

## 2 Produzione di materiale plastico

Il settore manifatturiero riveste un ruolo di grande importanza nell'economia di ogni paese, e l'Italia non fa eccezione. Comprendere l'evoluzione storica di questo settore può fornire preziose informazioni per prendere decisioni informate sul futuro e sviluppare strategie efficaci.

Questa sezione del progetto si propone di analizzare la serie storica del settore manifatturiero dei materiali plastici in Italia, concentrandosi sull'applicazione di reti neurali artificiali per sviluppare previsioni. In particolare, il focus di questo lavoro sarà sull'indice di produzione nel settore dei materiali plastici in Italia. La serie storica analizzata coprirà un periodo che va dal 1° gennaio 2001 al 12 dicembre 2017, comprendendo le osservazioni sull'indice di produzione industriale registrato nel paese durante questo arco temporale.

L'indice di produzione industriale è un indicatore economico utilizzato per misurare la variazione della produzione nel settore manifatturiero, estrattivo e delle utility di un paese. Esso fornisce una misura quantitativa dell'attività produttiva delle industrie di un determinato paese o regione in un dato periodo di tempo. L'indice di produzione industriale nel settore dei materiali plastici è un indicatore specifico che misura la variazione della produzione nell'industria manifatturiera dei materiali plastici. Questo indice si concentra esclusivamente sulla produzione di beni e prodotti derivati dalla lavorazione delle materie plastiche, come imballaggi in plastica, tubi, film plastici, bottiglie, contenitori, componenti automobilistici in plastica e molti altri manufatti realizzati con questo materiale. L'obiettivo di questo indice è quello di monitorare e misurare l'attività produttiva specifica di settore, consentendo di valutare l'andamento e le tendenze di crescita o contrazione di questa industria in un determinato periodo di tempo.

### 2.1 La serie storica

La serie storica registra osservazioni relative al periodo che va dal 1° gennaio 2001 al 12 dicembre 2017, le osservazioni riguardano l'indice di produzione industriale registrato nel paese durante questo arco temporale.

In particolare la serie storica fornita dalla piattaforma [EuroStat](#), risulta univariata in quanto presenta osservazioni relative ad una sola variabile: l'indice di produzione manifatturiero dei materiali plastici. La serie presenta frequenza mensile e risulta equispaziata, dunque il campione si compone di 204 osservazioni. Il campione fornito dalla piattaforma Eurostat è stato coinvolto in una procedura di destagionalizzazione e rimozione degli effetti di calendario.

Di seguito viene mostrato il grafico della serie temporale:

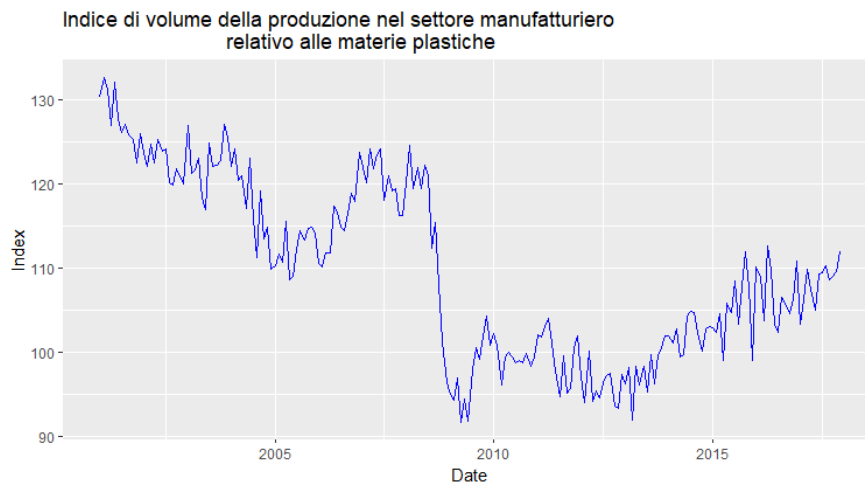


Figure 12

Come possiamo osservare dal grafico, l'indice di produzione nel settore dei materiali plastici ha mostrato nel corso del tempo diverse fluttuazioni. Dal 2001 al 2005 l'indice di produzione ha eviden-

ziato complessivamente un trend decrescente, dunque la produzione nel settore dei materiali plastici è diminuita nel corso di questi anni. Dal 2006 al 2008, l'andamento della produzione ha registrato una flessione positiva, indicando una ripresa nel settore. Durante gli anni della crisi finanziaria, il settore dei materiali plastici ha subito un impatto significativo, l'indice di produzione è calato a causa della contrazione economica generale, con una riduzione della domanda e una riduzione della produzione. Dopo gli anni della crisi economica, la produzione nel settore dei materiali plastici ha attraversato diverse fasi. Inizialmente, si è verificata una fase stagnante, con una produzione che ha mantenuto un livello costante. Successivamente, si è registrata una crescita, con l'indice di produzione che ha mostrato un trend positivo, indicando un aumento della produzione nel settore.

Chiaramente, essendo stata la serie sottoposta in precedenza a una procedura di destagionalizzazione e rimozione degli effetti di calendario, non sono presenti evidenze di stagionalità nella serie storica dell'indice di produzione nel settore dei materiali plastici.

Prima di procedere con la definizione e stima della rete neurale è stata valutata la non linearità del processo generatore dei dati attraverso il test di linearità di White e quello di Terasvistra.

In entrambi i test le ipotesi sono:

- $H_0$ : il processo generatore dei dati è lineare
- $H_1$ : il processo generatore dei dati non è lineare

I test confermano la non linearità del processo che ha generato la serie temporale, in particolare con un livello di  $\alpha$  fissato al 5% e considerando fino al terzo ritardo temporale.

## 2.2 Modellazione della serie

In molte applicazioni della realtà il processo generatore delle serie temporali risulta molto complesso e parzialmente/totalmente sconosciuto. In questo contesto possono essere molto utili le reti neurali, in quanto nell'applicazione pratica non richiedono precise specificazioni della natura delle relazioni che intercorre tra le variabili indipendenti e l'outcome. Chiaramente ciò è possibile tramite l'Universal Approximation Theorems che permettono di approssimare bene funzioni estremamente complesse come quella generatrice dei dati. La Single layer feed-forward neural network è la rete più utilizzata per la modellazione e previsioni di serie temporali. Il modello è composto da tre strati di unità di elaborazione collegate aciclicamente in cui le informazioni dallo strato di input confluiscono allo strato di output, passando per tramite lo strato "hidden".

La relazione tra l'output  $Y_t$  e gli input  $y_{t-1}$  è data da :

$$Y_t = f_r(y_{t-1}, \theta) + \epsilon$$

$$= c_0 + \sum_{k=1}^r c_k \psi \left( \sum_{j=1}^p w_{kj} Y_{t-j} + w_{k0} \right) + \epsilon \quad \text{dove } y_{t-1} = Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_{t-p}$$

I termini  $w_{kj}$  e  $c_k$  sono i parametri del modello,  $p$  è il numero di nodi di input,  $r$  il numero di nodi che compongono il layer nascosto.

Prima di procedere con la stima dei modelli, i dati sono stati processati ed in particolare sono stati normalizzati attraverso la tecnica di standardizzazione.

Successivamente è stata svolta l'identificazione del modello, ovvero identificare il modello che riesce a stimare al meglio l'indice di produzione industriale nel settore delle materie plastiche. I parametri fondamentali da identificare per trovare l'architettura ottima dell'ANN sono il numero di nodi di input  $p$ , ovvero il numero di osservazioni passate che verranno utilizzate come valori di input, ed il numero di nodi  $q$  nell'hidden layer. Diversi approcci sono presenti in letteratura per la stima del numero di nodi  $p$  e  $q$ , nonostante ciò nessuno di questi metodi però garantisce soluzione ottimale per tutti i problemi di previsioni reali.

La procedura impiegata solitamente è quindi quella di testare numerose reti aventi differenti configurazioni per input e dimensione del layer intermedio, stimare l'errore di previsione per ciascuno di essi, valutarlo e selezionare la rete con performance migliori. Fatta questa premessa è stata definita una griglia di valori  $p$  e  $q$ , il numero di nodi di input da valutare è stato fissato nell'intervallo  $1 \leq p \leq 7$

mentre la dimensione dell'hidden layer è stato fissato nell'intervallo  $3 \geq q \leq 7$ .

Le 35 reti neurali stimate sono state valutate per la loro capacità previsiva one-step ahead, attraverso la Cross-validation "rolling window" tecnica che permette di effettuare una valutazione iterativa del modello utilizzando finestre di dati sempre più grandi. Questo significa che i modelli vengono addestrati su un subset iniziale dei dati e vengono quindi testati sulla prima osservazione immediatamente successiva all'ultima presente nel set di addestramento.

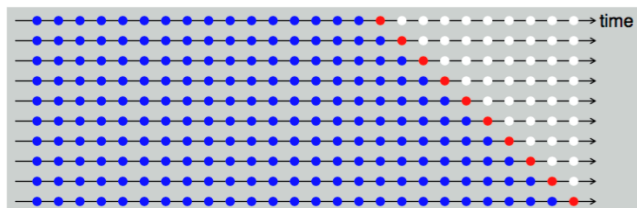


Figure 13

Il processo viene ripetuto iterativamente incrementando il set del train fino a coprire l'intero set di dati disponibili. La figura mostrata in precedenza rappresenta un esempio di Cv-rolling window. Le performance dei modelli sono state valutate in termini di Root Mean Squared Error(RMSE) e Mean Absolute Error(MAE). In particolare la selezione del modello più accurato per prevedere l'indice di produzione ha considerato la minimizzazione delle metriche prese in considerazione.

p	q	RMSE	MAE
3	4	4.080991	3.140689
4	3	4.527530	3.255614
5	3	4.526797	3.274444

Figure 14

La figura mostra la composizione delle reti neurali che minimizzano le metriche considerate: si osservano performance migliori sia in termini di MAE che di RMSE nella rete neurale che prende in input osservazioni fino al terzo lag con la dimensione del layer nascosto pari a quattro neuroni.

Il modello selezionato risulta avere 3 nodi nello strato dedicato all'input e 4 nodi nello strato intermedio, l'immagine successiva mostra graficamente le previsioni one-step ahead del modello ottenute durante la fase di CV-rolling window.

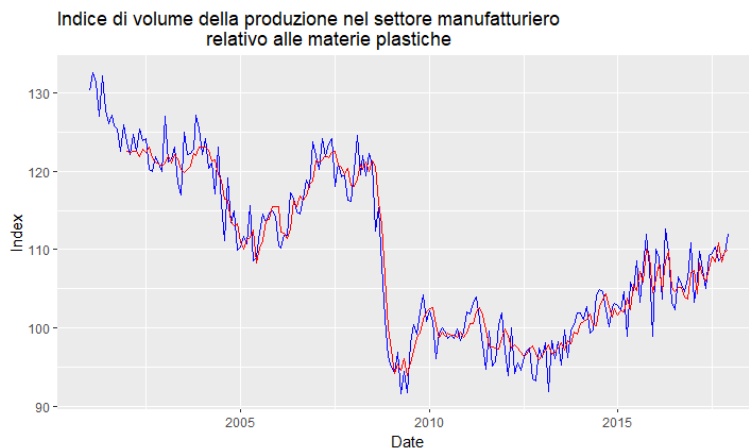


Figure 15

Dalla figura si nota che il modello (linea rossa) in alcuni punti tenda a sovrastimare l'indice e in altri punti lo sottostima, inoltre sebbene non in modo eccessivo in determinati livelli della serie il modello presenta problemi di overfitting.

Dopo la selezione dell'architettura della rete neurale, il modello è stato ristimato utilizzando i dati disponibili dal 1° gennaio 2005 al 12 dicembre 2016. Successivamente, il modello addestrato è stato utilizzato per fare previsioni sull'andamento dell'indice di produzione per l'anno 2017. Le previsioni ottenute sono state confrontate con le osservazioni effettive dell'indice nel corso di quell'anno, allo scopo di valutare l'accuratezza e l'efficacia del modello nel fare previsioni. Di seguito viene mostrato il grafico che confronta le previsioni ottenute dal modello con le vere osservazioni dell'indice di produzione nel corso dell'anno 2017.

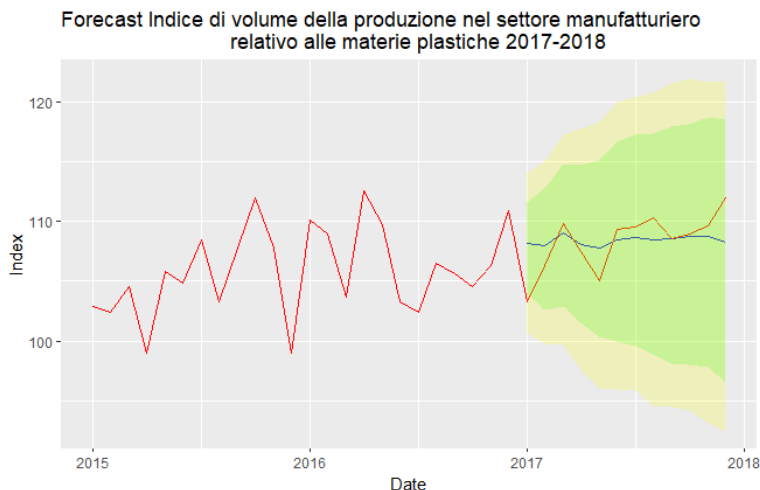


Figure 16

Il forecasting della serie storica sviluppato tramite la rete neurale prevede abbastanza bene il valore reale dell'indice di produzione. Nel plot vengono rappresentati anche gli intervalli di previsione al 95% e all'80%, che sono stati calcolati utilizzando la tecnica di simulazione tramite bootstrap dei residui .