

MICHELE SCARDI*

Analisi dei dati e modelli predittivi: prospettive per la qualità totale del prodotto in acquacoltura

L'informazione che può essere acquisita al fine di rappresentare gli stati ed i processi essenziali di una filiera produttiva, in acquicoltura come in altri settori, costituisce un coacervo di dati inevitabilmente molto eterogeneo per fonti, codifica, accuratezza, aggregazione spaziale e/o temporale ed attendibilità. Inoltre, in molti casi l'informazione non è disponibile in forma quantitativa, ma è piuttosto rappresentata da descrittori tipo nominale, qualitativo o semi-quantitativo. In questo campo, dunque, a monte di ogni attività di analisi dei dati disponibili si rende necessaria una fase di pre-trattamento dei dati stessi, mirata soprattutto alla loro ricodifica e normalizzazione. È molto importante, in particolare, che quest'ultima sia attuata nelle forme più opportune al fine di rendere omogenee le scale delle diverse grandezze analizzate. Una particolare attenzione deve anche essere posta nella selezione e nella codifica dell'informazione derivata da giudizi esperti, che può giocare un ruolo di grande importanza in molte circostanze e soprattutto in quelle caratterizzate dalla mancanza o dalla scarsità di dati rilevati in maniera diretta. Inoltre, il giudizio esperto è essenziale ai fini della calibrazione di sistemi esperti mirati alla stima di giudizi di qualità. Una rigorosa analisi statistica delle relazioni che legano le grandezze disponibili, di qualunque natura esse siano, è di vitale importanza per selezionare un insieme di descrittori efficaci il più possibile compatto, escludendo ogni fonte di ridondanza e dunque semplificando al massimo l'acquisizione dell'informazione effettivamente rilevante. I dati disponibili, dopo essere stati opportunamente trattati e selezionati, possono quindi essere utilizzati al fine di mettere a punto un sistema esperto capace di associare l'informazione descrittiva di filiere e processi produttivi con quella relativa alla qualità dei pro-

* *Dipartimento di Biologia, Università di Roma "Tor Vergata"*

dotti, sia in termini nominali o qualitativi, sia, laddove possibile, in termini quantitativi. Questo obiettivo può essere perseguito selezionando i metodi di trattamento dei dati non solo fra le più appropriate metodiche statistiche mono- e multivariate, ma anche fra strumenti di nuova generazione mutuati dal campo dell'Intelligenza Artificiale e del *Machine Learning*. In particolare, ai fini della formulazione di previsioni relative ai giudizi di qualità, sono di particolare interesse soluzioni basate su reti neurali artificiali, *classification trees* o tecniche di *Case Based Reasoning*. Ovviamente, il metodo più appropriato può essere selezionato solo in funzione delle caratteristiche dell'informazione effettivamente disponibile. Le soluzioni così sviluppate possono facilmente essere restituite in forma comunque fruibile da un'utenza non tecnica, rendendo, laddove necessario, trasparenti all'utente finale gli algoritmi utilizzati mediante la realizzazione di interfacce *user-friendly*.