

**UNIVERSITÀ DELLA VALLE D'AOSTA  
UNIVERSITÉ DE LA VALLÉE D'AOSTE**

**DIPARTIMENTO DI SCIENZE ECONOMICHE E POLITICHE**

**CORSO DI LAUREA MAGISTRALE IN ECONOMIA E POLITICHE DEL  
TERRITORIO E DELL'IMPRESA**

**ANNO ACCADEMICO 2021/2022**

**TESI DI LAUREA**

**PREVEDERE LE VENDITE A SEGUITO DI MUTAMENTI  
TECNOLOGICI E DI PREFERENZE DEI CONSUMATORI:  
IL CASO DEL MERCATO DEI TELEVISORI**

**DOCENTE:** Prof. Claudio Brenna

**STUDENTE:** Désiré Bus  
20G01226



## RINGRAZIAMENTI

Prima di procedere con la trattazione, vorrei dedicare qualche riga a tutti coloro che mi sono stati vicini in questo percorso di crescita personale e professionale.

Un sentito grazie al mio relatore Brenna Claudio per la sua disponibilità e professionalità. Grazie per avermi fornito materiale utile alla stesura dell'elaborato e per avermi trasmesso la passione per la sua materia di insegnamento.

Senza il supporto morale dei miei genitori, in particolare di mia madre non avrei mai potuto raggiungere questo obiettivo. Grazie per esserci sempre stata anche nei momenti più difficili.

Mi è doveroso ringraziare Lorenzo, Pier Giorgio e Luigi per la loro disponibilità e per aver creduto in questa collaborazione tra azienda e studente che ha permesso di arricchire in modo significativo il presente elaborato.

Ringrazio alcuni colleghi e colleghe Alice, Davide, Elena, Simone, Susanna, Ilaria e Daria per il supporto morale e la loro collaborazione durante i periodi più intensi dovuti non solo allo studio, ma anche alla pandemia e per aver gioito insieme a me per i traguardi raggiunti durante il percorso accademico.



# INDICE

RINGRAZIAMENTI.....	3
INTRODUZIONE .....	7
1. SALES FORECASTING .....	1
1.1. Evoluzione.....	1
1.2. Scopo ed implicazioni del previsore .....	6
1.3. Elaborazione delle previsioni di vendita .....	11
1.4. Previsione, pianificazione e definizione degli obiettivi .....	12
1.5. Tecniche di previsione .....	14
2. METODI QUANTITATIVI.....	17
2.1. Big Data e previsioni.....	17
2.2. Le serie storiche .....	19
2.3. Processi stocastici e autocorrelazione .....	22
2.4. Il modello SARIMA.....	25
2.5. Le reti neurali .....	28
2.6. Long Short Term Memory .....	33
2.7. Confronto e valutazione dell'errore .....	38
3. Il CASO AZIENDALE .....	40
3.1. Analisi del settore.....	40
3.2. I dati .....	42
3.3. Confronto ed applicazione .....	44
3.4. Conclusioni .....	57
Appendice .....	58

INDICE DELLE FIGURE.....	59
BIBLIOGRAFIA .....	61

# INTRODUZIONE

Le previsioni de vendite sono uno strumento di fondamentale importanza per prevedere il fatturato futuro derivante dalle vendite che è di vitale importanza in quanto si ripercuote su tutti i settori aziendali. Il presente elaborato affronta il tema delle previsioni di vendita in una particolare situazione di variabilità dei dati attraverso l'applicazione di differenti metodi quantitativi, utilizzando in particolare i modelli SARIMA e SARIMAX (Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors), messi a confronto con un metodo denominato LSTM (Long Short Term Memory). L'elaborato è stato suddiviso in tre capitoli. Nel primo capitolo è esposta la letteratura delle previsioni delle vendite riguardante in particolare gli studi precedenti di tale materia, le preferenze delle imprese per quanto riguarda le metodologie applicate e infine lo scopo e le numerose interazioni che il soggetto previsore ha all'interno dell'ambiente aziendale. È stato inoltre preso in considerazione il processo di elaborazione delle previsioni che è stato illustrato attraverso un diagramma di flusso. Inoltre, sono state analizzate alcune problematiche sorte durante tale processo. In ultimo vengono illustrate le varie tecniche di previsione delle vendite fino ad oggi utilizzate.

Nella prima parte del secondo capitolo vengono illustrati i vantaggi che i Big Data permettono di apportare alle tecniche di previsione ed in seguito sono esposti a livello teorico le serie storiche ed i metodi di analisi utilizzati per la previsione delle vendite (SARIMA, SARIMAX, LSTM), includendo alcune metodologie di misurazione dell'errore.

Nel terzo ed ultimo capitolo vi è l'applicazione dei metodi citati in precedenza alle serie storiche (circa 40.000 osservazioni per serie) inerenti alla vendita di alcune tipologie di televisori.

È necessario specificare che i dati analizzati sono reali, recenti e che derivano da una collaborazione tra azienda e studente.

# 1. SALES FORECASTING

La traduzione del termine inglese sales forecasting si identifica nella lingua italiana come la previsione delle vendite. In generale è l'insieme delle metodologie che permettono di individuare con maggior precisione possibile, il volume di vendite per un determinato periodo futuro. Una definizione più puntuale è data dall'American Marketing Association: "La previsione delle vendite è una stima delle vendite, in unità monetarie o fisiche, per un determinato periodo futuro nell'ambito di un piano aziendale o di un programma proposto e sotto un presunto insieme di forze economiche e di altro tipo al di fuori dell'unità per la quale viene fatta la previsione".

## 1.1. Evoluzione

Le tecniche di previsione si suddividono in due categorie: qualitative e quantitative. La maggior parte delle tecniche e dei modelli per cercare di eseguire delle previsioni è iniziata ad emergere negli anni 60. Tutti gli scritti che sono stati prodotti in questi anni si concentrano perlopiù sulla ricerca del metodo di previsione più appropriato. In breve tempo prevedere le vendite è diventato uno degli aspetti cruciali della pianificazione aziendale.

Sono stati condotti alcuni studi che dimostrano come su 175 uomini d'affari del Midwest il 65% pensa che prevedere le vendite è molto importante per il successo dell'attività imprenditoriale, mentre il 28% afferma che le previsioni di vendita sono importanti ma non indispensabili<sup>1</sup>. Gli studi che hanno confrontato le tecniche qualitative e quantitative utilizzate per la previsione delle vendite sono

---

<sup>1</sup> DALRYMPLE, Douglas J. Sales forecasting methods and accuracy. Business Horizons, 1975, 18.6: 69-73.

rari, ma nel 1976 Mabert ha rilevato che la previsione dei tre anni futuri è più accurata attraverso i metodi di smorzamento esponenziale, di Box-Jenkins e di media armonica, delle previsioni eseguite attraverso il giudizio del reparto vendite; tuttavia la differenza è minima.

Allo stesso modo nel 1979 Walton illustra che la previsione delle vendite di tre prodotti industriali eseguita attraverso la regressione lineare ha sovraperformato il parere del reparto vendite. Le dieci più grandi aziende dell'Australia, inoltre, hanno abbandonato i modelli statistici di previsione dopo all'incirca un anno di utilizzo. Le previsioni eseguite mediante un giudizio soggettivo e quelle computate attraverso il computer hanno come base di partenza una serie storica di dati. Il confronto mostra che il giudizio è accurato tanto più il metodo o la tecnica quantitativa utilizzata è puntuale. Tuttavia, in entrambi i casi, la previsione è basata solamente sulle serie temporali ed il previsore non disponeva di informazioni aggiuntive<sup>2</sup>.

Prevedere le vendite è di fondamentale importanza per la programmazione della produzione e per il rifornimento del magazzino. Delle previsioni imprecise possono significare scorte in eccesso o merce invenduta. Non sorprende quindi che in Mahmoud et al. (1988) e Dalrymple (1987) vi sono degli studi che indicano l'accuratezza come il criterio più importante nella scelta di una strategia di previsione adeguata. L'accuratezza delle previsioni dovrebbe migliorare all'aumentare del numero di informazioni incluse nel processo di previsione. Vi sono però delle prove che indicano una relazione inversa tra una maggiore disponibilità di informazioni ed una diminuzione dell'accuratezza del giudizio del previsore facendone però aumentare la fiducia. D'altro canto questa relazione nasce dal fatto che l'essere umano ha delle capacità cognitive limitate e all'aumentare del numero di informazioni vi è una diminuzione della capacità di comprendere.

---

<sup>2</sup> Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. F. (1993). The M2-competition: A real-time judgmentally based forecasting study. *International Journal of forecasting*, 9(1), 5-22.

Le tecniche di previsione hanno subito un'evoluzione continua in concomitanza allo sviluppo di nuovi strumenti, come computer e software, che permettono di calcolare, utilizzando metodi statistici avanzati, delle previsioni sempre più puntuali. Contrariamente a quanto ci si potrebbe aspettare, vista l'ampia disponibilità di software, molti studi empirici mostrano come il giudizio soggettivo sia il metodo preferito per eseguire delle previsioni perché le suddette vengono percepite come precise e facili da utilizzare<sup>3</sup>.

Ci si interroga quindi, se il giudizio della direzione aziendale fornisca effettivamente stime previsionali accurate. Uno studio di ricerca sull'accuratezza delle previsioni attraverso il giudizio dei manager mostra risultati contrastanti<sup>4</sup>. Il divario si amplia quando il previsore non può accedere alle consuete conoscenze che potrebbero portare ad una sostanziale accuratezza delle previsioni (dati macroeconomici, dati dei concorrenti, piani di promozione e altre conoscenze contestuali), ma anche quando il previsore è soggetto alla pressione di definire degli obiettivi futuri<sup>5</sup>.

Vi sono altri studi che si concentrano sulla previsione attraverso un giudizio soggettivo nella vita reale e la maggior parte ha coinvolto degli analisti della sicurezza<sup>6 7</sup>. In questo caso si vuole dimostrare che la previsione attraverso un giudizio soggettivo è distorta e inefficiente, in quanto la causa è sistematica. I parametri possono essere stimati dai periodi precedenti per eliminare una parte sostanziale dell'errore sistematico nelle previsioni attuali, migliorando così l'accuratezza e anche l'errore non sistematico può essere ridotto. In questo modo si

---

<sup>3</sup> Diamantopoulos, A., & Winklhofer, H. (2003). Export sales forecasting by UK firms: Technique utilization and impact on forecast accuracy. *Journal of Business Research*, 56(1), 45-54.

<sup>4</sup> Webby, R., & O'Connor, M. (1996). Judgemental and statistical time series forecasting: a review of the literature. *International Journal of forecasting*, 12(1), 91-118.

<sup>5</sup> Winklhofer, H., Diamantopoulos, A., & Witt, S. F. (1996). Forecasting practice: A review of the empirical literature and an agenda for future research. *International Journal of forecasting*, 12(2), 193-221.

<sup>6</sup> Ali, A., Klein, A., & Rosenfeld, J. (1992). Analysts' use of information about permanent and transitory earnings components in forecasting annual EPS.

<sup>6</sup> *Accounting Review*, 183-198. Mendenhall, R. R. (1991). Evidence on the possible underweighting of earnings-related information. *Journal of Accounting Research*, 29(1), 170-179.

afferma che una stima previsionale supportata da modelli può effettivamente essere più precisa di una previsione eseguita attraverso un giudizio soggettivo<sup>8</sup>.

Più di recente negli Stati Uniti si è rimarcato che su 240 imprese solamente il 10.8% utilizza dei software al fine di eseguire delle previsioni di vendita e di questi il 61,2% indica di effettuare delle modifiche regolari alle previsioni generate dal software, in base al parere del management. Si rileva inoltre che il 41,1% esprime una bassa dipendenza dalle previsioni generate automaticamente. Alla domanda su come compensare gli eventi imprevisi il 69.9% delle imprese ha risposto sostenendo di affidarsi al giudizio soggettivo del management per generare una nuova previsione, mentre il 21,4% delle imprese usa il giudizio personale del management per correggerle. Solamente l'8,7% delle imprese attua una modifica ai parametri del modello.

La precisione di stima previsionale data dai modelli quantitativi e qualitativi è ancora oggi fonte di dibattito e ricerca. Alcuni studi dimostrano che utilizzando metodi di ricerca quantitativi si hanno delle previsioni più affidabili, mentre altri sostengono il contrario. Non vi è quindi una modalità chiaramente delineata dalla letteratura per scegliere un metodo quantitativo o qualitativo, tantomeno non è specificato quale modello all'interno di queste due categorie utilizzare. La scelta del metodo previsionale sembra quindi essere legata alle peculiarità del contesto di applicazione. Si possono però individuare, al fine di prevedere le vendite, tre fattori che possono determinare la prevalenza nell'applicazione di un approccio quantitativo<sup>9</sup>:

- 1) il possibile accesso alle serie storiche dei dati di vendita;
- 2) il possibile utilizzo di software adeguati alla finalità di previsione;
- 3) il grado di incertezza del settore d'impresa.

---

<sup>8</sup> Elgers, P. T., Lo, M. H., & Murray, D. (1995). Note on adjustments to analysts' earnings forecasts based upon systematic cross-sectional components of prior-period errors. *Management Science*, 41(8), 1392-1396.

<sup>9</sup> Sanders, N. R., & Manrodt, K. B. (2003). Forecasting software in practice: Use, satisfaction, and performance. *Interfaces*, 33(5), 90-93.

Il grado di incertezza del settore è il punto su cui fare affidamento per scegliere il metodo di analisi nel modo più appropriato. Si riscontra che vi è una relazione tra la variabilità della domanda e l'accuratezza dei modelli di previsione. La variabilità della domanda viene misurata attraverso il coefficiente di variazione delle vendite che è il rapporto tra la deviazione standard delle vendite di un prodotto ed il valore medio delle vendite in un dato periodo di tempo. Si ritiene che quando il trend delle vendite di un prodotto è stabile, e cioè non ha delle variazioni significative, la scelta deve ricadere sull'utilizzo di modelli quantitativi. Infatti, quando il coefficiente di variazione delle vendite ha dei valori bassi, le tecniche quantitative hanno un'accuratezza migliore<sup>10</sup>. Nelle serie storiche delle vendite di prodotti ad alta variabilità l'aggiunta di informazioni qualitative aumenta l'accuratezza della previsione. In questa categoria di prodotti la tecnica più appropriata è quella qualitativa quando il previsore possiede delle informazioni (competenze settoriali e informazioni contestuali) che sono in grado di spiegare la variabilità delle vendite e che non sono prese in considerazione dal modello quantitativo. Le imprese utilizzano anche metodi di previsione che prevedono l'utilizzo di entrambe le tecniche, sia quantitative che qualitative. Si identificano quattro modalità per le quali le tecniche qualitative si possono abbinare:

- 1) le previsioni attraverso metodi qualitativi vengono corrette attraverso le previsioni quantitative;
- 2) le previsioni di vendita attraverso metodi quantitativi vengono corrette attraverso le valutazioni soggettive del previsore;
- 3) si utilizzano le due tecniche separatamente ed il risultato viene assemblato per creare le previsioni;
- 4) il modello quantitativo viene creato sulla base delle valutazioni qualitative.

---

<sup>10</sup> Sanders, N. R., & Ritzman, L. P. (1995). Bringing judgment into combination forecasts. *Journal of Operations Management*, 13(4), 311-321.

La combinazione delle due differenti tecniche è più efficace quando la correlazione degli errori non è alta e quindi, apportando nelle due tecniche delle informazioni differenti<sup>11</sup>.

## **1.2. Scopo ed implicazioni del previsore**

Per eseguire delle previsioni di vendita è doveroso comprendere quali mansioni svolge il soggetto incaricato di tale compito. Il previsore o il gruppo di soggetti addetti alle previsioni non ha solamente il compito di eseguire delle previsioni di vendita orientate alla pianificazione del marketing e del reparto sales, ma interagisce con tutti o quasi tutti i reparti e le funzioni aziendali.

In molte imprese le previsioni di vendita sono parte integrante dell'incontro tra domanda e offerta e contribuiscono alla realizzazione del Sales and Operations Planning (S&OP). Il Sales and Operations Planning è il piano che permette al management di dirigere strategicamente il proprio business, al fine di raggiungere il vantaggio competitivo desiderato, attraverso una reale integrazione dei piani commerciali con la gestione dell'intera supply chain. Quindi, in altre parole questo piano guida l'intero processo di pianificazione della produzione e dell'approvvigionamento, rappresentando l'anello di giunzione tra il budget aziendale e la programmazione operativa che coinvolge tutta la direzione aziendale. Tale processo assicura che il piano aziendale contribuisca effettivamente al raggiungimento degli obiettivi aziendali.

Nella Figura 1 sotto riportata si osserva il previsore che interagisce a livello organizzativo con le diverse funzioni aziendali: marketing, produzione, controllo di gestione e management.

---

<sup>11</sup> Sanders, N. R., & Ritzman, L. P. (2004). Integrating judgmental and quantitative forecasts: methodologies for pooling marketing and operations information. *International Journal of Operations & Production Management*.

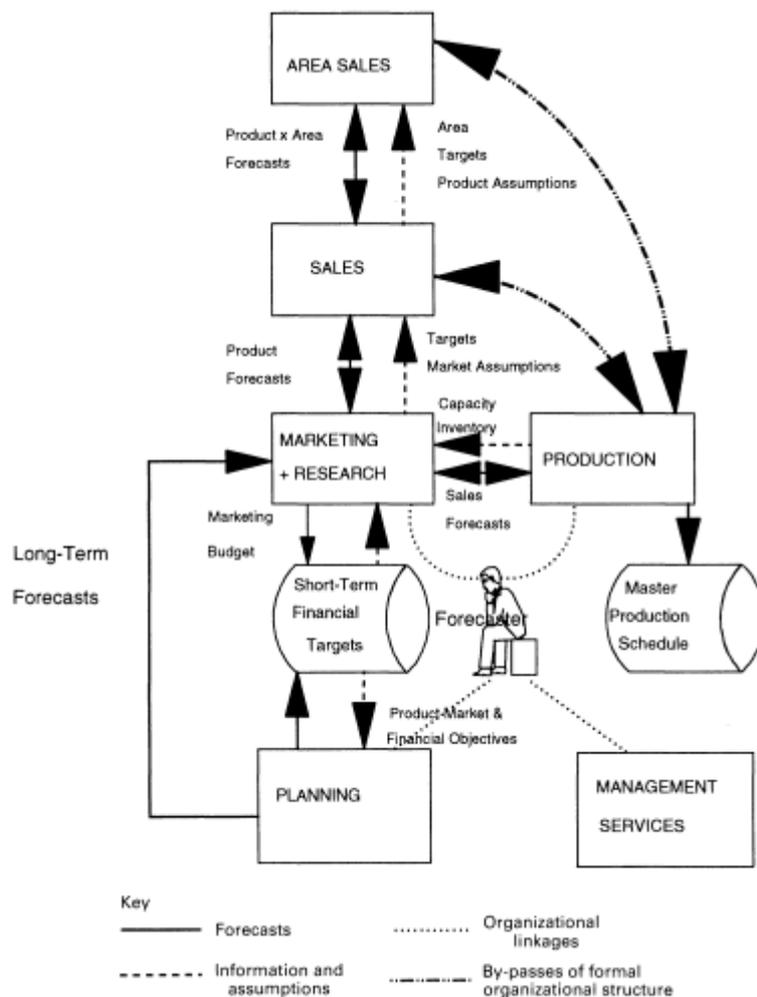


Figura 1: il posizionamento e le interazioni della figura del previsore all'interno di una impresa

Uno schema organizzativo differente potrebbe prendere in considerazione o ignorare dei dati oppure avere altri flussi di informazioni intra-organizzativi. Il diagramma mira a cogliere gran parte della complessità di qualsiasi sistema di previsione industriale anche se, in altre tipologie di imprese o settori, alcuni dei collegamenti ed alcune aree potrebbero non esistere. Il reparto vendite è responsabile della generazione e del mantenimento della domanda da parte di grandi clienti finali, di partner di canali all'ingrosso o al dettaglio, mentre il marketing è responsabile della domanda dei consumatori finali. Dell'approvvigionamento se ne occupano una serie di reparti: produzione, amministrazione, logistica o distribuzione, risorse umane e controllo di gestione.

D'altro canto, di tale compito ne sono complici una varietà di fornitori che devono per l'appunto fornire materie prime, componenti, prodotti finiti e imballaggi.

Il processo S&OP è una "scatola di giunzione" in cui le informazioni di una impresa possono fluire tra il lato della domanda e il lato dell'offerta. Come da Figura 1 l'input critico per il processo S&OP è la previsione di vendita, che come spiegato pocanzi, è la previsione della domanda futura.

La previsione di vendita proviene dal lato della domanda dell'impresa perché i reparti vendite e marketing sono responsabili della generazione della domanda e di avere una prospettiva ottimale su quale sarà la domanda futura ed è consuetudine che il previsore o gli addetti alla previsione siano inseriti all'interno di questi reparti. Oltre alle previsioni di vendita che provengono dal lato della domanda, un altro fattore critico è la pianificazione dell'approvvigionamento. Questo input è fornito dal lato dell'offerta dell'impresa e documenta le capacità di approvvigionamento nel lungo e nel breve termine. Il processo che si verifica all'interno del S&OP è la corrispondenza tra le proiezioni della domanda futura e cioè le previsioni di vendita, con le proiezioni dell'offerta futura ovvero, il piano della capacità di approvvigionamento. Il processo coinvolge il reparto vendite e il marketing che devono pianificare quello che dovrà essere venduto e commercializzato, tenendo in considerazione il fattore tempo e le capacità di fornitura dell'impresa. La pianificazione della domanda può comportare il ritiro dal commercio di uno o più prodotti a causa dei vincoli di capacità di approvvigionamento oppure può avere un impatto sullo spostamento della domanda da prodotti a basso margine ad alto margine di profitto. Il reparto marketing è responsabile di redigere dei piani annuali aggiornati mensilmente o trimestralmente con l'obiettivo del successo dei singoli prodotti o delle linee di prodotto che commercia. Questa programmazione coinvolge le modifiche del prodotto, gli sforzi promozionali, i canali di vendita e le politiche di prezzo. Al fine di sviluppare la pianificazione il marketing ha bisogno che le previsioni di vendita, tipicamente espresse nella valuta corrente, tengano conto di questi sforzi e dovranno essere eseguite a livello di prodotto, di linea di prodotto per un orizzonte temporale annuale, con aggiornamenti mensili o trimestrali.

Le esigenze del reparto vendite differiscono dal reparto marketing. Solitamente il reparto vendite richiede delle previsioni di vendita con un orizzonte di uno o due anni aggiornati mensilmente o trimestralmente, ma con la particolarità che l'aggregazione dei dati deve essere effettuata a livello territoriale. Tuttavia, l'organizzazione della modalità di vendita può essere definita geograficamente, per settore, per cliente, per prodotto, ed è questa aggregazione che aiuta a far comprendere al previsore quanto dettagliate dovranno essere le previsioni. L'orizzonte temporale è di un anno ma le previsioni devono essere aggiornate sulla base della tempistica del pagamento delle commissioni ai venditori. Se le commissioni vengono pagate trimestralmente le previsioni dovranno essere aggiornate ogni quattro mesi.

Il reparto del controllo di gestione e contabilità è un altro reparto che viene a contatto con il previsore. Il controllo di gestione ha degli orizzonti temporali che seguono il ciclo dell'analisi dei costi, profitti e del fabbisogno di capitale. Il piano di profitto che viene redatto da questo reparto è solitamente annuale ma l'orizzonte temporale può protrarsi fino a cinque anni e viene aggiornato trimestralmente o mensilmente. Le previsioni di vendita vengono aggregate a livello di linea di prodotto, reparto o all'impresa e l'orizzonte temporale come sopracitato è da uno a cinque anni con un aggiornamento coerente con la pianificazione del reparto in esame.

Un altro reparto in cui il previsore ha una evidente rilevanza, riguarda il soddisfacimento della produzione e l'acquisto di materie prime o prodotti finiti. In questo caso dovranno essere prodotte due tipi di previsioni: la prima a lungo termine, mentre la seconda a brevissimo termine. La previsione a lungo termine viene utilizzata per pianificare lo sviluppo di fornitori, impianti e attrezzature che possono richiedere diversi anni. Poiché questi piani a lungo termine dipendono dal mix di prodotti da effettuare nello stabilimento, la previsione deve essere a livello di singolo prodotto. L'orizzonte temporale dipende dal tempo necessario per acquisire nuovi fornitori, impianti ed attrezzature e quindi può variare da uno a tre anni, mentre l'aggiornamento di queste previsioni è tipicamente trimestrale. In contrasto con ciò, la previsione a breve termine si basa sulle esigenze del

programma di pianificazione della produzione, che può variare da uno a sei mesi (a seconda del ciclo di acquisto trasformazione e vendita delle materie prime) e richiede un dettaglio specifico dei prodotti. Pertanto, questa previsione per il reparto produzione è a livello di unità di prodotto, ha un orizzonte raramente superiore a sei mesi e aggiornamenti da giornalieri a mensili.

Coerentemente con il ciclo produttivo o di approvvigionamento, le esigenze previsionali della logistica sono inanellate ad una duplice tempistica. Nel lungo termine vi è la pianificazione degli impianti di stoccaggio in vari luoghi e quindi, le previsioni avranno come unità di misura il numero di prodotti, i metri cubi e le attrezzature necessarie per spostare i prodotti tra le strutture (previsioni in quintali). Ancora una volta l'orizzonte temporale è determinato dal tempo necessario per poter avere le strutture di stoccaggio pronte all'utilizzo. Una grande azienda chimica, ad esempio, ha bisogno di un orizzonte di pianificazione di 18 mesi per appaltare la costruzione di nuovi vagoni ferroviari al fine di spostare i suoi vari prodotti.

Tra le aziende, questi orizzonti a lungo termine possono variare da pochi mesi per le strutture in affitto o il trasporto tramite intermediari, a diversi anni per strutture personalizzate o attrezzature di trasporto costruite appositamente per l'azienda. Poiché entrambe le tipologie sono spesso utilizzate, l'intervallo è tipicamente mensile. Nel breve termine bisogna decidere quali prodotti spostare ed in che luogo, ma soprattutto quando bisogna spostarli. L'unità di misura utilizzata è il numero di prodotti, il metro cubo o il peso. L'orizzonte temporale sarà breve in quanto le tempistiche sono definite dal ciclo delle materie prime e semilavorati spostati da un impianto all'altro. Solitamente si parla di mesi, settimane o addirittura giorni, in base al ciclo produttivo applicato e gli aggiornamenti sono i medesimi<sup>12</sup>.

---

<sup>12</sup> Fildes, R., & Hastings, R. (1994). The organization and improvement of market forecasting. *Journal of the Operational Research Society*, 45(1), 1-16.

### 1.3. Elaborazione delle previsioni di vendita



Figura 2: trasmissione delle previsioni di vendita

Un esempio per comprendere meglio l'elaborazione delle previsioni di vendita e la loro trasmissione è dato dalla Figura 2. Il processo inizia con la creazione, attraverso il download della serie storica delle vendite, di un modello di previsione generato al computer. Le previsioni di vendita sono suddivise per prodotto, territorio geografico o settore e sono inviate tramite posta elettronica al reparto vendita. Ogni addetto alla vendita è chiamato a leggere e ad apportare eventuali modifiche o suggerimenti alle previsioni di vendita. Dopo aver eseguito

le appropriate osservazioni il reparto sales deve inviare la risposta al previsore che dovrà apportare le eventuali modifiche e successivamente inviare l'elaborato inerente al mercato o prodotto in esame al marketing manager. Quest'ultimo dovrà esaminare le proiezioni e applicare eventuali rettifiche o osservazioni e rinviarle al previsore. Dopo aver preso in considerazione le osservazioni del marketing manager il previsore deve inviare al management o alla direzione aziendale l'elaborato che dovrà apportare delle verifiche per ogni livello di divisione/reparto. Il previsore dopo aver ricevuto una risposta dal management e apportato le eventuali modifiche, dovrà inoltrare l'elaborato ad ogni livello ed in funzione a ciascuna area di pianificazione aziendale<sup>13</sup>.

#### **1.4. Previsione, pianificazione e definizione degli obiettivi**

Ogni giorno le persone eseguono delle previsioni, per esempio, se prendono un ombrello vuol dire che probabilmente potrebbe piovere o se pensano che farà caldo si metteranno una t-shirt e non una camicia. Orbene, questo non differisce dal comportamento delle imprese quando redigono dei piani finanziari sulla base delle vendite e dei costi per sostenere tali vendite. Un fenomeno su cui porre l'attenzione è la previsione di vendita accidentale. Questo fatto si verifica quando si è talmente concentrati nello sviluppare un piano che si suppone quali saranno le previsioni di vendita senza avere delle prove che attestino la veridicità delle previsioni. Un esempio calzante di previsione accidentale si rileva nel settore dei supermercati. Il proprietario di un supermercato ha affermato che l'anno precedente era stato previsto un aumento delle vendite del 5%. Analizzando il mercato, invece, si è potuto notare che il settore non stava crescendo e che il tentativo di aumentare la quota di mercato ha portato all'attuazione di sconti e promozioni che non hanno avuto alcun effetto. La dimensione del settore e la

---

<sup>13</sup> Mentzer, J. T., & Moon, M. A. (2005). Sales forecasting management. A demand management approach (2nd Edition). Thousand Oaks (London): Sage Publications.

quota di mercato sono invariate e quindi la previsione di un aumento del 5% delle vendite non è corretto<sup>14</sup>.

La risposta a questo problema è che il management voleva motivare tutti a raggiungere l'obiettivo pianificato, ma tale obiettivo era matematicamente irraggiungibile in quanto la previsione era accidentale e non informativa. Questo modo di gestire le informazioni ha confuso le previsioni con la pianificazione e la definizione di obiettivi. L'aumento delle vendite del 5% non era una previsione ma un obiettivo di vendita a lungo termine.

Non bisogna confondere le previsioni della domanda né con gli obiettivi di vendita da raggiungere o superare né tantomeno con la pianificazione operativa. Si può definire un obiettivo di vendita come un valore che è definito per motivare il personale addetto alla vendita mentre la previsione di vendita è una proiezione di ciò che potrebbe accadere in futuro sulla base dei dati passati. Del resto per generare degli obiettivi realistici si dovranno prendere in considerazione le previsioni. Un'altra funzione da sviscerare è quella della pianificazione operativa che è determinata dalle azioni che devono essere attuate per raggiungere o superare degli obiettivi. Alla luce di ciò sia la pianificazione che la definizione degli obiettivi si fondano sulle previsioni di vendita, ma tutte e tre hanno significato e metodo di valutazione differente. Le previsioni hanno come parametro di valutazione l'accuratezza mentre i piani operativi devono avere come scopo quello di soddisfare in maniera efficace ed efficiente la domanda futura attraverso il raggiungimento degli obiettivi definiti.

---

<sup>14</sup> Mentzer, J. T., & Moon, M. A. (2005). Sales forecasting management. A demand management approach (2nd Edition). Thousand Oaks (London): Sage Publications.

## 1.5. Tecniche di previsione

Vi sono vari metodi di previsione ma tutti possono essere ricondotti in tre categorie di seguito analizzate. La prima categoria è quella basata sui metodi estrapolativi fondati sull'analisi delle serie storiche. Questi metodi si fondano sull'idea che in futuro l'andamento della domanda si ripeterà come ha fatto nel passato. Per questo motivo si utilizzano solamente i dati storici delle vendite, senza prendere in esame altri tipi di variabili che potrebbero influenzare la domanda in futuro. Vi sono più di 70 metodi estrapolativi ma tutti hanno in comune il modo in cui la serie storica di dati viene scomposta. La serie storica, secondo questo metodo, è scomposta in quattro componenti:

- 1) Trend o componente tendenziale o dinamica è l'andamento medio di lungo periodo che è influenzato dalla domanda del prodotto sul mercato e dai mutamenti della capacità di competere dell'impresa in rapporto alla capacità di competere delle concorrenti sullo stesso mercato;
- 2) Ciclo o componente congiunturale che evidenzia l'andamento del contesto economico cioè l'andamento del Pil del paese di riferimento;
- 3) Stagionalità o componente stagionale che è una componente che mette in luce un andamento ricorrente, solitamente ogni 12 mesi, che può avere cause differenti come l'organizzazione sociale o il ciclo stagionale, ad esempio, la temperatura (in estate vi sarà un aumento del consumo di gelato e diminuirà il consumo di cioccolata calda, viceversa in inverno), le festività (nelle festività natalizie e pasquali vi sarà un aumento del consumo di dolci);
- 4) Componente di errore o residuale e cioè le variazioni che non sono spiegate dalle altre componenti e quindi casuali.

Le serie temporali permettono di scomporre la serie storica, di proiettare nel futuro alcune componenti e ricomporla, ottenendo così delle previsioni. Il metodo più utilizzato è chiamato media mobile semplice, in quanto il metodo è facilmente applicabile e automatizzabile. Questo metodo è stato rivalutato rispetto al passato

in quanto, anche se i modelli statistici più complessi si adattano molto bene all'andamento della serie storica, non sono in grado di prevedere in modo preciso gli andamenti futuri della domanda, al contrario della media mobile per cui le previsioni sono puntuali, ma ha un minor potere adattivo alla serie storica. Altre tecniche che ricadono in questo cluster sono: lo smorzamento esponenziale, la media mobile ponderata, la previsione mirata e lo smorzamento esponenziale con trend e stagionalità.

Il secondo raggruppamento è dato dai metodi causali di previsione che si basano sulla correlazione delle variabili prese in considerazione. Questi metodi si fondano sul fatto che la variabile dipendente, ovvero la domanda del bene venduto, possa essere spiegata dall'andamento di altre variabili chiamate indipendenti. Le variabili indipendenti possono essere suddivise in variabili esogene ed endogene. Le prime sono variabili che non possono essere modificate in alcun modo dal management, per esempio, l'età della popolazione, la tassazione sui prodotti, la temperatura, i conflitti militari, il comportamento dei concorrenti, il reddito pro capite, mentre, le variabili endogene, sono ad esempio il prezzo del prodotto, il marketing mix, le strategie di vendita, che possono essere manipolate dalla direzione aziendale. Per proiettare la domanda si utilizzano metodi di regressione semplice oppure multivariata. Quest'ultima prende in considerazione più di una variabile esplicativa che cerca di cogliere i motivi dell'andamento della serie storica. Il problema principale di questi modelli è la proiezione futura delle variabili esogene, più è precisa e più la previsione sarà affidabile. Attualmente la tecnologia ha permesso un'evoluzione di tale metodo di previsione, infatti, un esempio è la tecnica denominata sentiment analysis che permette, attraverso un algoritmo, di estrarre ed analizzare migliaia di recensioni (Big Data) rilasciate dai consumatori sugli e-commerce. In questo caso la variabile indipendente si identifica nella recensione del consumatore che, sempre più, ha una rilevante importanza nella scelta del prodotto da acquistare. Altri metodi prendono in considerazione invece i dati relativi alle ricerche che il consumatore esegue sul motore di ricerca prima di acquistare un prodotto.

L'ultimo cluster comprende i metodi qualitativi molto diffusi nelle piccole imprese e che si basano sulla raccolta di opinioni attraverso dei questionari. Il primo approccio, che è anche quello più atavico e di più semplice applicazione è quello che analizza le opinioni previsionali dei manager. L'insieme delle diverse previsioni viene tramutata in una media che risulta più precisa rispetto all'opinione del singolo manager. Vi sono anche varianti di questa tecnica che prendono in considerazione l'opinione di professionisti all'esterno dell'impresa per aumentare l'imparzialità e la precisione. Il secondo approccio si basa sull'analisi dell'opinione dei singoli venditori. Tutti gli esiti derivanti dalle domande che cercano di raccogliere l'opinione riguardante le vendite future vengono raggruppati al fine di ottenere delle previsioni di vendita. Il terzo ed ultimo approccio consiste nel registrare le aspettative dei clienti. Ai clienti viene somministrato un questionario con delle domande inerenti le modalità e le intenzioni dell'acquisto. Questo metodo è particolarmente indicato per il settore B2B visto che ogni cliente può incidere particolarmente sull'andamento dell'impresa; tale metodo può anche essere utilizzato quando si hanno a disposizione le serie storiche dei prodotti. Questo accade quando vi è un'immissione sul mercato di un nuovo prodotto o di un prodotto particolarmente innovativo<sup>15</sup>.

---

<sup>15</sup> Bellagamba, A. (2018) Previsione delle vendite e pianificazione della produzione: quale ruolo riveste la "forecast accuracy". *Electronic journal of management*, 2-9.

## 2. METODI QUANTITATIVI

Questo capitolo ha la funzione di illustrare alcuni metodi quantitativi che hanno lo scopo di elaborare delle previsioni attraverso l'analisi dei dati passati. Questi metodi possono essere utilizzati non solo per prevedere le vendite ma anche per prevedere il prezzo di alcuni strumenti finanziari, o per prevedere qualsiasi altro fenomeno che abbia una serie storica di dati. I metodi selezionati sono differenti sia per il metodo applicativo che per il funzionamento, ma tutti i metodi hanno in comune la materia d'origine, ovvero la statistica.

### 2.1. Big Data e previsioni

Il numero di dati può variare da caso a caso ma è rilevante ricordare che in un mondo iperconnesso il previsore ha a disposizione una grande quantità di dati. La traduzione afferente al termine inglese Big Data nella lingua italiana si identifica dall'espressione grandi dati o mega dati. I mega dati sono, in generale, una massiccia raccolta di dati eterogenei più o meno aggregati. La peculiarità è che per poter estrarre valore o conoscenza da essi, è necessario l'utilizzo di metodi statistici o informatici. Il termine Big Data, infatti, si riferisce alla capacità della scienza dei dati di estrapolare conoscenza e di trovare delle relazioni che siano utili nel massimizzare un processo, le prestazioni di un'organizzazione, nell'eseguire previsioni o nel trovare correlazioni tra i dati. La dimensione di un insieme di dati è definita dal volume, dalla velocità e dalla varietà (le cosiddette 3V). Secondo le ultime ricerche vi è stato un aumento del volume di dati utilizzati durante il processo di previsione. È interessante notare come in passato le tecniche di previsione di data mining venivano utilizzate solamente su dei dati statici e non sulle serie storiche e che si imputi una colpa della recente crisi finanziaria all'utilizzo di sistemi e di modelli finanziari che non sono in grado di rapportarsi

con una mole così considerevole di dati generando delle previsioni imprecise. Attualmente vi sono alcune applicazioni che sfruttano sistemi computazionali avanzati e Big Data con successo come ad esempio le previsioni meteorologiche (anche se oltre una settimana sono ancora imprecise), l'industria della moda sta utilizzando queste tecniche per prevedere il futuro della moda attraverso i dati dei social media, il trasporto aereo utilizza i metodi di previsione per avere una stima del numero di passeggeri e dei prezzi futuri del carburante, le compagnie che vendono energia elettrica attuano un sistema di previsione per comprendere la variazione della domanda in base alle stagioni e all'orario, la compagnia di intrattenimento Netflix utilizza metodi di previsione per comprendere se iniziare a produrre o meno una serie tv mentre più di recente si è visto applicare questi metodi per prevedere la pandemia Covid-19 ed i suoi effetti a livello economico<sup>16</sup>.

---

<sup>16</sup>Hassani, H., & Silva, E. S. (2015). Forecasting with big data: A review. *Annals of Data Science*, 2(1), 5-19.

## 2.2. Le serie storiche

Una serie storica consiste in un insieme di osservazioni su un certo fenomeno ordinate nel tempo. Di norma le osservazioni sono misurate e registrate ad intervalli di tempo equidistanti (giornalmente, mensilmente, trimestralmente ecc...) a seconda dell'oggetto di studio osservato anche se in alcuni casi questo non avviene. Una serie storica è definita come:

$$\{X_t ; t = 1, \dots, n\}$$

dove il pedice  $t$  indica il tempo a cui il dato  $X$  si riferisce. Le osservazioni di una serie storica sono tra loro dipendenti ed è questa peculiarità che è oggetto di studi<sup>17</sup>.

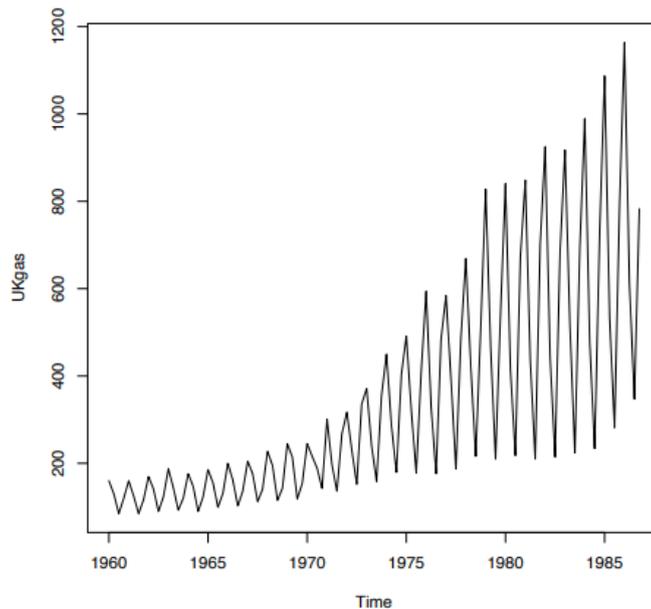


Figura 3: serie storica del consumo di gas nel Regno Unito

---

<sup>17</sup> Dagum, E. B. (2001). Analisi delle serie storiche: modellistica, previsione e scomposizione. Springer Science & Business Media.

Un esempio di serie storica è raffigurato in Figura 3 dove si può osservare una serie storica del consumo di gas, è evidente una componente ciclica con una frequenza annuale facilmente individuabile: il numero di massimi è 27 ed i minimi sono 26 su un totale di 26 anni di misurazione. È fattuale l'influenza che il cambio di temperatura ha sul consumo e l'aumento dell'oscillazione stagionale all'aumentare della serie stessa.

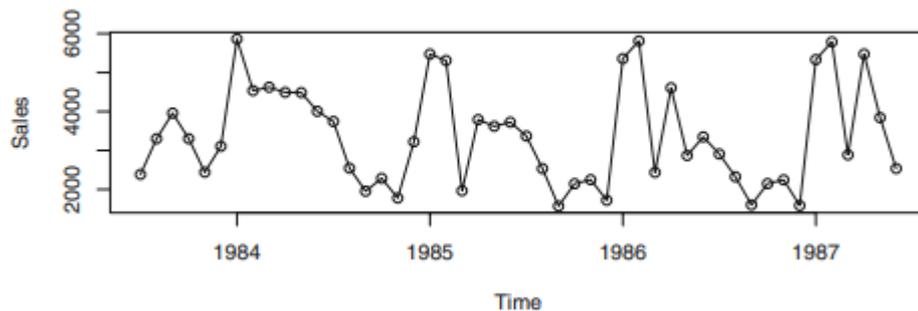
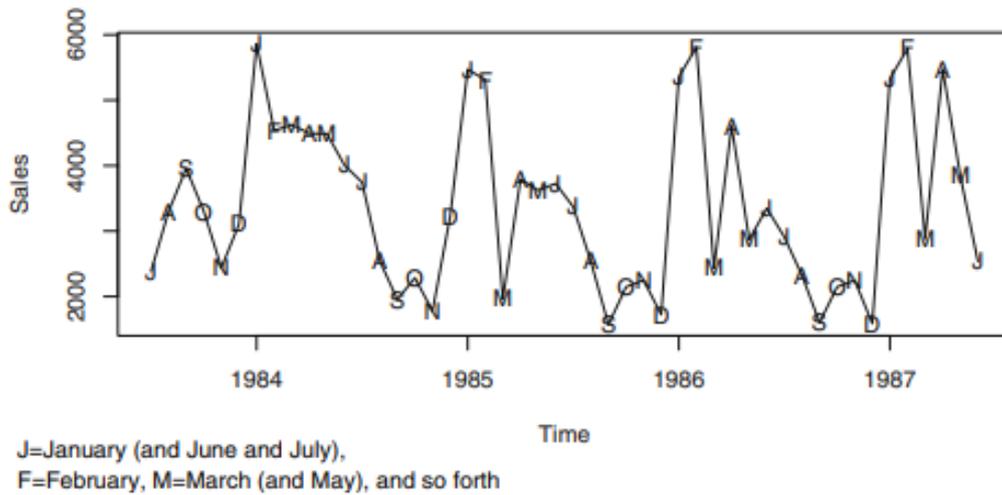


Figura 4: serie storica delle vendite di filtri per l'olio della John Deer

Un altro esempio di serie storica è dato dalla Figura 4 dove si può osservare le vendite di filtri per l'olio per trattori dell'azienda John Deer<sup>18</sup>. In questo caso ad un occhio non esperto è difficile identificare la ciclicità delle vendite. Pertanto si è pensato di aggiungere al grafico le mensilità così da avere una visione più chiara del fenomeno.

---

<sup>18</sup> L'azienda John Deere è una delle maggiori produttrici di macchine agricole al mondo, fu fondata nel 1837 da John Deere un fabbro pioniere. Ha sede nello stato dell'Illinois (USA) nella cittadina di Moline e durante la grande depressione divenne famosa a causa della sua politica di restituzione dei trattori dai clienti in difficoltà.



**Figura 5: serie storica delle vendite di filtri per l'olio della John Deer**

In Figura 5 è possibile notare come, grazie a questo minimo accorgimento, l'individuazione di una ripetizione ciclica delle vendite e come esse aumentino in modo esponenziale durante il mese di dicembre<sup>19</sup>. Si deve analizzare ulteriormente il fenomeno della stagionalità attraverso metodi statistici per comprendere se il ciclo di vendita è dovuto al caso oppure se vi sia correlazione.

---

<sup>19</sup> Cryer, J. D., & Chan, K. S. (2008). Time series analysis: with applications in R (Vol. 2). New York: Springer.

### 2.3. Processi stocastici e autocorrelazione

I processi stocastici<sup>20</sup> si definiscono come una successione di variabili casuali ordinate nel tempo con delle relazioni di dipendenza interna:

$$\gamma = \{Y_t : -\infty < t < +\infty\}$$

è possibile che un esperimento su  $\gamma$  sia in grado di far scaturire una successione di numeri in cui ogni  $y_t \in \{y_t : -\infty < t < +\infty\}$  e quindi ogni numero è un esperimento derivante dalla variabile casuale  $Y_t$ . Una successione numerica scaturita dall'esperimento viene denominata traiettoria e visto che il processo è stocastico e non deterministico per ogni esperimento su  $Y_t$  si avranno delle traiettorie, in generale tante quante volte viene eseguito l'esperimento, tutte diverse tra loro e in possesso di una maggiore o minore probabilità di realizzarsi.

Si dimostra che la distribuzione di probabilità di un processo stocastico ha la caratteristica dalle distribuzioni di probabilità di  $(Y_{t_1}, \dots, Y_{t_k})$  per qualsiasi  $k$  e per qualsiasi scelta di  $t_1, \dots, t_k$ <sup>21</sup>. In un contesto reale la stima di tutte queste distribuzioni risulta inefficace ed inefficiente e perciò è normale prendere in considerazione solamente alcune funzioni di stima:

$$\text{Media } \eta_t = E(Y_t)$$

$$\text{Varianza } \sigma_t^2 = \text{var}(Y_t)$$

$$\text{Autocorrelazione } \rho(t', t'') = \gamma(t', t'') / \sigma_{t'} \sigma_{t''}$$

Per poter stimare in modo accettabile un dato futuro è necessario aggiungere l'ipotesi di stazionarietà. La stazionarietà si può definire come:

---

<sup>20</sup> Nel calcolo delle probabilità il termine stocastico è sinonimo di casuale e aleatorio, con cui si caratterizzano gli enti matematici collegati a eventi aleatori; più in generale, gli strumenti, le teorie e i modelli stocastici sono atti a descrivere e studiare situazioni che variano in base a leggi probabilistiche (e non deterministiche), come, per esempio, tutti i fenomeni naturali, in quanto in essi è presente, sia per la loro stessa natura, sia per gli errori di osservazione, una componente casuale o accidentale.

<sup>21</sup> Masarotto, G. (2005). *Analisi delle Serie Temporal* con R. Facoltà di Scienze Statistiche Università degli Studi di Padova.

Media:  $E(Y_t) = \eta$

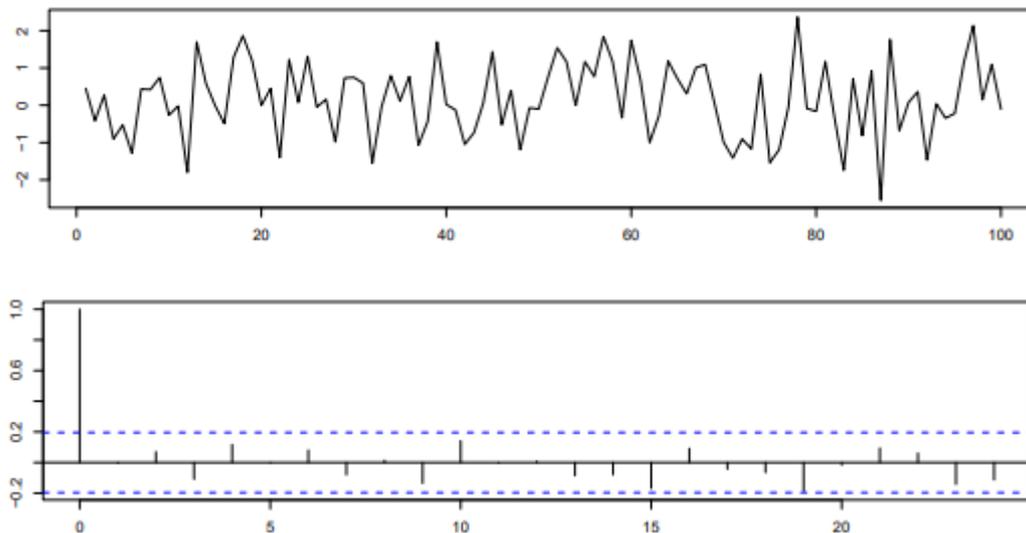
Varianza:  $\text{var}(Y_t) = \sigma^2$

Covarianza:  $\text{cov}(Y_{t+h}, Y_t) = \gamma(h)$

Correlazione:  $\text{corr}(Y_{t+h}, Y_t) = \rho(h)$

Per qualsiasi  $t$  e  $h$ .

Pertanto, se il processo è stazionario la media e la varianza non cambiano nel tempo e le autocorrelazioni sono una funzione della distanza nel tempo tra le due variabili casuali coinvolte. Se il processo è stocastico e stazionario l'autocorrelazione viene utilizzata come metodo per individuare la variabilità della serie storica in relazione alle osservazioni passate. Il metodo di studio delle serie storiche basato sui metodi stocastici non ha come scopo quello scomporre la serie in trend, ciclo, stagionalità e componente erratica ma quello di identificare un modello probabilistico al fine di prevedere dati futuri.



**Figura 6: correlogramma della serie storica soprariportata**

Si può affermare che la serie storica non sia autocorrelata quando le autocorrelazioni ricadono nell'intervallo di confidenza:

$$\left[ \frac{-Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}; \frac{Z_{1-\frac{\alpha}{2}}}{\sqrt{n}} \right]$$

In Figura 6 è possibile osservare il correlogramma della serie storica che attua la tecnica dell'intervallo di confidenza ed espone il fatto che in questo caso l'autocorrelazione è dovuta al caso in quanto tutte le autocorrelazioni ricadono all'interno dell'intervallo. Questo metodo di analisi ci permette di comprendere quanto un'osservazione possa influire sulle future osservazioni, se la serie possiede una componente periodica e le sue relative frequenze. Nel correlogramma delle serie storiche è possibile osservare un 5% di osservazioni al di fuori dei limiti e cioè calcolato con una confidenza al 95%<sup>22</sup>.

---

<sup>22</sup> Masarotto, G. (2005). *Analisi delle Serie Temporal* con R. Facoltà di Scienze Statistiche Università degli Studi di Padova.

## 2.4. Il modello SARIMA

L'abbreviazione o acronimo SARIMA deriva dall'inglese "Seasonal Auto-Regressive Integrated Moving Average" ed è un modello che è costruito sulla base di un altro modello precedente chiamato ARIMA addizionato di una componente stagionale. Vi è quindi un'evoluzione che somma diversi fattori: AR(p) è la componente auto regressiva, MA(q) che è la media mobile. Sommando queste due metodologie si ottiene il modello ARMA(p, q)<sup>23</sup>. Questa tipologia di metodo è possibile utilizzarla solamente se la serie storica è stazionaria altrimenti bisognerà attuare un processo in grado di renderla tale. Questo procedimento è stato studiato da (Box e Jenkins, 1973) ed è rappresentabile attraverso la seguente formula:

$$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d)X_t$$

La formula dimostra che una serie storica viene resa stazionaria attraverso la sottrazione dalla stessa serie del lag o ritardo (se è mensile si avrà lag 12 se è trimestrale lag 4 ecc...) e questo procedimento deve essere effettuato finché la serie storica diventerà stazionaria. Attraverso tale processo si ottiene il modello denominato ARIMA (p,d,q) o modello auto regressivo a media mobile differenziato. La componente stagionale è fondamentale nell'analisi delle serie storiche ai fini previsionali per ottenere un risultato, e quindi una previsione, che prenda in considerazione le variazioni cicliche nel tempo e che altrimenti non potrebbero essere esterne nel modello da altri fattori. Il modello ARIMA è stato quindi esteso attraverso l'aggiunta del parametro stagionale. Il modello SARIMA è sintetizzato nella seguente funzione<sup>24</sup>:

---

<sup>23</sup> Arunraj, N. S., Ahrens, D., & Fernandes, M. (2016). Application of SARIMAX model to forecast daily sales in food retail industry. *International Journal of Operations Research and Information Systems (IJORIS)*, 7(2), 1-21.

<sup>24</sup> Divisekara, R. W., Jayasinghe, G. J. M. S. R., & Kumari, K. W. S. N. (2021). Forecasting the red lentils commodity market price using SARIMA models. *SN Business & Economics*, 1(1), 1-13.

$$\text{SARIMA} = (p, d, q)(P, D, Q)[S]$$

Costituita dalle seguenti componenti:

- p e P (stagionale): indicano il numero di termini autocorrelati e cioè i ritardi o “lags” della serie stazionaria osservabili dalla funzione di autocorrelazione parziale;
- d e D (stagionale): indicano il numero di volte che deve essere applicato il processo di differenziazione per rendere una serie stazionaria;
- q e Q (stagionale): indicano il numero di termini della media mobile autocorrelati e cioè i ritardi o “lags” della serie storica osservabili dalla funzione di autocorrelazione;
- S: indica la stagionalità nei dati.

Questo tipo di modello non considera altri tipi di fattori esogeni che potrebbero essere correlati con l’esito della serie storica ma utilizza solamente i dati passati per trovare i parametri che siano in grado di predire la serie. Il modello SARIMA possiede due particolari peculiarità: il modello assume che i dati consecutivi siano correlati (come nel modello ARIMA) ma anche che vi sia un ritardo nelle osservazioni che sia correlate e cioè che vi sia stagionalità, dando la possibilità di analizzare anche una stagionalità non periodica e quindi all’utilizzo anche quando la serie storica non è stazionaria.

Il modello SARIMA può essere visto anche come la seguente equazione:

$$\begin{aligned} & (1 - \phi_1 B^w - \phi_2 B^{2w} \dots - \phi_p B^{pw}) \times (1 - \varphi_1 B - \varphi_2 B^2 \dots - \varphi_p B^p) \times (1 - B^w) \\ & \times (1 - B)^D \times (1 - B)^d Q_n(t) \\ & = (1 - \vartheta_1 B^w - \vartheta_2 B^{2w} \dots - \vartheta_{1Q} B^{Qw}) \\ & \times (1 - \theta_1 B - \theta_2 B^2 \dots - \theta_q B^q) e(t) \end{aligned}$$

Dove  $\phi$  sono i parametri delle autocorrelazioni non stagionali e  $\theta$  sono i parametri non stagionali della media mobile,  $\varphi$  sono i parametri stagionali delle

autocorrelazioni, mentre  $\vartheta$  sono i parametri stagionali della media mobile,  $w$  è la frequenza e  $B$  sono i differenziali delle variabili<sup>25</sup>. Questo modello ha la caratteristica di avere una stagionalità variabile ed è differente dai coefficienti di stagionalità utilizzati per filtrare l'impatto della stagionalità sulla serie nell'approccio classico. Il modello si è ulteriormente evoluto al fine di ottenere delle previsioni più precise costituendo il modello SARIMAX (with eXogenous factors) che si differenzia rispetto al suo predecessore in quanto si possono aggiungere una o più variabili esogene. L'aggiunzione di una variabile esogena ha lo scopo di esplicitare parte dell'andamento della variabile endogena che la stessa non è in grado di decifrare come, ad esempio, l'effetto del cambiamento di temperatura spiega in parte l'aumento del consumo di bevande calde. Nel capitolo successivo verranno confrontati i risultati dei due metodi attraverso un'analisi di alcune serie storiche.

---

<sup>25</sup> Divisekara, R. W., Jayasinghe, G. J. M. S. R., & Kumari, K. W. S. N. (2021). Forecasting the red lentils commodity market price using SARIMA models. *SN Business & Economics*, 1(1), 1-13.

## 2.5. Le reti neurali

Alla base del funzionamento celebrale, è presente una cellula che permette la comprensione di tutto ciò che avviene nello spazio circostante e di acquisire nuove nozioni. Questa cellula fondamentale è il neurone, presente nella mente per oltre 100 miliardi ed ognuno interconnesso ad altri 10000<sup>26</sup> e connessi tra loro grazie a dei processi elettrochimici. Il neurone artificiale non è figlio dell'ultima rivoluzione tecnologica, bensì ha origini remote<sup>27</sup> si pensi che il primo neurone artificiale venne ideato durante la Seconda guerra mondiale.

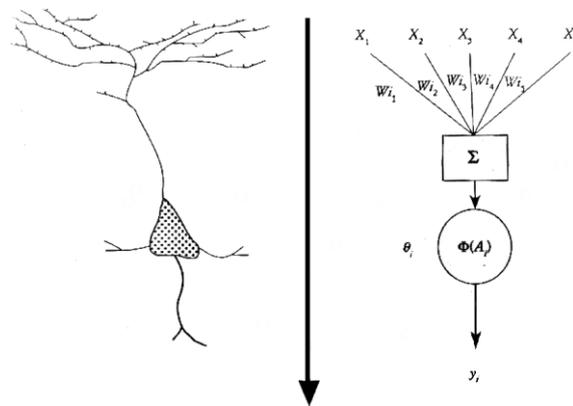


Figura 7: confronto tra un neurone biologico e uno artificiale

Il neurone artificiale in Figura 7, prende ispirazione dal neurone biologico e per questo motivo vi è una certa somiglianza tra i due, infatti, è un elemento semplice che se messo assieme ad altri riesce a risolvere problemi complessi. Questo elemento è alla base del riconoscimento vocale, dei veicoli a guida autonoma e della generazione del linguaggio. Il neurone artificiale in Figura 8 è in grado di

---

<sup>26</sup> Marieb, E., Hoehn, K., Martini, F. H., Nath, J. L., Bartholomew, E. F., McKinley, M. P., ... & Russ, A. (2015). *Essentials of Human Anatomy*.

<sup>27</sup> Fitch, F. B. (1944). Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biophysics*, vol. 5 (1943), pp. 115–133. *The Journal of Symbolic Logic*, 9(2), 49-50.

risolvere un problema utilizzando un set di dati. Le grandezze hanno i seguenti significati:

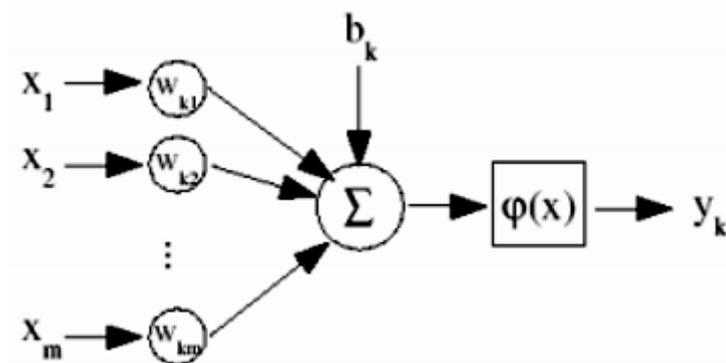


Figura 8 : rappresentazione di un neurone artificiale

- $W_{kn}^{28}$  sono i pesi sinaptici. Sono parametri che definiscono l'efficacia delle connessioni sinaptiche dei dendriti. In un primo istante saranno computati con casualità, poi modificati attraverso un algoritmo.
- $b_k$  è una costante c.d. "bias". Serve a definire l'output assieme alla somma pesata degli input per i pesi sinaptici. Per questo motivo se il termine costante è grande si avrà maggior probabilità di ottenere un output pari a uno viceversa uguale a zero. Ad esempio, se volessimo risolvere un problema di classificazione vegetale tra due specie (A,B), se il risultato fosse uguale ad 1 la pianta apparterebbe alla specie A, altrimenti se fosse vicino a zero apparterebbe alla specie B.
- $\varphi(x)$  è la funzione di attivazione. Nel neurone artificiale la trasmissione dell'informazione avviene tramite la funzione di attivazione. Le funzioni di attivazione possono essere di diverse tipologie a seconda del problema da risolvere. Quelle più comunemente utilizzate sono la funzione sigmoideale e la funzione three-shold detta anche a "gradino". L'output corrisponde alla risposta del problema che abbiamo somministrato al neurone o alla rete neurale.

---

<sup>28</sup> Il pedice  $n$  è in riferimento al neurone analizzato all'interno di una rete neurale, mentre  $k$  è per distinguere il peso sinaptico all'interno del neurone.

- $y_k$  è l'output identificato da  $y_k = \varphi(u_k + b_k)$ . Esso è dato dalla somma pesata degli input ( $\Sigma = u_k$ ) per i pesi sinaptici, addizionata dal parametro "bias", il tutto moltiplicato per la funzione di attivazione.

La rete neurale artificiale, o dall'inglese Artificial Neural Network (ANN), è un modello matematico guidato dallo stesso funzionamento della rete neurale biologica ed è composto dall'interconnessione tra più neuroni artificiali. È un sistema adattativo che modifica la propria struttura, in particolare le interconnessioni o pesi attraverso un algoritmo, utilizzando i dati interni o esterni che scorrono durante il processo di apprendimento. Per quanto concerne l'analisi dei dati le ANN vengono utilizzate nella classificazione, nella regressione e nel clustering<sup>29</sup>. Le reti neurali sono altresì utilizzate nei campi più disparati, dal controllo veicolare, alla generazione del linguaggio, al rilevamento di mine sottomarine, all'applicazione in finanza.

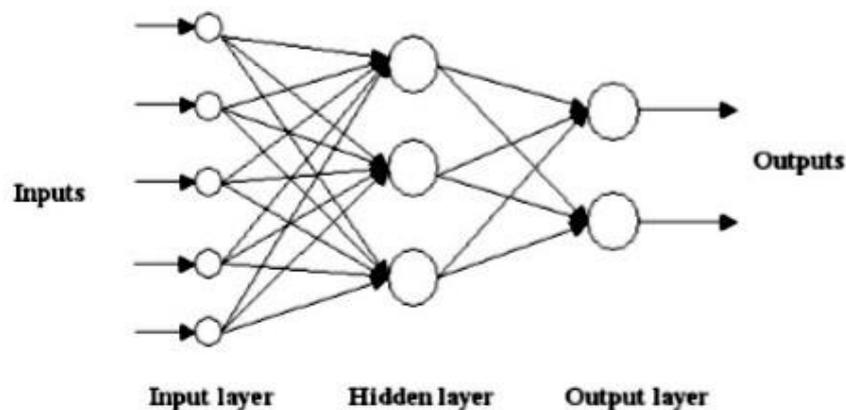


Figura 9: struttura di una rete neurale artificiale

---

<sup>29</sup>Gurney, K. (2018). An introduction to neural networks. CRC press.  
Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

La rete in Figura 9 è una rete multistrato perché suddivisibile in layer o livelli che hanno funzioni simili. L'input layer è il primo livello che i dati da analizzare incontrano mentre l'output layer restituisce le risposte al supervisore. Il numero di neuroni input è pari al numero di variabili da analizzare mentre il numero di neuroni output coincide al numero di risposte da ottenere. Tra l'input e l'output layer è possibile che ci sia uno o più hidden layer (livelli nascosti)<sup>30</sup>. I neuroni nascosti interagiscono solamente con gli altri neuroni senza avere alcun contatto con il supervisore. Solitamente per una rete neurale è richiesto il livello input e output, mentre il livello nascosto è facoltativo, infatti, è possibile avere layer che fungono sia da input che da output<sup>31</sup>. Il termine rete neurale o ANN si è evoluto fino a comprendere una vasta classe di modelli e metodi di apprendimento<sup>32</sup>. Le reti neurali artificiali si possono distinguere per il tipo di apprendimento (supervisionato, per rinforzo e non supervisionato) e per la tipologia di struttura (ricorrenti e non ricorrenti).

Prima di esporre il modello Long Short Term Memory è necessario capire in che modo apprende una rete neurale artificiale. Il tipo di apprendimento più utilizzato è l'apprendimento supervisionato. Quest'ultimo è caratterizzato dal fornire alla rete delle soluzioni dette etichette ( $y$ ) nell'insieme dei dati di addestramento ( $x$ ). L'insieme di questi dati viene chiamato training set, in modo tale che l'output sia generato dalla funzione obiettivo  $y = f(x)$ . Attraverso l'algoritmo di apprendimento si genera una funzione ipotesi  $y = h(x)$  che approssima la funzione obiettivo. Ogni volta che il training set viene immesso nella rete il risultato migliorerà. In altre parole la rete neurale analizza un training set cioè un insieme di input ai quali corrispondono output noti e apprende come calcolare nuove regole di associazione input-output. Questo processo termina quando l'errore

---

<sup>30</sup> Demuth, H. B., Beale, M. H., De Jess, O., and Hagan, M. T. (2014). *Neural network design*. Martin Hagan.

<sup>31</sup> Chung, S., Lee, D. D., & Sompolinsky, H. (2015). Classification of Manifolds by Single-Layer Neural Networks. *CoRR*.

<sup>32</sup> Nasser, I. M., & Abu-Naser, S. S. (2019). Artificial neural network for predicting animals category.

dell'output del training set viene reso accettabile. L'apprendimento supervisionato ha la capacità di addestrare la rete e se effettuato correttamente, la rete artificiale sarà in grado di risolvere dei problemi come il clustering o il forecasting analizzando un dataset. Si osserva che il processo di apprendimento di una rete neurale artificiale è molto simile a quello umano per cui per apprendere è necessaria la ripetizione<sup>33</sup>.

L'algoritmo più utilizzato per minimizzare l'errore (dato dalla differenza tra l'output ed il valore target) è quello di backpropagation o generalized delta rule. L'algoritmo di retro propagazione è in grado di attribuire l'errore dell'output ad ogni nodo o neurone in modo proporzionale ai suoi pesi sinaptici e attraverso la tecnica della discesa del gradiente minimizzare quest'ultimo, modificando i rispettivi pesi. La tecnica della discesa del gradiente è la minimizzazione della funzione di errore attraverso l'utilizzo della chain rule e la velocità di minimizzazione è data dal tasso di apprendimento. Possiamo riassumere tutto il processo di apprendimento nei seguenti punti<sup>34</sup>:

- il calcolo dell'output per un dato input procede in avanti propagando l'errore;
- calcola l'errore E tra l'output della rete ed il valore target;
- propaga all'indietro l'errore modificando i pesi in rapporto all'influenza che hanno avuto nella formazione di E.

Prima di utilizzare la rete per l'analisi dei dati e dopo aver effettuato questo procedimento utilizzando il training set, è possibile confermare ulteriormente l'apprendimento con il test set che avrà dati differenti (appartenenti alla stessa tipologia) rispetto al training set.

---

<sup>33</sup> Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

<sup>34</sup> Nielsen, M. A. (2015). Neural networks and deep learning (Vol. 25). San Francisco, CA, USA: Determination press.

## 2.6. Long Short Term Memory

Il modello LSTM è una estensione di una rete neurale ricorrente (RNN) che ha la capacità di apprendere la dinamica della serie storica tenendo conto del fattore temporale. Le reti neurali ricorrenti, nate negli anni '80 hanno una propria struttura o meglio architettura<sup>35</sup>.

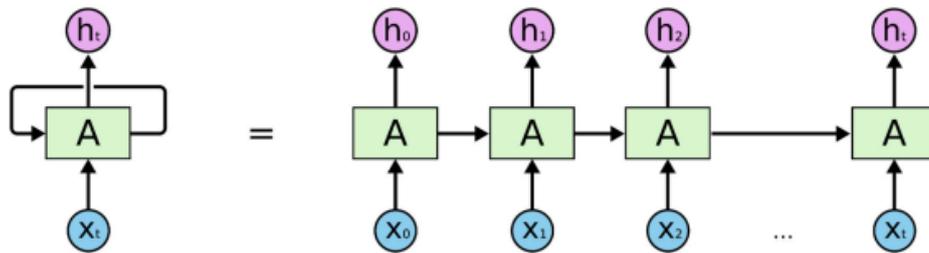


Figura 10: l'architettura di una rete neurale ricorrente

Nella Figura 10, a sinistra, è possibile osservare l'architettura di una RNN che si può anche visualizzare come una ripetizione della trasmissione dell'output della rete alla stessa in modo continuativo. Questa tipologia di rete neurale si contraddistingue dalla tipologia feedforward<sup>36</sup> - in cui i dati in uscita sono identificabili come la risposta data dalla computazione dell'algoritmo scelto - in quanto l'output calcolato dalle stesse viene reintrodotta all'interno della rete che può avvenire sia dall'output all'input e sia tra neuroni adiacenti. Questo processo automatico di reintroduzione dei dati è ciò che permette a questa tipologia di rete neurale di essere utilizzata per effettuare delle previsioni utilizzando dati passati inerenti ad una ampia finestra temporale. In Figura 10 è possibile osservare come la rete sia composta da una catena di molti moduli o celle che hanno la capacità di

---

<sup>35</sup> Son, N., Yang, S., & Na, J. (2020). Deep neural network and long short-term memory for electric power load forecasting. *Applied Sciences*, 10(18), 6489.

<sup>36</sup> La struttura delle reti neurali feedforward ha un'architettura semplice e funziona attraverso l'inserimento degli input da parte del supervisore e generando un output che non viene reimesso nella rete; hanno una capacità di calcolo inferiore alle reti feed-back.

memorizzare una parte di informazioni dalle precedenti fasi di elaborazione e quindi il risultato della fase T-1 viene reimmesso nella rete per influenzare il risultato della fase T e ciò avviene per ogni fase successiva. Le grandezze hanno il seguente significato:

- $X_t$  è l'input al passaggio t;
- $H_t$  è l'output al passaggio t;
- A è l'algoritmo di apprendimento che vi è all'interno della cella.

Il processo computazionale avviene attraverso un algoritmo di retro-propagazione attraverso il tempo (BPTT), che tuttavia, in questa tipologia di reti, fa insorgere delle problematiche riguardanti l'aggiornamento dei pesi <sup>37</sup>portando a dei risultati insoddisfacenti. La risoluzione del problema è stata affrontata trasformando la RNN in LSTM<sup>38</sup>.

---

<sup>37</sup> Il processo BPTT usa un modello working-backward: partendo dall'output finale della rete neurale quest'ultima modifica i pesi di ciascun neurone in base alla proporzione di calcolo dell'errore in uscita dal neurone stesso. Questo processo avviene per ogni strato della rete. La trasmissione ciclica di informazioni causa degli aggiornamenti molto elevati ai pesi della rete che porta ad una instabilità di tutta la rete.

<sup>38</sup> Zaytar, M. A., & El Amrani, C. (2016). Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *International Journal of Computer Applications*, 143(11), 7-11.

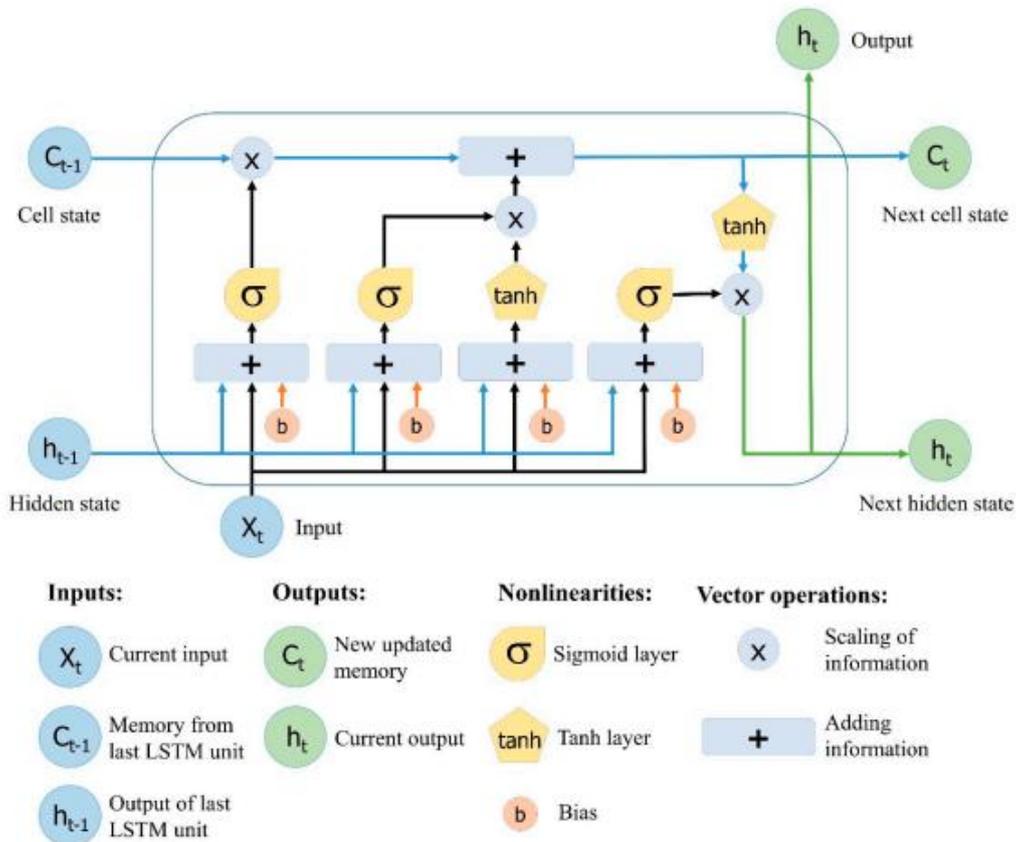


Figura 11: Long short term memory

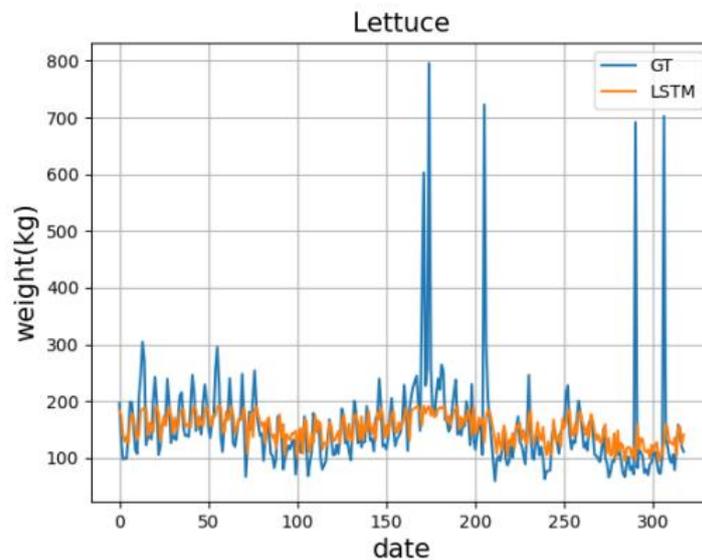
La Long Short-Term Memory è un'evoluzione della RNN che aggiunge ulteriori interazioni per ogni neurone artificiale. Questa tipologia di rete è in grado di affrontare problemi di previsione grazie alla capacità di imparare le correlazioni tra i dati ed avere una memoria a lungo termine. Come si può osservare dalla Figura 11 questa tipologia di rete ha la peculiarità di avere 3 input:  $X_t$  sono i dati all'istante di calcolo  $t$ ,  $C_{t-1}$  è la memoria della cella precedente e  $h_{t-1}$  è l'output della cella precedente.

In Figura 11 è possibile osservare la struttura di una rete neurale LSTM e si può notare come sia composta da quattro hidden state (4 neuroni artificiali) che computano grazie a delle funzioni. Ogni neurone è composto da funzioni matematiche denominate “porte” o “cancelli” che hanno scopi diversi: dimenticare, ingresso, aggiornamento e porte di uscita. La prima porta “dimenticare” ha lo scopo di identificare le informazioni superflue, ometterle e funziona grazie ad una funzione sigmoideale  $\sigma$  (0-1) mentre la porta di “aggiornamento” serve per aggiornare lo stato della cella. Quest’ultimo passaggio è suddiviso in due fasi: la funzione sigmoideale  $\sigma$  (0-1) decide le informazioni da ignorare o da aggiornare mentre la funzione  $\tanh$  (-1;+1) attribuisce un peso e quindi più o meno importanza alle informazioni passate. I due valori in uscita da questo procedimento vengono moltiplicati, creando una “memoria recente”, sommati (+) ai valori della cella precedente “memoria passata” ( $C_{t-1}$ ) ed utilizzati per aggiornare la cella, ottenendo così una “nuova memoria”. Nell’ultimo passaggio i valori output ( $h_t$ ) derivano dai valori della “nuova memoria” filtrati da una funzione  $\tanh$  (-1; +1) e successivamente vengono moltiplicati per i valori input filtrati da una funzione sigmoideale (0-1). Come si può vedere ci sono due tipologie di output: la prima della quale si è già esposto il metodo computazionale è la “nuova memoria” che andrà a formare la “vecchia memoria” nella cella successiva mentre la seconda tipologia è il risultato e quindi in questo caso la previsione<sup>39</sup>.

---

<sup>39</sup> Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of long short-term memory (LSTM) neural network for flood forecasting. *Water*, 11(7), 1387.

Un esempio per comprendere meglio il modello LSTM è dato dallo studio sottostante che cerca di predire il volume di prodotti agricoli utilizzando i dati dal primo giugno 2014 al 31 dicembre 2019.



**Figura 12: test del modello per prevedere il volume (Kg) di lattughe vendute**

La LSTM è un algoritmo adattivo, una cella di memoria che è in grado di comprendere l'andamento delle vendite passate, creare un modello di regressione che ha la capacità di prevedere l'andamento futuro dei dati. In Figura 12 è possibile osservare l'andamento della linea blu che interpreta le vendite passate di lattuga in Kg mentre la linea arancione è il test del modello LST. Al fine di ottenere questo risultato bisogna attuare diversi passaggi preliminari. Dopo aver visualizzato la serie storica di dati bisogna trattare i dati normalizzandoli, in seguito bisogna costruire un training set che serve per allenare la rete neurale e poi specificare i parametri del modello LSTM come: il numero di neuroni, di hidden layer, di neuroni output e di epoche di apprendimento. Dopo aver effettuato l'allenamento e quindi prodotto un modello predittivo che ad ogni inserimento del training diminuisce l'errore bisogna visualizzare l'output. Infine, è necessario verificare l'errore di stima del modello e nel caso apportare le necessarie modifiche ai parametri del LSTM.

## 2.7. Confronto e valutazione dell'errore

È possibile confrontare i metodi utilizzati attraverso delle misure comuni dell'errore di stima della previsione. La prima misurazione comune a tutti i modelli è l'errore quadratico medio (RMSE) o deviazione standard dato da:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{U}_i - U)^2}$$

La deviazione standard (radice quadrica della varianza) permette di calcolare l'errore del modello attraverso la differenza degli scarti dalla retta di regressione elevati al quadrato. Questo procedimento impedisce che gli scarti positivi e negativi si annullino a vicenda rendendo così non nulla la somma degli stessi inoltre bisogna calcolarne la media. Il valore ottenuto è chiamato varianza ed è posto sotto la radice quadrata in quanto l'elevazione al quadrato ha alterato l'unità di misura originaria. Il RMSE può adottare valori maggiori o uguali a 0 ma questa ampiezza di valori non è facilmente interpretabile. Un valore pari a 0 coincide con l'ipotesi in cui tutte le osservazioni sono uguali alla media aritmetica delle osservazioni. In statistica è cosa comune osservare che il divisore N è sostituito con N-1 quando la deviazione standard è calcolata a partire da un campione estratto da una popolazione. L'effetto derivante da questa modifica si attenua all'aumentare delle dimensioni del campione.

Un'altra misura della bontà dei modelli è  $R^2$  corretto che spiega il potere predittivo del modello utilizzato.

$$R^2_{adj} = 1 - \frac{s^2}{s_y^2}$$

Dove  $s^2$  è l'errore quadratico medio mentre  $s_y^2$  è la varianza campionaria di  $y$ .  $R^2$  corretto aumenta all'aumentare della bontà del modello e assume valori da 0 a 1. Questo indicatore è molto utile nel caso in cui i modelli siano integrati con

variabili esogene che possano spiegare parte della variabilità delle osservazioni; infatti, è noto che se si introduce una variabile esplicativa l'indicatore aumenta mentre se la variabile non è utile nel migliorare il modello l'indicatore diminuisce.

Solitamente all'affiancamento degli indicatori si cerca di comprendere se il modello di previsione è corretto attraverso la visualizzazione dei residui, che sono la differenza tra i valori osservati e quelli predetti dal modello. Se l'andamento dei residui è casuale vuol dire che il modello ha una buona capacità di previsione al contrario, se vi è una tendenza stagionale il modello non è in grado di approssimare in maniera corretta l'andamento di fondo delle osservazioni.

## 3. II CASO AZIENDALE

La Tv Square Spa (nome fittizio) è stata fondata negli anni '50 e ad oggi conta all'incirca 40 punti vendita sul territorio italiano. L'azienda è specializzata nella distribuzione di elettronica di consumo, elettrodomestici e di servizi ad essi collegati. La Tv Square ricade sicuramente nella tipologia della grande impresa in quanto possiede quasi mille collaboratori e un fatturato di oltre 300 milioni di euro annui.

### 3.1. Analisi del settore

Il televisore<sup>40</sup> ha avuto nel corso degli anni un'evoluzione radicale. In origine questo elettrodomestico era basato sulla tecnologia del tubo catodico che per riprodurre un'immagine al suo interno utilizzava dei fasci di elettroni convogliati su una superficie fotosensibile frontale la cui diagonale viene misurata in pollici. Fino agli anni '70 in Italia la televisione era trasmessa in bianco e nero e solamente nel 1977 ci fu la possibilità di vedere tutti i canali a colori. La rivoluzione tecnologica ha visto l'introduzione di nuove tipologie di schermi, con una maggior risoluzione e con un ingombro molto inferiore<sup>41</sup>. Il televisore al plasma, nato negli anni '90, sfruttava dei gas nobili imprigionati tra due pannelli per poter visualizzare le immagini e venne adottato come tecnologia in sostituzione del tubo catodico anche se il costo di questa tipologia di televisore era

---

<sup>40</sup> Apparecchio ricevente televisivo, in passato solo in bianco e nero, attualmente quasi esclusivamente a colori, di varie misure (*t. a 16, 20, 22, ..., 40, 46, 52 pollici*, a seconda della misura della diagonale dello schermo) e tipologie tecniche.

<sup>41</sup> Si passa da televisori con uno schermo di 4/3 a 16/9 e dopo agli schermi misurati in pollici. Il cambio di proporzioni è dovuto soprattutto al modo di fare televisione (proiettare le pellicole cinematografiche al cinema è diverso rispetto a proiettarle in Tv) visto che anche i film e gli schermi si stavano evolvendo in favore dei consumatori, tenendo conto di come una persona vede e di come si potesse fare per mantenere sullo schermo le stesse proporzioni della pellicola.

notevolmente superiore. Dopo l'introduzione degli schermi a cristalli liquidi, in precedenza utilizzati per gli schermi dei computers e delle calcolatrici, i televisori al plasma vennero in breve tempo sostituiti. I televisori a cristalli liquidi avevano il vantaggio di avere un costo di produzione inferiore grazie al fatto che sono prodotti attraverso il taglio, tramite laser, di pannelli di cristalli liquidi e quindi senza bisogno di avere linee produttive differenti per ogni diversa grandezza dello schermo. Altri vantaggi sono dati dal minore spazio occupato, tanto da poterli fissare alla parete e dalla miglior qualità di fruizione. L'attuale tipologia di televisori più venduta è la cosiddetta LED (diodo ad emissioni luminose), caratterizzati dall'alta definizione e da un ampio schermo (anche di 65 pollici) rendendo la visione di trasmissioni e film sempre più nitida e realistica. La tecnologia a led permette una visione in 4k o ultra-HD che permette una visione quattro volte più nitida rispetto al full HD e recentemente è apparso sul mercato anche lo schermo 16K che presenta quattro volte più pixel di uno schermo 4K. Un'altra rivoluzione importante che ha impattato sul comportamento dei consumatori è stato il cambio di frequenze da standard al digitale terrestre (2012) ed ora al nuovo digitale terrestre (2023). Il cambio di frequenze è conseguenza di normative elargite a livello europeo e statale. Le frequenze sono state cambiate in maniera graduale, rinviando più volte le date dello spegnimento e rendendo obsoleti gli apparecchi non in grado di decodificare il nuovo segnale. Inoltre in contemporanea sono stati approvati dei bonus per incentivare i consumatori ad acquistare nuovi televisori per facilitare la transizione al nuovo digitale terrestre. Inizialmente il bonus era sottoposto ad una politica per cui i consumatori potevano accedervi solamente in base ad una soglia ISEE mentre successivamente (da agosto 2021) il vincolo ISEE è stato abolito in favore dell'erogazione del "bonus rottamazione" per cui si configura come uno sconto del 20% sul prezzo d'acquisto fino a un massimo di 100€ previa la rottamazione di un televisore acquistato prima del 22 dicembre 2018. Un altro fattore che ha contribuito al cambiamento tecnologico è stato l'avvento della Smart TV che consente all'utente di visualizzare dei contenuti del web non solo direttamente dal televisore, ma anche collegandosi dal telefono o dal computer.

Alcuni fattori come l'avvento della pandemia COVID-19 iniziata nel dicembre 2019 mentre in alcuni settori ha influito negativamente in altri ha avuto degli effetti che si sono rivelati positivi. Dopo aver effettuato delle analisi interne, durante il periodo pandemico, l'azienda ha confermato con sorpresa che la pandemia stava portando ad un aumento significativo delle vendite dei televisori e così facendo ha potuto fare fronte alle richieste dei consumatori.

### 3.2. I dati

I dati utilizzati nell'indagine si riferiscono a televisori, decoder e cam venduti negli anni 2007-2011 e 2012-2022 e sono stati forniti in formato "CSV". I dati sono stati suddivisi per tipologie di prodotto: Tv tubo catodico in bianco e nero, Tv tubo catodico a colori, Tv Lcd, Tv al Plasma, Tv Oled, Cam e Decoder.

**Tabella 1: Visualizzazione dei dati inviati dall'azienda (Tv Oled)**

CDPAR	DEPAR	CDMAG	QTMOV	DAMOV	CDCAU
TVOHIS55A80G	OLED 4K HISENSE 55A80G	1	1.0	20211209	33
TVOHIS55A80G	OLED 4K HISENSE 55A80G	2	1.0	20211218	33
TVOHIS55A80G	OLED 4K HISENSE 55A80G	3	1.0	20211121	3
TVOHIS55A80G	OLED 4K HISENSE 55A80G	3	1.0	20211211	33

Nella Tabella 1 soprariportata è possibile osservare i le prime quattro righe dei dati inerenti la vendita dei televisori Oled. Le etichette hanno i seguenti significati: CDPAR (codice articolo); DEPAR (descrizione articolo); CDMAG (codice filiale); QTMOV (quantità venduta); DAMOV (data); CDCAU (causale di vendita). Per effettuare le analisi seguenti sono stati utilizzati in un formato ridotto (ogni serie era composta da all'incirca 40000 osservazioni), prendendo in considerazione solamente la quantità venduta e la data. Per ottenere un risultato soddisfacente si è dovuto aggregare i dati a livello settimanale o mensile. Così facendo si sono perse delle informazioni sull'andamento della serie storica (variabilità) ma si è potuto ottenere una serie storica più omogenea. Per effettuare l'analisi dei dati è stata utilizzata la piattaforma messa a disposizione di Google Drive denominata Google Colaboratory dove è possibile utilizzare il linguaggio di programmazione Python per applicare i modelli quantitativi sopraesposti.

### 3.3. Confronto ed applicazione

Dopo aver aggregato i dati a livello mensile sono state visualizzate le serie storiche per osservare l'andamento delle vendite nel tempo.

Figure

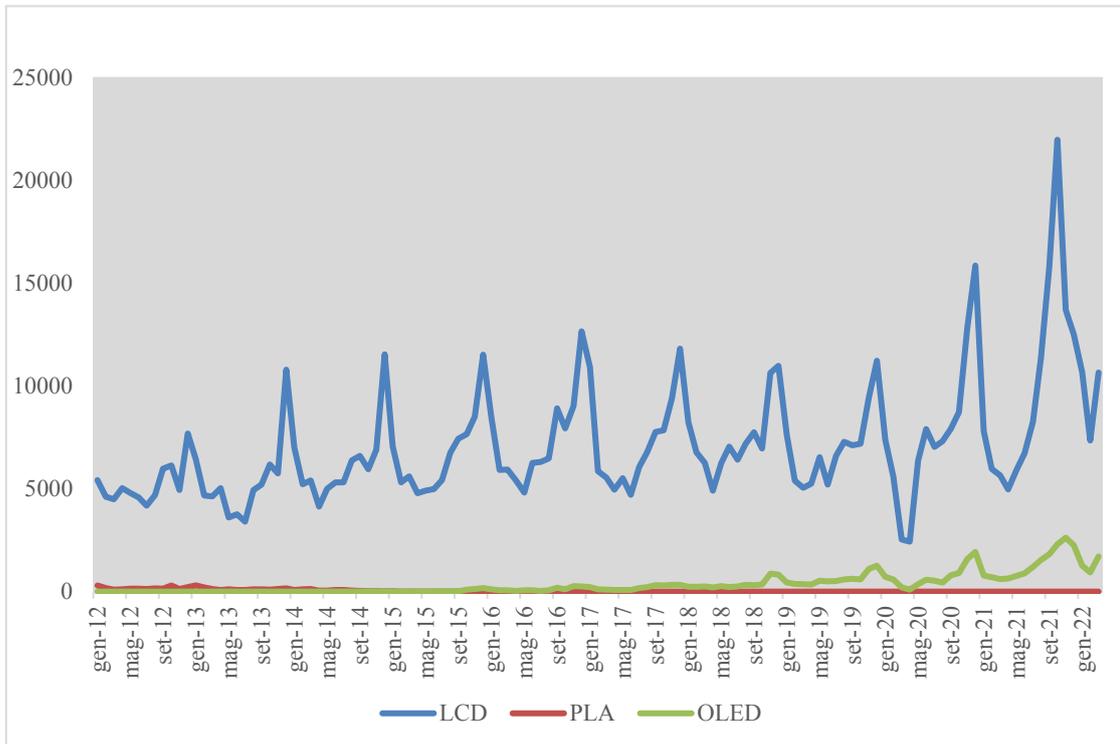
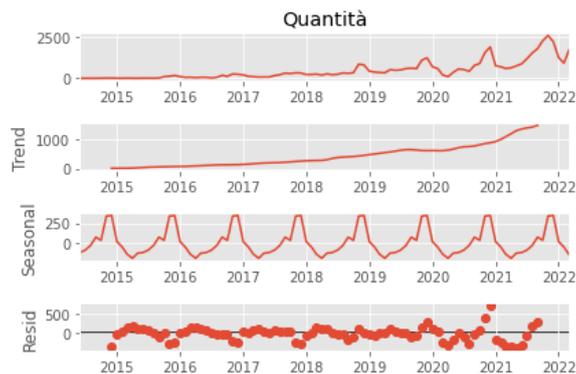
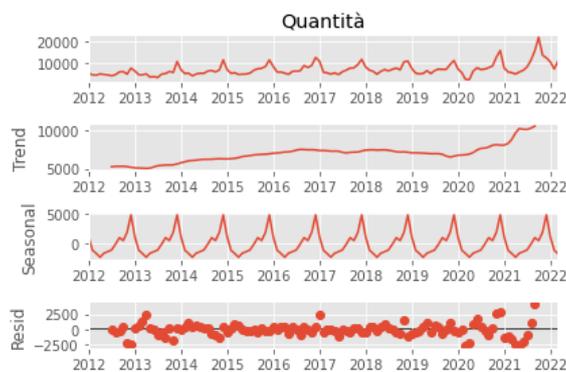


Figure 13: serie storica dei televisori LCD (in blu), serie storica dei televisori al plasma (in rosso) e serie storica dei televisori OLED (in verde)

La Figura 13 è utile a comprendere i volumi di vendita e l'importanza di ogni tipologia di televisore. Sicuramente la tipologia predominante è ancora quella Lcd seguita dagli Oled, mentre le vendite dei televisori al Plasma sono velocemente diminuite fino ad estinguersi. In generale si osserva una tendenza delle vendite in notevole aumento dall'anno 2020 fino a marzo 2022, sia dei televisori Oled che dei televisori Lcd.

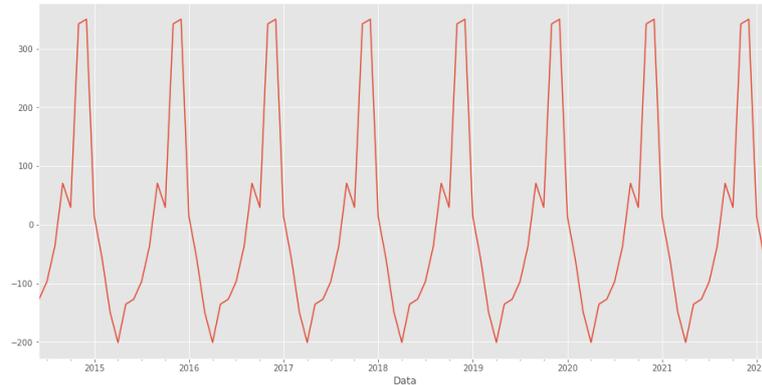


**Figure 14: scomposizione della serie storica Oled**



**Figura15: scomposizione della serie storica Lcd**

Per ogni serie storica Oled e Lcd è stato applicato il metodo della scomposizione (osservabile in Figura 14 e Figura 15 ) composto dal trend che è l'andamento medio di lungo periodo che in questo caso ha una tendenza positiva, dalla stagionalità che può avere cause differenti come l'organizzazione sociale o il ciclo stagionale; in questo caso tale ciclo mostra dei picchi ogni 12 mesi ed una componente di errore o residuale e cioè delle variazioni che non sono spiegate dalle altre componenti e quindi casuali. Il modello è composto da  $(T+S+E)$  oppure da  $(T*S*E)$  nel caso in cui la serie storica si definisce stazionaria (se la serie storica oscilla sempre all'interno dello stesso range) o evolutiva (se la serie storica ha un range crescente o decrescente).



**Figura 16: ingrandimento della tendenza stagionale della serie storica Oled**

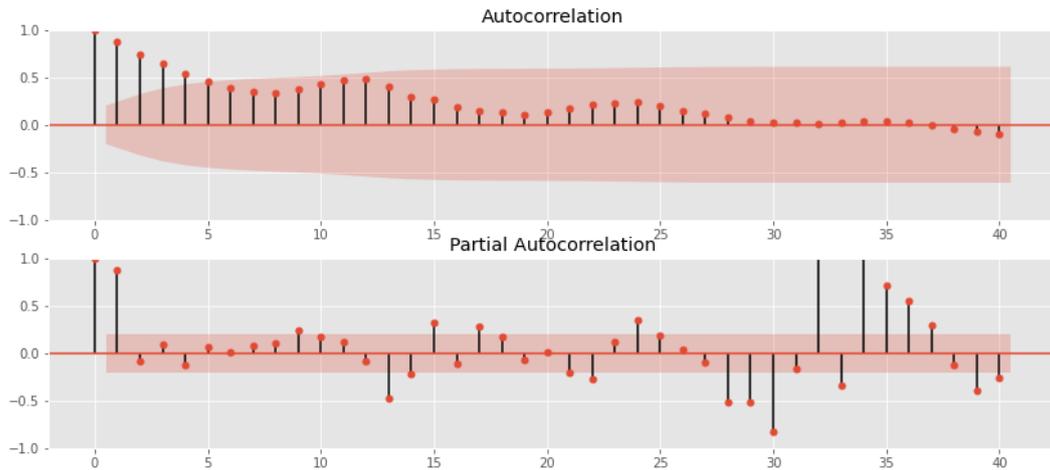
Generalmente sia per la tipologia Oled che Lcd è presente una componente stagionale che mostra un aumento significativo delle vendite dal mese di novembre fino ad arrivare al massimo livello nel mese di dicembre. Tale livello potrebbe essere connesso agli acquisti natalizi e l'intuizione è confermata dal confronto con i referenti aziendali. Per avere una conferma della corretta applicazione dei metodi SARIMA e SARIMAX (anche se in presenza di stagionalità) è necessaria la verifica della stazionarietà della serie storica attraverso l'ADF test. Di seguito è esposto il codice utilizzato nel test applicato alla serie storica dei televisori.

```
#Test stazionarietà(ADF Test)
ad_fuller_result = adfuller(df['Quantità'])
print(f'ADF Statistic: {ad_fuller_result[0]}')
print(f'p-value: {ad_fuller_result[1]}')
```

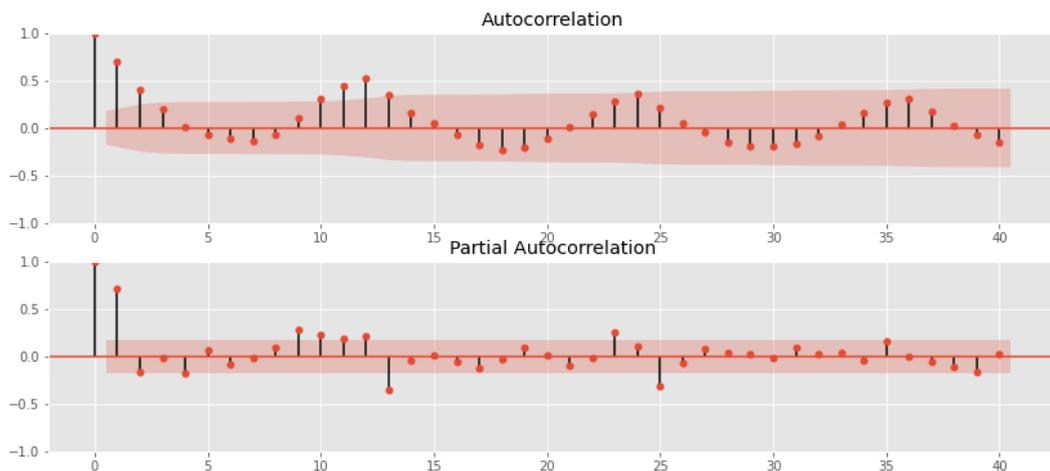
```
ADF Statistic Oled: 3.1371359983207463
p-value oled: 1.0
```

```
ADF Statistic Lcd: 1.4987082816598645
p-value lcd: 0.9975195024619786
```

Il test ha confermato che entrambe le serie storiche non sono stazionarie (P-value > 0.05) e che è possibile utilizzare i metodi sopracitati.



**Figura 17: funzioni di autocorrelazione e autocorrelazione parziale della serie storica dei televisori Oled**



**Figura 18: funzioni di autocorrelazione e autocorrelazione parziale della serie storica dei televisori Lcd**

Un altro test necessario è quello inerente all'autocorrelazione e all'autocorrelazione parziale delle due serie storiche in esame. L'autocorrelazione indica se l'osservazione presente è correlata all'osservazione precedente (che è influenzata da quella ancora antecedente ecc...) e serve a confermare la presenza di stagionalità, mentre l'autocorrelazione parziale è più puntuale, in quanto cerca di verificare la correlazione escludendo l'influenza derivante dalle osservazioni antecedenti la precedente. In Figura 17 è possibile osservare in rosa il limite della significatività delle autocorrelazioni e delle autocorrelazioni parziali. In entrambi i

casi si possono notare dei risultati statisticamente significativi. Nella autocorrelazione dei televisori Oled è possibile osservare che solamente i primi dati sono statisticamente significativi, ma osservando l' autocorrelazione parziale si possono notare dei valori statisticamente significativi all'incirca ogni 12 mesi. Per i televisori Lcd è più semplice individuare delle autocorrelazioni al limite della significatività, ma che vengono meglio evidenziate nel grafico di autocorrelazione parziale.

Come si può osservare tutti e tre i metodi di previsione prendono in considerazione la ciclicità della serie storica passata per prevedere i dati futuri. Le previsioni del metodo SARIMA sono simili alla tendenza dei dati passati. Il modello utilizza i seguenti parametri:

$$\text{SARIMA} = (p, d, q)(P, D, Q)[S]$$

Questo modello dà rilevanza alla variabile endogena costituita dalle vendite effettuate mensilmente e alla presenza di stagionalità [S]. I parametri P,p; D,d e Q,q si riferiscono al numero di termini autocorrelati della serie stazionaria osservabili dalla funzione di autocorrelazione parziale, al numero di volte che deve essere applicato il processo di differenziazione per rendere una serie stazionaria e i parametri q e Q al numero di termini della media mobile autocorrelati e cioè i ritardi o "lags" della serie storica osservabili dalla funzione di autocorrelazione. Il modello è riassumibile nella seguente funzione:

$$Y = \alpha + \beta X + \varepsilon$$

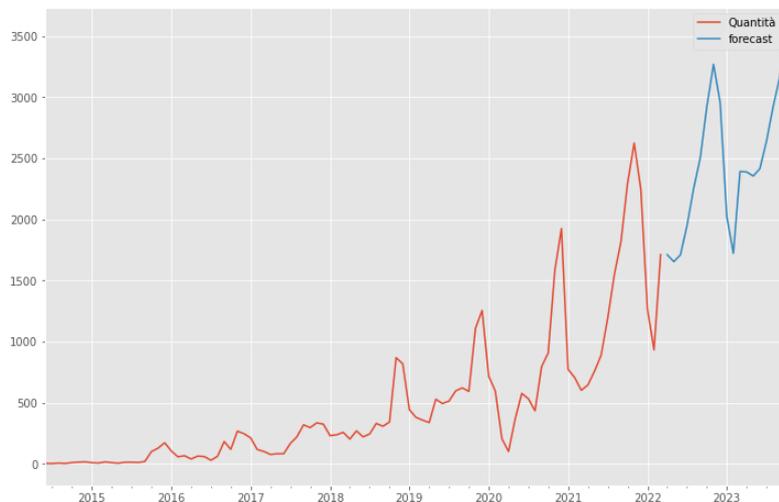
- $\alpha$  è l'intercetta della retta di regressione
- $\beta X$  è il coefficiente angolare che moltiplica la variabile endogena X
- $\varepsilon$  è l'errore
- Nell'elaborato presente l'equazione è:

$$Y = \alpha + \beta \text{Vendite} + \varepsilon$$

Il modello SARIMA è stato applicato utilizzando il seguente codice:

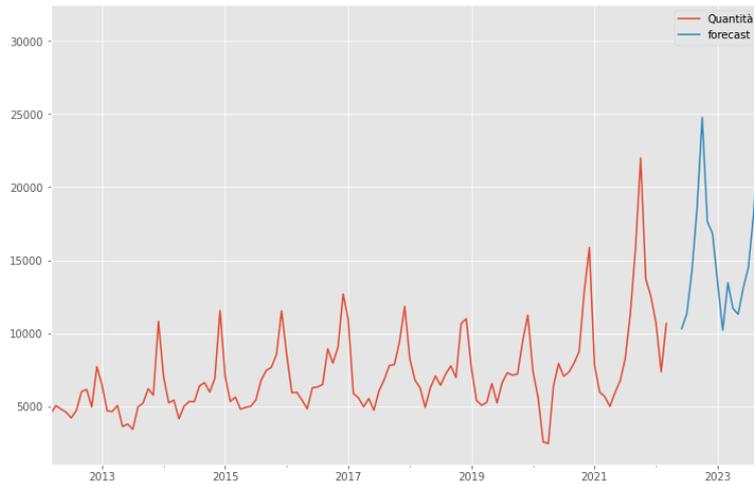
```
#Modello SARIMA
model=sm.tsa.statespace.SARIMAX(df['Quantità'],order=(1,0,2),s
easonal_order=(1, 2, 1, 12) )
res=model.fit()
print(res.summary())
```

Dove i parametri  $p$  e  $P$  (sono il numero di termini della serie stazionaria osservabili nell'autocorrelazione parziale),  $d$  e  $D$  (indicano il numero di volte per cui il processo di differenziazione deve essere applicato),  $q$  e  $Q$  (indicano il numero di termini autocorrelati osservati nel correlogramma) ed  $S$  (indica la stagionalità dei dati). Il parametro più importante che indica se i parametri sono corretti è il valore AIAC osservabile nell'output della Figura 26 in appendice, per cui la combinazione di valori che dimostra di eseguire un modello con il minor valore AIAC è definito il migliore possibile.



**Figura 19: previsione dei televisori Oled con il modello Sarima**

In Figura 19 è possibile osservare le previsioni con il modello SARIMA inerenti le vendite dei televisori Oled. Le previsioni mostrano un picco di vendite mai registrato in precedenza a dicembre 2022 di 3500 unità.



**Figura 20: previsione dei televisori Lcd con il modello Sarima**

In Figura 20 si possono osservare le previsioni dei televisori Lcd e si rileva che a dicembre 2022 si osserva un picco di vendite di 25000 unità.

Il modello SARIMAX ha gli stessi parametri del modello SARIMA: SARIMA = (p, d, q)(P, D, Q)[S]. La differenza tra i due modelli è che il modello SARIMAX oltre lo studio della variabile endogena (le vendite) e della stagionalità prende in considerazione una o più variabili esogene, ovvero delle variabili che aiutano a spiegare meglio l'andamento delle vendite e quindi il fine è quello di ottenere delle previsioni più precise. Il modello è identificabile nella seguente equazione:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

$\alpha$  indica il valore teorico di Y quando tutti i k regressori  $X_k$  valgono 0.

- $\beta_i$  indica l'entità della variazione teorica di Y ad una variazione unitaria del corrispondente regressore  $X_i$ , tenendo sotto controllo tutte le altre variabili esplicative, ovvero quando tutti gli altri regressori sono tenuti fissi.
- I k  $\beta_i$  sono chiamati coefficienti di regressione parziale: essi rappresentano infatti l'impatto del singolo regressore associato sulla Y, tenendo costanti tutti gli altri regressori.

- $\varepsilon$  rappresenta quella parte di  $Y$  che non può essere spiegata congiuntamente da tutti i regressori inclusi nel modello.

L'equazione del presente caso può essere riassunta come:

$$Y = \alpha + \beta_1 \text{PIL} + \beta_2 \text{Dummy Covid} + \beta_3 \text{Dummy Incentivi} + \beta_4 \text{Feb} + \dots + \beta_{15} \text{Dic} + \varepsilon$$

Per il modello SARIMAX è stato utilizzato il codice seguente:

```
#Modello SARIMAX
model=sm.tsa.statespace.SARIMAX(df['Quantità'],exog=df[['exog'
]],order=(1,0,2),seasonal_order=(1, 2, 1, 12) )
res=model.fit()
print(res.summary())
```

In tale codice i valori dei parametri sono stati scelti utilizzando la stessa tecnica del modello SARIMA. Si può osservare che nel codice è presente l'inserimento delle variabili esogene. Le variabili esogene sono il valore del PIL, la dummy Covid e la dummy Incentivi<sup>42</sup>. In particolar modo è risultato interessante osservare che, inserendo la variabile esogena della categoria di televisori concorrenti si ha un coefficiente positivo, in quanto le due serie sono correlate positivamente ed è quindi palese che l'andamento delle vendite di una categoria non è utile a spiegare l'andamento della categoria concorrente. Infatti, la crescita di entrambe le tipologie di televisori è spiegata in modo migliore dalle altre variabili.

---

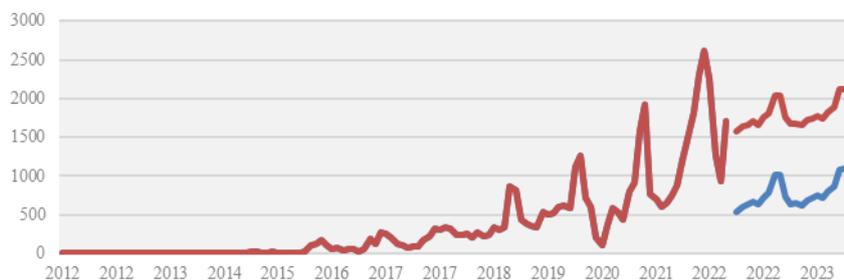
<sup>42</sup> Le dummy sono variabili semplici dove vi sono solamente due tipologie di dati: 1 quando il fenomeno si verifica e 0 quando è assente.

OUTPUT RIEPILOGO									
statistica della regressione									
R multiplo	0,82496811								
R al quadrat	0,68057239								
R al quadrat	0,63916511								
Errore stand	1718,35698								
Osservazion	123								
ANALISI VARIANZA									
	<i>gdl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>significatività F</i>				
Regressione	14	679442037	48531574,1	16,4360554	9,0258E-21				
Residuo	108	318897076	2952750,71						
Totale	122	998339114							
	<i>Coefficienti</i>	<i>errore standa</i>	<i>Stat t</i>	<i>p</i>	<i>inferiore 95%</i>	<i>superiore 95%</i>	<i>inferiore 95,0%</i>	<i>superiore 95,0%</i>	
Intercetta	-7113,1287	4243,29641	-1,6763214	0,09656778	-15524,078	1297,8209	-15524,078	1297,8209	
PIL	0,00856887	0,00248521	3,44793973	0,00080628	0,00364274	0,01349499	0,00364274	0,01349499	
LokD	2240,88878	746,626481	3,00135186	0,00333881	760,945564	3720,83199	760,945564	3720,83199	
Inc/Fre	3096,69521	527,638073	5,86897604	4,8845E-08	2050,82501	4142,5654	2050,82501	4142,5654	
feb	-2344,9549	736,964306	-3,1819111	0,00191076	-3805,746	-884,16377	-3805,746	-884,16377	
mar	-2473,8141	736,85943	-3,3572402	0,00108782	-3934,3973	-1013,2309	-3934,3973	-1013,2309	
apr	-2969,4476	754,589381	-3,9351834	0,00014739	-4465,1747	-1473,7205	-4465,1747	-1473,7205	
mag	-2285,2476	754,589381	-3,0284651	0,00307485	-3780,9747	-789,52053	-3780,9747	-789,52053	
giu	-2006,8476	754,589381	-2,6595227	0,00901621	-3502,5747	-511,12053	-3502,5747	-511,12053	
lug	-1758,3476	754,589381	-2,3302045	0,02165373	-3254,0747	-262,62053	-3254,0747	-262,62053	
ago	-959,13647	751,271019	-1,276685	0,20445176	-2448,286	530,013026	-2448,286	530,013026	
set	173,663534	751,271019	0,23115963	0,81762815	-1315,486	1662,81303	-1315,486	1662,81303	
ott	776,563534	751,271019	1,0336663	0,30360129	-712,58596	2265,71303	-712,58596	2265,71303	
nov	1237,66353	751,271019	1,64742617	0,10237706	-251,48596	2726,81303	-251,48596	2726,81303	
dic	3782,46353	751,271019	5,03475236	1,9265E-06	2293,31404	5271,61303	2293,31404	5271,61303	

Figura 21: output del modello Sarimax applicato alla serie storica Oled

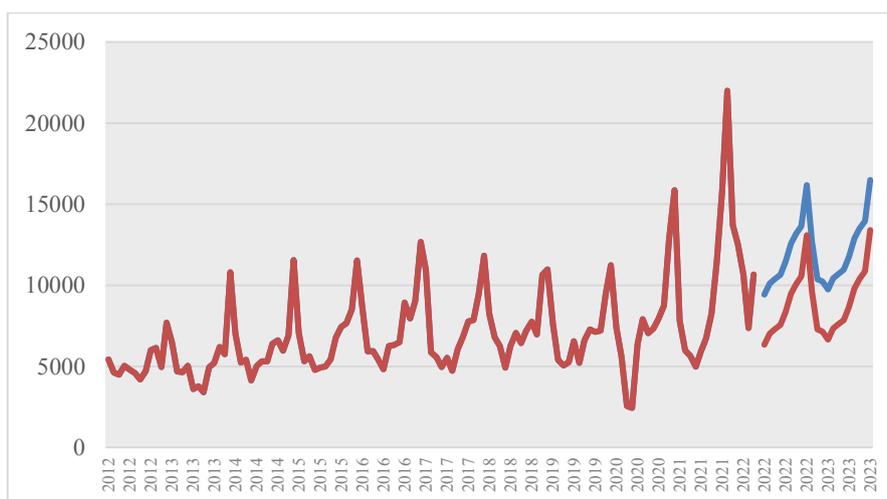
La variabile che ha destato più interesse per l'impatto che ha avuto sull'andamento delle vendite è la variabile denominata 'incentivi' che si riferisce agli incentivi elargiti per quanto riguarda la rottamazione dei vecchi apparecchi televisivi, a favore di quelli più recenti, secondo l'output del modello osservabile in Figura 21 è la variabile più statisticamente significativa (parametro evidenziato in verde). Come osservabile dalla Figura 21 anche la variabile PIL è statisticamente significativa (parametro evidenziato in verde) e quindi serve a spiegare una parte dell'andamento delle vendite. L'ultimo regressore utilizzato è la dummy Covid o più precisamente lock down. Anche quest'ultimo regressore è statisticamente significativo e ciò a conferma delle analisi effettuate dall'azienda dell'impatto positivo che questo fenomeno ha avuto sulle vendite.

In Figura 22 e Figura 23 è possibile osservare come l'inserimento o l'omissione della variabile incentivi modifichi la previsione.



**Figura 22: previsioni dei televisori Oled con il modello Sarimax**

Nella Figura 22 si osserva in rosso la previsione effettuata con i regressori Pil, Covid e Incentivi. I valori del mese di dicembre 2022 sono notevolmente inferiori rispetto al modello SARIMA. Questo è deducibile dal fatto che il Pil ha un andamento più lineare in confronto alla sola variabile endogena che fa abbassare il valore massimo dei televisori a 2000 unità vendute. In blu è tracciata la previsione dei televisori Oled senza l'utilizzo della variabile Incentivi. Questo sta a significare che per la tipologia di televisori Oled, secondo le previsioni, gli incentivi impattano del 100% sulle vendite.



**Figura 23: previsioni dei televisori Lcd con il modello Sarimax**

Nella Figura 23 si osserva in blu la previsione effettuata con i regressori Pil, Covid e Incentivi. I valori del mese di dicembre 2022 sono notevolmente inferiori rispetto al modello SARIMA per lo stesso motivo del grafico precedente ed in tale grafico si ha un volume di vendite massimo nel mese di dicembre 2022 di 16000 unità vendute. In rosso è tracciata la previsione dei televisori Lcd senza l'utilizzo della variabile Incentivi. Si può notare, osservando il grafico, che l'impatto che ha la variabile incentivi sulla categoria Lcd, secondo il modello di previsione, è del 23% e cioè inferiore rispetto a quella degli Oled del 77%.

Si deduce che l'andamento economico, pandemico e gli incentivi hanno un effetto positivo per entrambe le categorie. Il rallentamento delle vendite dovute all'arresto dell'erogazione del bonus rottamazione (causa esaurimento fondi) e l'eventuale acquisto obbligatorio entro il 2023 potrebbero portare ad un aumento del volume di vendita e quindi del fatturato anche nei mesi successivi a dicembre 2022. Come ultimo metodo di previsione si è deciso di applicare il modello LSTM. Il modello in questione è una tipologia di rete neurale artificiale ricorrente con una tipologia di apprendimento supervisionata. Questa tipologia di algoritmo utilizza solamente i dati inerenti le vendite passate per prevedere i volumi di vendita futuri e ciò lo rende confrontabile con il modello SARIMA. Il modello LSTM è stato applicato con il seguente codice:

```
model = tf.keras.models.Sequential()  
model.add(tf.keras.layers.LSTM(50, activation = 'relu', input_  
shape=(10, 1)))  
model.add(tf.keras.layers.Dense(10))  
model.compile(loss = 'mse', optimizer='adam')  
model.fit(X_train, Y_train, epochs=200)
```

```
Epoch1/2004/4[=====]-2s 8ms/step-  
loss:0.0971  
Epoch200/2004/4[=====]-0s 10ms/step-  
loss: 0.0095
```

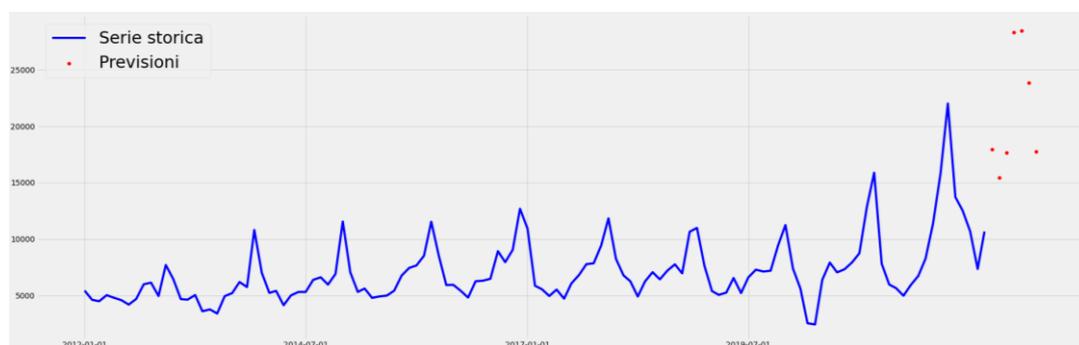
Il modello è costituito da 50 neuroni distribuiti in 10 layers e 200 epoche di apprendimento della rete neurale artificiale che ha permesso un abbassamento rilevante della varianza (il dato mean squared error è passato da 0.09 a 0.009).

In Figura 24 e Figura 25 è possibile osservare il risultato del modello LSTM.



**Figura 24: previsione dei televisori Oled con il metodo LSTM**

Come per gli altri due metodi utilizzati si evidenzia un rallentamento delle vendite nei mesi successivi a marzo e una ripresa verso novembre e dicembre 2023. Confrontando la Figura 24 con le previsioni del modello SARIMA che prevede un volume massimo di 3500 televisori venduti il LSTM definisce una stasi nei mesi estivi, per poi arrivare al massimo nel mese di dicembre, con un volume di vendita previsto di 5000 televisori stimando un valore del 43% più elevato rispetto al modello SARIMA.



**Figura 25: previsioni dei televisori Lcd con il modello LSTM**

Per i televisori Lcd le cifre sono sempre superiori a quelle stimate dal modello SARIMA, più precisamente il LSTM stima un volume di vendite massime nel mese di dicembre di 30000 unità mentre il volume di vendita stimato dal modello SARIMA è di 25000 unità. Il LSTM prevede un volume di vendite superiore del 20% rispetto al modello SARIMA. Il problema di sovrastima dei volumi di vendita è comune a questa tipologia di algoritmo in quanto è molto sensibile agli outlier, ovvero ai valori anomali, cioè al raggiungimento di volumi venduti che mai si furono verificati negli anni passati.

### **3.4. Conclusioni**

L'utilizzo di nuovi software e di dati sempre più precisi ha permesso un'evoluzione dei metodi di previsione che permetterà nel prossimo futuro di ottenere delle informazioni future sempre più precise. I metodi utilizzati hanno reso possibile individuare uno schema ripetitivo (trend e stagionalità) ed effettuare delle previsioni sempre più realistiche. Delle quattro tipologie di televisori si sono effettuate le analisi dei prodotti che hanno una rilevanza maggiore per quanto riguarda il volume di vendita. I televisori Lcd e Oled hanno in comune un aumento delle vendite verso la fine di ogni anno che è anche stato evidenziato dalle previsioni dei mesi di novembre e dicembre 2022. L'analisi inoltre ha dimostrato che i metodi quantitativi sono più efficaci se interpretati ed integrati con le informazioni del settore di appartenenza del prodotto. Alcune variazioni delle serie storiche sono state causate dal Covid-19 e dagli incentivi forniti per la rottamazione del vecchio televisore. Questi fenomeni hanno influito positivamente sulle vendite ma non erano certamente prevedibili ed è arduo prevedere esattamente come la pandemia evolverà. Per quanto riguarda gli "incentivi rottamazione" saranno elargiti fino a dicembre 2022. Un ulteriore dato di fondamentale importanza è lo switch-off delle frequenze standard in favore della trasmissione delle frequenze in alta definizione, motivo per cui è stato elargito il bonus e il conseguente aumento delle vendite. Si prevede che questo cambiamento delle frequenze avverrà entro il 2023 su tutto il territorio nazionale e che sia l'impulso principale che farà aumentare le vendite a fine 2022.

## Appendice

SARIMAX Results						
Dep. Variable:	Quantità		No. Observations:	121		
Model:	SARIMAX(2, 0, 2)x(1, 2, [1], 12)		Log Likelihood	-870.791		
Date:	Thu, 30 Jun 2022		AIC	1755.581		
Time:	14:09:00		BIC	1773.604		
Sample:	03-01-2012		HQIC	1762.869		
	- 03-01-2022					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.7913	0.270	2.929	0.003	0.262	1.321
ar.L2	0.2077	0.267	0.779	0.436	-0.315	0.731
ma.L1	-0.2354	0.271	-0.870	0.385	-0.766	0.295
ma.L2	-0.5452	0.189	-2.881	0.004	-0.916	-0.174
ar.S.L12	-0.2781	0.208	-1.336	0.182	-0.686	0.130
ma.S.L12	-0.7626	0.240	-3.175	0.001	-1.233	-0.292
sigma2	3.507e+06	2.75e+05	12.760	0.000	2.97e+06	4.05e+06
Ljung-Box (L1) (Q):			0.01	Jarque-Bera (JB):	153.96	
Prob(Q):			0.93	Prob(JB):	0.00	
Heteroskedasticity (H):			6.74	Skew:	0.98	
Prob(H) (two-sided):			0.00	Kurtosis:	8.85	

Figura 26: output modello Sarima applicato alla serie storica Oled

## INDICE DELLE FIGURE

Figura 1: il posizionamento e le interazioni della figura del previsore all'interno di una impresa .....	7
Figura 2: trasmissione delle previsioni di vendita .....	11
Figura 3: serie storica del consumo di gas nel Regno Unito .....	19
Figura 4: serie storica delle vendite di filtri per l'olio della John Deer .....	20
Figura 5: serie storica delle vendite di filtri per l'olio della John Deer .....	21
Figura 6: correlogramma della serie storica soprariportata .....	23
Figura 7: confronto tra un neurone biologico e uno artificiale .....	28
Figura 8 : rappresentazione di un neurone artificiale .....	29
Figura 9: struttura di una rete neurale artificiale.....	30
Figura 10: l'architettura di una rete neurale ricorrente .....	33
Figura 11: Long short term memory .....	35
Figura 12: test del modello per prevedere il volume (Kg) di lattughe vendute .....	37
Figura 13: serie storica dei televisori LCD (in blu), la serie storica dei televisori al plasma (in rosso) e la serie storica dei televisori OLED (in verde) .....	44
Figura 14: scomposizione della serie storica Oled .....	45
Figura 15: scomposizione della serie storica Lcd .....	45
Figura 16: ingrandimento della tendenza stagionale della serie storica Oled.....	46
Figura 17: funzioni di autocorrelazione e autocorrelazione parziale della serie storica dei televisori Oled .....	47
Figura 18: funzioni di autocorrelazione e autocorrelazione parziale della serie storica dei televisori Lcd.....	47
Figura 19: previsione dei televisori Oled con il modello Sarima .....	50
Figura 20: previsione dei televisori Lcd con il modello Sarima.....	51
Figura 21: output del modello Sarimax applicato alla serie storica dei televisori Oled.....	53
Figura 22: previsioni dei televisori Oled con il modello Sarimax.....	54

Figura 23:previsioni dei televisori Lcd con il modello Sarimax .....	54
Figura 24:previsione dei televisori Oled con il metodo LSTM.....	56
Figura 25: previsioni dei televisori Lcd con il modello LSTM .....	56
Figura 26: output modello Sarima applicato alla serie storica Oled.....	59

## BIBLIOGRAFIA

Accounting Review, 183-198. Mendenhall, R. R. (1991). Evidence on the possible underweighting of earnings-related information. *Journal of Accounting Research*, 29(1), 170-179.

Ali, A., Klein, A., & Rosenfeld, J. (1992). Analysts' use of information about permanent and transitory earnings components in forecasting annual EPS.

Arunraj, N. S., Ahrens, D., & Fernandes, M. (2016). Application of SARIMAX model to forecast daily sales in food retail industry. *International Journal of Operations Research and Information Systems (IJORIS)*, 7(2), 1-21.

Bellagamba, A. (2018) Previsione delle vendite e pianificazione della produzione: quale ruolo riveste la “forecast accuracy”. *Electronic journal of management*, 2-9.

Chung, S., Lee, D. D., & Sompolinsky, H. (2015). Classification of Manifolds by Single-Layer Neural Networks. *CoRR*.

Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R (Vol. 2)*. New York: Springer.

Dagum, E. B. (2001). *Analisi delle serie storiche: modellistica, previsione e scomposizione*. Springer Science & Business Media.

DALRYMPLE, Douglas J. Sales forecasting methods and accuracy. *Business Horizons*, 1975, 18.6: 69-73.

Demuth, H. B., Beale, M. H., De Jess, O., and Hagan, M. T. (2014). *Neural network design*. Martin Hagan.

Diamantopoulos, A., & Winklhofer, H. (2003). Export sales forecasting by UK firms: Technique utilization and impact on forecast accuracy. *Journal of Business Research*, 56(1), 45-54.

Divisekara, R. W., Jayasinghe, G. J. M. S. R., & Kumari, K. W. S. N. (2021). Forecasting the red lentils commodity market price using SARIMA models. *SN Business & Economics*, 1(1), 1-13.

Elgers, P. T., Lo, M. H., & Murray, D. (1995). Note on adjustments to analysts' earnings forecasts based upon systematic cross-sectional components of prior-period errors. *Management Science*, 41(8), 1392-1396.

Fildes, R., & Hastings, R. (1994). The organization and improvement of market forecasting. *Journal of the Operational Research Society*, 45(1), 1-16.

Fitch, F. B. (1944). Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biophysics*, vol. 5 (1943), pp. 115–133. *The Journal of Symbolic Logic*, 9(2), 49-50.

Ghosh, P., Neufeld, A., & Sahoo, J. K. (2022). Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. *Finance Research Letters*, 46, 102280.

Gurney, K. (2018). *An introduction to neural networks*. CRC press.

Hassani, H., & Silva, E. S. (2015). Forecasting with big data: A review. *Annals of Data Science*, 2(1), 5-19.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of long short-term memory (LSTM) neural network for flood forecasting. *Water*, 11(7), 1387.

Marieb, E., Hoehn, K., Martini, F. H., Nath, J. L., Bartholomew, E. F., McKinley, M. P., ... & Russ, A. (2015). *Essentials of Human Anatomy*.

- Masarotto, G. (2005). *Analisi delle Serie Temporali con R. Facoltà di Scienze Statistiche Università degli Studi di Padova.*
- Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. F. (1993). The M2-competition: A real-time judgmentally based forecasting study. *International Journal of forecasting*, 9(1), 5-22.
- Meng, M., & Song, C. (2020). Daily photovoltaic power generation forecasting model based on random forest algorithm for north China in winter. *Sustainability*, 12(6), 2247.
- Mentzer, J. T., & Moon, M. A. (2005). *Sales forecasting management. A demand management approach (2nd Edition).* Thousand Oaks (London): Sage Publications.
- Morgan, J. (2014). *Classification and regression tree analysis. Boston: Boston University, 298.*
- Nasser, I. M., & Abu-Naser, S. S. (2019). Artificial neural network for predicting animals category.
- Nielsen, M. A. (2015). *Neural networks and deep learning (Vol. 25).* San Francisco, CA, USA: Determination press.
- Sanders, N. R., & Manrodt, K. B. (2003). Forecasting software in practice: Use, satisfaction, and performance. *Interfaces*, 33(5), 90-93.
- Sanders, N. R., & Ritzman, L. P. (2004). Integrating judgmental and quantitative forecasts: methodologies for pooling marketing and operations information. *International Journal of Operations & Production Management.*
- Son, N., Yang, S., & Na, J. (2020). Deep neural network and long short-term memory for electric power load forecasting. *Applied Sciences*, 10(18), 6489.
- Webby, R., & O'Connor, M. (1996). Judgmental and statistical time series forecasting: a review of the literature. *International Journal of forecasting*, 12(1), 91-118.

Winklhofer, H., Diamantopoulos, A., & Witt, S. F. (1996). Forecasting practice: A review of the empirical literature and an agenda for future research. *International Journal of forecasting*, 12(2), 193-221.

Zaytar, M. A., & El Amrani, C. (2016). Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *International Journal of Computer Applications*, 143(11), 7-11.