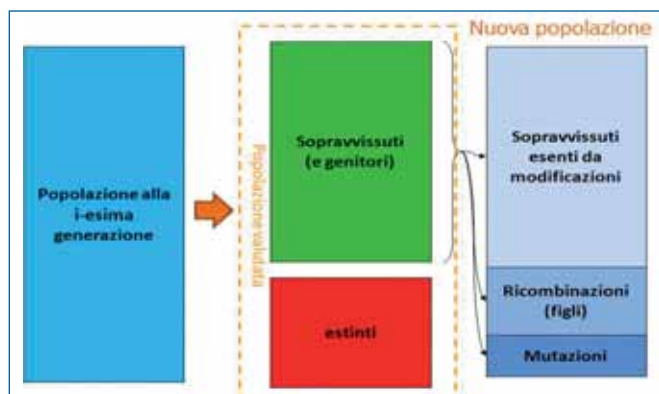
Lo svantaggio principale è dato dalla difficoltà del tuning iniziale dei fuzzy set relativi alle variabili d'ingresso, cosa che peraltro può essere superata acquisendo "conoscenza" dagli esperti (cioè dal personale tecnico). In alternativa, sono stati proposti svariati algoritmi di unificazione fra reti neurali e logica fuzzy [8] [9], nei quali i parametri del FIS vengono ottimizzati sfruttando un set adeguato di dati. Si parla in tal caso più propriamente di sistemi neuro-fuzzy.

#### Algoritmi genetici

Gli Algoritmi Genetici (AG), sono una tecnica euristica di ricerca globale di soluzioni esatte o approssimate a problemi di ottimizzazione, introdotta, fra gli altri, da J. Holland [10] negli anni 70, e costituiscono una particolare classe dei cosiddetti *Algoritmi Evolutivi* basati su concetti mutuati dalla genetica e dalla biologia (quali selezione, mutazione ed incrocio). Possono essere applicati efficientemente nella soluzione di problemi multi-dimensionali, multi-modalità e complessi.

Un AG è un algoritmo iterativo nel quale una *popolazione* costituita da rappresentazioni astratte (chiamate cromosomi o genoma) di soluzioni candidate (chiamate individui o fenotipi) di un problema di ottimizzazione evolve verso soluzioni migliori. Nella loro forma tradizionale, i cromosomi sono costituiti da stringhe di bit, ma sono possibili anche altre codifiche, come vettori di numeri interi o reali, e genericamente rap-

presentano un punto nello spazio di ricerca. Il primo passo nella implementazione di un AG è rappresentato dunque dalla definizione di una opportuna funzione di codifica, che mappa lo spazio delle soluzioni nello spazio dei cromosomi (che è poi lo spazio di ricerca). L'evoluzione parte normalmente da una popolazione di individui generata casualmente che successivamente muta al passare delle *generazioni* (termine che indica le iterazioni dell'AG). Ad ogni passo (cfr figura 3), la popolazione viene suddivisa in almeno tre gruppi, ciascuno dei quali subirà un tipo di ricombinazione diversa o non la subirà affatto: *mutazione*, *cross-over* e *selezione*. La prima prevede che venga selezionato e mutato un elemento (detto gene) scelto a caso dal vettore che rappresenta l'individuo, mentre la seconda prevede che venga creato un nuovo individuo (detto figlio) combinando i geni di due individui (detti genitori), infine gli individui migliori (elite) vengono selezionati e passano alla successiva generazione immutati (selezione). Gli individui così ottenuti vanno a costituire una nuova popolazione.



**Figura 3 - Schema della generazione di una nuova popolazione tramite AG**

Al termine di ogni generazione, ciascun individuo viene valutato rispetto ad una funzione obiettivo, detta *funzione di fitness*, e tipicamente l'algoritmo termina quando viene raggiunto il numero massimo di generazioni oppure quando uno o più individui raggiungono un livello di fitness soddisfacente. Ogni operatore ha una sua funzione specifica e necessaria alla corretta evoluzione della popolazione: la mutazione introduce individui nuovi nella popolazione, per permettere l'esplorazione dello spazio di ricerca. Il crossover permette di generare individui con caratteristiche migliori (ossia un più elevato valore della funzione di fitness) rispetto ai loro genitori. Infine la selezione consente di non perdere il migliore "pool genetico" presente nella popolazione.

Al variare delle funzioni di fitness, crossover e mutazione si hanno AG diversi che possono così essere adattati alla soluzione di moltissimi problemi tra cui, in particolare, quelli caratterizzati da funzioni obiettivo discontinue, non derivabili, stocastiche, o fortemente non lineari.

## Esempi applicativi

Molteplici sono i campi di applicazione delle tecniche di intelli-

genza artificiale: esse si prestano particolarmente bene ai casi in cui manca un modello matematico che rappresenta il problema in oggetto, ma anche ad applicazioni in cui un modello matematico esiste ma occorre ottimizzarne i parametri in modo che i valori forniti in uscita dal modello si accordino al meglio con alcuni dati sperimentali. Naturalmente esistono tecniche di ottimizzazione di tipo tradizionale: le tecniche di intelligenza artificiale sono da applicare qualora tali tecniche si siano dimostrate infruttuose o non applicabili. Ad esempio le tecniche di ottimizzazione basate sul gradiente richiedono il calcolo della derivata del modello, che, in caso di modelli complessi e articolati, non è sempre agevole.

### *Predizione della durezza di acciai microlegati al boro*

La temprabilità è una caratteristica fondamentale dell'acciaio: nella pratica operativa la durezza di un acciaio viene valutata tramite un test standard chiamato test Jominy, durante il quale un provino di forma cilindrica viene sottoposto ad un trattamento termico standard che influenza la microstruttura dell'acciaio (e, di conseguenza, la sua temprabilità) lungo la dimensione maggiore del provino. Viene costruita una curva misurando la durezza in gradi Rockwell (HRC) in punti prefissati, che è chiamata *profilo di durezza Jominy* e caratterizza la temprabilità di un prodotto in acciaio.

Sin dagli anni '40 il boro viene impiegato come elemento microlegante nella produzione di acciaio per accrescere la durezza dei prodotti destinati ai trattamenti termici. I benefici economici dell'impiego del boro per accrescere la durezza sono evidenti, se si considera che l'aggiunta di modeste quantità (dai 5 ai 30 ppm) di tale elemento permettono di ottenere caratteristiche di notevole profondità di tempratura senza l'impiego di altre costose ferroleghie (quali, ad esempio, cromo, molibdeno, vanadio, nichel ecc.). D'altro canto, non è sempre facile stimare la temprabilità degli acciai al boro, a causa della influenza sia della composizione chimica che del processo di lavorazione. Di conseguenza gli strumenti statistici standard normalmente impiegati con successo per la predizione del profilo di temprabilità Jominy per acciai di tipo tradizionale non forniscono prestazioni soddisfacenti nel caso degli acciai al boro [11]. Per tale motivo sono stati sviluppati vari approcci basati su reti neurali per la predizione del profilo Jominy di questo tipo di acciai. I più semplici sono basati su reti di tipo Multi-Layer Perceptron (MLP) a due strati le cui uscite corrispondono ai diversi punti del profilo stesso [12] e presentano un numero notevole di inconvenienti: approcci più avanzati [13] utilizzano la caratterizzazione parametrica del profilo Jominy e le reti neurali per associare la chimica dell'acciaio ai parametri della curva che riproduce il profilo di durezza, come schematizzato nella figura 4(a). Questa soluzione tuttavia non dà risultati soddisfacenti quando si hanno profili di curve di forma molto variabile: in tale caso ha dato ottimi risultati l'approccio sequenziale, in cui la predizione di ciascun punto è funzione non solo della chimica dell'acciaio ma anche del valore predetto nei punti della curva precedenti (ossia corrispondenti a valori minori della distanza), come nella figura 4(b).

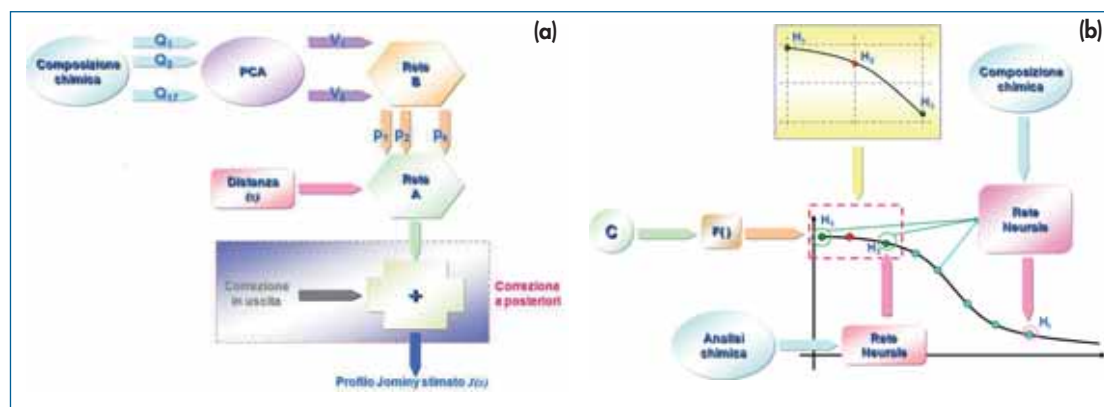


Figura 4 - Predizione delle curve Jominy per acciai micro legati al boro tramite caratterizzazione parametrica della curva (a) e tramite predizione sequenziale con una serie di reti neurali (b)

### Predizione del fenomeno del clogging

Nella produzione dell'acciaio, l'acciaio liquido è trasferito dalla siviera, tramite la paniera, agli altri macchinari utilizzati per le lavorazioni successive. Durante questo passaggio l'acciaio fuso transita attraverso degli ugelli della sezione di alcuni centimetri i quali, a causa dei depositi presenti nel materiale, possono otturarsi dando luogo al fenomeno chiamato *clogging* che impedisce il passaggio del materiale con effetti deleteri per l'intera produzione. Se previsto con il dovuto anticipo il clogging può essere evitato con l'aggiunta di materiale opportuno all'acciaio liquido che non compromette la qualità del prodotto [14] ma tale fenomeno non è ancora del tutto conosciuto e non esistono modelli per la predizione del clogging affidabili a causa del numero delle variabili e della complessità delle interazioni in gioco [15]. I metodi utilizzati ad oggi per la predizione del rischio di clogging sono basati sulla determinazione di zone sicure o insicure dello spazio degli input determinate empiricamente o con semplici metodi statistici lineari che non tengono conto delle complesse interazioni fra le variabili e la cui affidabilità è bassa. Una ulteriore difficoltà che è stata riscontrata dai metodi standard nella identificazione del clogging è dovuta all'esiguo numero di tali situazioni rispetto a quelle normali che ne rende difficile la caratterizzazione. I metodi standard, infatti quando utilizzati per la classificazione di tali dati, falliscono non rilevando proprio le situazioni di clogging.

Le reti neurali sono state usate per l'individuazione delle situazioni di clogging in base alla composizione chimica dell'acciaio prodotto e ad alcuni parametri inerenti al processo. Tale strumento si è rivelato particolarmente adatto al compito prefissato grazie alla capacità di modellare relazioni fortemente non lineari e le interazioni fra le variabili in gioco. In particolare è stata utilizzata con successo una rete neurale di tipo SOM la quale ha permesso la creazione di "classi" di patterns fra cui sono state individuate quelle a rischio di clogging [16]. I risultati ottenuti dal metodo basato sull'uso delle SOM, mostrati nella tabella, sono molto soddisfacenti e superano le limitazioni incontrate dagli altri metodi testati, infatti l'approccio descritto permette di riconoscere gran parte delle situazioni critiche generando una percentuale accettabile di falsi allarmi.

Metodo	Clogging individuati %	Falsi allarmi %	Corretti %
MLP	33,3	5,2	91,4
Rec BF	48,2	30,1	69,5
ArtMap	0	2,8	86,3
SOM 8x8	77,2	30,7	76,7

Tabella - confronto fra le prestazioni dei sistemi per la predizione del clogging basati su diversi tipi di rete neurale

### Filtraggio di database industriali tramite un FIS

Un outlier è un'osservazione che si discosta considerevolmente dalle altre osservazioni disponibili. La rilevazione degli outliers diventa essenziale quando i database da analizzare provengono da un contesto industriale. Infatti molto spesso i dati contengono diverse anomalie dovute per esempio a errori umani durante le registrazioni dei dati, condizioni operative anomale o malfunzionamenti dei sensori utilizzati. I metodi classici di rilevazione degli outliers [17] si possono classificare in quattro categorie:

- *Metodi basati sulla distanza.* Vengono definiti outliers quei dati che si discostano in termini di distanza dalle altre osservazioni [18].
- *Metodi basati sul clustering.* Dopo aver classificato l'intero database con metodi tradizionali di clustering, si definiscono outliers quei dati che rimangono isolati dai clusters formati.
- *Metodi basati sulla densità.* Si ricercano gli outliers in base al comportamento (densità) dei dati nell'intorno o vicinato.
- *Metodi basati sulla distribuzione.* Si definiscono outliers quei dati che deviano da un dato modello di distribuzione.

Non esiste il miglior metodo di rilevazione di outliers in quanto ciascuno di essi diventa esaustivo o inutile in base alla distribuzione del database e alla natura degli eventuali outliers.

La logica fuzzy è stata utilizzata in [19] per la realizzazione di un sistema automatico di identificazione degli outliers che prescinde dalla conoscenza a priori del database da analizzare e riesce a rilevare tutte le tipologie di outliers eventualmente presenti. Questo approccio si basa sulla combinazione dei metodi tradizionali utilizzando un FIS [20], come rappresentato nella figura 5.

Per ciascun campione vengono calcolate quattro variabili:

- *Distanza di Mahalanobis* [21]: considerando possibili outliers quei campioni più lontani dal centro di massa della distribuzione totale.
- *Clustering:* si effettua una clusterizzazione tramite Fuzzy C Means. [22] e il numero ottimale di clusters è calcolato in

- automatico secondo il metodo proposto in [23].
- *Densità*: si calcola la densità dei dati attraverso il metodo Local Outliers Factor [24].
  - *Distribuzione*: si utilizza il metodo del Test di Grubbs [25] basato sulla distribuzione gaussiana.

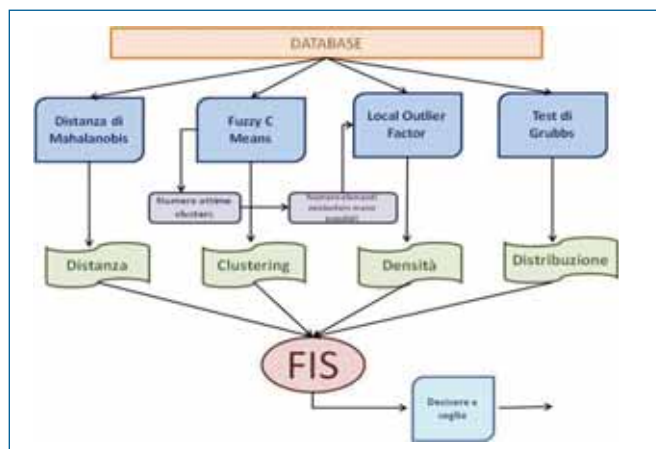


Figura 5 - Schema della generazione di una nuova popolazione tramite AG

Queste grandezze costituiscono l'input di un FIS che valuta il "grado di outlier" di ciascuna variabile. Infine viene utilizzato un decisore a soglia per estrarre gli outlier presenti nel database. Il metodo è completamente automatico, è quindi utilizzabile anche da utenti non esperti e presenta tutti i vantaggi relativi ai singoli metodi che sono combinati. Predizione caratteristiche acciai AI-killed e IF

La ricottura è un trattamento termico applicato all'acciaio durante le lavorazioni a freddo con lo scopo di migliorare la lavorabilità dell'acciaio, alleviare gli stress interni indotti da precedenti lavorazioni e affinare il grano ferritico. L'acciaio viene riscaldato, successivamente subisce il processo di "soaking" per un tempo sufficiente per permettere gli opportuni cambiamenti microstrutturali, dopo di che si ha un raffreddamento a velocità prefissata. Dal punto di vista termodinamico, durante la deformazione il metallo è in uno stato altamente energetico e di non equilibrio. Tende a ritornare all'equilibrio, se possibile, riducendo la sua energia libera tramite una riduzione della densità dei difetti attraverso il processo di ricristallizzazione.

Negli ultimi dieci anni sono stati sviluppati una gran quantità di modelli matematici atti a predire le proprietà meccaniche finali dell'acciaio laminato a freddo e sottoposto a ricottura (si veda ad esempio [26] [27]). La maggior parte di questi modelli sono basati sulla riproduzione del processo di ricristallizzazione e crescita del grano ferritico [28] e forniscono risultati affidabili per particolari e solita-

mente ristrette categorie di acciaio. Seguendo tale approccio, è stato sviluppato un modello matematico [29] per la predizione di alcune e proprietà meccaniche (carico di rottura e carico di snervamento) che impiega solamente dati di input disponibili on-line sulla linea di produzione, in modo che esso possa essere realmente impiegato sul campo evitando ulteriori analisi lunghe e complesse. Uno schema del modello è riportato nella figura 6(a). Tale modello contiene, nella parte che porta alla valutazione del diametro del grano ferritico, dei parametri interni per i quali si possono individuare con considerazioni fisiche dei range di variabilità, ma il cui valore esatto non è noto, anche perché le eventuali indicazioni che possono essere reperite a riguardo in letteratura si riferiscono a tipologie di acciaio diverse da quelle di interesse nel presente caso. Essendo il modello molto complesso (per esempio, alcuni moduli sfruttano algoritmi ricorsivi o contengono statement condizionali), l'impiego di tecniche standard di identificazione (come, ad esempio, quelle basate sul gradiente) risulta non praticabile. È stata pertanto organizzata una campagna di raccolta dati e di analisi, al fine di ottenere dati sperimentali per la ottimizzazione del valore dei parametri. Successivamente è stato applicato un AG che sfrutta come funzione di fitness una misura dell'errore fra i dati sperimentali ed i valori predetti. La figura 6(b) illustra i risultati ottenuti in termini di predizione delle dimensioni del grano ferritico, le quali si riflettono ovviamente in un considerevole miglioramento della predizione delle proprietà meccaniche finali.

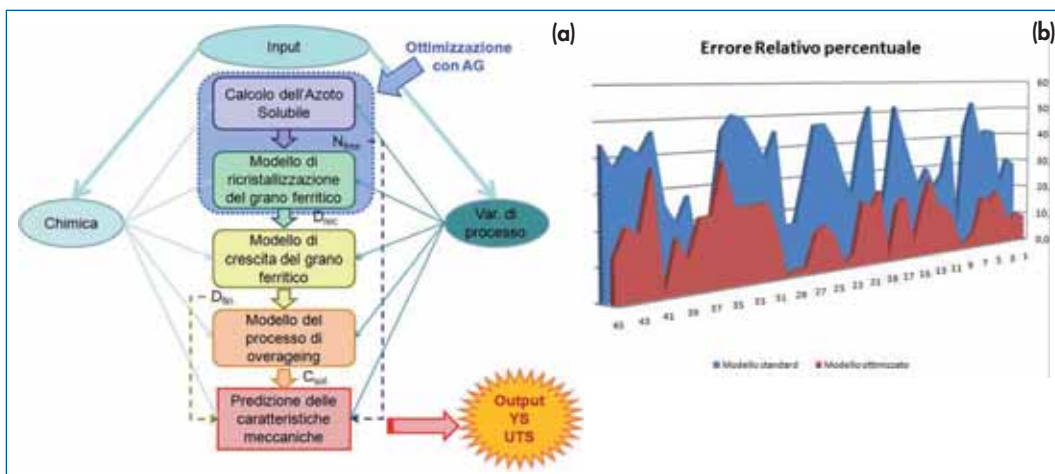


Figura 6 - Schema del modello matematico per la predizione delle proprietà meccaniche (a) e risultati ottenuti in termini di predizione del diametro del grano ferritico (b)

### Considerazioni conclusive

Le tecniche di Intelligenza artificiale costituiscono un potente strumento per la modellazione di fenomeni complessi e sui quali vi è scarsa conoscenza preconstituita ma è disponibile una grande quantità di dati sperimentali, ovvero nei quali la conoscenza pregressa esiste ma non può essere formalizzata in termini di modelli matematici. È questo un caso frequente nelle applicazioni industriali, in cui è possibile trovare personale tecnico molto esperto nella gestione di un determinato problema od impianto, ma che è in grado di trasferire tale esperienza in termini di regole comportamentali espres-

se in forma linguistica piuttosto che con un modello matematico. Per tale motivo, le tecniche di intelligenza artificiale trovano un vasto e sempre crescente impiego nelle applicazioni industriali e quindi anche nel settore siderurgico. Esistono tuttavia delle condizioni base per l'impiego delle tecniche di intelligenza artificiale: per impiegare le ANN, è necessario avere a disposizione dati sperimentali per l'addestramento, i quali devono essere in numero sufficiente ed opportunamente distribuiti nello spazio di ingresso affinché il risultato dell'addestramento della rete sia soddisfacente. In taluni casi particolari, come i problemi di classificazione con database "sbilanciati" (ossia in cui dati appartenenti ad una classe sono in numero molto minore rispetto alle altre classi) occorre sfruttare paradigmi particolari appositamente progettati, perché gli algoritmi classici non danno risultati accettabili.

Nel caso dei FIS, la disponibilità di dati non è strettamente necessaria purché si abbia una buona conoscenza dei meccanismi che regolano il problema in oggetto, in modo da poter definire correttamente sia gli insiemi fuzzy che la base delle regole, ma i dati sono sicuramente utili nel caso sia necessario effettuare un tuning fine di alcuni parametri interni al sistema.

Infine, l'eventuale conoscenza a priori del problema può (anzi deve) essere sfruttata nel progetto di un sistema che ambisca ad essere efficiente ed affidabile. Da qui nasce anche la possibilità di impiegare le tecniche di intelligenza artificiale unitamente e non in alternativa alle tecniche classiche della modellazione matematica e della statistica.

## Bibliografia

- [1] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Mc Millan College Publishing Company, New York, 1994.
- [2] Raul Rojas, *Neural Networks - A Systematic Introduction*, Springer-Verlag, Berlin, New-York, 1996.
- [3] T. Kohonen, "Self-Organizing Maps," *Springer Series in Information Sciences* Vol. 30, Springer, Berlin, Heidelberg, New York, 2001.
- [4] L.A. Zadeh, "Fuzzy set", *Information and Control*, Vol. 8, pp. 338-353, 1965.
- [5] E.H. Mamdani, S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic Controller", *International Journal of Man-machine Studies*, Vol. 7, pp. 113, 1975.
- [6] R.E. Bellman, L.A. Zadeh, "Decision-making in a fuzzy environment", *Management Science*, Vol. 17, No. 4, pp. 114-164, 1970.
- [7] H.X. Li, V.C. Yen, "Fuzzy Sets and Fuzzy Decision-Making", CRC Press, 1995.
- [8] J.S. R. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 23, No. 3, pp. 665-685, 1993.
- [9] L.M. Reyneri, "Unification of neural and wavelet networks and fuzzy systems," *IEEE Transactions on Neural Networks*, Vol 10, No. 4, pp. 801-814, July 1999.
- [10] J.H. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, 1975.
- [11] D.V. Doane, J.S. Kirkaldy, *Hardenability Concepts with Applications to Steel*, TMS-AIME, 1978.
- [12] W.G. Vermeulen, P.J. Van Der Wolk, A.P. De Weijer, S. Van Der Zwaag, "Prediction of Jominy Hardness Profiles of Steels Using Artificial Neural Networks," *Jour. Materials Eng. and Performances*, Vol. 5 No. 1, pp.57-63, 1996.
- [13] V. Colla, M. Sgarbi, M. Vannucci, R. Valentini, "Predizione dei profili di durezza Jominy di acciai microlegati al boro tramite reti neurali," *32° Convegno Nazionale dell'Associazione Italiana di Metallurgia*, Ferrara, 2008.
- [14] M.J. Heesom, "Physical and chemical aspects of nozzle blockage during continuous casting," *Proc. 1<sup>st</sup> Int. Calcium Treatment Sym.*, London, 1988
- [15] G.M. Faulring, J.W. Farrel, D.C. Hilty, "Steel flow through nozzles: influence of Calcium," *Ironmaking and Steelmaking*, February 1980.
- [16] V. Colla, M. Vannucci, S. Fera, R. Valentini, "Ca-treatment of Al-Killed Steels: inclusion modification and application of Artificial Neural Networks for the prediction of clogging", *5<sup>th</sup> European Oxygen Steelmaking Conference EOSC'06*, Aachen, pp. 387-394, 2006.
- [17] D. Hawkins, *Identification of outliers*, Chapman and Hall, London, 1980.
- [18] E.M. Knorr, R. Ng, "Algorithms for Mining Distance-Based Outliers in Large in Large Datasets," *Proceedings of VLDB*, pp. 392-403, 1988.
- [19] S. Cateni, V. Colla, M. Vannucci, "A fuzzy system for combining different outliers detection methods", *Int. Conf. on Artificial Intelligence & Applications AIA 2009*, Innsbruck (A), 2009.
- [20] S. Cateni, V. Colla, M. Vannucci, "Outlier detection methods for industrial applications," *Advances in Robotics, Automation and Control*, I-Tech Education and Publishing KG, pp. 265-282, 2008.
- [21] P.C. Mahalanobis, "On the generalized distance in statistics," *Proc. of the National Institute of Science of India*, pp.49-55, 1936.
- [22] D. Birant, A.v Kut, "Spatio-Temporal Detection in Large Databases," *28<sup>th</sup> Int. Conf. Information Technology Interfaces ITI 2006*, Croatia, 2006.
- [23] S. Ray, R.H. Turi, "Determination of number of clusters in k-means clustering and application in colour image segmentation," *4<sup>th</sup> Int. Conf. Advances in Pattern Recognition and Digital Techniques*, (ICAPRDT '99), Calcutta (india), pp. 137-43, 1999.
- [24] S. Papadimitriou, H. Kitawaga, P.B. Gibbons, C. Faloutsos, "LOCI: Fast Outlier Detection Using the Local correlation Integral," *19<sup>th</sup> Int. Conf. Data Eng. (ICDE)*, pp.315-326, 2002.
- [25] F.E. Grubbs: "Procedures for detecting outlying observations in samples" *Technometrics*, Vol. 11, pp. 1-21, 1969.
- [26] F.J. Humphreys, M. Hatherly, *Recovery after deformation in: Recrystallization and related annealing phenomena*, Pergamon Press, Oxford, 1996.
- [27] J. Kwiecinski, J.W. Wyrzykowski: "The effect of recovery annealing after small plastic deformations on the yield strength of polycrystalline aluminium," *Acta Metall. Mater.*, Vol. 41, No. 11, pp. 3089-3095, 1993.
- [28] M.M. Petite, A. Monsalve, I. Guiterrez, J. Zaitegui, J.I. Larburu, "Modelizacion de la evolution microestructural durante el recocido continuo de chapas de acero bajo en carbonio. Prediction de las propiedades mecanicas", *Rev. Metal.*, Vol. 34, No. 5, 1998.
- [29] V. Colla, R. Valentini, G. Bioli "Mechanical properties prediction for Aluminium-Killed and Interstitial-Free steels," *Revue de Metallurgie*, Special Issue JSI, pp. 100-101, Dic. 2004.