**A Comparative study of Machine learning Approaches for Demand Forecasting in online retail**

Muhammad Ahsan

A Thesis Submitted in Partial Fulfilment

of the requirements for the

Degree of

Master of Science in Data Analytics

Diagram

Description automatically generated with medium confidence

17 May 2024

Supervisor: Sam Weiss

**DECLARATION**

I, Muhammad Ahsan hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Signed: Muhammad Ahsan

Student Number: 2023132

Date: 17/05/2024

**Acknowledgements**

Dear Sam Weiss, I wish to convey my heartfelt appreciation for your unwavering support and guidance as my research supervisor. Your patience and constructive recommendations have been invaluable in shaping and enhancing this research project. Your willingness to generously invest your time has been instrumental in its development. Additionally, I extend my deepest gratitude to my family for their constant support and motivation throughout this journey.

Abstract:

The modern online retail industry sees demand forecasting as playing a crucial role in inventory management efficiency and customer satisfaction. Traditional forecasting is not effective in precisely predicting demand due to the increasing complexity and dynamic nature of online retail. Therefore, the research paper in question suggests a comparative analysis of machine-learning approaches to improving demand forecasting in online retail. This paper, therefore, aims at evaluating the performance of several distinct machine learning tools in predicting retail product demand. To reach the goal, it was essential to conduct a literature review and research different machine learning tools and the instances in which they could be used for demand forecasting. This research aims to use some of the traditional machine learning models and deep learning models to perform demand forecasting. Some of the models to be applied to transactional dataset (Online Retail II) include Auto Regressive Integrated Moving Average (ARIMA), Seasonal Auto Regressive Integrated Moving Average (SARIMA), Long Short-Term Memory (LSTM) and with a focus on Darts models, including NBEATSModel, BlockRNNTest, TiDEModel, and TCNModel. In addition, the research will also take into account some of the factors that might affect demand forecasting such as seasonality, trend and customer behaviour.

After conducting an extensive study and evaluation of different machine learning algorithms for demand forecasting in online retail, the results revealed some interesting insights. The ARIMA model provided unsatisfactory results because of neglection in seasonality and non-linear patterns in the data. SARIMA, with its ability to capture seasonal patterns, outperformed ARIMA in accurately predicting sales fluctuations over time. On the other hand, LSTM, being a deep learning model with memory retention capabilities, excelled in capturing complex temporal dependencies and non-linear trends within the data. However, it should be noted that SARIMA model presented a challenge in their predictions, often yielding negative values. This indicates that the forecasted demand occasionally fell below zero, which is not practically meaningful in the context of online retail. To address the limitations of traditional models, the study explored the application of the DARTS library, which encompasses advanced deep learning models designed specifically for time series forecasting. Among these, the NBeats, BlockRNNTest, TiDE, and TCN exhibited superior performance compared to traditional machine learning models. However, the TCN, stood out as the most accurate and superior model for demand forecasting in online retail. The comparative study highlighted the significance of leveraging advanced machine learning models such as TCN for accurate and reliable demand forecasting in the online retail industry.

**Table of Contents**

[Abstract: iv](#_Toc166638455)

[1. **Introduction**: 1](#_Toc166638456)

[1.1. Overview: 1](#_Toc166638457)

[1.2. Background: 1](#_Toc166638458)

[1.3. Research Objectives: 2](#_Toc166638459)

[1.4. Research Questions: 3](#_Toc166638460)

[1.5. Outline of the thesis: 4](#_Toc166638461)

[2. **Literature Review**: 4](#_Toc166638462)

[2.1. A Comparative Study of Demand Forecasting Models for a Multi-Channel Retail Company: 5](#_Toc166638463)

[2.2. Time-series forecasting of seasonal items sales using machine learning: 5](#_Toc166638464)

[2.3. Developing and Preliminary Testing of a Machine Learning-Based Platform for Sales Forecasting Using a Gradient Boosting Approach: 5](#_Toc166638465)

[2.4. Intelligent Sales Prediction Using Machine Learning Techniques: 6](#_Toc166638466)

[2.5. Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail: 6](#_Toc166638467)

[2.6. A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks: 7](#_Toc166638468)

[2.7. Conclusion: 7](#_Toc166638469)

[3. **Research Methodology**: 8](#_Toc166638470)

[3.1. CRISP-DM Methodology: 9](#_Toc166638471)

[3.1.1. Business Understanding: 10](#_Toc166638472)

[3.1.2. Data Understanding: 12](#_Toc166638473)

[3.1.3. Data Preparation: 13](#_Toc166638474)

[3.1.4. Modelling: 14](#_Toc166638475)

[4. **Evaluation and Analysis**: 25](#_Toc166638476)

[4.1. Understanding the data: 25](#_Toc166638477)

[4.2. Data Cleaning and Preprocessing: 26](#_Toc166638478)

[4.3. Assessment of Stationarity using Augmented Dickey-Fuller Test: 29](#_Toc166638479)

[4.4. Rescaling and Normalization: 30](#_Toc166638480)

[4.5. Model Performance: 32](#_Toc166638481)

[4.5.1. Parameters for Measuring Model Performance: 32](#_Toc166638482)

[4.6. Model Results: 33](#_Toc166638483)

[4.6.1. ARIMA model: 33](#_Toc166638484)

[4.6.2. SARIMA Model: 34](#_Toc166638485)

[4.6.3. Long Short-term Memory: 35](#_Toc166638486)

[4.6.4. NBeats Model: 36](#_Toc166638487)

[4.6.5. TiDE Model: 38](#_Toc166638488)

[4.6.6. BlockRNN Model: 40](#_Toc166638489)

[4.6.7. TCN Model: 42](#_Toc166638490)

[4.7. Comparison of Models: 44](#_Toc166638491)

[4.7.1. Time Taken by DART models: 44](#_Toc166638492)

[4.7.2. Results of DART Models with Top 10 Selling Products: 45](#_Toc166638493)

[5. **Conclusion, Limitations and Recommendations**: 47](#_Toc166638494)

[5.1. Conclusion: 47](#_Toc166638495)

[5.2. Future Recommendations: 48](#_Toc166638496)

[5.3. Limitations: 49](#_Toc166638497)

[6. References: 50](#_Toc166638498)

**Abbreviations:**

ARIMA AutoRegressive Integrated Moving Average

SARIMA Seasonal AutoRegressive Integrated Moving Average

LSTM Long Short-Term Memory

DARTS Differentiable Architecture Search

NBeats Neural Basis Expansion Analysis for Time Series

BlockRNN Block-wise Recursive Neural Network

TiDE Temporal Convolutional Network with Integral Dynamics

TCN Temporal Convolutional Network

MAE Mean Absolute Error

MSE Mean Squared Error

RMSE Root Mean Squared Error

CSV Comma-Separated Valu

Introduction:

* 1. Overview:

In simple terms, this project efforts on comparing different time series models. Why? The objective is to predict the quantity needed, helping businesses improve inventory management and reduce costs with this comparative study. By figuring out which model works best, this research aim is to improve the accuracy of predicting how much stock is required, ultimately making inventory management more efficient and cost-effective. The main objective is to provides the context for the comparative analysis and emphasizes the need to identify the best approach for improved predictive accuracy. This contrast is crucial for businesses which was observing to improve their operations and make better decisions about stocking their products.

* 1. Background:

The accurate forecasting of demand in online retail is vital for inventory management, resource allocation, and overall business demand forecasting plays a serious role in the supply chain management of online retail businesses. Controlling the retail market is the secret to sustainability in today's business world. By accurately predicting demand, online retailers can optimize inventory levels, minimize stockouts and overstocks, reduce costs, and enhance customer satisfaction (Ferreira, J, K., Lee, A, H, B. and Simchi‐Levi, D., 2016). Nevertheless, traditional forecasting methods may fall rapid in taking the complexity and dynamic nature of the online retail industry. To address this challenge, machine learning techniques have gained prominence as effective tools for demand forecasting in online retail. These techniques leverage the power of advanced algorithms and large datasets to extract patterns, trends, and insights that can improve the accuracy of demand forecasts **(**Ferreira, J, K., Lee, A, H, B. and Simchi‐Levi, D., 2016**)**. Several studies have investigated various analytical techniques, from traditional time-series-based models to advanced neural network architectures. This academic journey involves integrating a wide range of machine learning algorithms, including, but not limited to Support Vector Machine, K-Nearest Neighbour, Gaussian Nave Bayes, regressions, Random Forest, Decision Tree Regressor, and Extreme Gradient Boosting (XGBoost) models. Notably, the Gaussian Nave Bayes algorithm emerges as a standout performer, showcasing unparalleled accuracy in the demand estimation **(**Wang, J., Liu, Q, G. and Liu, L., 2019**)**. This revelation has promising implications for retailers navigating the dynamic online marketplace, enabling them to anticipate and meet product demand. As the online retail space continues to redefine consumer interactions, robust demand forecasting becomes increasingly important for businesses aiming to optimize their supply chain and operational efficiency. The other study provides a detailed look at the changing landscape of time series forecasting methods, highlighting the growing use of machine learning over traditional ARIMA models. The research compares the performance of AI algorithms with ARIMA and hybrid statistical-AI models, emphasizing the superior predictive accuracy of machine learning in retail sales forecasting **(**Kontopoulou, I, V. et al., 2023**)**. This comparative study aims to contribute to the existing literature by examining and comparing the performance of various machine learning approaches for demand forecasting in the online retail sector. On the other hand, the work introduces a novel forecasting method that combines Long Short-Term Memory (LSTM) networks and Random Forest (RF), showcasing superior performance compared to traditional forecasting techniques **(**Punia, S. et al., 2020**)**. This study highlights the robustness and suitability of advanced machine learning algorithms in multi-channel retail demand forecasting, further corroborating the trend observed in contemporary research. The Decision Tree Algorithm achieved nearly 71% overall accuracy, while the Generalized Linear Model attained 64% accuracy. Notably, this study identifies the Gradient Boosted Tree as the best-fit model for sales forecasting, highlighting its superior performance compared to other models evaluated. In conclusion, the use of machine learning approaches for demand forecasting in online retail has shown significant promise and outperformed traditional methods such as ARIMA. In spite of the advancements in machine learning techniques for demand forecasting in online retail, there was still unresolved challenges. For example, there is a need to leverage real-time data and update forecasting models in a changing environment (Ferreira, J, K., Lee, A, H, B. and Simchi‐Levi, D., 2016). Additionally, there is a lack of research on how to leverage demand and supply information throughout the product lifecycle, rather than relying on a single data point **(**Liu, N. et al., 2013**)**. So, this research will more explore the potential of machine learning approaches for demand forecasting in online retail and identify the most effective models for accurate predictions on time series dataset.

* 1. Research Objectives:

This research aims to hypothesis a strong and versatile framework for machine learning approaches modified to the complex domain of online retail forecasting. This research provides context for the comparative analysis and emphasizes the need to identify the best approach for improved predictive accuracy.

1. The first objective is to perform Exploratory Data Analysis on real-world online retail dataset (Online Retail II) to understand the characteristics of the data and identify significant predictor variable.
2. The second objective is to conduct demand forecasting using the DARTS library models, including NBeats, TiDE, TCN, and BlockRNN. This involves exploring their advantages, characteristics, and functionalities to determine their suitability for accurate demand prediction.
3. The third objective is to evaluate and compare the performance of DARTS library models to determine which one yields the most effective and accurate results for demand forecasting.
4. The final objective is to analyze the effectiveness of models in terms of generalization, runtime efficiency, and user-friendliness. Additionally, it involves identifying any limitations and proposing improvements for future research in online retail forecasting.

In struggling toward these objectives, this research attempt to provide a comprehensive guide for retail industry, offering practical insights into the utilization of machine learning approaches for demand forecasting. By conducting a comprehensive analysis of diverse forecasting methods, the study seeks to delineate the advantages and drawbacks of each approach. This endeavor empowers decision-makers to make informed selections that align closely with their unique business requirements. Additionally, the research seeks to excite further progress in the field by analytical areas where current knowledge is lacking and suggesting paths for future research, particularly in the exploration of advanced deep learning models and their application to complex time-series forecasting scenarios. Ultimately, the goal of this study is to contribute to the ongoing improvement of demand forecasting practices in online retail, facilitating more efficient operations, improved customer satisfaction, and informed decision-making in the ever-changing world of e-commerce.

* 1. Research Questions:

This study addresses the challenge of identifying the most suitable machine learning model for predicting demand in online retail. Various aspects such as accuracy, reliability, efficiency, and simplicity of implementation will be compared. The study hypothesizes that by comparing multiple machine learning paradigms including auto-regressive, traditional, and deep learning models - the best approach for accurately forecasting time series data related to online retail trends can be found. The objective is to provide businesses with a clear understanding of the most effective machine learning model to improve forecasting precision in online retail (Ensafi, Y. et al., 2022). This study aims to offer valuable insights to decision-makers, enabling them to make well-informed decisions and enhance their forecasting strategies within the dynamic online retail environment. Through the evaluation of various machine learning algorithms alongside traditional statistical methods for demand forecasting in online retail, the research endeavors to identify the most effective approaches for accurately predicting demand in this industry. Potential research inquiries in this investigation encompass:

1. What are the key characteristics of the “Online Retail II” dataset, and which variables serve as important predictors for demand forecasting in online retail?
2. How do DARTS library models, including NBeats, TiDE, TCN, and BlockRNN, compare in terms of their characteristics, and functionalities for accurate demand prediction in online retail?
3. Which DARTS library model demonstrates the highest effectiveness and accuracy for demand forecasting in online retail, and how does its performance compare to other models?
4. How do the selected models perform in terms of generalization, runtime efficiency, and user-friendliness when applied to real-world online retail forecasting scenarios? What are the limitations of these models, and how can future research in online retail forecasting be improved?
   1. Outline of the thesis:

The research journey begins with Chapter 1, the Introduction, which offers a comprehensive overview of the study's background, objectives, and research questions, setting the context for the investigation. In Chapter 2, the Literature Review, existing studies and scholarly works related to demand forecasting models and machine learning techniques are critically analysed, identifying key trends and gaps in the current research landscape. Chapter 3, the Research Methodology, outlines the research approach, detailing the CRISP-DM methodology employed throughout the study. Chapter 4, Evaluation and Analysis, rigorously evaluates the data and forecasting models used, covering various stages such as data cleaning, preprocessing, and model performance evaluation. Finally, Chapter 5, the Conclusion, Limitations, and Recommendations, synthesizes the key findings, discusses study limitations, and provides insightful recommendations for future research endeavours in demand forecasting and machine learning.

Literature Review:

Several studies have been presented on demand forecasting in the online retail sector. Principally, the studies was conducted to explore different machine learning algorithms like time-series-based models, regression-based models, and neural network models. Most of the machine learning models like Linear regression and decision trees get low accuracy as you will read in this literature review. The main goal of this research is to get high accuracy for demand forecasting by using different machine learning algorithms and recognize the best-performing models for demand forecasting. The choice of the algorithm depends on the characteristics of the online retail business because each ML algorithm offers unique advantages and limitations.

* 1. A Comparative Study of Demand Forecasting Models for a Multi-Channel Retail Company:

In this research paper by (Mitra et al., 2022), the authors conducted a comparative study on demand forecasting models for the retail sector, recognizing the pivotal role of sales forecasting and demand planning in optimizing supply chain performance. The study focused on several machine learning models, including Random Forest (RF), Extreme Gradient Boosting (XGBoost), Gradient Boosting, Adaptive Boosting (AdaBoost), and Artificial Neural Network (ANN). Notably, the hybrid model combining Random Forest, XGBoost, and Linear Regression (LR) emerged as the top performer regarding accuracy. The findings underscore the potential of machine learning approaches to enhance sales forecasting precision in retail. The proposed hybrid model outperforms other models and carries significant implications for optimizing capacity, labour, inventory, and overall supply chain efficiency, offering valuable insights for practitioners in the field.

* 1. Time-series forecasting of seasonal items sales using machine learning:

This research by (Ensafi et al., 2022) looked into predicting furniture sales using different models. They compared methods like SARIMA, Triple Exponential Smoothing, Prophet, LSTM, and CNN. The Stacked LSTM method was the most accurate, with Prophet and CNN performing well. The study emphasizes the importance of sales forecasting in the furniture industry. The literature review explores various approaches for accuracy, considering related papers in sales forecasting. To measure precision, they used methods like RMSE and MAPE. Overall, the research sheds light on effective ways to predict furniture sales, focusing on finding the most accurate forecasting model.

* 1. Developing and Preliminary Testing of a Machine Learning-Based Platform for Sales Forecasting Using a Gradient Boosting Approach:

This research by (Panarese et al., 2022) introduces a machine learning-based platform designed for enhancing sales forecasting in the retail sector. The study employs machine learning techniques to optimize sales forecasting by recognizing the retail industry's integration of IT innovation for improved customer experiences. The methods applied involve decision trees (DT) and boosting techniques, focusing on the XGBoost regression model. The results demonstrate the development of a prototype platform tailored for managing trading companies, showcasing the XGBoost algorithm as the top performer in sales forecasting. Notably, the XGBoost regression model contributes to a remarkable 15-20% improvement in forecasting accuracy, highlighting its efficacy in advancing the capabilities of sales prediction platforms within the retail sector. This research presents a promising step toward leveraging machine learning for more accurate and efficient sales forecasting, with the potential to significantly impact the operations of trading companies in the retail landscape.

* 1. Intelligent Sales Prediction Using Machine Learning Techniques:

This paper by Cheriyan et al. (2018) researches the realm of intelligent sales prediction, leveraging machine learning techniques and data mining to elevate the accuracy and efficiency of forecasting sales trends through a meticulous analysis of various data mining techniques, including the Generalized Linear Model, Decision Tree, and Gradient Boost Algorithm. The model performance results show that the Gradient Boost Algorithm achieves the highest overall accuracy of 98% in sales forecasting, followed by the Decision Tree Algorithm with approximately 71% accuracy and the Generalized Linear Model with 64% accuracy. The study reveals that the Gradient Boost Algorithm emerges as the top performer with an impressive 98% overall accuracy in sales forecasting. This research envisions future improvements in sales prediction models, aiming to explore additional machine learning algorithms for enhanced system intelligence without manual intervention. The significance of these findings lies in their contribution to the field of sales forecasting, showcasing the effectiveness of machine learning algorithms and data mining techniques in fortifying the accuracy and reliability of predictions.

* 1. Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail:

This paper by (Punia et al) introduces a novel forecasting method that combines long short-term memory (LSTM) networks and random forest (RF) for demand forecasting in multi-channel retail. The proposed method exhibits a unique ability to model complex relationships of both temporal and regression types, providing a competitive edge in accuracy compared to other forecasting methods. Evaluation of a real-world multivariate dataset from a multi-channel retailer, benchmarked against various techniques, including neural networks, multiple regression, ARIMAX, LSTM networks, and RF, reveals the new method's statistical superiority in bias, accuracy, and variance. The study not only ranks explanatory variables by their relative importance but also highlights the forecasting method's robustness for extended horizons across online and offline channels in multi-channel retail. This research contributes significantly to Operational Research Management Science (ORMS) and Predictive Analytics, opening avenues for future exploration into advanced neural networks, comparative analyses with other deep learning models, and testing the proposed method on data from non-traditional fulfilment channels.

* 1. A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks:

This paper by (Kontopoulou et al., 2023) navigates the evolving landscape of time series forecasting, contrasting traditional ARIMA models with contemporary artificial intelligence (AI) approaches. Acknowledging ARIMA models' historical prominence for their mathematical simplicity, the study emphasizes a recent shift towards AI techniques, revealing that, in general, AI algorithms outperform ARIMA models. Significant exceptions exist, prompting the exploration of hybrid statistical-AI models that combine the strengths of both approaches, consistently showcasing superior performance. In stock index forecasting, LSTM and XGBoost algorithms prove more effective than ARIMA, except for datasets with significantly lower values. The findings underscore the potential superiority of AI algorithms in time series forecasting, emphasizing the need for further evaluation and analysis. The paper suggests future research directions, focusing on understanding cases where ARIMA models excel and optimizing hybrid statistical-AI models for enhanced forecasting performance.

* 1. Conclusion:

In synthesis, the exploration of various methodologies for demand forecasting in the online retail domain underscores the crucial role of machine learning (ML) and artificial intelligence (AI) techniques in optimizing predictive accuracy. The literature review clarifies a important trend towards leveraging ML algorithms such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM) networks for enhanced sales prediction. These studies collectively advocate for the effectiveness of hybrid models, combining diverse ML approaches to harness the strengths of individual algorithms (Tsai, C. et al., 2013).

Firstly, while existing studies have showcased the efficiency of hybrid ML models, there remains a need to investigate deeper into the specific characteristics and challenges of online retail datasets (Adulyasak, Y. et al., 2023). This research aims to bridge this gap by performing thorough Exploratory Data Analysis (EDA) on the dataset (Online Retail II), thereby explaining the underlying patterns and identifying significant predictor variables. Secondly, while previous research has compared various ML techniques for demand forecasting, including ARIMA, SARIMA, LSTM, Random Forest and XGBoost methods, there is a lack of agreement on the most effective models for online retail demand prediction (Resul, T. and Sule, O., 2020). This study seeks to fill this gap by conducting a comprehensive comparison of these techniques with DARTS library model such as NBeats, RNN, TiDE, and TCN, considering their accuracy and robustness in addressing the unique challenges of online retail forecasting. Furthermore, while some studies have evaluated the performance of ML models, there is limited understanding into their generalization ability, runtime, and practical applicability in real-world online retail environments. This research endeavours to assess these aspects to control the scalability and convenience of different ML models for practical implementation (Ensafi, Y. et al., 2022). Lastly, while previous research has underwritten valuable findings, there remains a need for critical analysis of limitations encountered and recommendations for future development. This study aims to fulfil this objective by classifying areas for development, proposing potential extensions, and offering guidance for further exploration in online retail forecasting.

In essence, by addressing these gaps and objectives, this research paper aims to provide a comprehensive understanding of machine learning approaches for demand forecasting in online retail, offering practical insights for enhancing predictive accuracy and optimizing supply chain performance in the digital commerce landscape.

Research Methodology:

Over time, analysts and researchers have proposed various data mining and machine learning techniques to address prediction challenges that significantly impact daily operations. Between these challenges, forecasting product demand is vital for organizations and businesses, aiding them in making informed decisions essential for growth and advancement. This study centres on utilizing the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology as a administrative framework for data mining efforts, pointing to produce robust forecasts aligned with business goals. The CRISP-DM methodology delivers a structured approach that rationalizes the research process, facilitating more significant and advanced outcomes. Comprising six distinct stages, this methodology provides a systematic roadmap for the preparation and execution of the study (Plotnikova, V., Dumas, M. and Milani, P, F., 2000), thereby enhancing its clarity and coherence.

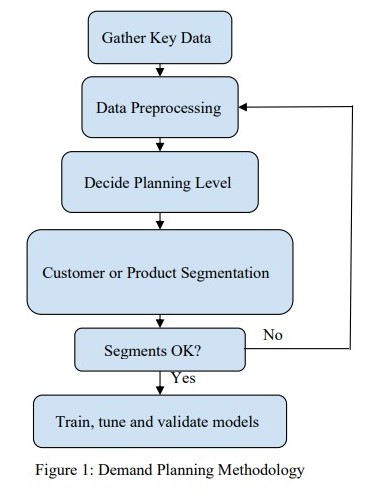


Figure 1: Demand Planning Methodology

* 1. CRISP-DM Methodology:

Data mining represents an advanced approach that requires a various range of skills and expertise. Currently, there is no commonly embraced framework for conducting data mining projects, highlighting the considerable influence of individuals or teams spearheading such initiatives (Cheriyan, S. et al., 2018). The necessity for a standardized methodology in data mining is apparent, one capable of guiding the translation of business or organizational challenges into actionable data mining tasks, suggesting appropriate data transformations and techniques, and providing a systematic means to evaluate feasibility and document acquired knowledge (Sun, J. and Sun, Y., 2018).

To ensure a thorough and effective research approach, this study will utilize the CRISP-DM methodology. It offers a structured framework with six key phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. By employing this methodology, the research guarantees systematic and cohesive execution of each step in the investigative process. This method serves as a reliable guide for navigating demand forecasting complexities in online retail by facilitating communication and collaboration among team members with diverse expertise and backgrounds. The structured approach improves research outcomes' reliability while providing clear direction towards achieving objectives related to demand forecasting within online retail. Additionally, it promotes efficient project integration through its core purpose of fostering effective communication and collaboration among team members with varying skills and experiences. Through adherence to this methodology, the study endeavors to gain clarity on business problems associated with demand forecasting in online retail sector goals – ultimately developing accurate prediction models using machine learning algorithms.The CRISP-DM data mining model offers a comprehensive framework for the entire lifecycle of a data mining project, outlining its stages, tasks, and outcomes (Schröer, C., Kruse, F. and Gómez, M, J., 2021). This lifecycle is structured into six phases, as depicted in Figure 1. The order of these phases is adaptable, with arrows indicating the primary dependencies between stages, signifying the result of each step or the subsequent key activity in the phase. The outer circle in Figure 1 represents the cyclical nature of data mining. Even post-implementation of the solution, data mining continues as an ongoing process. Information gained throughout the project, in conjunction with the chosen approach, may lead to new and often more precise business concerns. Subsequent data mining endeavours can then benefit from this accumulated experience. The stages of the CRISP-DM model have been utilized throughout this project and will now be explained within the context of demand forecasting (Schröer, C., Kruse, F. and Gómez, M, J., 2021).

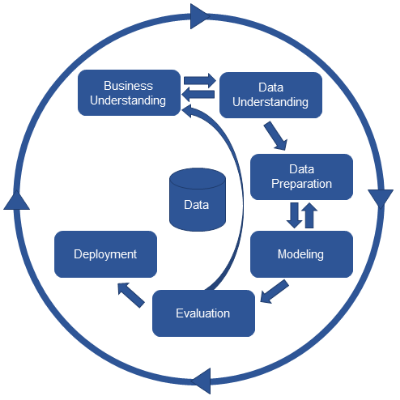


Figure 2: CRISP-DM workflow

The stages outlined in the CRISP-DM model have been systematically applied throughout the duration of this project. Hereafter, an explanation of these stages will be presented within the context of demand forecasting.

* + 1. Business Understanding:

In the preliminary phase of this research endeavour, the central objective revolves around attaining a comprehensive understanding of the business goals entwined with demand forecasting, with a specific lens focused on predicting the requisite quantity for a gift store's inventory. This endeavour is fundamentally rooted in the alignment of demand forecasting efforts with the overarching strategic objectives of the gift store. A multifaceted approach is adopted, entailing several key considerations integral to this phase (Fisher, L, M. and Vaidyanathan, R., 2014). Firstly, the process begins with a meticulous examination aimed at identifying and selecting the top 10 products within the gift store's inventory landscape that hold paramount significance. This selection is not arbitrary but guided by the recognition that understanding the demand dynamics surrounding these specific products will yield actionable insights pivotal for streamlining inventory management practices and optimizing sales strategies. Subsequently, a critical aspect entails the precise delineation of the key metrics indispensable for effective demand forecasting (Silva, C, J., Figueiredo, M. and Braga, C, A., 2019). Of particular significance is the task of accurately predicting the quantity required for the identified top 10 products. This necessitates a granular understanding of various metrics and indicators, including the frequency of demand forecasts and discerning any discernible patterns in seasonal variations.

Moreover, there exists a pressing need to delve into the potential business impact that accurate demand forecasting for the top 10 products can wield. This entails a comprehensive exploration of how such forecasting initiatives can catalyse transformative changes across critical business dimensions. Areas of consideration encompass the optimization of inventory management protocols and the strategic deployment of promotional tactics aimed at enhancing operational efficiency and bolstering revenue streams. Integral to this phase is also the active engagement with key stakeholders vested in the gift store's operations, including store managers and inventory personnel. By soliciting insights from these stakeholders, this research ensures alignment with their unique requirements and expectations, thereby fostering a collaborative environment conducive to the attainment of shared objectives (Drieniková, K. and Sakál, P., 2012). Furthermore, a thorough examination of historical sales data pertaining to the selected top 10 products is undertaken to unravel underlying patterns, trends, and fluctuations in demand. This historical backdrop serves as a foundational cornerstone upon which robust forecasting models can be constructed, leveraging insights gleaned from past performance to inform future projections. Lastly, due consideration is accorded to external factors capable of exerting influence on demand dynamics, such as seasonal events or holidays. Incorporating these external variables into the forecasting framework is deemed essential to fortify the predictive accuracy of the models, thereby ensuring their efficacy in real-world scenarios.

* + 1. Data Understanding:

Commencing with the basics of understanding data, the first step involves collecting it. Then, this study dives deeper to get to know the data better. This study looks for any issues like inconsistencies and explore the data to find initial insights or interesting groups of information (Kontopoulou, I, V. et al., 2023). The main aim is to draw conclusions and uncover hidden knowledge by thoroughly understanding the details and patterns within the dataset.

1. **Dataset Name**: Designated as the "Online Retail II Transaction Dataset."
2. **Source:** Originating from a UK-based and registered online retail entity, encompassing transactions spanning from 01/12/2009 to 09/12/2011.

**Link:** https://archive.ics.uci.edu/dataset/502/online+retail+ii

1. **Description:** This dataset summarizes real-time online retail transactions spanning a two-year duration, representative of a company specializing in the sale of unique all-occasion giftware. Particularly, a significant portion of its clientele comprises wholesalers.
2. **Dataset Characteristics**:

* Multivariate: Signifying the presence of multiple variables or features associated with each transaction.
* Sequential: Ordered chronologically, reflecting a time series of transactions.
* Time-Series: Captures transactional activities over a defined period, thereby classified as a time-series dataset.
* Subject Area**:** Business domain.

1. **Feature Type:**

* Integer: Encompassing integer values denoting quantities, customer IDs, and similar attributes.
* Real: Comprising real numbers indicative of prices, monetary values, and related metrics.

1. **Number of Instances:** A substantial dataset comprising 1,067,371 transactions.
2. **Number of Features:** While an exact count of features is unspecified, the dataset harbours multiple dimensions of information.
3. **Dataset Information:** Representing transactions conducted by customers, the dataset constitutes sequential and time-series data, relating interactions with the online retail establishment over a span of two years. Each transaction is anticipated to encompass of details encompassing product particulars, purchase quantities, pricing details, customer identifiers, and timestamps.
4. **Additional Information:**

* Specific to a UK-based online retail entity.
* Transactional period spans from 01/12/2009 to 09/12/2011.
* Specialization in unique all-occasion giftware.
* Majority of wholesale customers.

1. **Missing Values:** Affirmative, indicating the presence of missing data points within the dataset.
   * 1. Data Preparation:

The data preparation phase is crucial for refining collected data to ensure accurate demand forecasting for the gift store. It involves tasks aimed at improving dataset quality and relevance. This includes selecting suitable data types, identifying outliers, addressing missing values, ensuring non-negative quantity, integrating a holidays library, creating a new dataset, and refining attributes. The initial step in data preparation involves meticulous selection of appropriate data types for each variable within the dataset (Kramar, V. and Alchakov, V., 2023). This critical phase ensures that the data is structured in a manner that aligns seamlessly with the requirements of the forecasting model, thereby fostering streamlined and effective analytical procedures. This study used a dataset name as Online Retail ll which is shown in Figure 3:



Figure 3: Dataset before preprocessing

Beginning with the selection of appropriate data types for each variable, this step ensures that the dataset is structured in a manner helpful to efficient analysis and unified integration with forecasting models. Subsequently, outlier detection mechanisms are employed to identify and address inconsistent data points that could potentially skew model predictions. By implementing robust outlier handling techniques, the dataset's integrity is bolstered, thereby enhancing the reliability of forecasting outcomes.

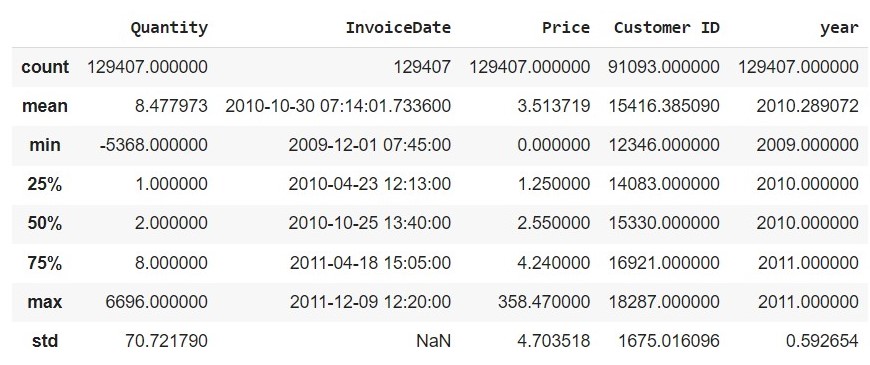


Figure 4 : Description of the dataset

* + 1. Modelling:

During the modelling stage, a particular selection and implementation of various modelling techniques were undertaken, aiming to optimize their performance through careful parameter tuning. This phase commonly involves encountering multiple techniques suitable for addressing the same data mining problem, each necessitating specific data formats for effective application (Gupta, D, S., Fuehrer, F. and Jeyachandra, C, B., 2014). The processes of data planning and modelling were closely intertwined, as challenges related to data quality or the need for additional insights frequently emerge during the modelling process (Cappiello, C., Samá, W. and Vitali, M., 2018**).** In the context of this research, seven distinct types of modelling techniques were considered for prediction purposes such as ARIMA, SARIMA, LSTM and four models of DARTS library. These techniques encompass a diverse range of methodologies, each offering unique strengths and capabilities for addressing the complexities inherent in the dataset. Through a systematic exploration of these techniques, the aim is to identify the most suitable approach that yields robust and accurate predictions in the domain of online retail demand forecasting (Wang, J., Liu, Q, G. and Liu, L., 2019). The ARIMA model provided unsatisfactory results because of neglection in seasonality and non-linear patterns in the data. SARIMA, with its ability to capture seasonal patterns, outperformed ARIMA in accurately predicting sales fluctuations over time. On the other hand, LSTM, being a deep learning model with memory retention capabilities, excelled in capturing complex temporal dependencies and non-linear trends within the data. However, it should be noted that SARIMA model presented a challenge in their predictions, often yielding negative values. This indicates that the forecasted demand occasionally fell below zero, which is not practically meaningful in the context of online retail. To address the limitations of traditional models, the study explored the application of the DARTS library, which encompasses advanced deep learning models designed specifically for time series forecasting. Among these NBeats, BlockRNNTest, TiDE, and TCN exhibited superior performance compared to traditional machine learning models. However, the Time-series Dense Encoder model (TiDE), stood out as the most accurate and superior model for demand forecasting in online retail. The comparative analysis emphasized the importance of utilizing advanced machine learning models like TiDE for precise and dependable demand prediction in the e-commerce sector. By adopting the exceptional predictive abilities of the TiDE model, online retailers can enhance decision-making, optimize inventory management, and ultimately boost profitability. This study examines seven different approaches to demand forecasting in online retail.

#### Autoregressive Integrated Moving Average (ARIMA) Model:

The ARIMA model is a widely recognized method in time series forecasting that excels at detecting trends, patterns, and seasonal variations in sequential data. As part of the autoregressive model family, ARIMA incorporates elements of moving averages, making it an effective choice for analyzing time-dependent datasets. This study explores the application of the Autoregressive Integrated Moving Average model to predict demand in online retail (Ferreira, J., Lee, A., & Simchi‐Levi, D., 2016). ARIMA is a commonly employed technique for forecasting time series renowned for its capacity to capture trends, patterns, and seasonality within sequential data.

**The model comprises three key components**:

Autoregressive, Integrated (I), and Moving Average are crucial factors that influence time series analysis. The Autoregressive component characterizes the relationship between a data point and its previous values, assessing the extent to which current data points depend on historical observations. On the other hand, the Integrated component encompasses differencing in time series data to ensure stationarity by tackling trends or seasonality. The Moving Average component captures the dependency between an observation and residual errors from a moving average of past observations, aiding in capturing short-term fluctuations (Ahmed, K, N. et al., 2010).

In constructing the ARIMA model, the order of parameters (p, d, q) plays a crucial role. The AR order (p) determines the number of lag observations included in the model, while the Integrated order (d) denotes the number of differencing operations applied to achieve stationarity. The MA order (q) signifies the size of the moving average window, indicating the extent of the moving average component. The implementation involves three main steps: identification, estimation, and diagnostic checking. Identification entails defining the values of p, d, and q based on data investigation, autocorrelation functions, and partial autocorrelation functions. Estimation involves estimating the model parameters using statistical methods such as maximum likelihood estimation. Analytical checking appraises the model's goodness of fit and identifies any inadequacies in capturing patterns or residuals (Ma, T. et al., 2014). While ARIMA offers adaptability and the ability to integrate trends and seasonality, its effectiveness can be limited by confident factors. ARIMA accepts that the time series is stationary, dictating preprocessing steps like differencing.

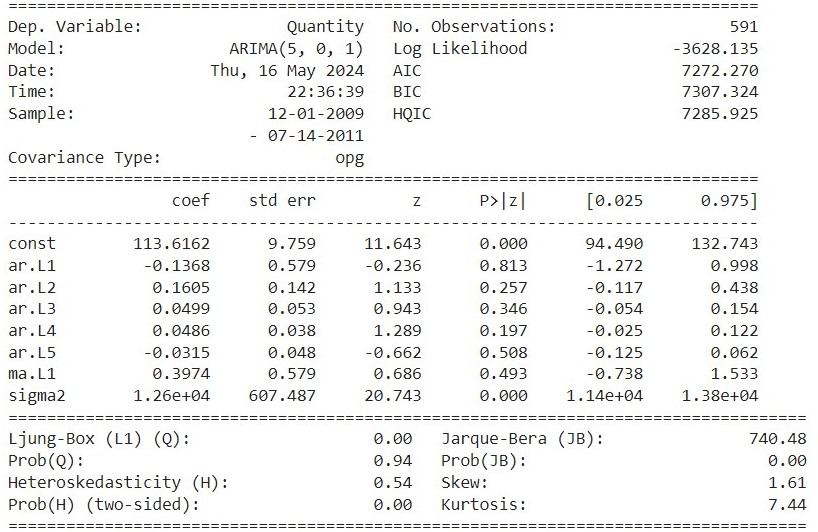


Figure 5: ARIMA Model Training Results

Besides, careful selection of parameters (p, d, q) is central for optimal model performance. Nevertheless, in the context of demand forecasting in online retail, the results obtained from the ARIMA model may not be satisfactory, as evidenced by the subpar prediction accuracy observed in this study. This lack could stem from the model's inability to capture complex patterns and nonlinear relationships inherent in online retail data. Therefore, it is important to delve deeper into more advanced forecasting methods such as Seasonal Autoregressive Integrated Moving Average and Long Short-Term Memory networks. SARIMA builds upon ARIMA by incorporating seasonal elements, making it better at capturing the seasonal fluctuations in online retail demand. On the other hand, LSTM, a form of recurrent neural network, excels at capturing time-based dependencies and complex relationships, potentially providing better results in demand forecasting applications.

In summary, this study highlights the limitations of the ARIMA model in online retail demand forecasting and underlines the need to explore more advanced techniques like SARIMA, LSTM and DARTS library models to improve prediction accuracy and effectiveness.

#### Seasonal Autoregressive Integrated Moving Average (SARIMA) Model:

The SARIMA model stands as a development of the ARIMA framework, personalized to differentiate and model seasonal patterns entrenched within time series data. Through the combination of supplementary seasonal components, SARIMA emerges as a pivotal asset for predictive analytics, particularly in scenarios where datasets exhibit repetitive patterns at consistent intervals. SARIMA model represents a notable advancement over the ARIMA framework, specifically engineered to discern and model seasonal patterns within time series data. By introducing additional seasonal components, SARIMA becomes a crucial tool for forecasting, particularly in scenarios where data exhibits recurring patterns at fixed intervals (Pongdatu, N, A, G. and Putra, H, Y., 2018).

**Comprising various components:**

Autoregressive (AR), Integrated (I), Moving Average (MA), Seasonal Autoregressive (SAR), Seasonal Integrated (SI), and Seasonal Moving Average (SMA), SARIMA offers a comprehensive approach to catching complex dependencies within time series data. Model creation involves thorough identification of parameters (p, d, q, P, D, Q, and s) through data exploration and analysis, followed by parameter estimation using statistical methods and diagnostic checking to assess model performance (Kontopoulou, I, V. et al., 2023). One of the key advantages of SARIMA lies in its seasonal sensitivity, allowing it to effectively capture and model seasonal fluctuations inherent in the data. This capability makes SARIMA particularly advantageous for time series with evident seasonal patterns, leading to improved forecasting accuracy in such scenarios (Kontopoulou, I, V. et al., 2023). However, despite its strengths, SARIMA models may encounter challenges, one notable issue being the occurrence of negative values in demand forecasting predictions. This phenomenon could stem from various factors, including the model's inability to constrain predictions within realistic bounds or insufficient consideration of external factors influencing demand fluctuations.

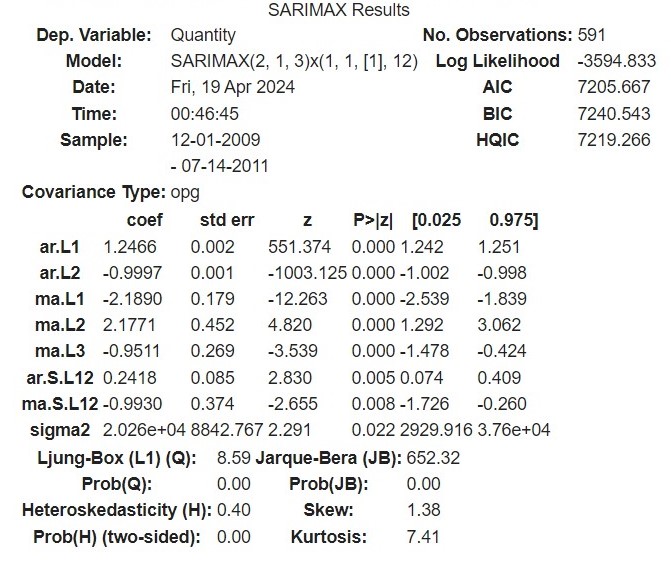


Figure 6: SARIMA model Training Results

Additionally, the intricacy of SARIMA models, particularly those with higher orders, requires careful attention when being used, as excessively complex models can result in overfitting or computational inefficiency. In conclusion, while SARIMA offers clear advantages in capturing seasonal patterns and improving the accuracy of time series data forecasts, it is essential to address issues such as negative forecasted values and model complexity to ensure reliable and efficient demand prediction in real-world scenarios. Further research into refining SARIMA techniques and exploring alternative forecasting approaches may offer promising prospects for addressing these challenges and enhancing predictive abilities in practical contexts.

#### Long Short-term Memory (LSTM):

The LSTM model is a powerful deep learning architecture, especially effective for analysing sequential data and forecasting tasks such as demand forecasting in online retail. As a type of recurrent neural network, it differentiates itself by its ability to capture long-term dependencies and complex patterns within sequential data, making it an effective tool for demand forecasting in online retail and other time series prediction tasks (Ensafi, Y. et al., 2022). With specialized memory cells designed to retain information over extended sequences, LSTM excels in handling complex, nonlinear patterns without the need for explicit feature engineering. Its adaptability to varying trends and seasonality in non-stationary data sets further enhances its utility (Kontopoulou, I, V. et al., 2023). However, LSTM's computational intensity and data hunger pose challenges, requiring significant computational resources and large datasets for effective training. The model's inherent complexity limits interpretability and necessitates careful hyperparameter tuning. Despite these drawbacks, LSTM's ability to automatically learn relevant features and outperform traditional forecasting methods in many scenarios underscores its importance in modern predictive analytics. By leveraging recurrent neural networks (RNNs) with specialized memory cells, LSTM models excel in capturing long-term dependencies and complex patterns within time series data (Kramar, V. and Alchakov, V., 2023). Data preprocessing for LSTM involves generating input-output pairs by moving a window of constant time steps across the sequential data.

The architecture of the LSTM model includes an LSTM layer and a fully connected dense layer. The LSTM layer, featuring 100 units and utilizing a ReLU activation function, handles input sequences, while the dense layer generates the ultimate output. The model is trained using the Adam optimizer and mean squared error loss function. During training, the model learns to predict future values based on historical input sequences. The training process involves multiple epochs, where the model iteratively updates its parameters to minimize the difference between predicted and actual values. The validation split allows monitoring of model performance on unseen data during training. The trained LSTM model determines promising results, as portrayed by the loss values during training and validation. The achieved loss values indicate that the model effectively learns the underlying patterns in the training data and generalizes well to unseen validation data (Shih, Y. and Lin, M., 2019).

Comparing LSTM with traditional methods like ARIMA and SARIMA, LSTM shows several advantages. Firstly, LSTM can capture complicated patterns and dependencies in the data without the need for explicit feature engineering (Siami‐Namini, S., Tavakoli, N. and Namin, S, A., 2019). Additionally, LSTM models often outperform traditional methods when dealing with nonlinear and non-stationary data. Importantly, LSTM models tend to provide more accurate predictions without encountering issues such as negative prediction values, as observed in some cases with ARIMA and SARIMA models (Siami‐Namini, S., Tavakoli, N. and Namin, S, A., 2019). However, LSTM models also have their limitations. They typically require more computational resources and training data compared to traditional methods. Moreover, LSTM models may be more challenging to interpret due to their inherent complexity. Additionally, LSTM models may suffer from overfitting, especially when trained on small datasets or when the model architecture is overly complex. In this research, the LSTM model emerges as a powerful tool for demand forecasting in online retail, offering superior performance in capturing complex patterns and providing accurate predictions (Shih, Y. and Lin, M., 2019). While LSTM surpasses traditional methods like ARIMA and SARIMA in many aspects, careful consideration of computational resources, data requirements, and potential overfitting is necessary when adopting LSTM for forecasting tasks.

This study has also absorbed on **Differentiable Architecture Search (DARTS)** library which presents a groundbreaking approach to automatically discovering best neural network architectures for various tasks, including time series forecasting. Leveraging the power of differentiable optimization, DARTS explores a vast search space of network architectures to find designs that maximize performance on the given task. This automated approach significantly reduces the need for manual architecture engineering, accelerating the model development process and potentially leading to more efficient and effective forecasting models (Nasir, A, N. and Jeong, S., 2021). In the context of demand forecasting, several models trained using the DARTS library have verified remarkable performance. Notably, models such as NBeats (Neural Basis Expansion for Time Series), TiDE (Time-series Dense Encoder), TCN (Temporal Convolutional Network), and Block Recurrent Neural Network (RNN) have showcased their effectiveness in capturing intricate temporal patterns and dependencies within sequential data.

#### Neural Basis Expansion for Time Series (NBeats):

This research used NBeats model, for designed time series forecasting, addresses the shortcomings of conventional methods by harnessing the capabilities of neural networks to discern intricate patterns and dependencies within sequential data. Its architecture comprises stacked blocks, each housing a fully connected neural network, arranged hierarchically to capture patterns at various levels of granularity. This design allows NBeats to effectively capture complex temporal relationships inherent in time series data. To extract and model the complex temporal patterns, NBeats utilizes generic blocks and theta layers (Ensafi, Y. et al., 2022). The generic blocks were responsible for capturing intricate temporal patterns, while the theta layers capture the global trend present in the time series data. This dual-layer approach enhances the model's ability to discern both local and global patterns, contributing to its robust forecasting performance.

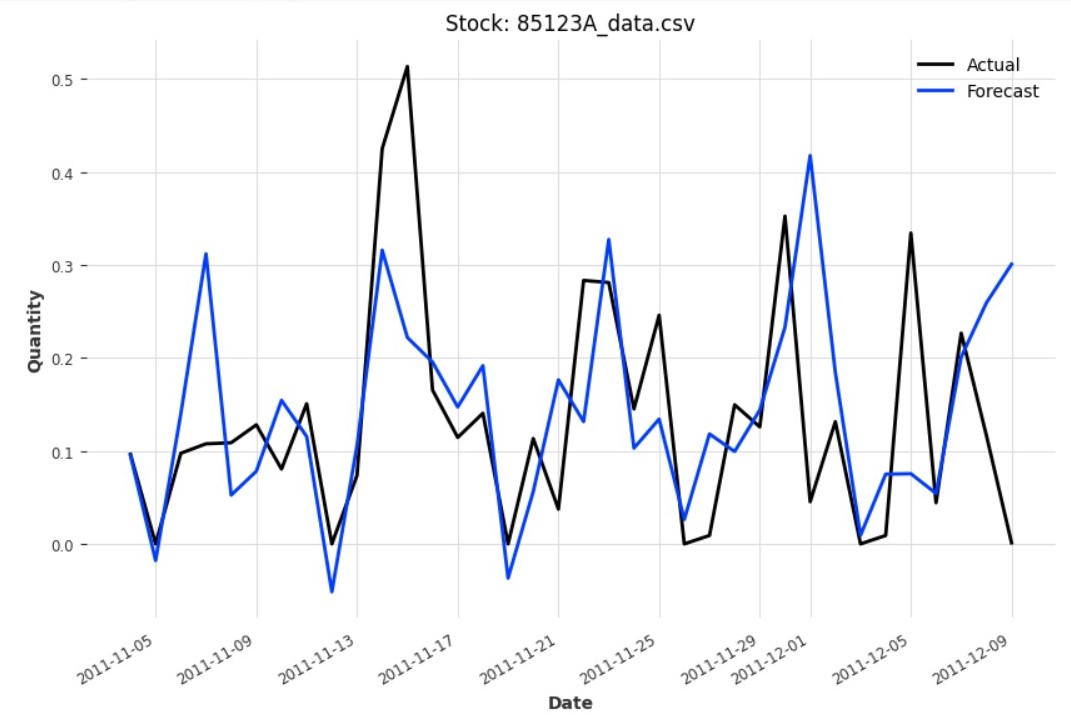


Figure 7: Some result of NBeats model

One of the notable features of the NBeats model is its ability to handle multi-horizon forecasting, enabling it to predict multiple future time steps in a single pass. This capability is particularly advantageous in scenarios where forecasting multiple future time steps is essential for decision-making. NBeats is trained using backpropagation to optimize model weights and minimize forecast error. During training, the model simultaneously produces both back cast (reconstruction of past values) and forecast (prediction of future values), facilitating a comprehensive understanding of its information processing over time. Key advantages of the NBeats model include its adaptability, interpretability, and scalability. It can adapt to various forecasting tasks without requiring task-specific tuning, providing interpretable representations of temporal patterns. Moreover, the model scales well for large-scale forecasting tasks, making it suitable for real-world applications.

However, it is crucial to consider the complexity of the NBeats model, especially concerning model size, based on dataset and task complexity. Additionally, sufficient representative training data is essential to ensure effective generalization to unseen patterns and optimize the model's forecasting performance.

#### Time-series Dense Encoder (TiDE):

The Time-series Dense Encoder (TiDE) model, important of transformers although without attention mechanisms, adopts a multilayer perceptron (MLP)-based encoder-decoder architecture. Comprising residual blocks within both the encoder and decoder components, TiDE effectively handles past and future covariates, as well as static covariates, facilitating probabilistic forecasting tasks. The model's architecture allows users to adjust parameters such as the number of encoder and decoder layers, hidden size, and temporal width, thereby offering flexibility to tailor the model to specific forecasting requirements. TiDE model incorporates training parameters for batch size, number of epochs, loss function, and optimizer configuration, enhancing its adaptability and usability. TiDE's compatibility with the Lightning framework further streamlines its integration into existing machine learning pipelines. Leveraging the framework's capabilities, users can seamlessly train and evaluate the TiDE model, benefiting from its modular design and efficient implementation.

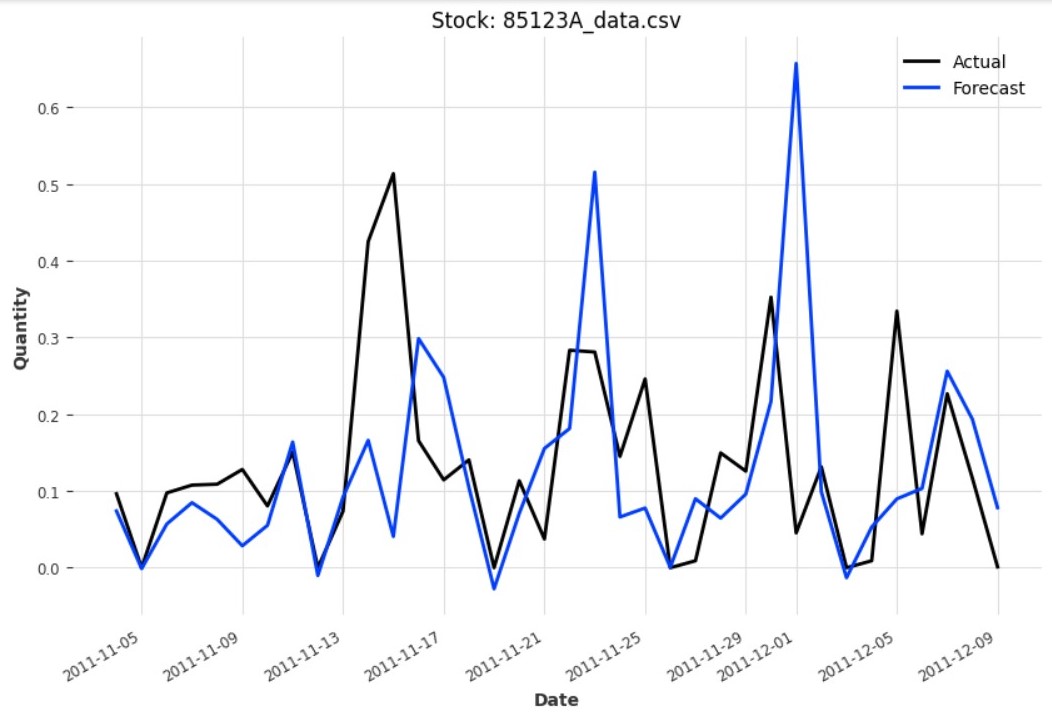


Figure 8: Some result of TiDE model

TiDE has shown remarkable results in time series forecasting, as confirmed by its performance metrics. In this study, the model attains a Mean Absolute Error of 0.07, Mean Squared Error of 0.01, and Root Mean Squared Error of 0.12, representing accurate predictions across numerous forecast horizons. With its adaptable design and competitive performance, TiDE provides a compelling solution for time series forecasting applications. Its capability to handle diverse data types including past and future covariates and support for probabilistic forecasting makes it suitable for a wide range of forecast tasks across different domains. Additionally, TiDE's compatibility with deep learning streamlines the training and deployment process, making it more accessible and user-friendly for researchers and practitioners alike.

#### Block Recurrent Neural Network (BRNN):

In the area of demand forecasting, the BlockRNN model also conduct as a robust solution for capturing sequential dependencies and generating accurate predictions. Leveraging a recurrent neural network (RNN) architecture, the BlockRNN model propose flexibility through alternatives such as Vanilla RNN, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU), enabling researchers to experiment with different architectures to identify the most suitable one for the forecasting task at hand. The BlockRNN model's architecture consists of an RNN encoder followed by a fully connected network, enabling efficient processing of fixed-length input chunks to produce precise forecasts. This architecture is particularly effective in capturing the sequential nature of demand data and extracting relevant patterns for forecasting purposes. A notable aspect of the BlockRNN model is its support for past covariates and customizable parameters, allowing for fine-tuning to specific forecasting requirements. This adaptability enhances the model's effectiveness in predicting demand fluctuations across different contexts and datasets.

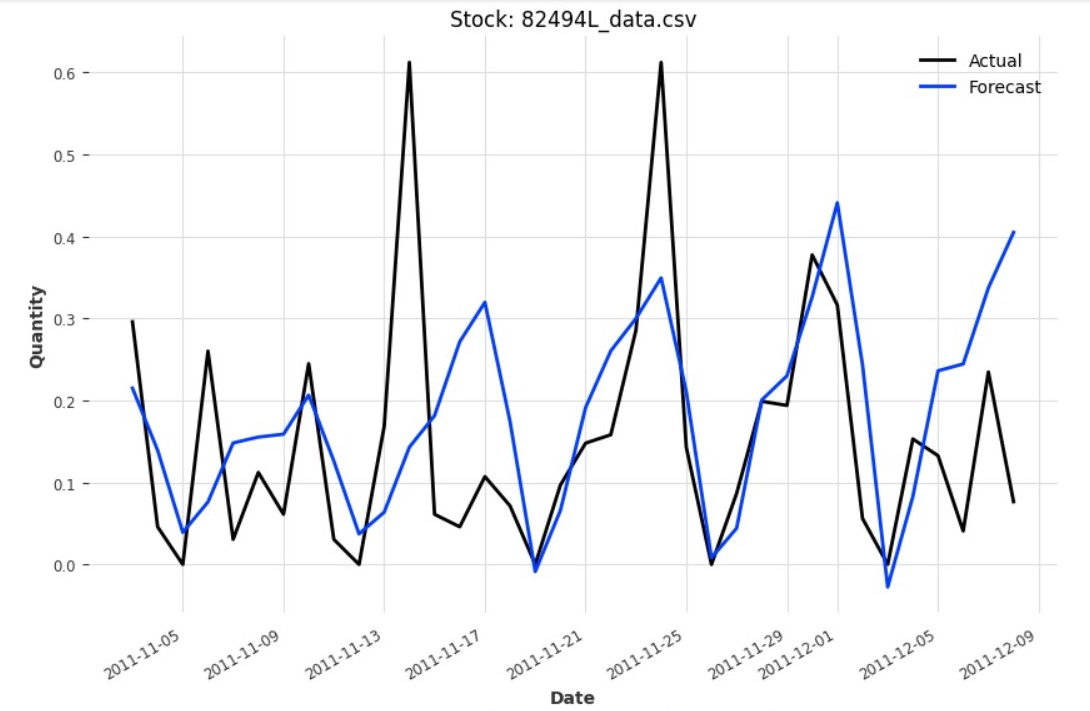


Figure 9: Some result of BRNN\_test model

Furthermore, the BlockRNN model incorporates features such as reversible instance normalization, automatic covariate generation, and tensor board logging, which collectively contribute to its robustness and ease of use in demand forecasting research. These features streamline the training and evaluation process, facilitating efficient experimentation and model optimization. Through industrious training and optimization efforts, the BlockRNN model has demonstrated promising results in demand forecasting tasks. The model attains competitive performance metrics in this research, including a MAE of 0.13, MSE of 0.03, and RMSE of 0.18, signifying its capability to accurately forecast demand patterns across numerous product categories. In summary, the BlockRNN model stands as a powerful tool in the empire of demand forecasting, offering a versatile architecture, customizable parameters, and robust performance. Its ability to capture temporal dependencies and adapt to diverse forecasting requirements makes it a valuable asset for researchers and practitioners seeking accurate and reliable demand forecasting solutions.

#### Temporal Convolutional Network (TCN):

The Temporal Convolutional Network (TCN) model has emerged as a significant player in demand forecasting research, offering a specialized architecture tailored for sequential data analysis. Leveraging dilated convolutions, TCN effectively captures temporal dependencies across various time scales, facilitating precise predictions of demand patterns. One of TCN's strengths lies in its support for past covariates and customizable parameters, providing researchers with the flexibility to adapt the model to different forecasting scenarios. This adaptability enables the model to handle diverse datasets and capture nuanced patterns inherent in demand data.

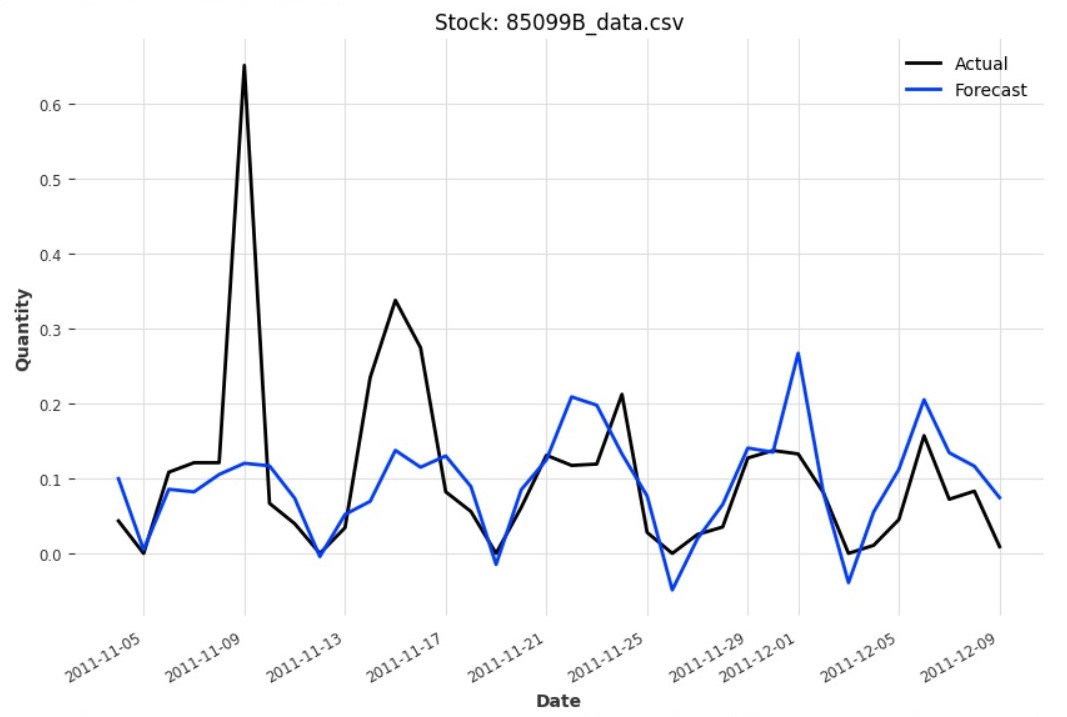


Figure 10: Some result of TCN model

Moreover, TCN integrates training parameters and functionalities for automatic covariate generation, streamlining the experimentation and evaluation processes. These features contribute to the model's effectiveness in demand forecasting, allowing researchers to focus on model optimization and performance evaluation. With its focus on capturing complex temporal patterns, TCN demonstrates robust performance across diverse product categories. The model accomplishes competitive performance metrics, including a MAE of 0.06, MSE of 0.01, and RMSE of 0.11. These metrics emphasize TCN's relevance and utility in advancing demand forecasting research.

In summary, the Temporal Convolutional Network (TCN) model stands out as a formidable tool in demand forecasting, offering a specialized architecture, customizable parameters, and robust performance across various datasets. Its ability to capture temporal dependencies and adapt to different forecasting scenarios makes it a valuable asset for researchers and practitioners seeking accurate and reliable demand forecasting solutions.

Evaluation and Analysis:

* 1. Understanding the data:

Succeeding the data collection phase, an exploratory data analysis was shown to gain insights into the dataset's key features and characteristics. This involved the evaluation and classification of the data through visualizations to identify patterns, trends, and potential outliers. Then, data cleaning and preparation procedures were performed to ensure the dataset's quality and suitability for model implementation. The initial step in the exploratory data analysis (EDA) process involved examining the dataset's structure and content (Al-Razgan, M., Al-Khalifa, S, A. and Al‐Khalifa, S, H., 2013). The figure below presents the first 5 rows of the dataset, providing an initial glimpse into its composition.

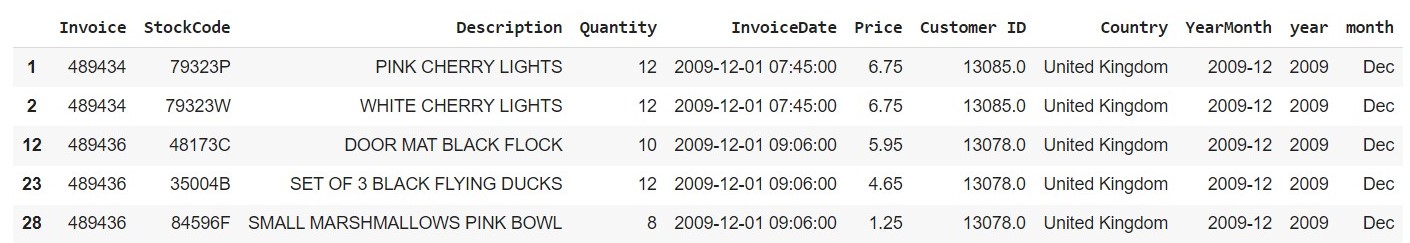


Figure 11 :First Five Rows of dataset

In the initial phase of this study, the focus was on data collection, setting the groundwork for a comprehensive exploration of the dataset's intricacies. Referred to as the "Online Retail II Transaction Dataset," this dataset was sourced from a UK-based and registered online retail entity, reporting transactions from December 1st, 2009, to December 9th, 2011. The dataset summarizes an excess of real-time online retail transactions, spanning a duration of two years, and serves as a representation of a company specializing in the sale of unique all-occasion giftware, with a notable clientele comprising wholesalers. Characterized as multivariate, sequential, and time-series data, this dataset encompasses a range of features associated with each transaction, offering insights into various aspects of the business domain (Aswal, D., 2021). These features include integer values denoting quantities and customer IDs, as well as real numbers representing prices and financial values. With an extensive size comprising 1,067,371 transactions, the dataset harbours multiple dimensions of information, although an exact count of features remains unspecified. Each transaction within the dataset reflects a sequential and time-series record of customer interactions with the online retail establishment over the two-year period. It is anticipated that each transaction entry comprises details such as product particulars, purchase quantities, pricing information, customer identifiers, and timestamps. Additionally, the dataset is supplemented with contextual information specific to the UK-based online retail entity, including its specialization in unique all-occasion giftware and its predominant customer base of wholesale clients. An important aspect highlighted during the initial exploration of the dataset is the presence of missing values, indicating the need for thorough data cleaning and preprocessing procedures to ensure the integrity and reliability of the dataset for subsequent analysis. Overall, this comprehensive understanding of the dataset's characteristics and nuances serves as a crucial foundation for further analysis and exploration in this study.

* 1. Data Cleaning and Preprocessing:

The code provided in the data preparation phase clarifies the practical implementation of these preparatory tasks. Through a series of data cleaning and preprocessing steps, including outlier removal, handling of missing values, and attribute creation, the dataset is meticulously curated to align with the objectives of demand forecasting (Kumar, A., Shankar, R. and Aljohani, R, N., 2020). Each code snippet is accompanied by explanatory comments, elucidating the rationale behind the respective data manipulation techniques employed. This transparent approach not only facilitates reproducibility but also underscores the rigor and thoroughness inherent in the data preparation process. The dataset description provides a comprehensive overview of the variables included in the dataset, offering insights into their respective distributions and characteristics. Examining the "Quantity" variable reveals a wide range of values, from negative to positive integers, indicating both purchases and potentially returns. The presence of negative quantities suggests a need for further investigation into the data collection process to understand the context of such occurrences. Upon detection of such values, a notification is printed, indicating the need for further analysis. Then, corrective measures were implemented to address inconsistencies in the dataset. Rows with positive quantities, indicative of purchases, were analysed to identify entries where the corresponding invoice contains the 'C' identifier, denoting cancellations or returns. These entries were deemed inconsistent and subsequently removed to maintain data coherence. Furthermore, attention is directed towards rows featuring negative quantities, potentially signifying returns, yet lacking the 'C' identifier in the corresponding invoice. Such inconsistencies were rectified by eliminating these entries, ensuring that negative quantities were appropriately associated with cancellation transactions. This particular approach to data validation and cleaning safeguards against inaccuracies and inconsistencies, thereby bolstering the dataset's suitability for subsequent analyses, including demand forecasting within online retail (Gupta, P. et al., 2021). Moving on to the "InvoiceDate" variable, the provided mean date and time indicate the central tendency of transaction timestamps, albeit with an unspecified standard deviation, which warrants caution regarding potential inconsistencies or missing values in this temporal data.

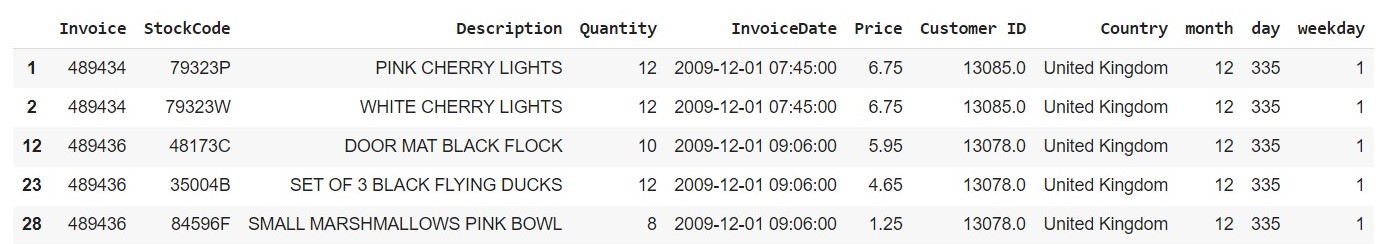


Figure 12: Cleaned Data for Training

Regarding "Price," the dataset showcases a diverse range of item prices, as reflected in the minimum, maximum, and mean values. The presence of zero-priced items raises questions about their nature, potentially indicating promotional offers or data entry errors that require scrutiny during the data preparation phase. "Customer ID" serves as a unique identifier for customers, with the count of unique IDs suggesting the breadth of the customer base. The variability in customer IDs may also indicate varying purchasing patterns or levels of engagement across different customer segments. Then the preprocessing steps applied to the "Description" column, the aim is to enhance the clarity and consistency of product names for improved understanding and analysis. The string manipulation functions were utilized to standardize the format of product names. Specifically, the “replace()” method is employed to remove any periods ('.') present in the product names, ensuring uniformity. Additionally, the names were converted to uppercase using the “upper()” function to standardize the casing. The strip() method is then applied to remove any leading or trailing white spaces from the names. Following this standardization process, rows containing product names with a question mark ('?') were identified and subsequently removed from the dataset. This step aims to eliminate entries with ambiguous or potentially erroneous product names, thereby enhancing the overall quality of the dataset (**Verstraete, G., Aghezzaf, E. and Desmet, B., 2019**). Missing product names (identified by null values in the "Description" column) were addressed through a data imputation process based on the associated "StockCode." For each missing product name, the most frequent product name corresponding to the same "StockCode" is determined using the model() function. If a suitable replacement is found, it is assigned to the missing entry. However, if no matching product name is available for a given "StockCode," the corresponding row is removed from the dataset. To ensure consistency, the "Description" column is explicitly converted to a string data type, facilitating uniform handling and compatibility with subsequent analyses. Overall, these preprocessing steps contribute to refining the "Description" column, thereby enhancing the interpretability and reliability of the dataset for further analysis, such as demand forecasting in the context of online retail.

Furthermore, particular attention is directed towards handling missing and incorrect values within the dataset. Techniques such as data imputation and removal of incomplete data points were applied to moderate the adverse effects of missing or erroneous data on model performance. Moreover, a rigorous rule is enforced to ensure that quantity values remain non-negative throughout the dataset (Kontopoulou, I, V. et al., 2023). This normalization step aligns with the practicality of product quantities and contributes to the coherence of the dataset. An integral aspect of data preparation involves incorporating contextual information, such as the impact of holidays on sales, into the dataset. Leveraging a holidays library, this study improves the dataset with pertinent temporal context, thereby empowering the forecasting model with insights into seasonal variations in demand. Moreover, holiday data is incorporated into the dataset to account for their influence on sales and demand patterns in online retail. Initially, holiday information is retrieved for selected countries between 2009 and 2011, with dates and holiday names appended to a consolidated list. This data is then organized into a DataFrame and filtered based on the specified date range, ensuring relevance to the dataset. Duplicate entries were removed to maintain data integrity. Alongside, date-related features, including month, day of the year, and weekday, were extracted from transaction dates to provide temporal context. Then, the holiday data is merged with the main dataset, enabling the identification of transaction dates coinciding with holidays. New columns representing each unique holiday were created, with values indicating holiday observance. Finally, redundant columns were eliminated, resulting in a refined dataset enriched with holiday information, poised for subsequent analysis, particularly in demand forecasting within online retail. Then, the "Year" variable highlights the temporal distribution of transactions, with the mean year indicating the central tendency of transaction years within the dataset (Cheriyan, S. et al., 2018).

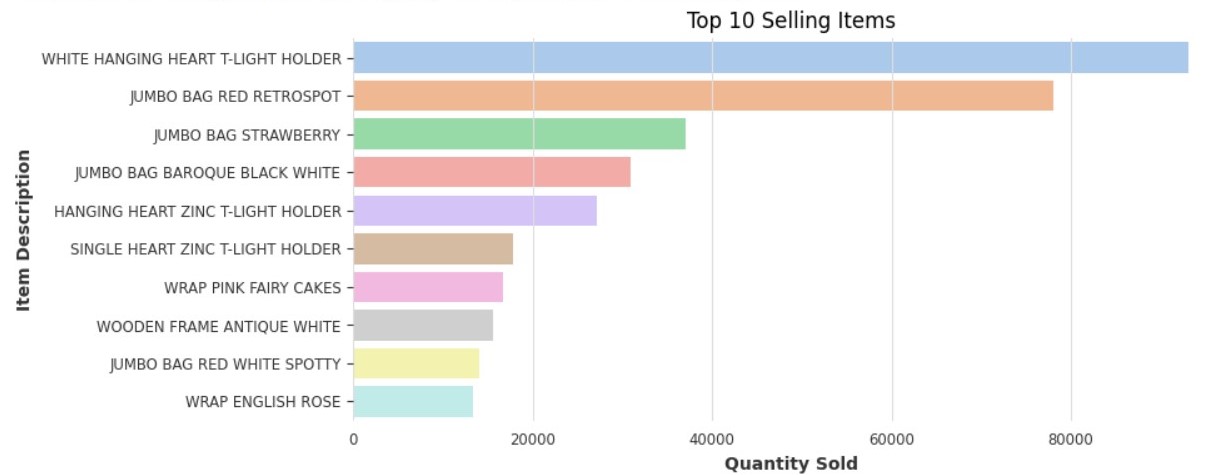


Figure 13: Top 10 Selling Products

Directing through the extensive array of products within retail datasets can be intimidating. To streamline the analysis and ensure a focused approach, this study deliberately narrowed its scope to the top 10 selling products as you can see the top selling products in the Figure 13. This strategic decision was driven by the understanding that these products exercise significant influence over overall sales and customer demand. By concentrating the efforts on these key performers, this research aimed to extract insights that were directly relevant to business strategies and decision-making processes. This focused approach not only allowed for efficient resource allocation but also ensured that the forecasting models were trained on data that truly mattered. Simultaneously training all Darts library models on this select group of products eased a comprehensive comparative analysis, providing valuable insights into the strengths and weaknesses of each model in real-world retail contexts. By improving in on the top 10 products, this study aimed to maximize the relevance and applicability of findings to the broader retail landscape, ultimately contributing to more informed decision-making and enhanced business performance.

* 1. Assessment of Stationarity using Augmented Dickey-Fuller Test:

To certify the correctness of the time series data for ARIMA modelling, an Augmented Dickey-Fuller test was conducted in this study. The ADF test is a statistical method normally employed to assess the stationarity of a time series. The Python code snippet utilized the adfuller function from the statsmodels.tsa.stattools module to perform the ADF test on the given time series data. Upon execution of the ADF test, the resulting statistics and p-value were obtained. The ADF statistics represent the test statistic value computed during the test, while the p-value indicates the significance level of the test. A p-value below a predetermined threshold (typically 0.05) suggests strong evidence against the null hypothesis, indicating that the data is stationary (Ensafi, Y. et al., 2022). Conversely, a p-value exceeding the threshold implies weak evidence against the null hypothesis, indicating non-stationarity in the data. In this breakdown, the accomplished p-value from the ADF test was compared against the significance level of 0.05. If the p-value was less than 0.05, it indicated solid evidence against the null hypothesis and confirmed stationary data. If the p-value was greater than 0.05, it suggested weak evidence against the null hypothesis, indicating non-stationarity in the data. Based on the ADF test results, it was determined that the time series data under consideration was already stationary, as evidenced by the p-value obtained (Kramar, V. and Alchakov, V., 2023). Therefore, there was no need for further differencing to achieve stationarity. This finding provided valuable insights into the initial state of the data and informed subsequent steps in the analysis, including the selection of appropriate ARIMA parameters.

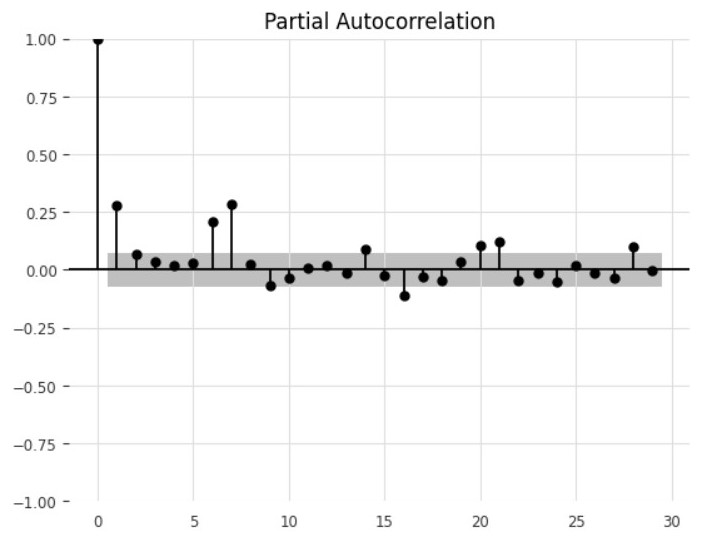
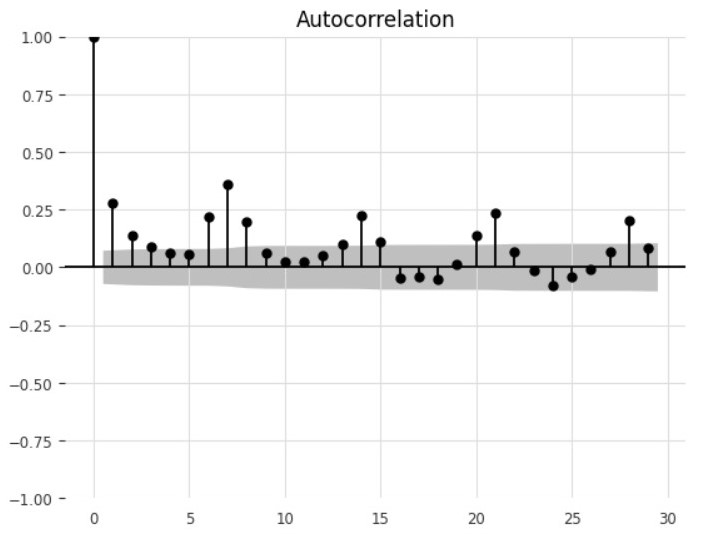


Figure 14: Auto Corelatioin Graph Figure 15: Partial Auto Corelation Graph

In addition to the ADF test, auto-correlation function (ACF) and partial auto-correlation function (PACF) plots which are shown in Figure 15 & 16 were generated to visualize the autocorrelation structure of the time series data. These plots aided in identifying potential patterns and informing the parameter selection process for the ARIMA model. Ultimately, the chosen ARIMA parameters (5, 1, 0) were based on a comprehensive analysis of the data's stationarity and autocorrelation characteristics, ensuring the robustness and effectiveness of the forecasting model.

* 1. Rescaling and Normalization:

Before implementing machine learning algorithms, it is imperative to undertake data rescaling as a fundamental aspect of data processing. Datasets often encompass attributes with varying scales for different quantities, thereby influencing the outcomes of machine learning algorithms. Rescaling aims to mitigate this variance by transforming all features of the dataset into a uniform scale. One dominant rescaling technique is normalization, which contains scaling values to fit within a encoded range, typically between 0 and 1. This process, also known as Min-Max scaling, guarantees that the values were shifted and rescaled proportionally, maintaining their relative distribution while adhering to the specified range boundaries. The normalization formula, which serves as the mathematical underpinning of this process, facilitates the systematic adjustment of values to achieve uniformity and consistency across the dataset. By employing normalization, data analysts and machine learning practitioners can enhance the efficacy and interpretability of their models by reducing the influence of varying scales on algorithmic performance.

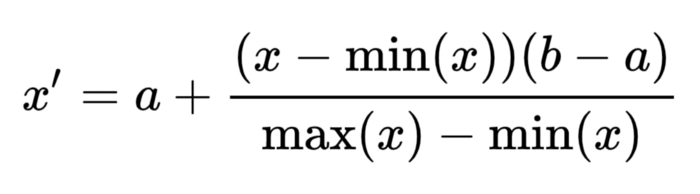


Figure 16: Formula of Rescaling and Normalization

The figure 14 shows a visual representation of scaled data, as depicted by the array of values ranging between 0 and 1. Scaling data is a critical preprocessing step in machine learning, aimed at standardizing the feature values to a uniform scale. This ensures that all features contribute equally to the learning process, preventing any undue influence from features with larger magnitude values. In the context of the provided array, the scaling process, likely achieved through techniques such as normalization or Min-Max scaling, has transformed the original data into a standardized format suitable for machine learning algorithms. By visually presenting the scaled data in the figure 17, this research effectively conveys the importance of preprocessing techniques in enhancing the performance and interpretability of machine learning models.

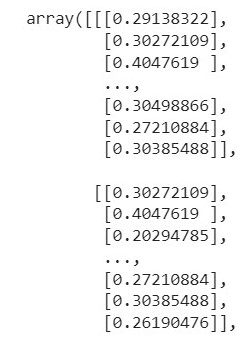


Figure 17 : Scaled and Normalized Dataset for training models

* 1. Model Performance:

In measuring the efficiency of forecasting models, a combination of metrics is commonly used, focusing on both computational efficiency and prediction accuracy. One important metric is the model training time, which offers insights into the computational efficiency of the training process. By defining the duration required for the model to learn from the provided data, this research gain an understanding of the computational resources needed for model training. Additionally, key error metrics such as MAE, MSE, and RMSE play a crucial role in evaluating the accuracy and precision of the model's predictions. These metrics provide quantitative measures of the inequality between predicted and actual values, allowing this study to assess the performance of the forecasting model comprehensively. By analysing both computational efficiency and prediction accuracy metrics, this research can make informed decisions regarding the suitability and effectiveness of the forecasting model for real-world applications (Ensafi, Y. et al., 2022).

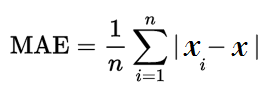
* + 1. Parameters for Measuring Model Performance:

Certainly, here are the parameters for measuring model performance rewritten in a more elaborate manner:

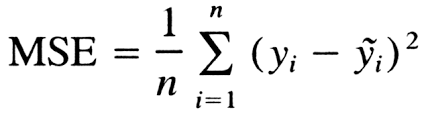
1. **Time Elapsed**: The "Time Elapsed" parameter indicates the duration taken by the forecasting model to undertake training and generate predictions. This metric, calculated as the difference between the end time and the start time, provides valuable insights into the computational efficiency of the model. By understanding the time required for model training, researchers gain essential information for optimizing computational resources and enhancing overall efficiency (Marzi, F. et al., 2023).

Time\_Elapsed = End\_Time – Start\_Time

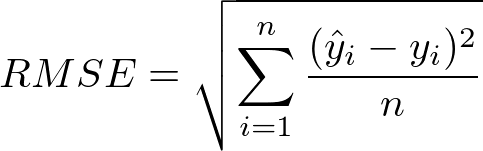
1. **Mean Absolute Error (MAE):** The MAE serves as a central quantity of prediction accuracy, measuring the average absolute differences between predicted and actual values. This metric is calculated by adding the absolute differences between predicted and actual values and dividing by the total number of observations. MAE offers a straightforward and intuitive metric for assessing the model's accuracy, with lower values indicating better performance (Marzi, F. et al., 2023).



1. **Mean Squared Error (MSE):** MSE computes the average of squared differences between predicted and actual values. By squaring the differences, adding them up, and dividing by the total number of observations, MSE provides insights into the overall scale of errors in the model's predictions. Similar to MAE, lower MSE values indicate improved prediction accuracy, with each error being weighted by its squared value (Marzi, F. et al., 2023).



1. **Root Mean Squared Error (RMSE):** RMSE is resultant from the square root of MSE and represents a brief measure of the typical size of errors in the model's predictions. By taking the square root of MSE, RMSE offers a more understandable metric for assessing prediction accuracy. RMSE is widely used in forecasting tasks to evaluate the model's performance and suitability for real-world applications (Marzi, F. et al., 2023).

****

These parameters collectively provide a comprehensive framework for evaluating the performance of forecasting models, encompassing both computational efficiency and prediction accuracy. By analysing these metrics, this research can make informed decisions regarding the effectiveness and reliability of forecasting models in practical settings.

* 1. Model Results:
     1. ARIMA model:

During the analysis of the ARIMA model, which is commonly used for forecasting future values, this research encountered significant challenges, especially when focusing on a specific stock code, '85123A'. As you know that, this research has trained the model on only one product to assess its performance. However, this study found that the model faced considerable obstacles, possibly due to its outdated nature and reliance on specific assumptions. In an effort to improve the model's accuracy, this research took a different approach by fine-tuning its parameters (p, d, q). The chosen parameters were p=5, d=0, q=1.

**Methodology for Selecting 𝑝, d, and q:**

To determine the appropriate values for p, d, and q, the following steps were employed:

**Stationarity Check:** The time series data was examined for stationarity. Stationarity implies that the statistical properties of the series, such as mean and variance, remain constant over time.In this case, the data was already stationary, so no differencing was required d=0.

**Autocorrelation Function and Partial Autocorrelation Function Plots:**

The ACF and PACF graphs of the stationary series were plotted in Figure 14 and Figure 15 to identify potential values for p and q. The **ACF plot** helps in identifying the value of q. It shows the correlation of the time series with its own lagged values. For the stationary series, the ACF plot displayed a significant spike at lag 1, suggesting that a moving average order (q) of 1 could be appropriate. The **PACF plot** assists in determining the value of p. It shows the correlation of the time series with its own lagged values, controlling for the values of the intermediate lags. The PACF plot exhibited significant spikes up to lag 5, indicating that including up to 5 lagged observations could be beneficial. Hence, p=5 was selected.

**Interpretation of ACF and PACF Plots:**

**ACF Plot**: The ACF plot in Figure 14 showed a significant spike at lag 1, indicating the presence of a moving average component. This led to choosing q=1.

**PACF Plot:** The PACF plot revealed significant spikes at lags 1, 2, 3, 4, and 5, but not beyond. This pattern suggests that the series has significant partial correlations up to lag 5. Therefore, p=5 was chosen to incorporate these lag effects into the model.

This study conducted thorough tests for check the stationarity and parameters values to enhance the model's predictive capabilities. Despite the diligent efforts, this study was unable to achieve the desired results with ARIMA model, as shown in the figure below.

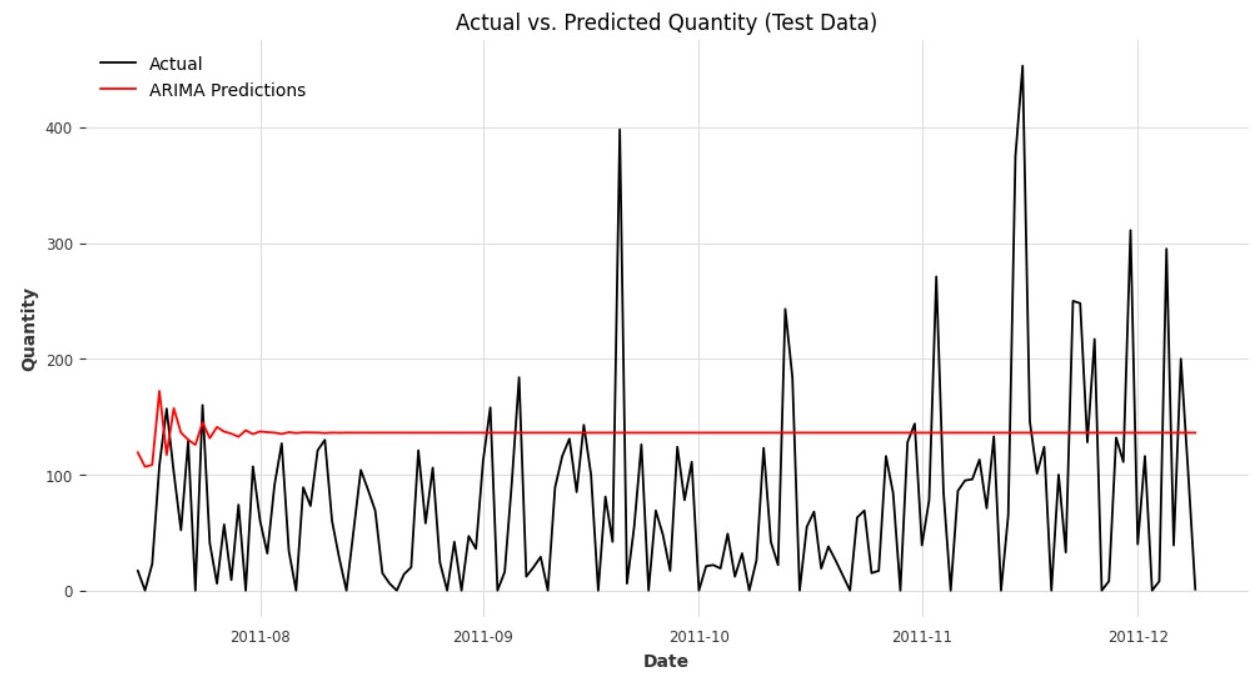


Figure 18: ARIMA model result

* + 1. SARIMA Model:

Comparable to the ARIMA model, the SARIMA model met outstanding challenges during the analysis. Despite its potential, the SARIMA model showed performance issues when relying exclusively on the important features of the data, mirroring the challenges observed with the ARIMA model. However, an important revelation surfaced when this research introduced external factors into the analysis. Including exogenous data, such as a custom-created dataset representing holidays, yielded promising results for the SARIMA model. The inclusion of external factors, particularly holidays, appeared to positively influence the model's predictive capabilities (Atique, S. et al., 2020). This was demonstrated by the improved performance established by the SARIMA model when external factors were considered, as illustrated in the accompanying figure. These results underline the importance of accounting for external effects in time series forecasting, as they can significantly impact the accuracy and reliability of predictive models. The SARIMA model is characterized by two sets of parameters: non-seasonal parameters (p,d,q) and seasonal parameters (P,D,Q,s). This study chooses p=2, d=1, q=3 for the non-seasonal part, and P=1, D=1, Q=1, and s=12 for the seasonal parts.

**Non-seasonal Parameters**: The non-seasonal parameters p, d, and q were determined based on the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots of the differenced series. The ACF and PACF plots suggested p=2 and q=3 based on significant spikes up to lag 2 in PACF and a decay in ACF after lag 3.

**Seasonal Parameters:** The seasonal parameters P, D, Q, and s were selected to account for the seasonal component in the data. The seasonal ACF and PACF plots, considering a seasonal period of 12 (months in a year), indicated significant spikes at lag 12 in both ACF and PACF. The choice of P=1, D=1, and Q=1 was made based on these plots to capture the seasonal patterns effectively.

Moving forward, further exploration of incorporating data into forecasting models may offered valuable insights and enhance predictive accuracy in real-world applications. This indicates that external factors, like holidays, positively influenced the model's predictive capabilities as shown in the figure below:

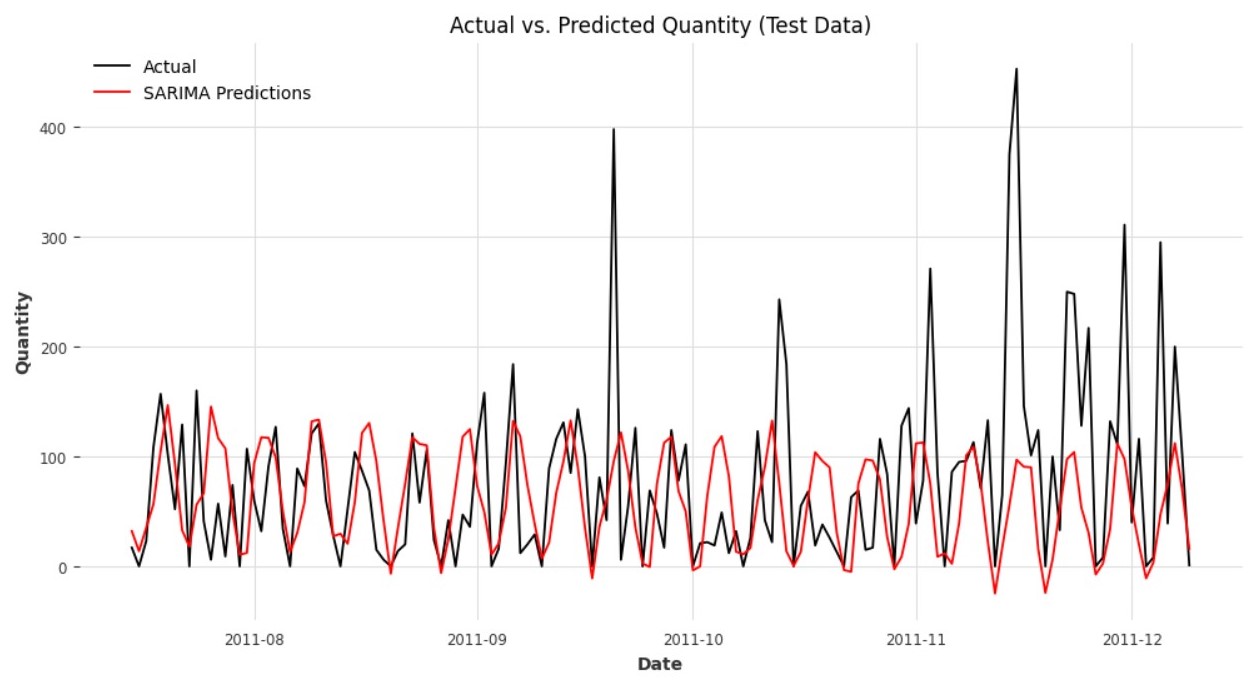


Figure 19: SARIMA model result

However, the scenario shifted when attempting to predict for the product. Despite the successful integration of exogenous data, this research found that the SARIMA model struggled to deliver optimal results for this specific forecasting task. The complexities associated with predicting top-selling products may surpass the capabilities of the SARIMA model, suggesting the necessity for alternative approaches or more sophisticated models tailored to the unique characteristics of high-demand items.

* + 1. Long Short-term Memory:

This research working on the LSTM model to forecast the demand of an online retail store using a retail dataset. The LSTM model planning consists of an LSTM layer followed by a fully connected dense layer. The LSTM layer, with 100 units and a ReLU activation function, processes input sequences, while the dense layer produces the final output. The model is trained using the Adam optimizer and mean squared error loss function. During training, the model learns to predict future values based on historical input sequences. The training process involves multiple epochs, where the model iteratively updates its parameters to minimize the difference between predicted and actual values. The validation split allows monitoring of model performance on unseen data during training. The trained LSTM model determines promising results, as portrayed by the loss values during training and validation but results were not satisfactory for accurate prediction. The achieved loss values indicate that the model effectively learns the underlying patterns in the training data and generalizes well to unseen validation data (Lee, J. et al., 2021). The incorporation of holiday data proved beneficial in enhancing the dataset and understanding the impact of holidays on demand patterns. This research incorporated holiday data into the traditional models as well. However, despite this augmentation, the traditional models failed to produce satisfactory results. Therefore, there remains a pressing need to explore and develop forecasting models that can offer improved accuracy and effectiveness in predicting demand for diverse product categories.

The LSTM model confirmed capable performance in predicting future demand compared to traditional models such as ARIMA and SARIMA. Nevertheless, it is noteworthy to mention that the results obtained from the LSTM model were not entirely satisfactory. Despite its ability to capture complex temporal dependencies in the data, the LSTM model encountered challenges in accurately forecasting demand patterns, particularly for high-demand items. As you can see the result in below Figure 19.

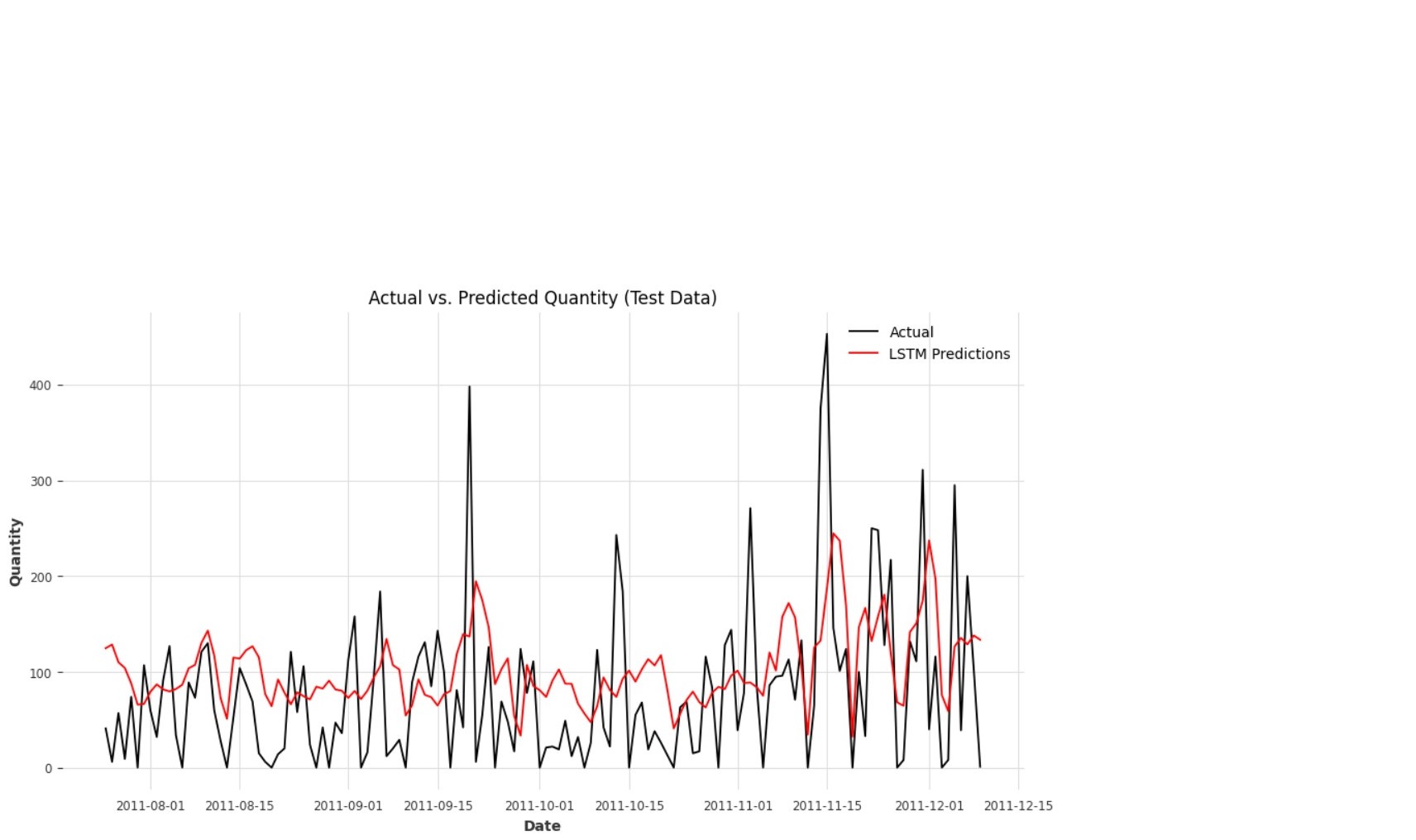


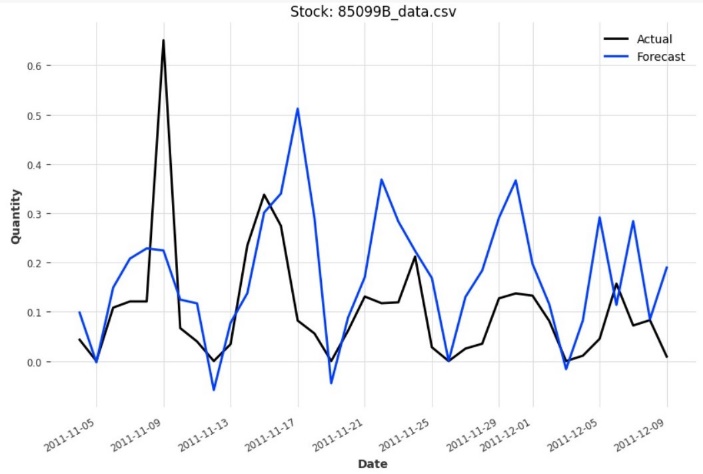
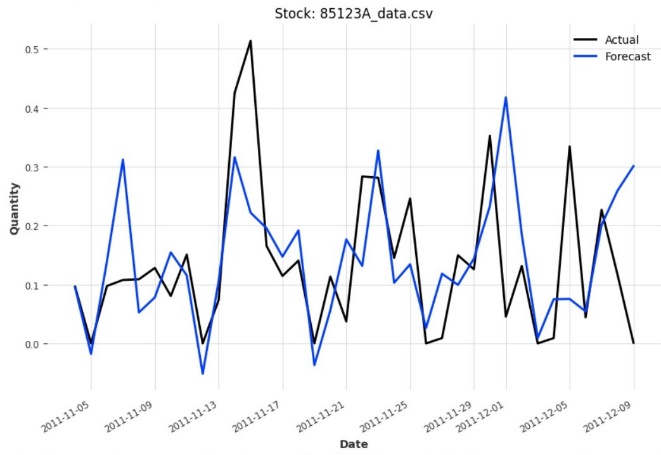
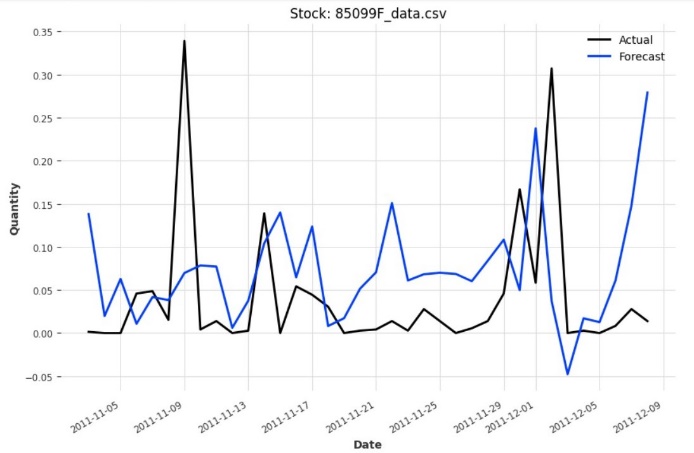
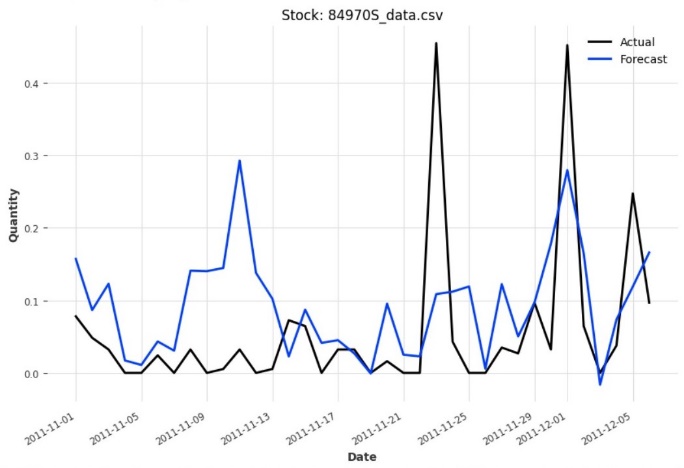
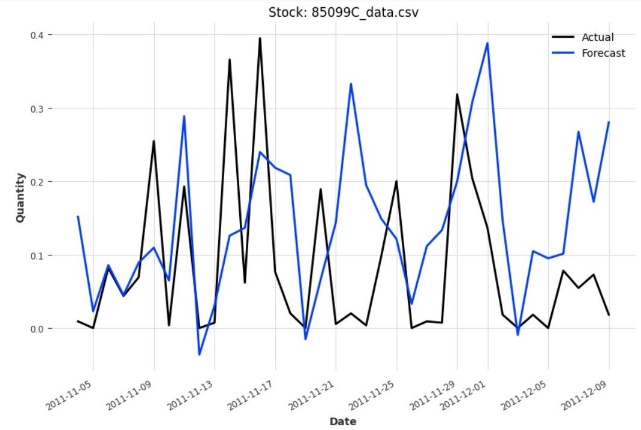
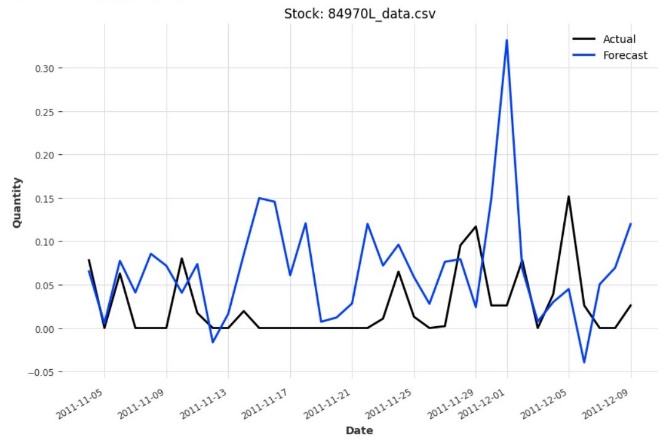
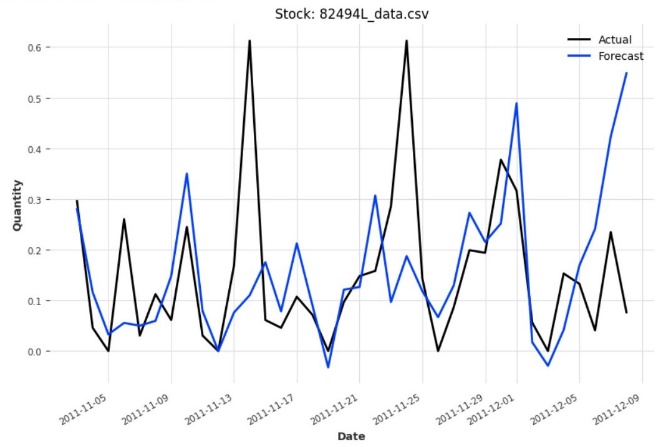
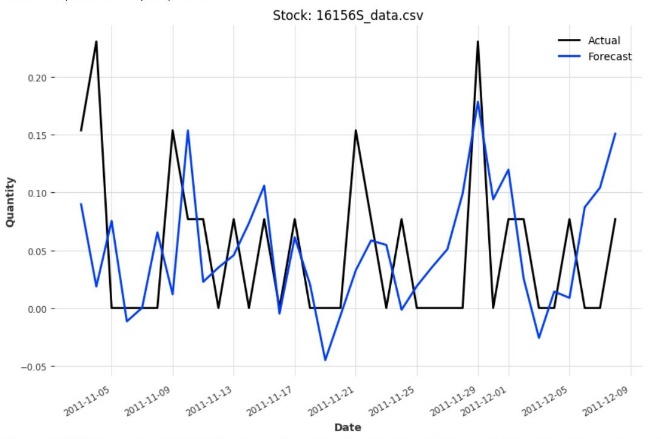
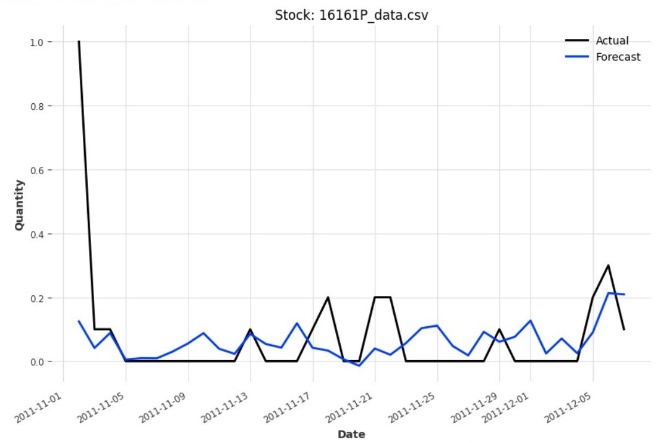
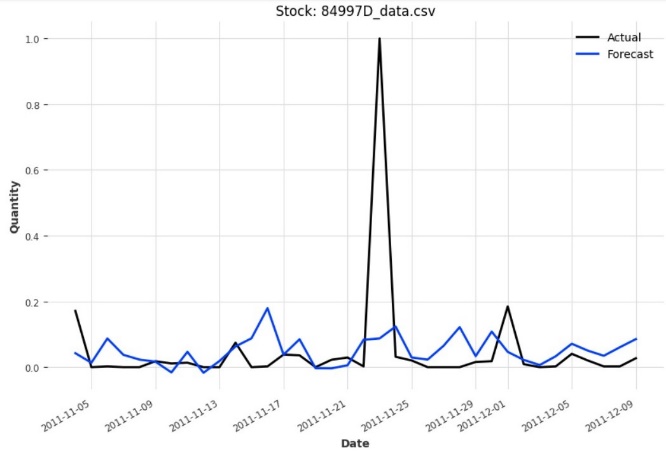
Figure 20: LSTM Model Result

However, it is essential to note that while LSTM outperformed these models, the accuracy of the results was not optimal. The primary objective of this research is to identify the most effective model for demand forecasting, one that can provide accurate predictions across various products. Such predictions were crucial for businesses to determine the appropriate stock levels required to meet demand effectively.

* + 1. NBeats Model:

This study evaluated the performance of the NBeats model, a deep learning architecture implemented through the DARTS library, for demand forecasting in the context of an online retail dataset. The NBeats model was trained on a dataset comprising the top 10 selling products, aiming to provide accurate predictions of future demand patterns. The evaluation of the NBeats model revealed promising results, showcasing its effectiveness in forecasting demand for online retail products.

The NBeats model was trained on the top 10 selling products' dataframe, allowing for individualized forecasting for each product. The results of these forecasts were plotted separately, providing a visual representation of the model's performance for each product category. The plotted results highlighted the NBeats model's capacity to capture the unique demand patterns of different products, thereby facilitating informed decision-making regarding inventory management and stock optimization. The plots of all top 10 selling products results are shown in below as:

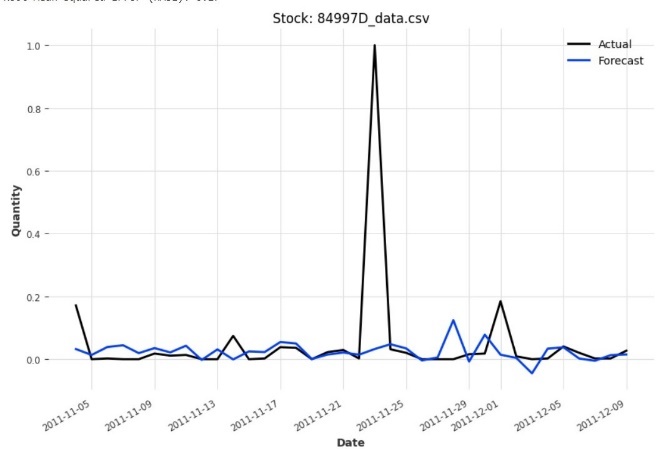
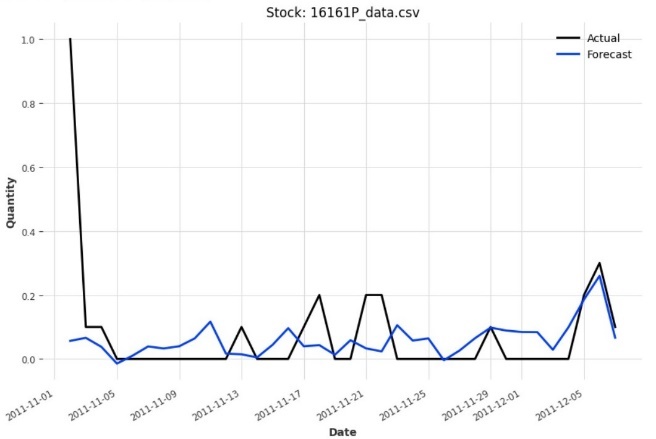
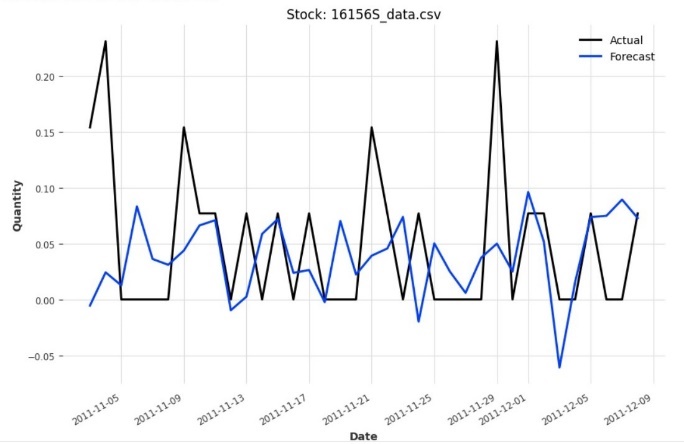
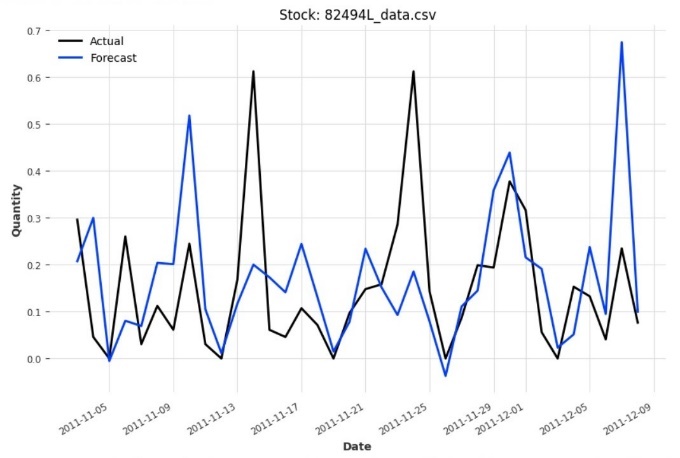
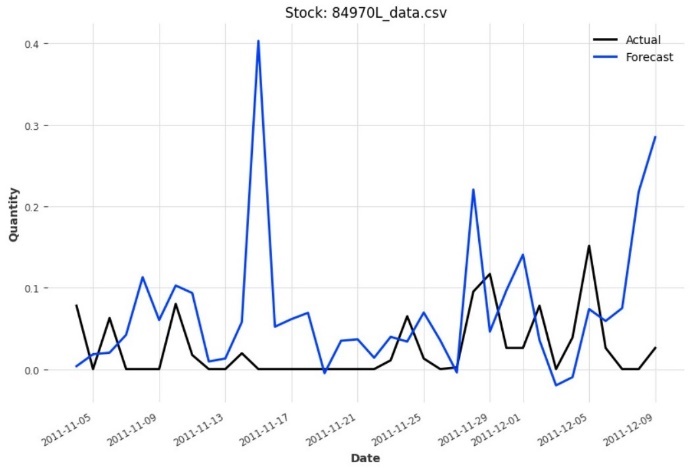
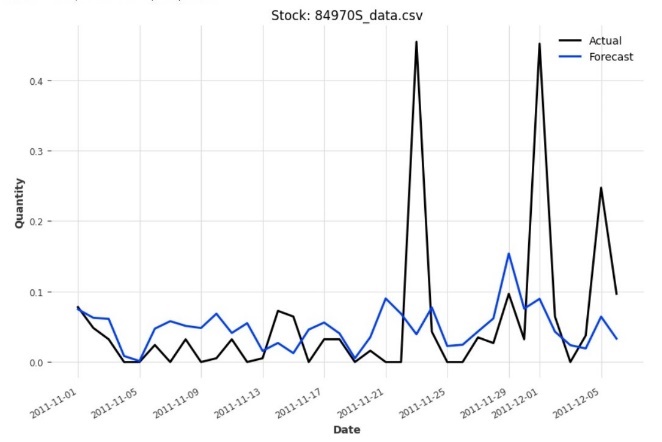
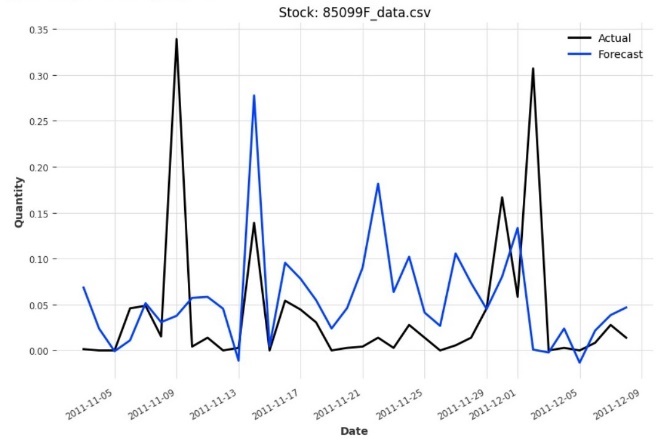
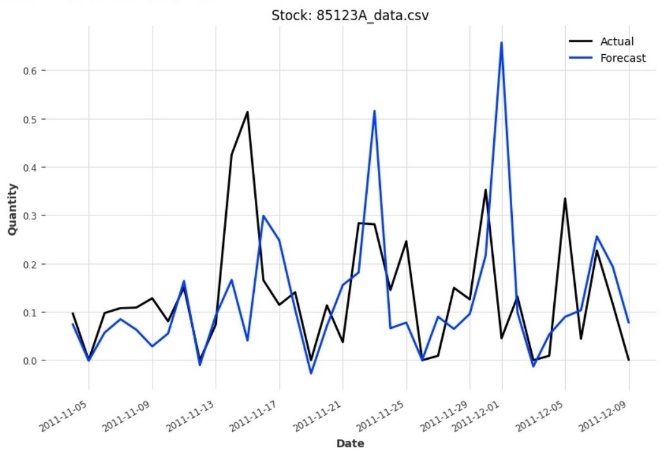
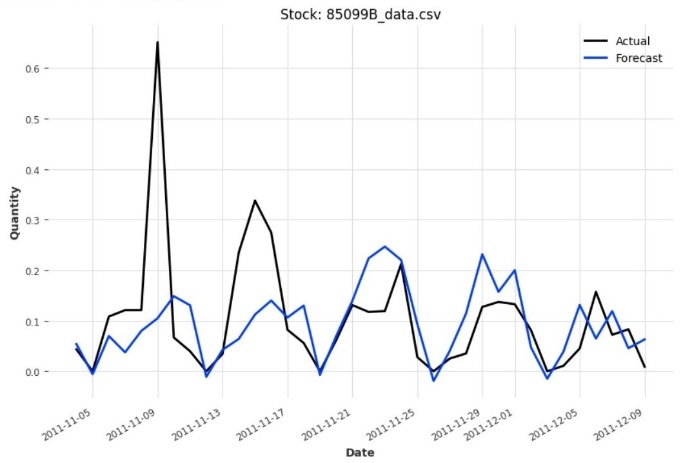


The metrics used for model evaluation, including MAE, MSE, and RMSE, were hired to assess the accuracy of the NBeats model predictions. The gained results verified a low MAE of 0.11, representing that, on average, the model's predictions deviated from the actual demand by a minimal margin. Additionally, the MSE of 0.02 and RMSE of 0.15 further underscored the model's ability to produce accurate forecasts with minimal error.

Overall, the evaluation of the NBeats model revealed its superiority over traditional forecasting models and even outperformed the LSTM model but it was observed to be time-consuming in the training phase. The NBeats model's ability to provide accurate demand forecasts for multiple products simultaneously demonstrates its potential it was observed to be time-consuming in the training phase as a powerful tool for demand forecasting in online retail settings. These findings contribute to advancing the understanding of deep learning techniques in demand forecasting and underscore the significance of leveraging advanced modelling approaches for improved business decision-making.

* + 1. TiDE Model:

In this research investigation, the TiDE (Time-series Dense Encoder) model, implemented through the DART library, emerged as a viable alternative for demand forecasting in the online retail domain. This model, while exhibiting a faster training time compared to the NBeats model, delivered predictions within an acceptable range of accuracy. To determine its performance, this research evaluated the TiDE model across multiple product categories, aggregating the evaluation metrics from each product for comprehensive analysis.

The valuation metrics used for the TiDE model included MAE, MSE, and RMSE. These metrics offer insights into the model's predictive accuracy and the magnitude of errors in its forecasts. By assembling the evaluation results from various product datasets, this research derived an average performance assessment for the TiDE model. Across the evaluated product categories, the TiDE model exhibited consistent performance, with an average MAE of 0.07, MSE of 0.02, and RMSE of 0.12. These metrics indicate that, on average, the TiDE model's predictions deviated from the actual demand by a relatively small margin, suggesting a satisfactory level of accuracy in forecasting demand patterns.

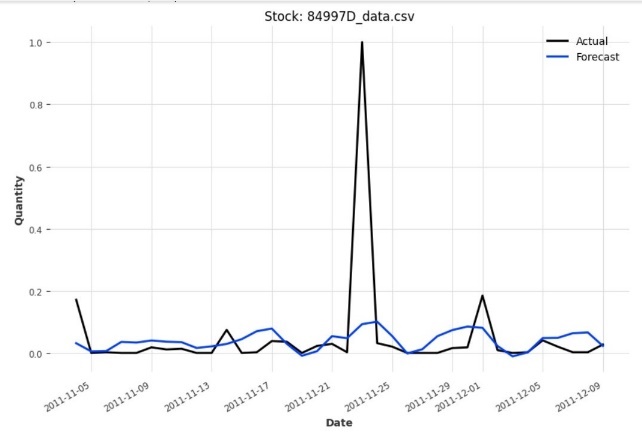
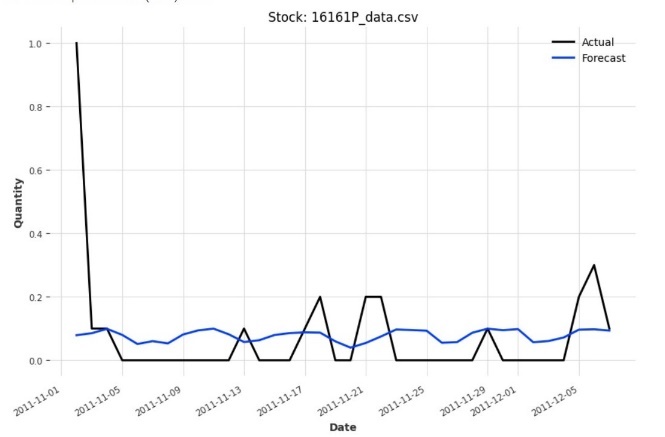
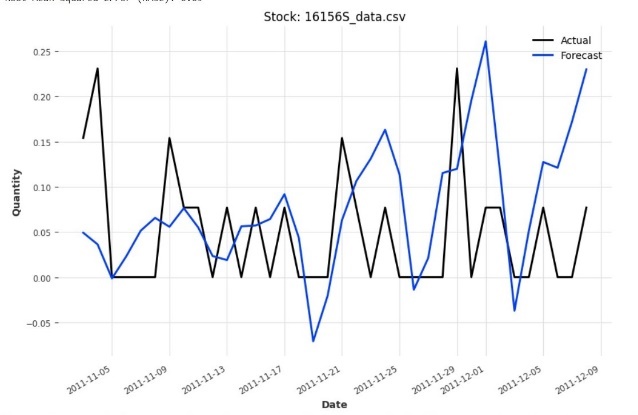
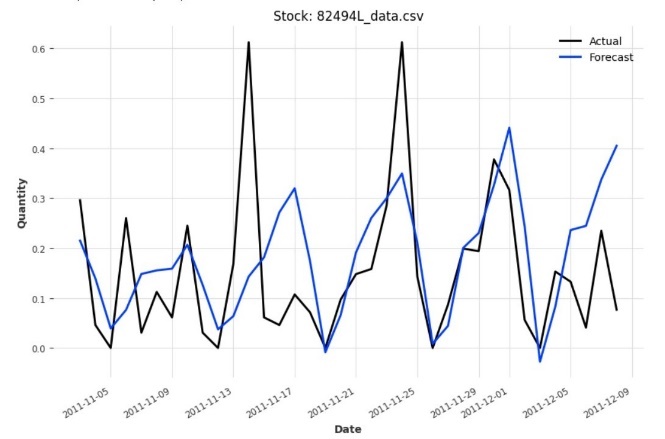
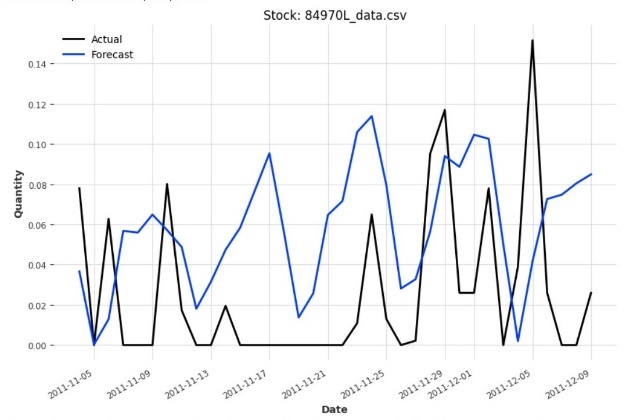
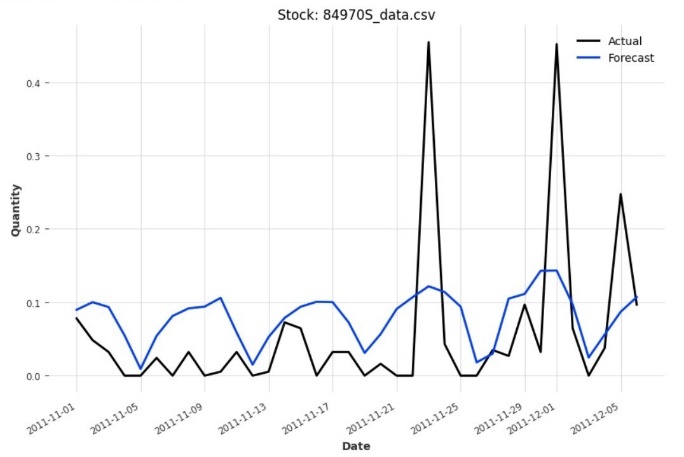
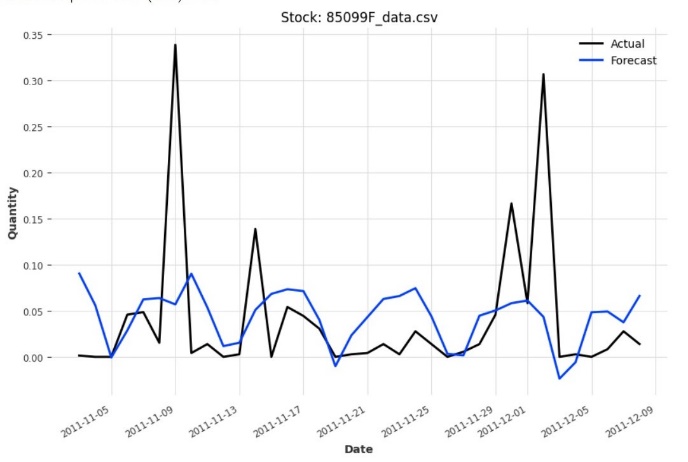
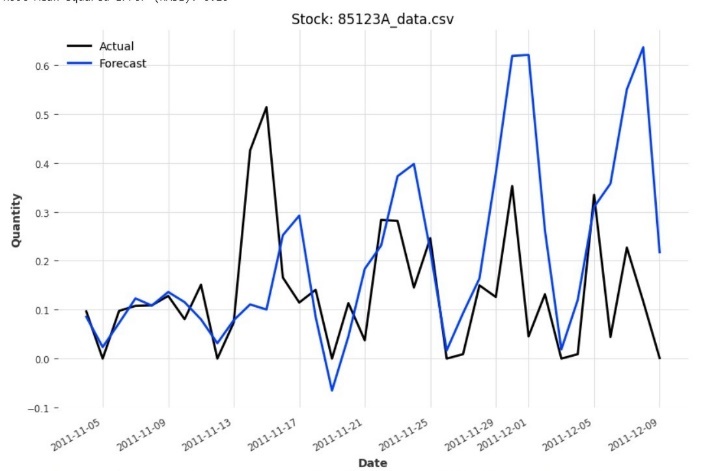
Furthermore, the TiDE model's ability to generate forecasts for diverse product categories underscores its versatility and applicability in real-world scenarios. By accurately predicting demand fluctuations for multiple products simultaneously, the TiDE model provides valuable insights for inventory management and strategic decision-making in online retail operations.

In summary, while the TiDE model surpass the NBeats model in terms of predictive accuracy, its competitive performance and faster training time make it a compelling option for demand forecasting tasks in the online retail sector. These findings contribute to advancing the understanding of deep learning-based approaches in demand forecasting and highlight the significance of leveraging advanced modelling techniques to optimize business operations.

* + 1. BlockRNN Model:

In the exploration of demand forecasting models, the BlockRNN (Block Recurrent Neural Network) model, implemented through the DARTS library, emerged as a robust and efficient solution. This model showcased commendable performance, characterized by rapid training times and accurate prediction capabilities across various product categories.

The valuation metrics, including MAE, MSE, and RMSE, were hired to assess the BlockRNN model's predictive accuracy. By combining the evaluation results from multiple product datasets, this study obtained a comprehensive understanding of the model's performance. Across the evaluated product categories, the BlockRNN model consistently delivered favourable results. With an average MAE of 0.09, MSE of 0.02, and RMSE of 0.15, the model verified its capacity to generate detailed forecasts with minimal error. These metrics underscore the BlockRNN model's effectiveness in capturing demand patterns and providing actionable insights for inventory management and decision-making processes.



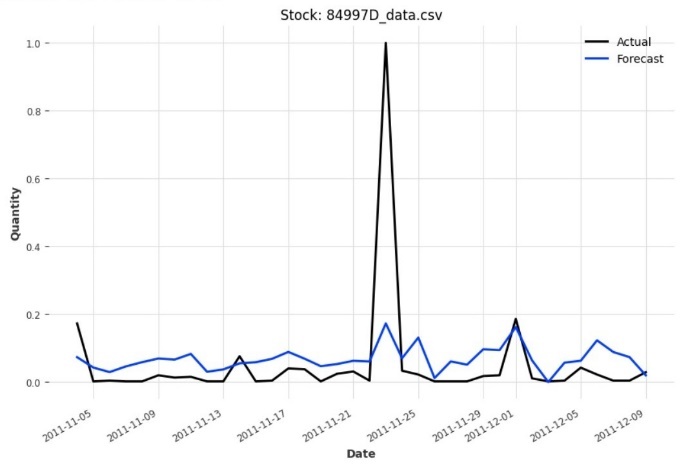
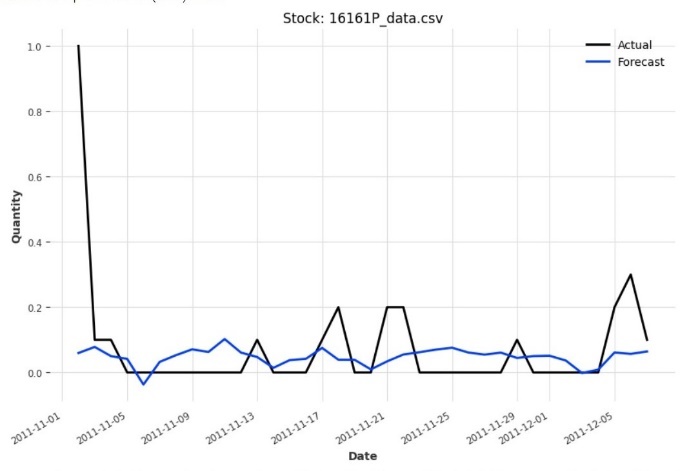
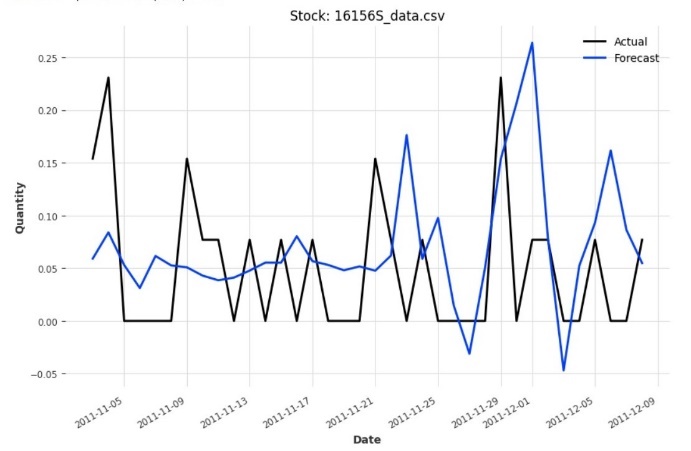
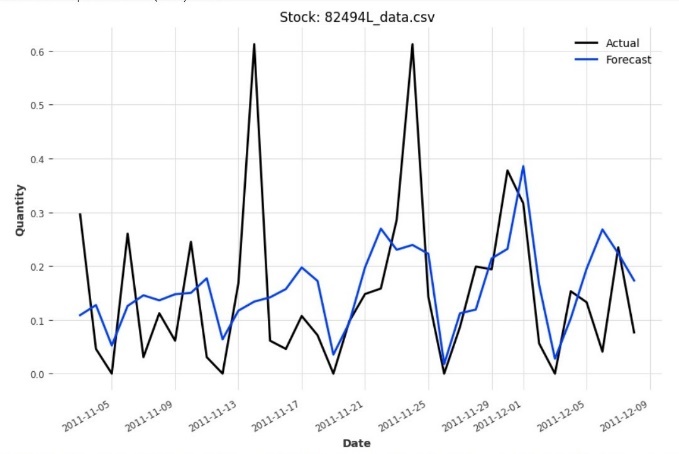
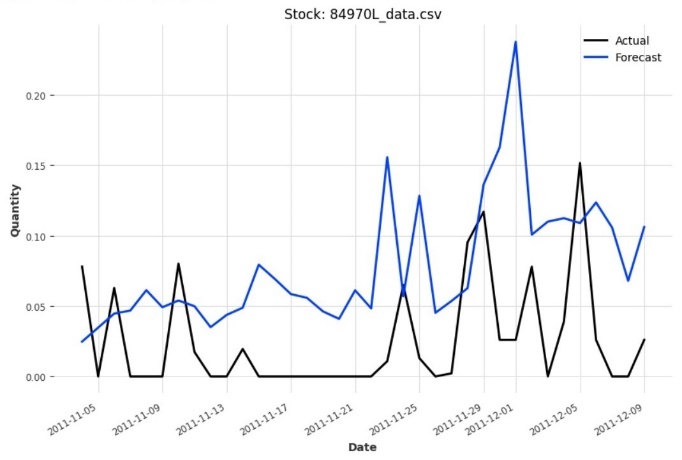
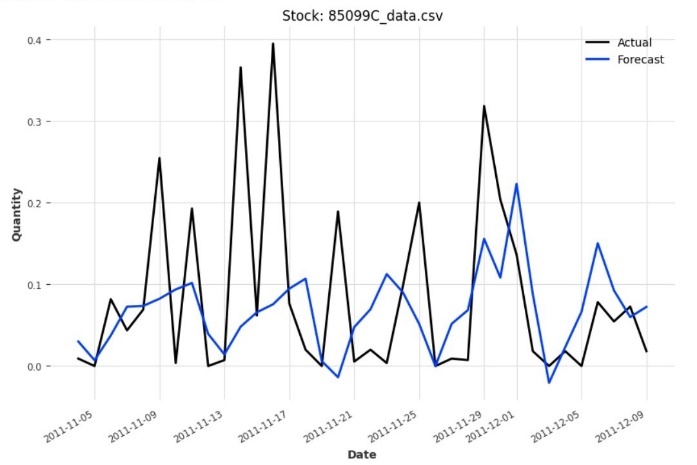
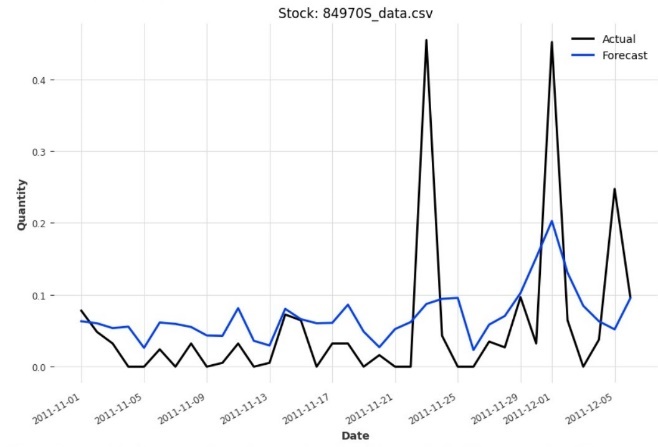
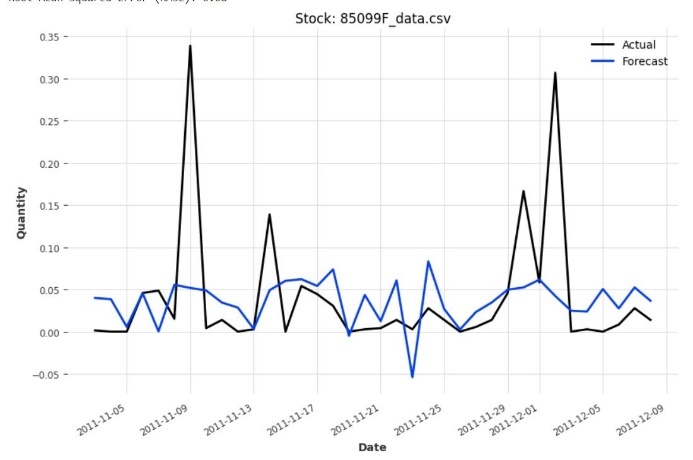
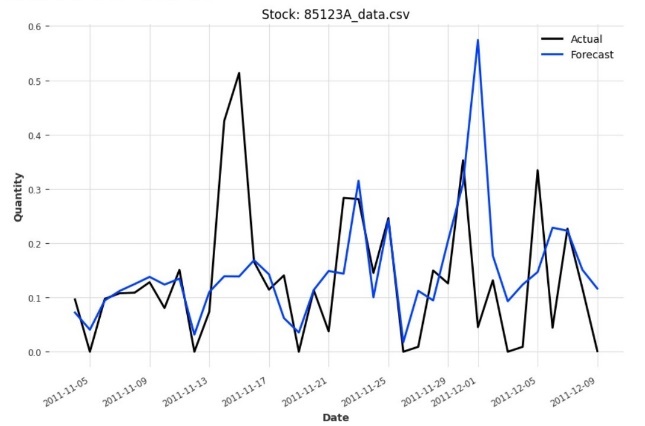
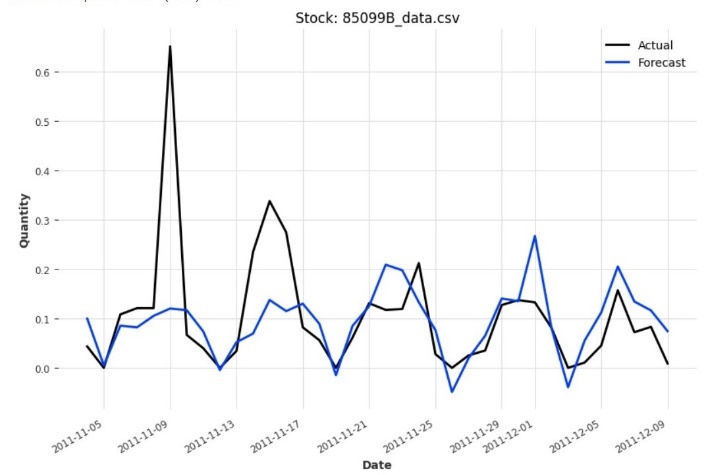
Furthermore, the BlockRNN model's swift training times contribute to its practicality and scalability, making it suitable for real-world applications with large datasets and stringent time constraints. Its combination of quick training and accurate predictions positions the BlockRNN model as a valuable tool for demand forecasting in online retail settings.

In summary, the BlockRNN model's performance highlights its viability as a competitive solution for demand forecasting tasks. Its well-organized training process and consistent prediction capabilities make it a hopeful option for businesses seeking accurate and timely insights into demand patterns. These findings contribute to the ongoing advancement of forecasting methodologies and underscore the importance of leveraging innovative models for improved decision-making in retail operations.

* + 1. TCN Model:

In the examination of demand forecasting models, the Temporal Convolutional Network (TCN) model, implemented through the DARTS library, emerged as a promising solution with commendable prediction accuracy. Despite requiring longer training times compared to the NBeats model, the TCN model demonstrated robust performance across various product categories.

The assessment metrics, including MAE, MSE, and Root RMSE, were used to gauge the TCN model's predictive capabilities. By analysing the evaluation results obtained from multiple product datasets, this study obtained a comprehensive understanding of the model's performance. Across the evaluated product categories, the TCN model consistently delivered accurate predictions. With an average MAE of 0.06, MSE of 0.01, and RMSE of 0.11, the model revealed a high level of precision in forecasting demand patterns. These metrics underscore the TCN model's effectiveness in capturing complex temporal dependencies and generating reliable forecasts will surpass all DARTs library model at demand forecasting in online retail.



Despite the longer training times required by the TCN model, its superior prediction accuracy makes it a compelling option for demand forecasting tasks in the online retail sector. Its ability to accurately predict demand fluctuations across diverse product categories highlights its versatility and applicability shows this is the best model for demand forecasting in online retail.

In summary, the TCN model's performance underscores its potential as a valuable tool for demand forecasting. Its commendable prediction accuracy, although with longer training times, positions it as a promising alternative to traditional forecasting methods has surpassed other DARTS model. These findings contribute to advancing the understanding of deep learning-based approaches in demand forecasting and emphasize the importance of leveraging innovative models for improved decision-making in retail operations.

This exploration with the Dart library unveiled a spectrum of forecasting models with varying trade-offs between training time and prediction quality. The versatility of the Dart library, allowing concurrent predictions for multiple products, makes it a valuable tool in navigating the diverse challenges of forecasting different items within a unified framework.

* 1. Comparison of Models:
     1. Time Taken by DART models:

The following table displays the time taken to train the models for top 10 products.

|  |  |
| --- | --- |
| Model | Time |
| Nbeats | 0:05:27.92 |
| TiDE | 0:02:45.96 |
| Block RNN | 0:02:03.77 |
| TCN | 0:06:55.35 |

* + 1. Results of DART Models with Top 10 Selling Products:

Before delving into the detailed evaluation results of the demand forecasting models, it's essential to preface with an overview of the models used and their respective performances. This research employed four distinct models, namely the NBeats model, BlockRNN model, TiDE model, and TCN model, all implemented through the DARTS library. These models were trained on a diverse range of product datasets from the online retail domain in Top 10 Selling products, each representing unique demand patterns and characteristics. The evaluation metrics employed for model assessment include Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). These metrics provide valuable insights into the accuracy and precision of the models' predictions. As this study proceed to examine the evaluation results for each model across different product categories, this research aims to discern patterns, identify strengths and weaknesses, and ultimately determine the most effective model for demand forecasting in the online retail sector. The results are shown below in Figure 20. The following table shows about each product model and their MAE, MSE, and RMSE.

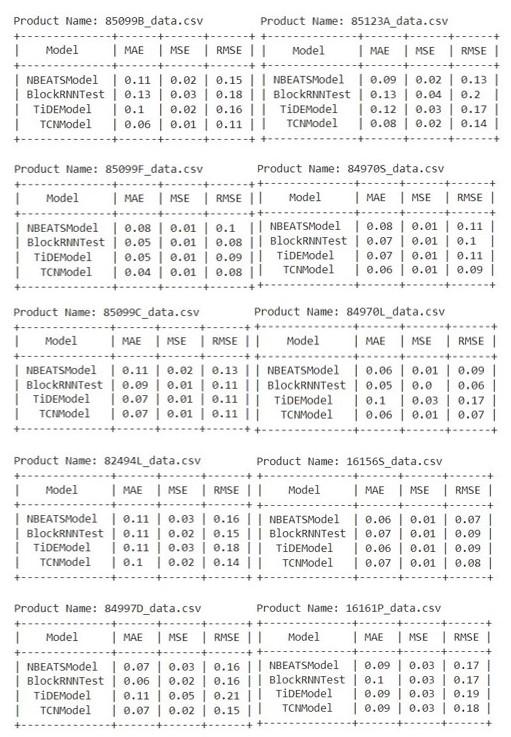


Figure 21: DARTS Model Results

After conducting a comprehensive analysis of the MAE, MSE, and RMSE, alongside an evaluation of the overall prediction performance across all products, it became evident that the TCN model emerged as the front runner in the race of accurate predictions of demand forecasting. Its predictions showed an admirable level of accuracy, as evidenced by the error metrics, indicating its ability to effectively capture the complicated distinctions present in the data. The plot of the DARTS model is shown in the Figure 22 below:

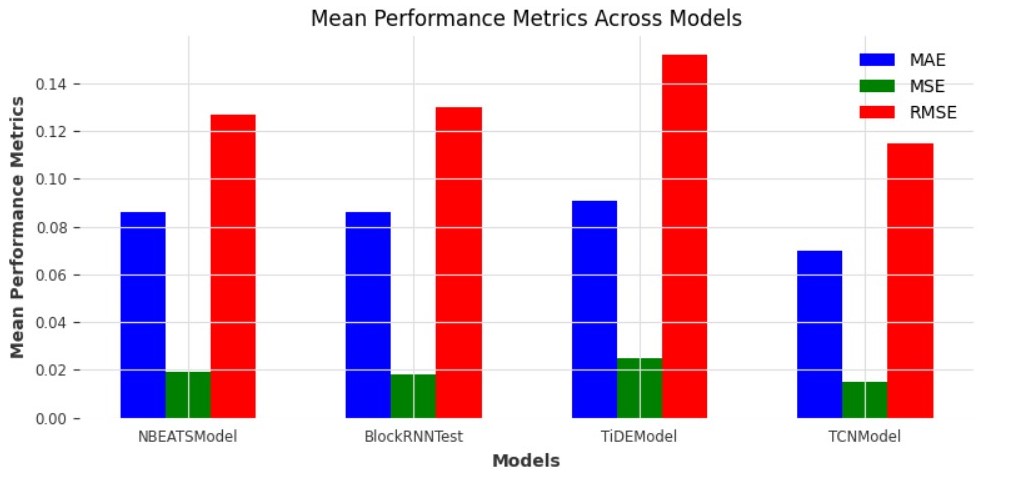


Figure 22: DARTS Model Visualized Result

However, it is essential to note that despite its longer training time, the TCN model surpassed its counterparts in terms of prediction accuracy alone. The TCN model demonstrated superior performance in terms of MAE, MSE, and RMSE metrics, suggesting that it produced the most precise predictions, although at the cost of longer training duration. This comparison underscores the trade-off between training time and prediction accuracy, emphasizing the importance of considering both factors when selecting the most suitable model for demand forecasting tasks.

Conclusion, Limitations and Recommendations:

* 1. Conclusion:

This research project investigated deeply into forecasting product demand for a gift store, employing a diverse array of time series forecasting models and methodologies. Traditional models like ARIMA and SARIMA, despite their widespread use, struggled to capture the complex nuances present in the dataset, resulting in inaccurate predictions. This highlighted the need for more cultured approaches to address the challenges posed by the dataset's complexity. The exploration of the Dart library, a specialized tool personalized for forecasting tasks, introduced a novel approach to the project. This library facilitated simultaneous prediction of multiple items within a single model, thereby enhancing the efficiency and accuracy of predictions. Through careful evaluation of four models from the Dart library NBeats, TiDe, BlockRNN, and TCN using metrics such as MAE, MSE, RMSE, and real-time prediction analysis, insights into their performance were gained.

Among these models, TCN emerged as the top performer, demonstrating exceptional prediction accuracy. Despite its quick training and reliable predictions, NBeats showcased superior accuracy, although with longer training durations, highlighting the trade-off between real-time efficiency and prediction accuracy. The project underscored the importance of tailoring model selection to suit the unique characteristics and requirements of the forecasting task, emphasizing the need for a nuanced approach to model selection. As the field of forecasting continues to evolve, implementation innovative models such as those offered by the Dart library presents hopeful opportunities for enhancing prediction quality. Continuous exploration of new methodologies remains crucial for staying at the forefront of forecasting capabilities.

In conclusion, this project serves as a valuable exploration into the dynamic landscape of time series forecasting, highlighting the importance of a nuanced model selection process to achieve accurate and efficient predictions across diverse scenarios. Concluded continuous exploration and innovation, this research can ensure the advancement of forecasting capabilities to meet the evolving needs of various industries.

* 1. Future Recommendations:

As this research draw the curtains on this forecasting project, this research sight into a horizon dotted with possibilities for future exploration and refinement, offering avenues to bolster the reliability and utility of forecasting models. Looking ahead, one avenue worthy of exploration is ensemble modelling, where the consolidation of diverse forecasting models' strengths could yield a more robust and accurate overall forecast (Lee, J. et al., 2021). By combining predictions from various models, this study could potentially moderate the limitations of individual models and enhance the overall predictive performance. Additionally, the exploration of feature engineering could shed light on additional factors influencing product demand, concrete the way for a richer understanding of demand patterns. The prospect of dynamic model selection presents another interesting possibility, where algorithms could be developed to assess dataset characteristics and dynamically select the most suitable forecasting model. Such an approach could enhance adaptability to changing data dynamics and potentially improve forecasting accuracy (Liu, N. et al., 2013). Moreover, the integration of external datasets, such as weather data or economic indicators, holds promise for enriching forecasting models with valuable insights into demand patterns.

In the realm of real-time forecasting, another area prepared for further investigation is hyperparameter tuning, particularly for the models employed in the Dart library. A deeper dive into fine-tuning these parameters could unlock additional performance gains, further enhancing the predictive capabilities and efficiency of the models. The development of models capable of swiftly adapting to evolving trends and patterns in the data is paramount. Such models could provide timely and relevant forecasts, empowering decision-makers with actionable insights. To facilitate decision-making, the creation of user-friendly interfaces and visualization tools could help stakeholders interact with and interpret forecasting results more effectively. Furthermore, benchmarking forecasting models against industry standards or best practices is essential to ensure their credibility and alignment with industry norms. Collaborating with industry experts to validate models against established benchmarks can strengthen confidence in their performance and applicability.

In conclusion, this project sets the stage for ongoing research endeavours aimed at advancing forecasting methodologies and addressing real-world demand prediction challenges. Through continued exploration of these avenues, researchers can contribute to the continual evolution and refinement of forecasting techniques, development their efficacy and relevance across diverse industry contexts.

* 1. Limitations:

While this comparative study provides valuable insights into machine learning approaches for demand forecasting in online retail, it's essential to acknowledge several limitations that may impact the scope and generalizability of the findings. Firstly, the accuracy of forecasting models heavily relies on the quality and availability of historical data. Incomplete or inaccurate data may introduce biases and limit the reliability of predictions, potentially affecting the validity of comparative analyses across different machine learning approaches. Moreover, the forecasting models evaluated in this study, including ARIMA, SARIMA, and those from the Dart library, operate based on certain assumptions about underlying patterns in the data. Variations from these assumptions could impact the models' accuracy and comparability (Masum, S., Liu, Y. and Chiverton, J., (no date)).

Secondly, the performance of machine learning models, particularly those requiring parameter tuning, may be sensitive to the choice of parameters. Suboptimal parameter selection can lead to less accurate predictions, potentially skewing the comparative evaluation of different approaches. Additionally, the focus of this research on key features of the dataset may limit consideration for external factors that could influence product demand. Incorporating a broader range of external variables, such as economic indicators or promotional events, could enhance the completeness and accuracy of the comparative analysis.

Moreover, the emphasis on historical data in this research may restrict the ability to make long-term forecasts. Generalizing beyond the observed time frame introduces uncertainties that may affect the reliability of comparative evaluations across different machine learning approaches. Additionally, certain machine learning models, particularly those with significant computational demands, may pose challenges in resource-intensive environments. Consideration of computational constraints is vital for practical applicability and scalability of the comparative analysis. The dynamic nature of online retail markets introduces complexities that machine learning models may not fully capture. Unexpected events or shifts in consumer behaviour can significantly impact product demand, potentially influencing the comparative evaluation of different approaches. Lastly, while evaluation metrics such as MAE, MSE, and RMSE provide quantitative insights into model performance, they may not capture all sides of forecasting accuracy. Additional metrics or qualitative assessments may be needed for a comprehensive evaluation of different machine learning approaches. Acknowledging these limitations is crucial for interpreting the findings of this research accurately and guiding future enhancements. Continuous refinement and adaptation of machine learning models will be necessary to address these limitations and advance the effectiveness of demand forecasting in online retail.

References:

Adulyasak, Y. et al. (2023) "Retail Analytics in the New Normal: The Influence of Artificial Intelligence and the Covid-19 Pandemic," arXiv (Cornell University). Available at: https://doi.org/10.48550/arxiv.2312.00046.

Ensafi, Y. et al. (2022) Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. Available at: https://doi.org/10.1016/j.jjimei.2022.100058.

Ferreira, J, K., Lee, A, H, B. and Simchi‐Levi, D. (2016) Analytics for an Online Retailer: Demand Forecasting and Price Optimization. Available at: https://doi.org/10.1287/msom.2015.0561.

Kontopoulou, I, V. et al. (2023) A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. Available at: https://doi.org/10.3390/fi15080255.

Liu, N. et al. (2013) Sales Forecasting for Fashion Retailing Service Industry: A Review. Available at: https://doi.org/10.1155/2013/738675.

Punia, S. et al. (2020) Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail. Available at: https://doi.org/10.1080/00207543.2020.1735666.

Wang, J., Liu, Q, G. and Liu, L. (2019) A Selection of Advanced Technologies for Demand Forecasting in the Retail Industry. Available at: <https://doi.org/10.1109/icbda.2019.8713196>.

Resul, T. and Sule, O. (2020) Demand Prediction Using Machine Learning Methods and Stacked Generalization. Available at: <https://doi.org/10.48550/arXiv.2009.09756>.

Tsai, C. et al. (2013) "A comparative study of hybrid machine learning techniques for customer lifetime value prediction," Kybernetes, 42(3),p. 357-370. Available at: <https://doi.org/10.1108/03684921311323626>.

Kumar, A., Shankar, R. and Aljohani, R, N. (2020) A big data driven framework for demand-driven forecasting with effects of marketing-mix variables. Available at: <https://www.sciencedirect.com/science/article/pii/S0019850118306606>.

Gupta, P. et al. (2021) Implementation of Demand Forecasting – A Comparative Approach. Available at: <https://doi.org/10.1088/1742-6596/1714/1/012003>.

Verstraete, G., Aghezzaf, E. and Desmet, B. (2019) A data-driven framework for predicting weather impact on high-volume low-margin retail products. Available at: <https://doi.org/10.1016/j.jretconser.2019.02.019>.

Cappiello, C., Samá, W. and Vitali, M. (2018) Quality awareness for a Successful Big Data Exploitation. Available at: <https://doi.org/10.1145/3216122.3216124>.

Gupta, D, S., Fuehrer, F. and Jeyachandra, C, B. (2014) Production Forecasting in Unconventional Resources using Data Mining and Time Series Analysis. Available at: <https://doi.org/10.2118/171588-ms>.

Plotnikova, V., Dumas, M. and Milani, P, F. (2000) Adapting the CRISP-DM Data Mining Process: A Case Study in the Financial Services Domain. Available at: <https://www.semanticscholar.org/paper/Adapting-the-CRISP-DM-Data-Mining-Process%3A-A-Case-Plotnikova-Dumas/a4e44d5ce058e6b815631947db23044a22f8b114>.

Sun, J. and Sun, Y. (2018) "Analysis of Data Mining Based on Object Oriented Analysis Method," Journal of physics. Conference series, 1069,p. 012041-012041. Available at: https://doi.org/10.1088/1742-6596/1069/1/012041.

Schröer, C., Kruse, F. and Gómez, M, J. (2021) A Systematic Literature Review on Applying CRISP-DM Process Model. Available at: https://doi.org/10.1016/j.procs.2021.01.199.

Fisher, L, M. and Vaidyanathan, R. (2014) A Demand Estimation Procedure for Retail Assortment Optimization with Results from Implementations. Available at: https://doi.org/10.1287/mnsc.2014.1904

Drieniková, K. and Sakál, P. (2012) Respecting Stakeholders and Their Engagement to Decision Making – The Way of Successful Corporate Social Responsibility Strategy. Available at: https://sciendo.com/article/10.2478/v10186-012-0028-z.

Kramar, V. and Alchakov, V. (2023) Time-Series Forecasting of Seasonal Data Using Machine Learning Methods. Available at: https://doi.org/10.3390/a16050248.

Ferreira, J, K., Lee, A, H, B. and Simchi‐Levi, D. (2016) Analytics for an Online Retailer: Demand Forecasting and Price Optimization. Available at: https://doi.org/10.1287/msom.2015.0561.

Ahmed, K, N. et al. (2010) An Empirical Comparison of Machine Learning Models for Time Series Forecasting. Available at: https://doi.org/10.1080/07474938.2010.481556.

Ma, T. et al. (2014) Robust Logistic and Probit Methods for Binary and Multinomial Regression. Available at: https://doi.org/10.4172/2155-6180.1000202.

Pongdatu, N, A, G. and Putra, H, Y. (2018) Seasonal Time Series Forecasting using SARIMA and Holt Winter’s Exponential Smoothing. Available at: https://hcvalidate.perfdrive.com/fb803c746e9148689b3984a31fccd902//.

Kramar, V. and Alchakov, V. (2023) Time-Series Forecasting of Seasonal Data Using Machine Learning Methods. Available at: https://doi.org/10.3390/a16050248.

Shih, Y. and Lin, M. (2019) A LSTM Approach for Sales Forecasting of Goods with Short-Term Demands in E-Commerce. Available at: https://doi.org/10.1007/978-3-030-14799-0\_21.

Siami‐Namini, S., Tavakoli, N. and Namin, S, A. (2019) "A Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM, and BiLSTM," arXiv (Cornell University). Available at: https://doi.org/10.48550/arxiv.1911.09512.

Longpre, S. et al. (2016) A Way out of the Odyssey: Analyzing and Combining Recent Insights for LSTMs. Available at: https://doi.org/10.48550/arxiv.1611.05104.

Nasir, A, N. and Jeong, S. (2021) Fast Content Delivery Using a Testbed-Based Information-Centric Network. Available at: https://doi.org/10.1109/access.2021.3096042.

Al-Razgan, M., Al-Khalifa, S, A. and Al‐Khalifa, S, H. (2013) Educational Data Mining: A Systematic Review of the Published Literature 2006-2013. Available at: https://doi.org/10.1007/978-981-4585-18-7\_80.

Marzi, F. et al. (2023) "Towards a Prediction of Machine Learning Training Time to Support Continuous Learning Systems Development," arXiv (Cornell University). Available at: https://doi.org/10.48550/arxiv.2309.11226.

Atique, S. et al. (2020) Time series forecasting of total daily solar energy generation: A comparative analysis between ARIMA and machine learning techniques. Available at: https://doi.org/10.1109/greentech46478.2020.9289796.

Masum, S., Liu, Y. and Chiverton, J. (no date) Comparative analysis of the outcomes of differing time series forecasting strategies. Available at: <https://ieeexplore.ieee.org/document/8393069/>.

Dataset: <https://archive.ics.uci.edu/dataset/502/online+retail+ii>

GitHub: <https://github.com/muhammadahsan2023/Capstone_Project>