**Arab American University**

**Faculty of Graduate Studies**

**Department of Natural, Engineering and Technology Sciences**

**Master Program in Data Science and Business Analytics**

# Products Sales Forecasting Using Statistical And Machine-Learning Models – A Case Study

**Isra Abdul-Ellah Khanfar**
**202112678**

**Supervision Committee:**

**Prof. Mohammed Awad**

**Prof. Yousef Daraghma**

**Dr. Ahmed Ewais**

**This Thesis Was Submitted in Partial Fulfillment of the Requirements for the Master Degree in Data Science and Business Analytics.**

**Palestine, September/2024**

**Arab American University**

**Faculty of Graduate Studies**

**Department of Natural, Engineering and Technology Sciences**

**Master Program in Data Science and Business Analytics**

# Thesis Approval
# Products Sales Forecasting Using Statistical And Machine-Learning Models – A Case Study

Isra Abdul-Ellah Khanfar
202112678

This thesis was defended successfully on 26/9/2024 and approved by:

Thesis Committee Members:

| | Name | Title | Signature |
|---|---|---|---|
| 1. | Prof. Mohammed Awad. | Main Supervisor | |
| 2. | Prof. Yousef Daraghma. | Members of Supervision Committee | |
| 3. | Dr. Ahmed Ewais. | Members of Supervision Committee | |

Palestine, September/2024

## Declaration

I declare that, except where explicit reference is made to the contribution of others, this thesis is substantially my own work and has not been submitted for any other degree at the Arab American University or any other institution.

Student Name: Isra Abdul-Ellah Khanfar.

Student ID: 202112678.

Signature: Isra Khanfar.

Date of Submitting the Final Version of the Thesis: 19 Mar 2025.

## Dedication

This thesis is dedicated to my beloved parents, who have been my greatest source of inspiration and endless support. To my brothers, Dr.Nehad Khanfar and Abdul-hameed Khanfar, who were my light in the middle of the darkness, and those who paved this path for me and walked with me till the end. For my brothers and sisters unwavering belief in me of strength.

Student Name:

Isra Abdul-Ellah Abdul-Hameed Khanfar.

## Acknowledgments

I would like to express my sincere thanks to my supervisor, Prof. Mohammed Awad, for his guidance, valuable insights throughout the challenges of this thesis. His insights and advices have been invaluable. I am profoundly grateful. I would like to extend my sincere gratitude to the Arab American University and the supportive academic environment and for the opportunities for growth and learning.

## Products Sales Forecasting Using Statistical and Machine- Learning Models – A Case Study

**Isra Abdul-Ellah Abdul-Hameed Khanfar**

**Prof. Mohammed Awad, Dr. Ahmed Ewais, Dr. Yousef-Awwad Daraghma**

## Abstract

Sales forecasting is considered a pivotal tool to manage businesses from various disciplines, and a foundation to build an effective planning process in the company. Business owners' priority appears mainly in making accurate sales estimates to limit the challenges of underestimating or overestimating sales that affect their business costs and operations. In this thesis, two statistical models were applied, which are Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA), and four Neural Networks (NNs) which are Recurrent Neural Networks (RNNs), Long-Short Term Model (LSTM) Multi-Layer Perceptron Neural Networks (MLPNNs) and Radial Basis Function Neural Networks (RBFNNs). A statistical model has been combined with each neural network model to build four hybrid models. This study examines the efficiency in predicting sales and capturing patterns for five products as pure models and as hybrid models. Two scenarios were followed to apply these models. The first scenario was a combination of sales for five products, and the second scenario was based on each product level (individually). Models performance evaluation were used (MSE, RMSE, and MAE). Final results have shown that forecasting sales individually for each product is better than forecasting sales for all products as a combination. Results have shown that hybrid models of ARIMA-MLPNNs significantly improve prediction accuracy compared to individual statistical models, four neural networks models, and other hybrid models for combined products sales. The ARIMA-MLPNNs hybrid model has achieved an RMSE of 131.64 followed by the ARIMA-LSTM demonstrated an RMSE of 447.68, which has achieved better performance than the individual statistical model of SARIMA, four neural networks and other hybrid models. For individual product sales, the ARIMA-MLPNNs model has achieved RMSE of 31.13 for dairies, 19.54 for ice-cream, 51.74 for drinks, 60.99 for snacks and chips, and 74.21 for cleaning materials, while ARIMA-LSTM demonstrated better performance than individual statistical model SARIMA, four neural networks and other hybrid models with an RMSE of 80.54 for dairies, 50.64 for ice-cream, 169.13 for drinks, 188.03 for snacks and chips and 167.93 for cleaning materials. These findings suggest that hybrid models can provide more accurate predictions for products sales forecasting.

Keywords:
Sales Forecasting, Statistical Models, Neural Network Models, Hybrid Models, Error Metrics.

# Table of Contents

## List of Tables

## List of Figures

## List of Appendices

## List of Definitions of Abbreviations

| Abbreviations | Title |
|---|---|
| AI | Artificial Intelligence |
| ANNs | Artificial Neural Networks |
| ARIMA | Autoregressive Integrated Moving Average |
| ARIMA-LSTM | Autoregressive Integrated Moving Average-Long Short-Term Memory |
| ARIMA-MLPNNs | Autoregressive Integrated Moving Average-Multi-Layer Perceptron Neural Networks |
| ARIMA-RBFNNs | Autoregressive Integrated Moving Average-Radial-Basis Function Neural Networks |
| ARIMA-RNNs | Autoregressive Integrated Moving Average-Recurrent Neural Networks |
| d | The number of differences needed to make the data stationary |
| D | The number of seasonal differences to remove seasonal trends |
| LSTM | Long Short-Term Memory |
| MAE | Mean Absolute Error |
| ML | Machine Learning |
| MLPNNs | Multi-Layer Perceptron Neural Networks |
| MSE | Mean Square Error |
| NNs | Neural Networks |
| ADF | Augmented Dicky-Fuller |
| p | The number of lag observations included in the model (lag order) |
| P | The number of seasonal lags used for the seasonal autoregressive order |
| q | The number of lagged forecast errors |
| Q | The number of lagged forecast errors in the seasonal part |
| RBFNNs | Radial-Basis Function Neural Networks |

| | |
|---|---|
| RMSE | Root Mean Square Error |
| RNNs | Recurrent Neural Networks |
| S | Length of the seasonal cycle |
| SARIMA | Seasonal Autoregressive Integrated Moving Average |

# Chapter One: Introduction

## 1. Introduction

Sales forecasting is considered a substantial tool for managing business, and a foundation to build an effective planning process in the company. It is considered a business discipline, which can be found in various sectors such as marketing, operations, strategic planning, and finance to predict sales not only for production or inventory (Florance et al., 1993). If managers and business owners do not have forecasting capabilities, they will base their decisions on their personal experience or on a series of events that previously occurred during last year, last month, or at any previous time frame (Chase et al., 2014; Georgoff & Murdick, 1986; Staff, 2024).For retail businesses, the priority of business owners lies primarily in their pursuit of accurate estimates of sales through sales volumes. This will necessarily help them to face the most two important challenges in sales such as underestimated sales which results in the product being out of stock and overestimation sales which might cause the shelf life of products, which increases the cost of storage, products, and operations (Eglite & Birzniece, 2022), where the importance of obtaining accurate sales forecasting allows firms monitoring their costs (E. & K., 2023). Since meeting customer demands is the most fundamental goal that successful businesses seek to achieve at the right time and in the right place, businesses must invest in planning and forecasting processes, technology systems, methods and metrics, inventories, and business analytics for better demand satisfaction (Lawless & Mark, 2014).Accurate sales forecasting has advantages that expand to include aspects such as efficient customer service, which directly affects business performance (Hyndman & Athanasopoulos, 2018). Sales forecasting can also be used in significant managerial decision-making within companies (Moon, 2018). Management can take advantage of employing sales forecasting to help those forecast sales and then take advantage of the forecasted information to develop their plans for resources and capacity to meet the demand efficiently (E. & K., 2023).This study aims to develop and evaluate a varied array of sales forecasting models to capture the complexities of supermarket product sales dynamics and consumer behavior. Time series forecasting plays a significant role in sales forecasting by using historical data to predict future sales trends. It helps businesses to distinguish patterns such as seasonality, trends, and cyclical fluctuations and facilitates more accurate sales forecasting. Using techniques like statistical and machine learning models, helps businesses to make strategic decisions about inventory management, customer needs, and resource allocation. Therefore, time series forecasting not only boosts operational efficiency but also participates in reducing costs and maximizing profits by adjusting supply with expected demand.

Sales volumes are the primary focus of any supermarket, identifying accurate sales predictions is significant to the success of every business, and it necessarily leads to the ultimate goal of any supermarket, which is making profits. Moreover, the volume of sales data has been constantly increasing in recent years, and, certainly, supermarket owners cannot work with this huge amount of data smoothly without accurate sales forecasting help (Silvia Priscila et al., 2023). The forecasting of supermarket sales can be achieved using many ways, but historically, many supermarkets relied on traditional statistical models (Silvia Priscila et al., 2023). Time series forecasting is a popular concept, and studies have been conducted since the first published forecasting studies (de Almeida, 2021), a signified milestone in time series methods introduced by (George & Jenkins, 1976)book, which presented ARIMA models with its multiple time series processing techniques, namely autoregressive, integration, and MA modeling. Machine learning (ML) is an additional field that investigates time series forecasting (Marie-Aude Aufaure & Zimányi, 2013).Machine learning (ML) is a field within artificial intelligence (AI) that suggests algorithms to discover patterns in data, learning about these patterns without finding a closed-form function explicitly to describe them (Nilsson, 2003; de Almeida, 2021).Furthermore, ML techniques work well with non-linear series and commonly achieve effective results virtually (Zhang, 2003).

In this thesis, data was gained from the supermarket point of sale (POS), including historical sales records for five types of products which are: dairy, ice cream, drinks, snacks and chips, and cleaning materials products. The dataset contains 4875 records spanning from (1-3-2021) to (31-10-2023) with 975 data points for each category, providing a strong foundation for analysis. The preprocessing steps including checking the dataset from duplicates and null values, and date values were converted into date format to ensure accurate time-series analysis, followed with exploratory data analysis (EDA) using statistical and visualization techniques to understand the data’s underlying patterns and trends.

The following approach includes the employment of Neural Networks (NNs) which are advanced machine learning algorithms, traditional statistical models for time series forecasting, and hybrid techniques that merge the strengths of both approaches. In the machine learning domain, four algorithms will be explored such as Recurrent Neural Networks (RNNs), Long-short Term Memory Networks (LSTM), Multi-Layer Perceptron Neural Networks (MLPNNs), and Radial-Basis Function Neural Networks (RBFNNs). These models excel in handling complex, nonlinear relationships within the data, making them featured in capturing accurate patterns and trends in sales data, and are better suitable for long-term forecasting due to their strength to capture complex, nonlinear relationships and dependencies in the data over extended time horizons. On the other hand, statistical methods used for time horizons such as Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) techniques will be used to capture the underlying trends, seasonality, and cyclicality in the sales data, also they are effective for short-term forecasting due to their power to capture short-term trends and patterns in data. Furthermore, the hybrid approaches that combine the advantages of both neural networks (NNs) and statistical methods will be explored to provide the best forecasting results. Ensemble methods like weighted averages, will be used to combine predictions from multiple models to improve overall accuracy and robustness. This study is preparing to provide a remarkable step to the small-medium businesses sector in the local community by developing a product sales forecasting model based on time series analysis through the utilization of machine learning, statistical models, and hybrid models which takes advantage of their joint capabilities in obtaining more accurate sales forecasting model depending on real and local sales data. Through developing and comparing these varied forecasting models as individual and combined models, the research seeks to provide clear insights into their performance in predicting sales at both aggregated sales for all categories and based on each category level. Through accurate evaluation and comparison, this study seeks to recognize the most effective approach or combination of approaches for optimizing sales forecasting in small-medium businesses. By adapting models to the Palestinian supermarket industry, which belongs to the small-medium business, the research seeks to expand the knowledge by highlighting the significance of studying the product's sales data behavior through the time series analysis of main patterns or components ordinarily observed in sales data.

Using error metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), to assess and compare the accuracy and robustness of each model in capturing the underlying patterns and trends in forecasting sales at both aggregated and category levels.

## 1.1 Problem Statement

Supermarkets provide customers with day-to-day products, chiefly food products such as dairy, drinks, snacks, and chips products, and many other types like cleaning materials. Occasionally, supermarkets might have too many of some products and a supply shortfall of another, like having too many products of drinks products but not enough dairy products. This occurs because it is difficult for supermarket owners to predict what shoppers will buy and how much they will buy. Small-medium businesses such as supermarkets in the Palestinian local community, rely on analyzing typical Point of Sale (POS) records using their personal experience to study sales behavior over various time horizons. By scrutinizing these records, owners can review records into past sales, allowing them to construct expectations for future sales performance. These POS records serve as a primary dataset, allowing owners to make informed decisions and projections regarding inventory management, marketing strategies, and overall business operations. Hence, there exists an imperative demand for the development of accurate sales forecasting models that carefully capture underlying trends and patterns and non-linear dependences, that way furnishing business owners with the ability to make accurate predictions regarding their product sales within short-term or long-term time horizons. This necessity emphasizes the requirement for utilizing and employing advanced analytical techniques, particularly time-series analysis, to effectively model and forecast product sales. Through the utilization of such methodologies, businesses of such sizes can attain enhanced vision, enabling informed decision-making, optimized resource allocation, and strategic planning.

### 1.2 Objectives

The primary objective of this study is to develop and evaluate forecasting models using statistical models, advanced machine-learning models, and hybrid models that combine the strengths of statistical and advanced machine-learning models to identify the most effective approach for accurate product sales forecasting as aggregated sales or based on each product level (individually). The final model will forecast a sequence of time series that represents a curve of the local supermarket product sales during the periods from the past to the future. The study aims to achieve the following objectives:

- Developing and building a sales forecasting model using advanced machine learning and statistical techniques to enhance local supermarket inventory management and improve decision-making.
- Processing products sales data from local supermarket point of sales (POS), and evaluate the performance of different advanced machine-learning, statistical techniques, and hybrid models.

- Investigating and experiment with state-of-the-art machine learning and statistical techniques for forecasting sales to assess their suitability for the target problem.
- Studying the influence of local factors such as holidays, weekends, or internal local events on sales.

## 1.3 Contribution

This study contributes to the improvement of sales forecasting methodologies by exploring the effectiveness of statistical models (SM), machine learning (ML), and hybrid models in predicting product sales which are expected to have the following contributions:

1. The primary contribution of this study lies mainly in its novel contribution to the field of supermarket sales forecasting in Palestine and the implication of practical validation using a real-world dataset. Applying traditional statistical models for time series forecasting and advanced machine-learning models to capture underlying patterns and trends, short-term fluctuations, and long-term trends to provide experimental findings of the model's effectiveness in an actual business frame.
2. Suggesting a hybrid approach of machine-learning and statistical techniques to take advantage of both approaches, is a promising strategy for building a strong and accurate sales forecasting model. The adjustability to data dynamics using a hybrid approach can efficiently confirm the nature of sales data, through capturing the short-term fluctuations and long-term trends, as well as any unexpected changes or irregularities that may happen in the sales patterns.
3. The results will be compared using accuracy measurement methods such as MSE, RMSE, and MAE. Analyzing results and selecting the best methods of prediction.
4. The study results will enhance sales forecasting strategies and guide decision-making for local small-medium business owners such as supermarkets. A local supermarket located in Ramallah, has been taken as a case study. Moreover, it will help them to bypass the problems of inventory issues of “overstocking” or “understocking”, and minimize costs linked with good inventory management.

## 1.4 Thesis Structure

The remainder of this thesis is organized as follows: In the next chapter (Chapter 2), an introduction of the thesis subject background, which has started with a general background of sales forecasting definition and significance in business, time series analysis for sales forecasting, in addition to related works. In (Chapter 3), which is the main chapter, the proposed methodology and applied models will be introduced, the datasets description under two scenarios, datasets preparation and exploratory data analysis (EDA), also the applied and developed models will be introduced, and finally the evaluation criteria for models performance. In (Chapter 4), the experiments results will be presented for models as individual and hybrid models, in addition to the future forecasting for next month. (Chapter 5) is dedicated to the thesis conclusions, challenges and limitations, and proposes some future research directions to be conducted in the future.

# Chapter Two: Literature Review

## 2. Background

Small and medium businesses in the retail sector strive for a harmonious progression through the adoption of clear and accurate business strategies to ensure their success and continuity, as well as to gain their profit objectives. Sales forecasting for small and medium businesses, such as supermarkets, where the investigation of product sales is essential for creating informed estimates about future sales levels. This helps direct demand planning and efficiently manage inventory levels and cash flows. By leveraging advanced machine learning models, such as Artificial Neural Networks (ANNs), businesses can use historical data to generate more precise sales forecasts, capture complex non-linear relationships and dependencies over extended time horizons, and apply statistical methods to capture underlying trends, seasonality, and cyclicality. Those approaches are effective for short-term forecasting.

### 2.1 Sales Forecasting in Business

Many studies and researches have provided studied sales forecasting. For instance, some definitions focus on sales forecasting as predictions, while others focus on factors and business challenges. In previous works, the terms “sales prediction” and “demand forecast” have been used synonymously with “sales forecasting”. The researchers (Eglite & Birzniece, 2022) used the term “sales forecasting” because it is related to time series historical data for calculating the future value of sales. This goes with the definition agreed upon by researchers, who describe sales forecasting as the process of estimating product sales over some future period (Hibon & Evgeniou, 2005; Polanski & Stoja, 2017).Other researchers have expanded on this definition by noting that sales forecasting involves estimates or predictions of sales activities for the forecast period, which depend on past sales performance of the product or service (Chase, 2013; *Sales Forecasting*, n.d.). Others, such as (*Sales Forecasting Management: A Demand Management Approach*, 2005; Singh, 2016) introduced sales forecasts as projections of the future expected demand based on a set of environmental conditions. Regarding factors that affect sales forecasting, some researchers have suggested that these factors, which influence forecasting for predicting future demand (or sales), are based on the past and continue to impact the present and future still have an impact in (Chase, 2013). From a holistic point of view, forecasting is a process in which information gained from historical data is used to make predictions or estimates about the future (Haataja, 2016). This guides decision-makers on what their next step should be and helps in planning for the future. Sales forecasting should minimize uncertainty in management regarding strategic decisions and resource allocation (Haataja, 2016). One of the forecasting approaches assumed that past patterns would continue. However, this approach was based simply on historical data and did not account for any changes that occurred in the market, customers, technologies, or within the company itself. Obtaining more data will develop forecasting accuracy. A good approach should consider the whole business, including departments such as marketing. Knowing the demand drivers that lead to sales is important for sales forecasts (Florance et al., 1993).The time horizon of sales forecasting represents the duration over which how long-range forecasts need to be made. The time horizon varies between short-term and long-term sales forecasts, which depend on the needs of the business unit. For example, sales managers ask for short-term sales forecasts because their focus is on shorter time frames (Georgoff & Murdick, 1986).

### 2.2 Time Series Analysis for Sales Forecasting

Several industries and organizations usually use time series data, which describes any information collected over a regular interval of time within their operations (*A Guide to Time Series Analysis in Python | Built In*, n.d.). A time series is a collection of sequential data sampled in a specific time unit, which is used to record a process output to analyze its evolution (Hyndman & Athanasopoulos, 2018).In another words, it is a sequence of records sorted by a time parameter, which may be measured continuously or discretely. Continuous time series are listed instantaneously and steadily, while measurements that are made at regular intervals are described as discrete time series data (Granger & Newbold, 2014).Time series data are used widely in many fields such as stock market price variation (Leung et al., 2014; Mondal et al., 2014), energy consumption rates, social media engagement metrics, and retail demand, among others (*A Guide to Time Series Analysis in Python | Built In*, n.d.). In time series analysis, the main patterns or components ordinarily observed in data are known as trend, seasonal, and cyclic (Hyndman & Athanasopoulos, 2018). Each component will be explained below:

Trend: A *trend* appears when there is a long-term increase or decrease in the data, and it does not have to be linear. Occasionally the trend is described as “changing direction” when it might behave from an increasing trend to a decreasing trend (Hyndman & Athanasopoulos, 2018).

Seasonal: A *seasonal* pattern appears when a time series is influenced by seasonal factors such as the time of the year or the day of the week. Seasonality is characterized by a fixed and known frequency. The frequency is unchanging and connected with aspects of the calendar (Hyndman & Athanasopoulos, 2018).

Cyclic: A *cycle* appears when the data display rises and falls, which are not of a fixed frequency. The fluctuations that are usually appeared due to economic conditions are often related to the "business cycle". The interval of these fluctuations is usually at least 2 years (Hyndman & Athanasopoulos, 2018).

### 2.3 Related Work

The main focus of this study is to provide product sales forecasting by introducing the best forecasting performance. Each time series consists of complex linear and non-linear patterns which is difficult to forecast. As a result, there are various techniques to solve time-series forecasting problems (Zhao & Wang, 2017). For better dealing with time-series forecasting, each case might be solved with a diverse approach. Moving Average (MA) is a simple prediction technique for time-series projections without clear seasonal pattern (Chopra & Meindl, 2019).In many time-series cases, another sophisticated version of (MA) was used, which is known as Autoregressive Integrated Moving Average (ARIMA). In 1980s, the general perception was that the ARIMA model provided more accurate forecasts than other econometric models for direct and short-term forecasts (O'Donovan, 1983). This insight confirmed when ARIMA model forecasted the income items more accurately than Census X-11 and random walk models (Dugan et al., 1994). The main restriction of ARIMA models is the linear relation between independent and dependent variables (Dorne, n.d.). However, comparing ARIMA approaches advantages with neural networks, it offers better understanding of the studied phenomenon, and analyze the coefficients which provides the significance of each independent variable with the related dependent variable. This has allowed the model to produce the knowledge that explain the complex interdependencies regarding the considered time series (Aburto & Weber, 2007).ARIMA model with other traditional time-series forecasting methods such as winter's exponential smoothing and multivariate regression selected for aggregate retail sales data due to its ability to model trend and seasonal fluctuations (Alon et al., 2001). For the same reason of repeatable fluctuation patterns for retail sales of five different categories of women's footwear, the state space models and ARIMA models were applied (Ramos et al., 2015). A multivariate ARIMA performed proficiently for demand forecasting on perishable goods (Huber et al., 2017). Moreover, the ARIMA model was applied for cryptocurrency price forecasting based on social media impact (Tandon et al., 2021).

Researchers continue to search for models that fill the gap in traditional statistical models, which might help to provide better forecasting performance. ANNs are the new competitors in forecasting trends and seasonal data. ANNs appeared with a bright future for identifying and modeling data patterns that are not easily detectable by traditional statistical methods in some fields such as cognitive science, computer science, electrical engineering, finance and stock market (Abuzir & Baraka, 2019; Alon et al., 2001).As ANNs have restrictions on modeling contrasted with classical econometric methods, improved performance can be attained with enough data (Yin et al., 2020). Researchers examined the ability to model seasonal time series using neural networks (Zhang & Qi, 2005). Moreover, they proposed ANNs to inspect how and when seasonal patterns change over time (Franses & Draisma, 1997). The strength of ANNs over conventional econometric models is their ability to model complex, nonlinear relationships without any previous assumptions about the undelaying data-generating process (Alon et al., 2001). The data-driven nature of ANNs makes them more attractive in time series modeling and forecasting. ANNs models beat the traditional forecasting methods constraints such as misspecification, biased outliers, re-estimation, and assumption of linearity (Hill et al., 1996). Because of ANNs models techniques' resilience for discovering patterns in data, they have been widespread for sales forecasting (Tkáč & Verner, 2016). Even with the significant promise of ANNs in time-series forecasting, the observational results are to some extent mixed. On the 50-computation series, the researcher found that ARIMA model has a superior or equivalent mean absolute percentage error (MAPE) to that of ANNs, where the error is lower when trend and seasonal patterns are in the data (Alon et al., 2001). Others have found that ANNs excel the traditional methods of forecasting when forecasting quarterly and monthly data (Hill et al., 1996). Despite theoretical talking that ANNs outperforms the traditional time-series methods in forecasting a series with trend and seasonal patterns, researchers found that ANNs do not model the seasonal fluctuations in the data impressively (Nelson et al., 1994). One of the researchers conducted a comparison between ARIMA and MLPNNs, and found that MLPNNs can model non-linear processes, able to introduce more complex time series, and does not take into account specific features of the time series such as being stationary (Aburto & Weber, 2007) .

Researchers have found that on average, ANNs achieved good results across various forecasting periods and horizons, followed by Box-Jenkins and Winters exponential smoothing for aggregate retail sales forecasting monthly (Alon et al., 2001). Others applied more models of classical time-series forecasting techniques for retail store sales such as seasonal Autoregressive Integrated Average (SARIMA) and Triple Exponential Smoothing. Also, some advanced methods such as Prophet, Long Short-Term Memory (LSTM), and Conventional Neural Networks (CNNs), found that the Stacked LSTM has the most superior results(Ensafi et al., 2022). The Radial Basis Function Network (RBF NN) was applied to the time-series forecasting problem by improving the RBF center placement quality using the genetic algorithm technique, and it performed better than the (BP NN) model (Yan et al., 2005). Seasonal ARIMA (SARIMA) is another type of classical forecasting method. This technique has been applied effectively in different applications such as forecasting tourism demand (Goh & Law, 2002). Nevertheless, the researcher specified that SARIMA can have limitations in prediction because of its linear form and inability to discover nonlinear and highly volatile patterns (HAMZACEBI, 2008).

Some studies have involved artificial neural networks through building hybrid models for forecasting and concluded that hybrid forecasting methods are usually more proficient than pure static models or pure machine-learning models (Yin et al., 2020). A novel hybrid model combining ARIMA and ANN enhanced the forecasting accuracy for an area with limited air quality and meteorological data, and showed better results by either of using the models separately (Díaz-Robles et al., 2008). This research developed a total monthly sales forecasting model using a hybrid econometric-neural network model by integrating the structural features of econometric models with non-linear pattern recognition features of neural networks (Luxhøj et al., 1996).Another hybrid model of linear autoregressive integrated moving average (ARIMA) and nonlinear artificial neural network (ANN) models for the prediction of time series data, where the proposed hybrid model has higher prediction accuracy than applying each model individually (Babu & Reddy, 2014). Khandelwal et al. (2015) found that hybrid models of ARIMA and ANN based on DWT decomposition provided better prediction results on time-series data than using pure models separately. Two hybrid models were built to forecast the daily sales for perishable food in a German retail store, where hybrid models of (SARIMA-MLR) and (SARIMA –QR) provided better forecasts over seasonal naïve forecasting, traditional SARIMA, and multi-layer perceptron neural network (MLPNN) (Arunraj & Ahrens, 2015). Zhang, (2003) suggested a hybrid model by combining the seasonal ARIMA (SARIMA) model and the ANN model to predict seasonal time series.

Others have presented a hybrid intelligent system combining ARIMA models & neural networks for demand forecasting, which demonstrates enhancements in forecasting accuracy and suggests a replenishment system for a supermarket (Aburto & Weber, 2007).Another novel hybrid time-series prediction model based on recursive empirical model decomposition (REMD) and long-short term memory (LSTM), where results have shown that the prediction accuracy was enhanced by more than 20% compared with the LSTM algorithm(Yang et al., 2021).

### 2.4 Summary

In order to compare findings of different papers and reports on products sales forecasting in the literature, a summary table has been formed. This table presents a brief of the methodologies, key results and the used dataset of the reviewed papers, providing a clear and organized comparison. The following table underlines the main insights from the selected literature.

Table 2. 1**:** Papers and reports on products sales forecasting

| Year | Authors | Methodology | Result | Dataset |
|---|---|---|---|---|
| 2001 | Alon,et al. | Artificial neural networks compared with traditional methods, Winters exponential smoothing, Box}Jenkins, ARIMA model, and multivariate regression. | ANN performed the best. | US aggregate retail sales. |
| 2015 | Ramos,et al. | Apply ETS and ARIMA models. | Both models have quite similar forecasting performance. | Retail sales of five different categories of women footwear from the Portuguese retailer Foreva. |
| 2017 | Huber,et al. | Apply multivariate ARIMA models separately: (ARIMAX) at different levels and univariate ARIMA (1-WD) using weekday data and as a combination. | ARIMAX model outperforms the ARIMA (1-WD) model at store level with respect to MAPE &RMSE. | Point-of sales of an industrialized bakery chain. |

| | | | | |
|---|---|---|---|---|
| 2022 | Ensafia,et al. | SARIMA and Triple Exponential Smoothing models compared with Prophet, (LSTM), and (CNN). | Stacked LSTM method superior other methods. | A public dataset including the sales history of a retail store is investigated to forecast the sales of furniture. |
| 1996 | Luxhøj et al. | Develop a hybrid econometric-neural network model. | | An actual sales forecasting problem from a Danish company that produces consumer goods. |
| 2015 | Arunraj & Ahrens | Appling SARIMA, SARIMAX, and Quantile regression (QR), then developing two hybrid models (SARIMA-MLR) and (SARIMA-QR). | • SARIMA-MLR and -QR models yield better forecasts at out sample data comparing with seasonal naïve forecasting, traditional SARIMA, and MLPNN.<br>• SARIMA-QR model provides better prediction intervals and a deep insight into the effects of demand influencing | Daily sales data of banana measured in kilograms from a typical food retail store in Germany. |

| | | | | |
|---|---|---|---|---|
| | | | factors for different quantiles. | |
| 2007 | Aburto & Weber | Hybrid intelligent system of (ARIMA) models and (NNs). | Neural networks outperformed ARIMA models, and the proposed additive hybrid approach gave best results. | A Chilean supermarket dataset. |

A few researchers focused on the prediction of the product sales forecasting of small-medium businesses such as supermarkets in middle-east, particularly the Palestinian supermarkets, and after inspecting this particular research topic, it was observed that there are few researches applied and compared classical statistical models such as (ARIMA & SARIMA) with Artificial Neural Networks (ANNs) such as RNN, LSTM, MLPNN and RBFNN models, and hybrid models of both approaches based on the aggregating products sales and based on each product level to find the best approach that providing the best forecasting results. Therefore, there is a critical need for a more efficient model that combines the capabilities of classical statistical models and advanced machine-learning models to capture the linear and non-linear dependences to ensure a comprehensive analysis of the complex temporal patterns which might appear in the sales data. As a result, in this research, it is proposed to apply hybrid intelligent methods, which are (ARIMA-RNNs, ARIMA-LSTM, ARIMA-MLPNN, and ARIMA-RBFNN) to predict future sales based on the time series of (DUKKAN-11) supermarket in Ramallah, based on aggregate products sales and based on each product level separately. In addition, the performance of all applied models individually and as hybrid models will be evaluated according to the error metrics of (MSE, RMSE, and MAE) to recommend the best model.

# Chapter Three: Methodology

# 3. Proposed Methodology and Applied Models

### 3.1 Introduction

In this chapter, two classical statistical models were applied, such as ARIMA and SARIMA, as well as advanced machine-learning models such as Standard Recurrent Neural Networks (RNNs), Long-Short Term Memory (LSTM), Multilayer Perceptron (MLPNN) and Radial Basis Function Neural Networks (RBFNN). Then, hybrid modes that combine the best statistical model results with each advanced machine-learning model, such as Standard Recurrent Neural Networks (RNNs), Long-short Term Memory Networks (LSTM), Multilayer Perceptron (MLPNN), and Radial Basis Function Neural Networks (RBFNN)), based on the aggregated sales and based on each product level, as illustrated in Figure 3.1, to forecast the sales of the product based on time-series. Furthermore, a significant part of the analysis identifies and compares the importance of the predictive ability of each model, whether as a pure model or as a hybrid model, to decrease the prediction error factor. Which is in our work, these errors are measured using the mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE).

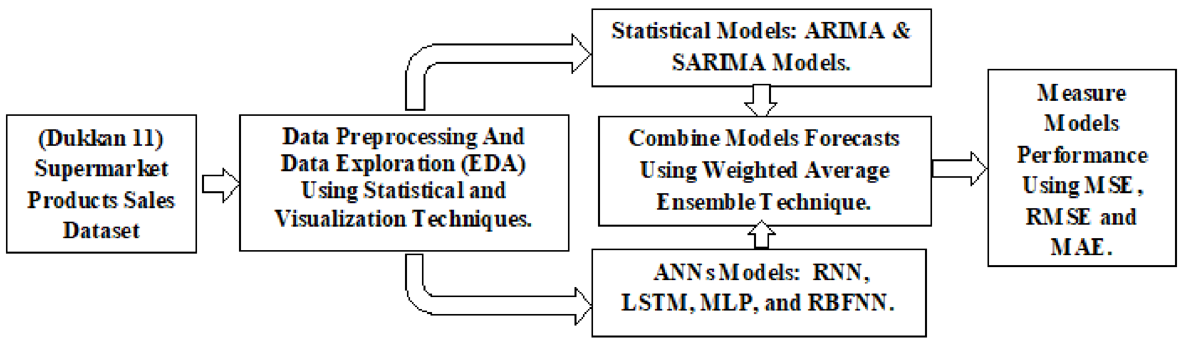


Figure 3. 1 General structure of all proposed hybrid models.

The general procedure used in performing experiments is illustrated in Figure 3.2. Where the input data represents a time series for (Dukkan 11) supermarket, this data is separated by aggregation sales for all categories, and for each category sales. The data then will be passed on a pre-processing step, where it will be checked for missing values, duplicated values, and variable types. Exploratory data analysis (EDA) will be applied using visualization techniques to study and identify the time series patterns, such as trends, seasonality, and cyclic patterns, which might appear in the sales data. In addition, a statistical analysis will be applied to identify patterns, trends, and seasonal variations within the sales data over time. This analysis aims to gain knowledge of sales performance, such as on weekends, and holidays. After that, for statistical models, the data will be checked for its stationary using the visualization technique and ADF statistical test. For advanced machine-learning models, the data will be split into training and testing sets. Finally, the model's performance will be evaluated using the error metrics including MSE, RMSE, and MAE for each approach, as pure models and as hybrid models.

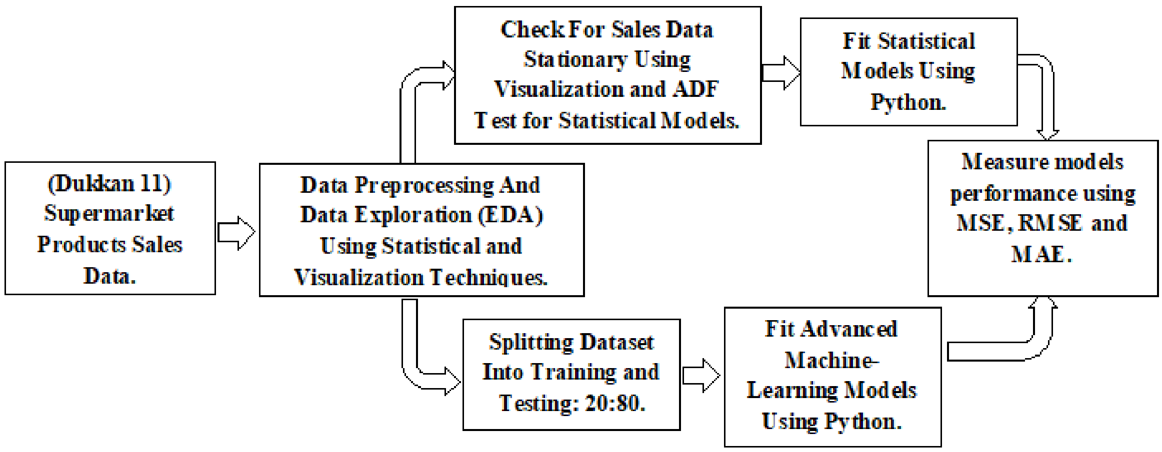


Figure 3. 2 The flow chart of the general procedure used in conducting experiments individually.

### 3.2 Datasets Description

In this study, a dataset has information for local supermarket sales mainly located in Ramallah, containing sales for the top five categories which are (drinks, dairies, snacks and chips, ice cream, and cleaning materials) based on daily records from (1-3-2021) to (31-10-2023). The dataset contains (4875) data points for five categories of sales, with (975) data points for each category. The study aims to manage the sales of categories in the dataset under two scenarios, the first scenario, combines the sales data across the five categories starting from (1-3-2021) to (31-10-2023) based on daily records with (975) data points. The second scenario involves building individual models for sales within each category, ensuring a concentrated analysis suitable to the unique characteristics and dynamics of each product category. The following descriptions explain the details of each scenario dataset:

#### 3.2.1 Combined Sales Dataset

The first scenario for the dataset of sales forecasting contains the integration sales of five categories (drinks, dairies, snacks and chips, ice cream, and cleaning materials) from (01-03-2021) to (31-10-2023). This approach enables analysis by taking into account the general behavior of sales data as combined sales with (975) daily records. The visualization analysis carried out on product sales data to discover insights and patterns that may not be clear from raw numerical data. The combined products sales for five categories (drinks, dairies, snacks, and chips, ice-cream, and cleaning materials) have revealed a notable pattern described by both trend and seasonality. An observable upward movement in sales, showing a clear trend in overall performance. In addition, the sales data displays fluctuations that match to specific periods, considering seasonality in consumer demand. These seasonal variations may link with predictable events such as holidays or cultural events, which help to understand consumer behavior and market dynamics. Figure 3.3 shows the combined sales data for five categories (drinks, dairies, snacks and chips, ice cream, and cleaning materials) together, with the general behavior of sales data.

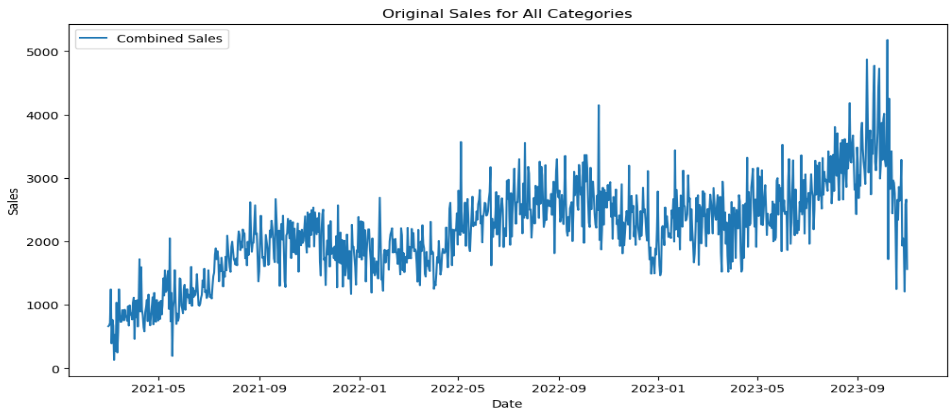


Figure 3. 3 Combined sales of five categories based on daily records.

### 3.2.2 Categories Sales Dataset

The second scenario involves the development of sales forecasting models, which have been made to fit each product category, using (975) daily sales records spanning from (01-03-2021) to (31-10-2023) for each category sales (drinks, dairies, snacks and chips, ice-cream and cleaning materials). This approach enables a focused analysis, wherein separate models are constructed for individual categories, taking into account the unique dynamics and sales patterns ingrained in each product grouping and getting insights into each category-specific trend. By dividing the dataset based on category and employing advanced analytical techniques such as time series analysis and machine learning algorithms, the goal is to accurately predict future sales performance for each category within the specified date range. The visualization analysis for drinks sales from (01-03-2021) to (31-10-2023) has shown an increasing trend and seasonality patterns specifically in May to December, which has been proven in the descriptive analysis based on numerical analysis for monthly sales. Figure 3.4 shows the sales data for the drink's category.

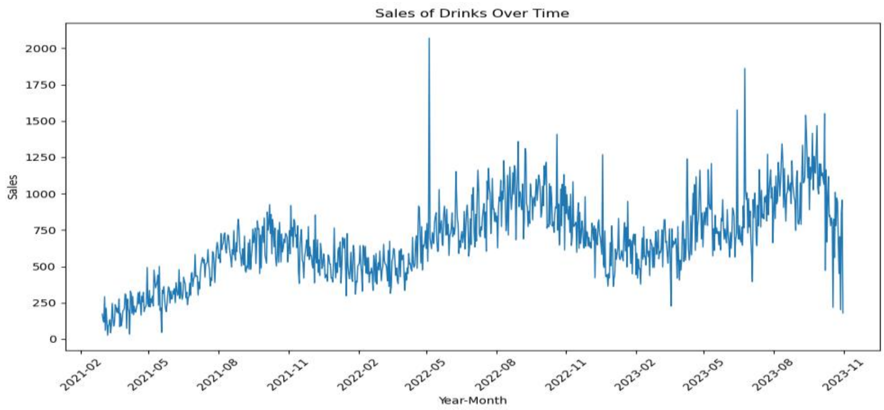


Figure 3. 4 The sales of drinks category based on daily records.

Dairies sales analysis has shown an increasing trend which has been proven in the descriptive analysis based on numerical analysis for monthly sales data. This increased consumption appeared in months from July to December due to summer holidays or outdoor activities, barbecues, and picnics. Moreover, the visualization has shown an interesting increased trend in weekends (Fridays and Saturdays). Figure 3.5 shows the monthly sales data for dairies category**.** Figure 3.6 shows the date of the sale for dairies category on weekends.

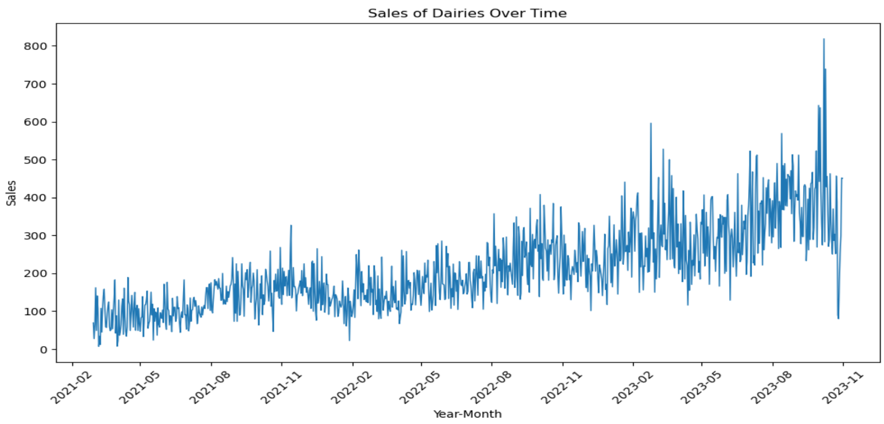


Figure 3. 5 The sales of dairies category based on daily records.

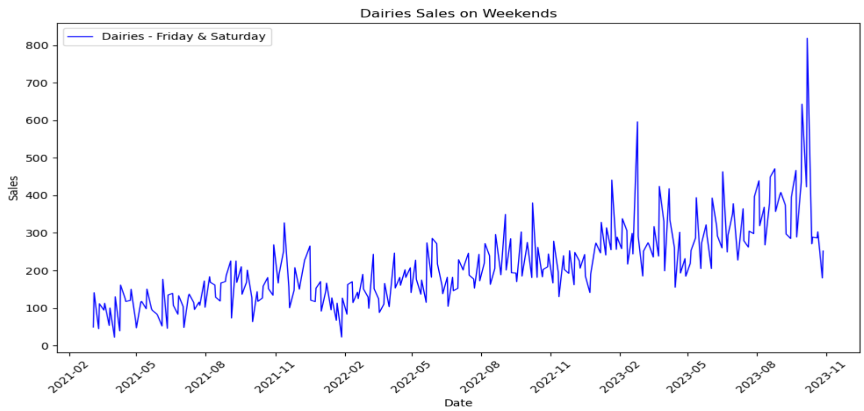


Figure 3. 6 The sales of dairies category based on weekends.

Snacks & chips sales has shown an increasing trend which has been proven in the descriptive analysis based on numerical analysis for monthly sales data and a slight seasonality in September to December. Figure 3.7 shows the monthly sales data for the snacks & chips category

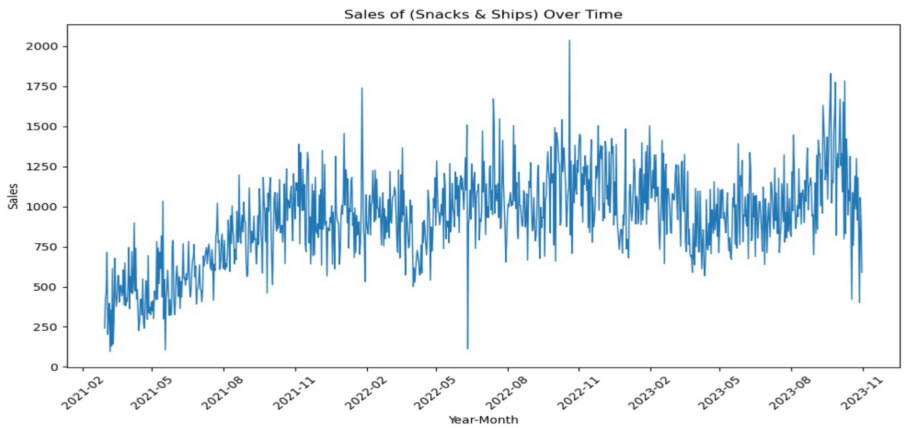


Figure 3. 7 The sales of (snacks & chips) category based on daily records.

The ice-cream category has shown a clear and strong seasonality in summer months especially from May to September over years. This type of product is available in the summer months and considered as preferable item because of high temperatures in summer. Figure 3.8 shows the monthly sales data for the ice-cream category.

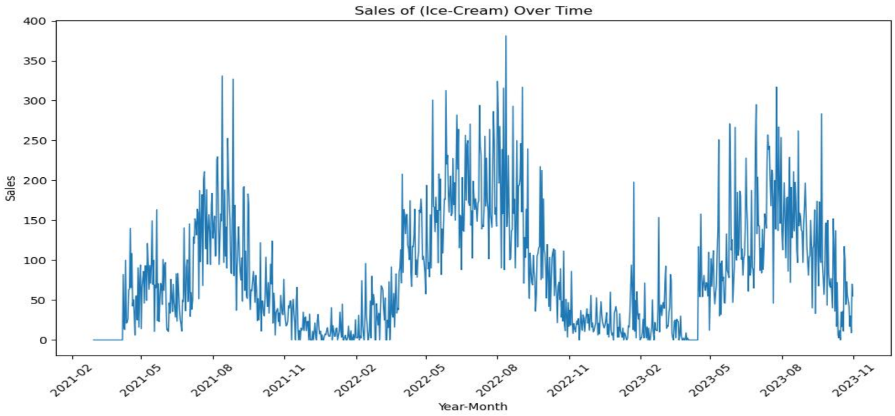


Figure 3. 8 The sales of Ice-cream category based on daily records.

Cleaning materials products witnessed a rise in sales in months between (May to Dec) with increasing trend and seasonality, especially in 2022 and 2023. Seasonality behavior is due to an annual tradition which is commonly known as “spring cleaning”, also it refers to seasonal cleaning routines, as the weather becomes warmer and people are motivated to refresh their living spaces. Many other possible reasons such as “back to school season” specifically in August and September. Moreover, as the year comes to a close, some individuals prefer to make end-of-year cleaning routines to start the new year with a fresh and organized home. Figure 3.9 shows the monthly sales data for the cleaning materials category.

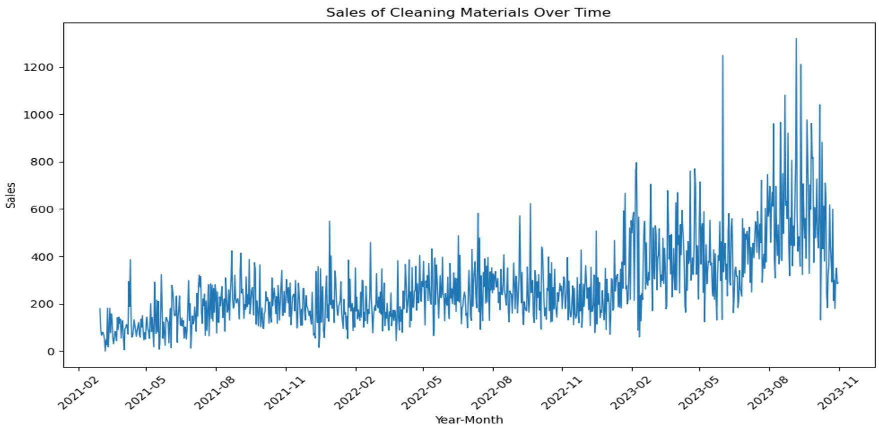


Figure 3. 9 The sales of (cleaning materials) category based on daily records.

### 3.2.3 Dataset Preparation and Data Extraction

Before analysis, several data preprocessing steps were performed to ensure data integrity. This includes an initial assessment to check for any duplicate entries and missing values in the dataset. Moreover, the datasets were filtered based on months, years, and weekend sales to include only those transactions occurring on holidays (Friday and Saturday) based on daily sales transactions.

### 3.2.4 Exploratory Data Analysis (EDA)

The models' application as pure models and as a combination of statistical models and advanced machine learning models using the weighted average technique, has preprocessing steps that will be applied to the loaded datasets (products sales of the supermarket), to be more suitable for model training. Moreover, an exploratory data analysis (EDA) using statistical analysis and visualization techniques will be applied to get more insights into sales data behavior to study the time series patterns such as (trends, seasonality, and cyclic) patterns and the consumers' behaviors.

## 3.3 Applied Models

This section will present and formulate the procedures of all applied methods in the proposed product sales forecasting models, which will come up with results that will be discussed in the next chapter based on error metrics. In detail, the process of applying the statistical models (ARIMA and SARIMA), and the advanced machine-learning models (RNNs, LESTM, MLPNNs, and RBFNNs) will be demonstrated, and finally the combination process of the selected models of the two approaches using the ensemble method of weighted average technique will be applied.

### 3.3.1 Products Sales Forecasting

There's a variety of statistical models like (SARIM and ARIMA) models and advanced machine learning models like artificial neural networks (ANNs), each suggesting diverse advantages in the context of sales forecasting. These models are carefully examined for their ability to predict sales over different time horizons, including short and long-term projections. The traditional statistical models excel at capturing linear relationships and patterns in sales data, where the underlying patterns are relatively stable. In contrast, ANNs capable of capturing complex, non-linear relationships and patterns in sales data that could exceed the limits of traditional statistical models. This led us to realize that many techniques might be used to solve time-series forecasting problems (Zhang & Kline, 2007). For instance, the advanced version of MA which is known as Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing method has been applied to forecast the retail sales of different women's footwear since it has products with repeatable fluctuations in their patterns (Ramos et al., 2015), also the traditional statistical model like (SARIMA) has shown constraints in prediction due to its linear form and inability to detect nonlinear and highly volatile patterns (HAMZACEBI, 2008).On the other hand, (ANN) models have been embraced as techniques for sales forecasting due to their flexibility in discovering patterns in data (Tkáč & Verner, 2016). Moreover, the data-driven nature of ANNs makes them more attractive in time series modeling and forecasting (Alon et al., 2001), and it beats the limitations of traditional forecasting methods such as misspecification, biased outliers, assumption of linearity, and re-estimation (Hill et al., 1996). A simplified overview of the traditional statistical models (ARIMA and SARIMA) and advanced machine-learning models such as (ANNs) techniques which have been used in the proposed hybrid models (SARIMA-RNN, SARIMA-LSTM, SARIMA-MLPNN, and SARIMA-RBFNN) are introduced in this chapter.

### 3.3.2 Time Series Forecasting

Anything noticed sequentially over time is a time series (Hyndman & Athanasopoulos, 2018).A time series is a collection of observations (Oi), each observation has been sampled at a specific time (Ti), where it will be displayed as a sequence of discrete-time data (Brockwell & Davis, 2016).Also, time series might be measured continuously, when time series are recorded instantaneously and steadily (Granger & Newbold, 2014).To build a time-series forecasting model, a time-series analysis must be performed to get insights into the data trends, seasonal patterns, and forecasts for future events. Because of this, time series data was studied differently from other data, as retail sales forecasting requires further data engineering associated with data granularity such as temporal granularity (e.g. days or months) and product hierarchy levels (e.g. total sales or by product category) (de Almeida, 2021). As time series forecasting needs sequential data to be considered, different techniques related to the analysis and prediction of time series were developed in the field of statistics (Alwan & Roberts, 1988).

### 3.3.3 Traditional Time-Series Methods

A benchmark in time series methods was released by Box and Jenkins book (Kleiner, 1977), which represented a three-step iterative model for the identification, estimation, and verification of time series. Also, it has contributed to the widespread adoption of autoregressive integrated moving average (ARIMA) models. ARIMA models consist of multiple time-series processing techniques, namely autoregressive (AR), integration (I), and moving average modeling (MA), as outlined below:

**Autoregressive (AR):** in an auto-regression model, the variable of interest is forecasted using a linear combination of past values of that variable. The concept of “auto-regression “shows that is a regression of the variable against itself. That is, the lagged values of the target variable as the input variables to forecast values for the future will be used. In general, AR (p) with p > 1 is a linear combination of AR components. Equation 3.1 Shows an AR model including lagged values, a constant value of $(c)$, and a white noise component $t$ . Parameters $\emptyset 1 \ldots \emptyset p$ of the p-order model must be fit to the data. An auto-regression model of order p (AR (p)) will look like this:

$$y_t = c + \emptyset y_{t-1} + \emptyset y_{t-2} + \cdots + \emptyset p y_{t-p} + \epsilon_t \qquad (3.1)$$

**Integration (I):** this order attempts to transform a non-stationary series into a stationary one through differentiation. In detail, consecutive values are subtracted as stated in equation 3.2**.** The differentiation aims to stabilize the average of the series, decreasing or even removing the trend and seasonality components. The number of time differentiation is employed as a hyper-parameter to be configured.

$$\mathrm{y}' = \mathrm{y_t} - \mathrm{y_{t-1}} \qquad (3.2)$$

Figure 3.10 shows how the differentiation in the combined sales data was applied and removed the trend and seasonality by applying a differencing factor (I), d=1. Figure 3.11, Figure 3.12, Figure 3.13, Figure 3.14, and Figure 3.15 show the differentiation of each category sales separately using a differencing factor(I), d=1.

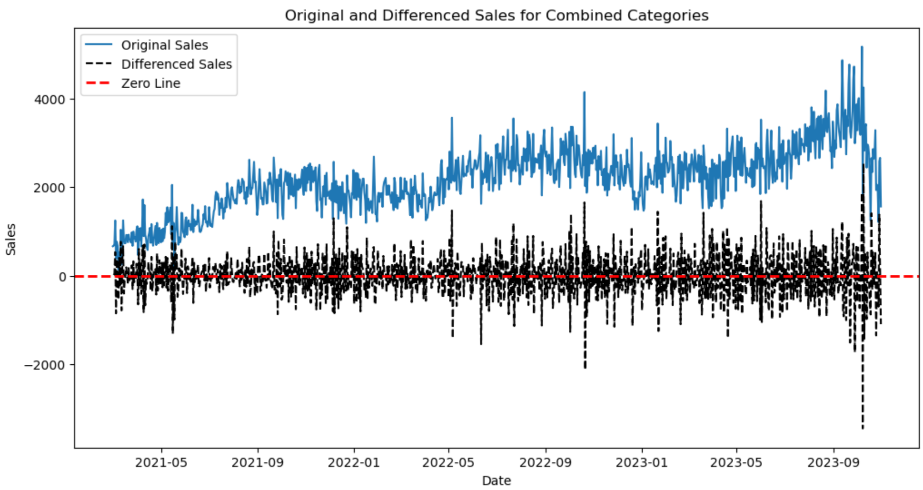


Figure 3. 10 The original and differenced combined sales with (I (d=1)).

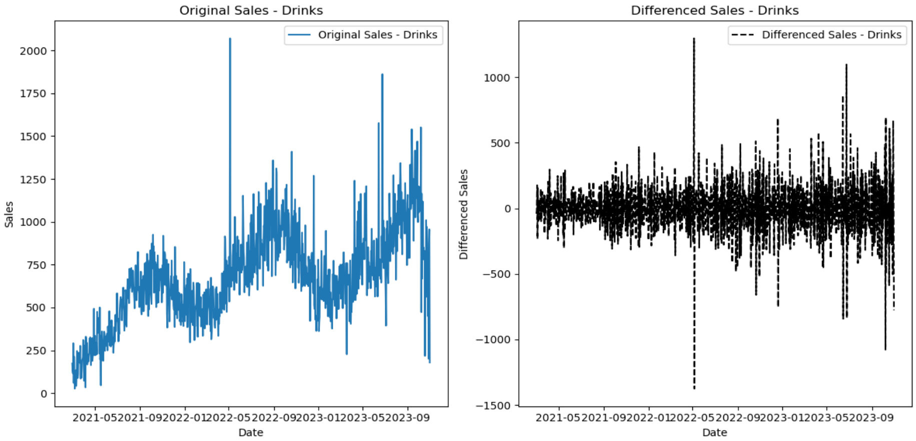


Figure 3. 11 The original and differenced for drinks category sales with (I (d=1)).

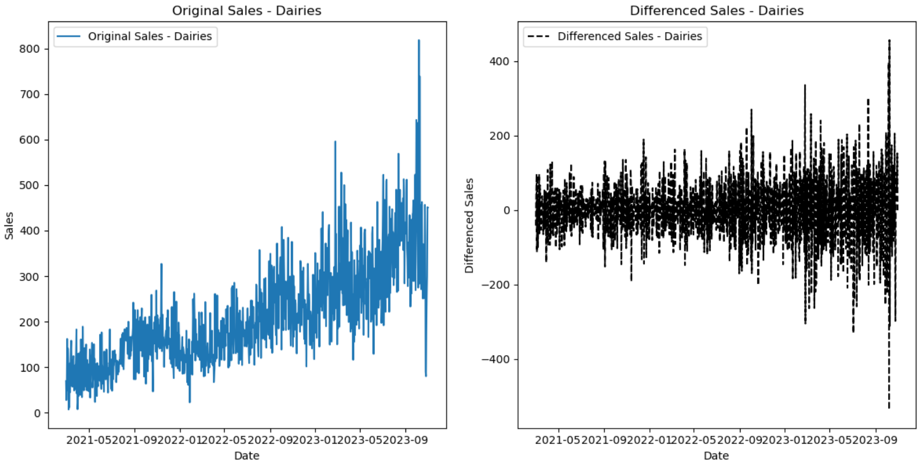


Figure 3. 12 The original and differenced for dairies category sales with (I (d=1)).

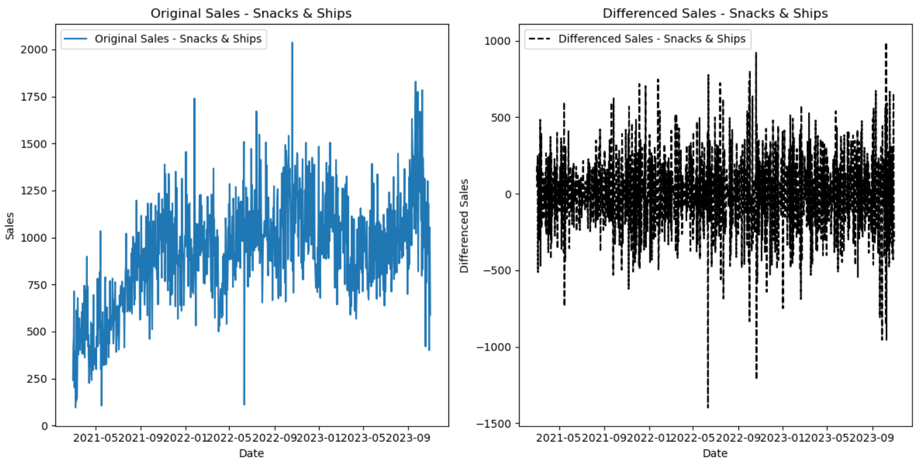


Figure 3. 13 The original and differenced for (snacks & chips) category sales with (I (d=1)).

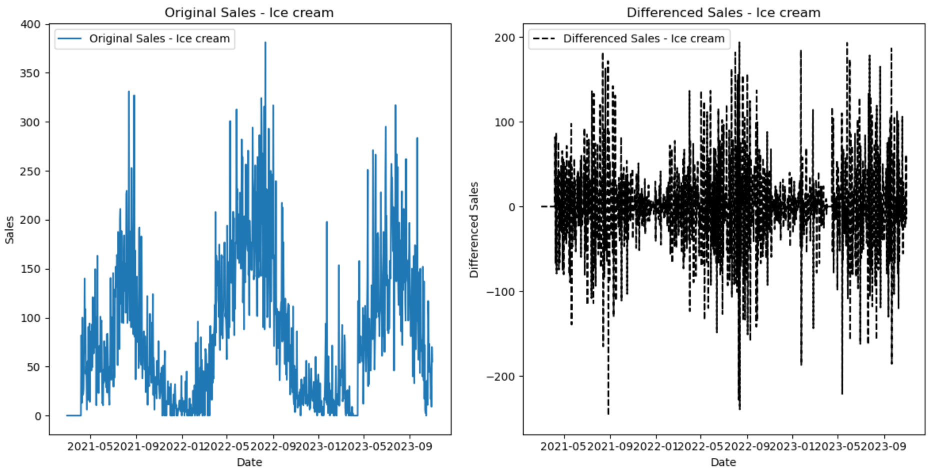


Figure 3. 14 The original and differenced for ice-cream category sales with (I (d=1)).

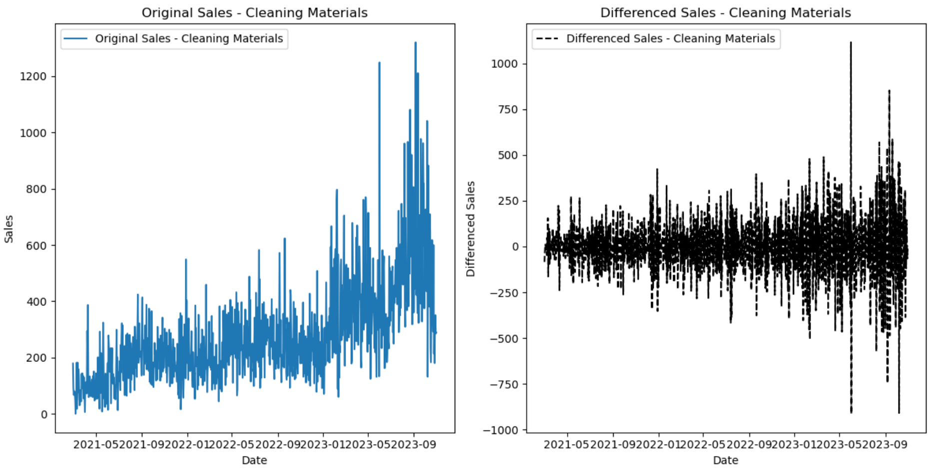


Figure 3. 15 The original and differenced for cleaning-materials category sales with (I (d=1)).

The difference between them is evidence, after differencing, it's clear that it's significantly more stationary than the original data. The ADF (Augmented Dicky-Fuller) test is a statistical test used to identify whether a unit root is present in a time series dataset, and helps to evaluate the stationary of a time series which is a crucial step for many time series analysis techniques.

**Moving Average (MA):** is a regression-like model, but it contains the past forecast errors in the estimation. Equation 3.3 represents an MA (q) model, where q is the order of the model and $\epsilon t$ is the white noise of previous forecast errors. In the equation, various patterns can be obtained when changing$\theta_1. \ \ . . \ \theta_q$values, while the other constant/variables will change the scale.

$$y_t = c + \epsilon_t + \theta_{1\in t-1} + \theta_{2\in t-2} + \cdots + \theta_q \in t - q \qquad (3.3)$$

The combination of all these models shapes the original ARIMA model. The ARIMA (p, d, and q) where p is the order of the autoregressive part, d is the degree of differentiation, and q is the order of the MA model. A significant shift in ARIMA models is the seasonal ARIMA (SARIMA), which involves the possibility of one or more seasonal components. To incorporate the seasonal part into the original ARIMA model, new hyper-parameters are considered. The ARIMA model becomes SARIMA (p, d, and q) (P, D, and Q) m, where m is the number of observations per year, and the hyper-parameters of (P, D and Q) are the seasonal equivalents for the original and non-seasonal hyper-parameters (p, d, and q) of ARIMA.

### 3.3.4 Artificial Neural Networks (ANNs)

ANNs are a type of generalized nonlinear and nonparametric models stimulated by studies of the human brain and nerve system which consists of cells and links, as links connect cells. Cells in ANNs known as artificial neurons (nodes) and their links are defined by a value called weight. In addition, ANNs simulate the knowledge-storing process of humans (de Almeida, 2021). ANNs surpass other classical econometric models because of their ability to model complex, non-linear relationships without any previous assumptions related to the underlying data-generating process (Alon et al., 2001). In addition, it also beat the biased outliers, miss-specification, and re-estimation (Hill et al., 1996). Moreover, the data-driven nature of ANNs makes them more attractive for time-series modeling and forecasting.

The feed-forward Network is one of the time-series prediction models that was inspired by the human brain technique. This model includes three layers: An input layer which contains the observations, hidden layers that operate the received information from the first layer, and the last layer which is the output layer that provides the actual prediction. This means that the input of each layer is the output of the previous layer, where the output of the first layer is produced by the weighted sum of inputs and adding the specific bias like $\alpha_0$and $\beta_{oj}$ to them (Ensafi et al., 2022; Loureiro et al., 2018). The model representation in the following equation:

$$y_t = \alpha_0 + \sum_{\mathrm{j}=1}^{\mathrm{n}} \alpha_j f\left(\sum_{i=1}^{m} \beta_{ij} y_{t-i} + \beta_{oj}\right) + \varepsilon_t \qquad (3.4)$$

The number of nodes in the input layer and hidden layers are symbolized by m and n respectively, and fitting the activation function such as RELU or Sigmoid. It’s worth noting that the process of forecasting using ANN requires selecting the architecture and its parameters of layers, number of units, and the assigned weights which influence the model forecasting accuracy (Ensafi et al., 2022).

Arithmetically, an artificial neuron is a function fi(x) which is calculated using the following equation:

$$y_i = f_i(x) = \sigma \sum (w_i * x) + b_i) \qquad (3.5)$$

The input data ($x$) is described as ($x$1, $x$2… $x$n), the weights ($w_i$) are described as ($w$1, 1, $w_i$,2,.. $w_i$, n), the neuron bias values are described as ( $b_i$), and the transfer (activation) function described as σ which can be: step, sigmoid function (or logistic), tanh and many other functions (Garcia-Pedrajas et al., 2003).The following equations represent step function, sigmoid function and tanh function equations respectively:

Step function:

$$\sigma(\mathrm{x}) = \begin{cases} 1, & \mathrm{x} \geq 0 \\ 0, & \mathrm{x} < 0 \end{cases} \qquad (3.6)$$

Sigmoid function:

$$\sigma(\mathrm{x}) = \frac{1}{1 + \mathrm{e}^{-\mathrm{x}}} \qquad (3.7)$$

Tanh function*:*

$$\sigma(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1 \qquad (3.8)$$

The neural network architecture may differ, involving single-layer, multilayer, or recurrent configurations, where each type is suitable for particular computational tasks.

#### 3.3.4.1 Recurrent Neural Network (RNNs) Model

The idea of RNNs is the beneficial use of sequential information. Traditional neural networks assumes that all inputs and outputs are independent of each other, but for time series forecasting tasks it will be a misguided idea. RNNs implement the same task for every element of a sequence, where the output being depend on the previous calculations. In another meaning, RNN have a “memory” that preserves information about what has calculated so far, which can be beneficial for time series forecasting (Denny, 2015).Each layer inputs pass to the hidden layer which has a recurrent loop to the back. Consequently, the function of the previous input, combined on the activation value of the past hidden layers which is presented as the output (Ensafi et al., 2022). Figure 3.16 shows the basic architecture of RNN network:

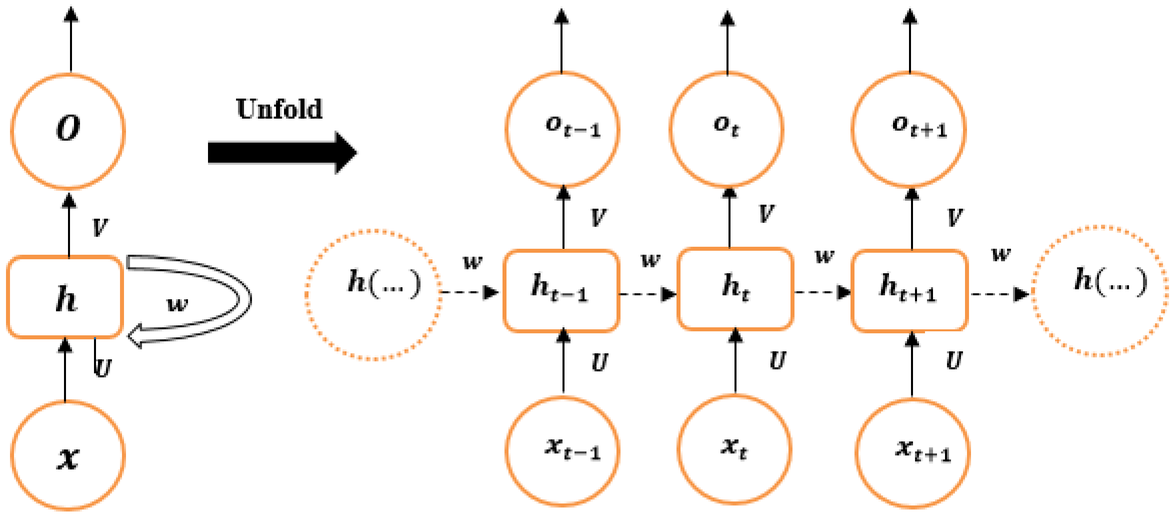


Figure 3. 16 The basic architecture of RNN network.

Input: $x_t$ Is the input to the network at time step $(t)$.

Hidden State: $h_{(t)}$Represents as hidden state at time $(t)$ and acts as "memory" of the network. The value of $h_{(t)}$is calculated based on the current input and previous time steps using the following equation:

$$h_{(t)} = f\left(U\, x_{(t)} + W\, h_{(t-1)}\right) \quad (3.9)$$

The function f is taken to be a non-linear transformation such as (tanh, ReLU).

Weights: The RNN has an input to hidden connections parametrized by a weight matrix $U$, hidden-to-hidden recurrent connection parametrized by a weight matrix $W$, and hidden-to-output connections parametrized by a weight matrix $V$.

Output: The output of the network represented as $o_{(t)}$, which is often subjected to non-linearity, specifically when the network contains more layers downstream.

#### 3.3.4.2 Long-Short Term Memory (LSTM) Model

Long-Short Term Memory (LSTM) model is a type of recurrent neural networks (RNNs), which developed to address the vanishing gradient problem and capture long-term dependencies in sequential data such as time series analysis tasks. LSTM architecture contains specific systems that allow it to store and restore information over long sequences. Figure 3.17 shows the basic architecture of LSTM network.

The key components of LSTM architecture illustrated as follows:

1. Cell State ($c_t$): represents the memory of the LSTM and store the information over long sequences, where this state can be updated, cleared or read from at each time step.
2. Hidden State ($h_{(t)}$): it acts as an intermediary between the cell state and the external world. Also, it can optionally remember or forget information from the cell state and generate the output.

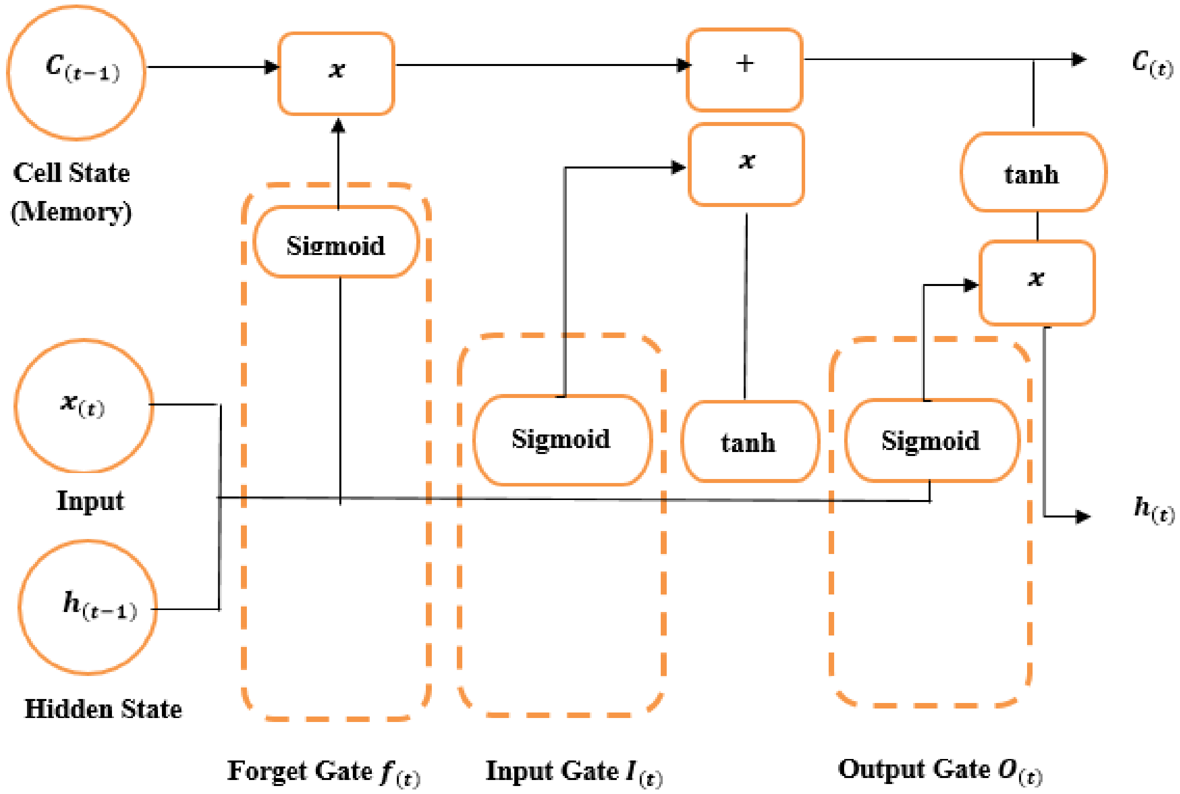


Figure 3. 17 The basic architecture of LSTM unit.

3. Input Gate ($i_t$): this gate monitors the information flow into the cell state. Also, it can learn to accept or reject incoming data.
4. Forget Gate ($f_{(t)}$): the forget gate identifies what information from the previous cell should be kept and what should be discarded. This gate allows the LSTM to "forget" unrelated information.
5. Output Gate($o_{(t)}$): this gate controls the information that is used to produce the output at each time step. Moreover, it has the decision what part of the cell state should be detected to the external world.

The following steps show how the state of LSTM works at time step t as ($h_{(t)}$,$c_{(t)}$), where $h_{(t)}$ is the hidden state and $c_{(t)}$, is the cell state.

Step one: the LSTM receives the input vector ($x_t$) and the previous state ($h_{(t-1)}$,$c_{(t-1)}$).

Step two: the forget gate ($f_{(t)}$) decides what information to reject from the cell state. Then, it uses the input vector ($x_t$) and the previous hidden state ($h_{(t-1)}$) to generate a number between (0 & 1) for number on the cell state ($c_{(t-1)}$), where the value of (1) means: "completely keep this", while the value of (0) means: "completely discard this". The following equation describe the forget gate calculations:

Forget Gate:

$$f_{(t)} = \sigma(wf\ .\left[h_{(t-1)}, x_{(t)}\right] + bf \qquad (3.10)$$

Where: $h_{(t-1)}$is the previous hidden state, $x_t$ is the input vector, $bf$is the bias, $Wf$ is weight between the gates (hidden and input layers).

Step three: in this step, the input gate $i_t$ selects what new information to store in the cell state. This gate has two parts: a sigmoid layer which called "input gate layer", which decide values to be updated, and a tanh layer, that creates a vector of new candidate values ($c_t$~) that might be added to the state. The following equations represents how the input gate decide values to be updated, and how a vector of new candidate values is created:

Input Gate:

$$i_{(t)} = \sigma\ (\ w_i\ .\left[h_{(t-1)}, x_{(t)}\right] + b_i \qquad (3.11)$$

Candidate Values (Cell State Update):

$$c_t\sim = \tanh(\ w_c\ .\ [\ h_{(t-1)}, x_{(t)}] + b_c \qquad (3.12)$$

Step Four: in this step, the old cell state $c_{(t-1)}$ will be updated to the new cell state $c_t$, where the old cell state is multiplied by $f_{(t)}$ to forget what it decides to forget earlier. After this, the new candidate values is added, and scaled by how much it decides to update each state value. The following equation represents how the old cell will be updated to the new cell state:

Cell State (Final Cell State):

$$c_t = f_t * c_{t-1} + i_t * c_t\sim \qquad (3.13)$$

Step five: at this step, the output should be decided. This output will be based on the cell state, but a filtered version. The sigmoid layer will decide what parts of the cell state is going to output. Then, it adds the cell state through tanh (to push values between -1 and 1), then multiplied it by the sigmoid gate output. The following equations how the output will be decided:

Output Gate*:*

$$o_t = \sigma \left( w_o . \left[ h_{(t-1)}, x_{(t)} \right] + b_o \right) \quad (3.14)$$

Hidden State:

$$h_t = o_t * \tanh(c_t) \quad (3.15)$$

#### 3.3.4.3 Multilayer Perceptron Neural Networks (MLPNNs) Model

A multilayer perceptron (MLPNN) is a feedforward ANN with at least three layers: an input layer, one or mode hidden layers, and an output layer. It performs various kind of tasks such as classification, regression and time-series forecasting. MLPNN performs a series of mathematical operations on input data to create a prediction or output, moreover, it consists of numerous layers of nodes, where each layer implementing a nonlinear modification on the input data. In more detail, at first, the input layer is formed of one or more nodes, where each node identical to a characteristic or input variable in the data. Then the input data is provided into the input layer, and each node calculates a weighted sum of the input values. Secondly, each node in the hidden layer obtains input from all the nodes in the previous layer and calculate a weighted sum of the inputs, and then handle it through an activation function to create the node’s output. Where each successive hidden layer transforms the data nonlinearly using an activation function such as sigmoid or ReLU functions. Finally, the last hidden layer outputs are provided into the output layer, where each node calculates the inputs weighted sum and operates them through an activation function to get the final prediction or output. It’s worth noting, that the weights in MLPNN are learned by backpropagation where the difference between the predicted and actual output is transferred back through the network, as the main aim of changing weights is to minimize the error. The mathematical formulas for MLPNN illustrated as:

The first hidden layer output:

$$z_1 = f(w_1 * x + b_1) \quad (3.16)$$

The second hidden layer's output:

$$z_2 = f\ (w_2 * z_1 + b_2) \quad (3.17)$$

The output layers output:

$$y = f\ (w_3 * z_2 + b_3) \quad (3.18)$$

Where *x* is the input vector, *w* is the weight matrix, *b* is the bias vector, and *f* is the activation function.

#### 3.3.4.4 Radial Basis Function Networks (RBFNNs) Model

Radial Basis Function Neural Network (RBFNN) has been vastly used for non-linear system recognition due to its simple topological structure and its ability to discover how learning proceeds in a straightforward way. The back-propagation neural network (BP NN) was used by many researchers for time-series forecasting, however it has some disadvantages such as, it heads for yielding local solutions, the learning rate is slow and the network structure is difficult to develop. The Radial Basis Function Neural Network (RBFNN) provides another solution for time series forecasting (Yan et al., 2005).The RBF were used by radial basis function neural networks (RBF NNs), which were introduced by Broomhead and Lowe, encompassing their primary functional approximation and time-series forecasting, in addition to classification or clustering tasks (Broomhead & Lowe, 1988). The RBFNN is formed of three-layers, shown in Figure 3.18 The input layer, hidden layer and output layer. The hidden layer of an RBFNN is nonlinear and hires radial basis functions as the activation functions, whereas the output layer is linear.

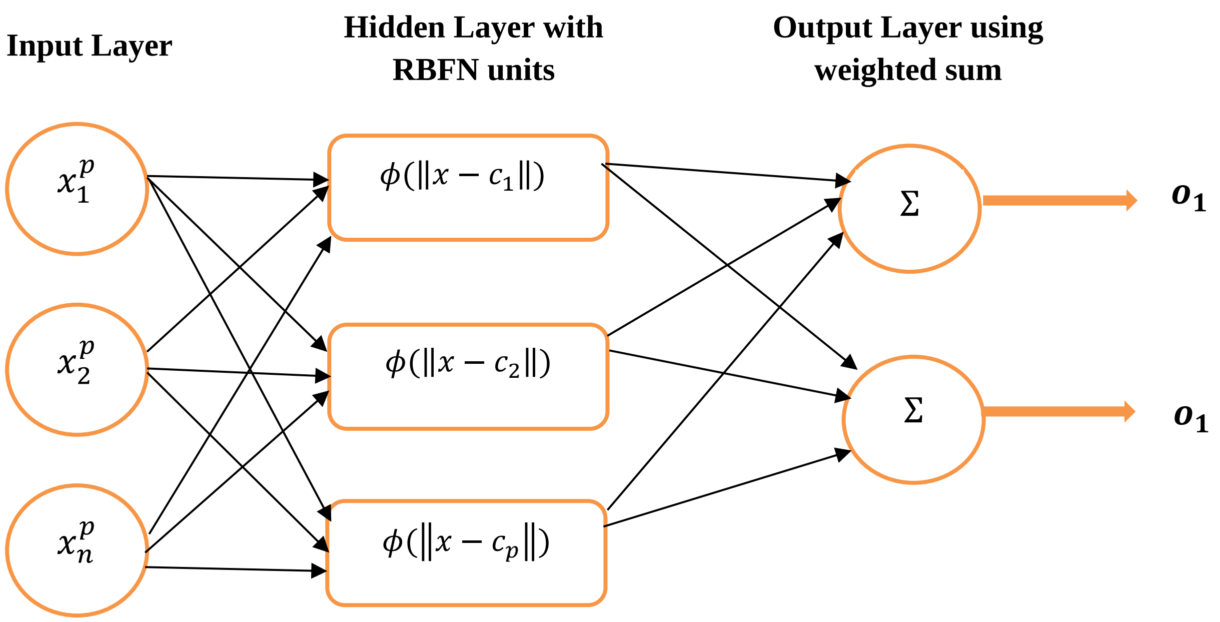


Figure 3. 18 The basic architecture of RBF neural network.

The mathematical formulation of the RBF network represented in the following equation:

$$o = g_{(x)} = \sum_{p=1}^{p} \lambda_p \varphi_{p\,(\|x - c_p\|)} \qquad (3.19)$$

Where $\|x - c_p\|$ represent the distance between the data point x and the RBF center $c_p$. ($\lambda_p$) is the weight linked with RBF center cp. Therefore, the RBF neural networks output is a weighted sum of the hidden layer's activation functions. The (RBFNN) model has indicated a promising results for time series forecasting for two different datasets of stick price and electric power load in (Rivas et al., 2004; Yan et al., 2005).

### 3.4 Developed Models

In this section, a group of models have been developed for the predictive analysis. These models include hybrid models, ensemble techniques and pseudo code for each model, prepared to cover the strengths of multiple individual models (statistical models & Ann's model) to reinforce predictive accuracy and robustness. Through combining different models, the goal is to enhance predictions performance and reliability. Below, the hybrid models developed in this study, along with the weighted average technique that was used to combine them and pseudo code for each, were discussed.

### 3.4.1 Hybrid Models

There is no single perfect forecasting model for a number of problems, which might provide a good accuracy of one problem, while it might be not good for another. Thus, it is possible to improve the forecasting accuracy by integrating and applying two or more models with different capabilities instead of single specific model with limited capabilities (Arunraj & Ahrens, 2015).The precept of Hybrid Models through combining the capabilities of classical statistical models and advanced machine-leaning models, helps to capture the linear dependencies and the non-linear dependencies to ensure a comprehensive analysis of complex temporal patterns which might be appear in sales data. Moreover, it can helps to develop an adaptive learning models which can update and refine itself based on new data, and guarantee the accuracy of the sales forecasting model to stay relevant in the dynamic daily sales environment such as “supermarkets”.

The development of a hybrid econometric-neural network model for forecasting total monthly sales, in order to integrate the structural characteristics of econometric models with the nonlinear patter recognition features of neural networks, where the forecasts from each individual sub-models are “averaged” to calculate the hybrid forecast (Luxhøj et al., 1996). By employing the strengths of different models, the hybrid forecasting method achieve more efficiently than the individual models for forecasting new products sales (Yin et al., 2020). The hybrid models took advantages of the unique capabilities of classic statistical models and advanced machine-learning models in linear and non-liner modeling over the time-series (Díaz-Robles et al., 2008).

#### 3.4.1.1 ARIMA-RNN

##### 3.4.1.1.1 Overview

The (ARIMA & RNNs) hybrid model combines the Autoregressive Integrated Moving Average (ARIMA) model with a Recurrent Neural Network (RNN). The ARIMA model is used to capture the linear components of the time series data such as trends and seasonality within the time series data, but it lacks when dealing with more complex, non-linear dependencies. On the other hand, RNNs exceed at capturing these non-linear dependencies because of their efficiency to preserve state information over time. By combining these two models, an investigation is conducted to handle both linear and non-linear aspects of the data, to improve the overall predictive performance.

#### 3.4.1.1.2 Methodology

1. **Data Preparation:** At first, the random seeds were set to 42, to ensure reproducibility, then converted the 'Date' index of the DataFrame to date time format, and extract the sales values from data frame.
2. **Train-Test Split:** Split the data into training and testing sets with an 80-20 ratio.
3. **ARIMA Modeling:** Train the ARIMA model on the training data, using the order set of ARIMA such as (6, 1, 1) which may be adjusted as needed, then forecasting for the length of the test data.
4. **RNN Modeling:** At first, generate sequences for the RNN model of sales data for training the model, which can also be adjusted as needed. Secondly, define the RNN architecture with one SimpleRNN layer with the required number of neurons (units) and the activation function "relu" followed by a Dense layer, then compile the model using the "Adam" optimizer and mean squared error loss function. Thirdly, train the RNN model on the training sequences and targets using 50 epochs with a batch size of 32.Forthly, generate sequence of test data for making predictions with the trained RNN model. Finally, make predictions on the test sequences using the trained RNN model.
5. **Combine Forecasts:** At first, assign weights to the ARIMA and RNN forecasts, as the weights can be adjusted based on model performance or trial and error. Secondly, calculate hybrid forecast by combining the forecasts using the weighted averaging technique.

#### 3.4.1.1.3 Pseudocode

The following pseudocode introduces the followed step in developing a hybrid model that combines the ARIMA model and Recurrent Neural Network (RNN). This approach aims to improve the forecasting performance by capturing both linear and non-linear patterns in the time series data.

# Function for ARIMA-RNN hybrid model (data, train_ratio, arima_order, sequence_ length, rnn_units, epochs, batch_size):

Step 1: Set random seeds for reproducibility

    Set random seeds to 42

Step 2: Prepare the data

    Convert "Date" index to date time

    Extract sales values from data

Step 3: Split the data into training and testing

train_size = train_ratio * length of sales values

Split sales values into train data and test data

Step 4: Train ARIMA model

arima_model = fit ARIMA (train_data, order=arima_order)

arima_result = arima_model. fit ()

Step 5: Prepare data for RNN model

Initialize train_sequences and train_targets

For i from 0 to length of train_data – sequence_length:

Append train_data [i: i + sequence_length] to train_sequences

Append train_data [i: i + sequence_length] to train_targets

Convert train_sequences and train_targets to arrays

Step 6: Define and compile RNN model

model = Sequential ([SimpleRNN(units = rnn_units, activation = 'relu', input_ shape = (sequence_length, 1)), Dense ( units = 1)

])

model. compile (optimizer = 'adam', loss = 'mean_squared_error')

Step 7: Train the RNN model

model.fit(train_sequences,train_targets, epochs=epochs, batch_size=batch_size, verbose = 1, shuffle = False)

Step 8: Make RNN predictions on test data

Initialize test_sequences

For i from 0 to length of test_data – sequence_length:

Append test_data [i: i+sequence_length] to test_sequences

Convert test_sequences to array

rnn_forecast = model. predict (test_sequences). flatten ()

Step 9: Make ARIMA predictions on test data

arima_forecast = arima_result. forecast (steps = length of test_data) [: length of test_data]

Step 10: Combine ARIMA & RNN forecast using weighted averaging

Define arima_weight and rnn_weight

min_len = minimum of length of arima_forecast and length of rnn_forecast

hybrid_forecast = (arima_forecast [:min_len] * arima_weight + rnn_forecast [:min_len] * rnn_weight) / (arima_weight + rnn_weight)

Step 11: Plot results

Plot actual data vs hybrid forecast

Step 12: Calculate and print error metrics

mse = Mean Square Error of actual and hybrid forecast

rmse = Root Mean Square Error of mse

mae = Mean Absolute Error of actual and hybrid forecast

Print mse, rmse, mae

End Function

#### 3.4.1.2 ARIMA-LSTM

##### 3.4.1.2.1 Overview

Combining (ARIMA with LSTM) covers the strengths of both linear and non-linear modeling techniques. ARIMA is known at capturing linear trends and seasonality within time-series data and its ability to handle short-term dependencies, enabling it for identifying and modeling predictable patterns. However, it considers weak when dealing with complex, non-linear dependencies. LSTMs, with their powerful capabilities to deal with long-term dependencies and maintain state information over time, surpass at capturing these non-linear dependencies. By combining these two models, the goal is to handle both linear and non-linear aspect of the data, to improve the overall predictive performance.

##### 3.4.1.2.2 Methodology

1. **Data Preparation:** At first, the random seeds were set to 42, to ensure reproducibility, then converted the 'Date' index of the DataFrame to date time format, and extract the sales values from data frame.
2. **Train-Test Split:** Split the data into training and testing sets with an 80-20 ratio.
3. **ARIMA Modeling:** Train the ARIMA model on the training data, using the order set of ARIMA such as (6, 1, 1) which may be adjusted as needed, then forecasting for the length of the test data.
4. **LSTM Modeling:** At first, define the sequence length for the LSTM model, which can be adjusted as needed. Secondly, generate sequences and identical targets from the training data, then ensure that the data is in the correct shape for input into the LSTM model. Thirdly, build the LSTM network with required and suitable number of neurons ( units ) and the activation function "relu" followed by adding a Dense layer, then compile the model using the "Adam" optimizer and mean squared error loss function.

Fourthly, train and fitting the LSTM model using the training sequences and targets. Finally, making predictions using the test sequences.

5. **Combine Forecasts:** At first, assign weights to the ARIMA and LSTM forecasts, as the weights can be adjusted based on model performance or trial and error. Secondly, calculate hybrid forecast by combining the forecasts using the weighted averaging technique.

#### 3.4.1.2.3 Pseudocode

The following pseudocode introduces the followed step in developing a hybrid model that combines the ARIMA model and Long Short-Term Memory (LSTM). This approach aims to improve the forecasting performance by capturing both linear and non-linear patterns in the time series data.

# Function for ARIMA-LSTM hybrid model (data, train_ratio, arima_order, sequence_ length, lstm_units, epochs, batch_size):

Step 1: Set random seeds for reproducibility

Set random seeds to 42

Step 2: Prepare the data

Convert "Date" index to date time

Extract sales values from data

Step 3: Split the data into training and testing

train_size = train_ratio * length of sales values

Split sales values into train data and test data

Step 4: ARIMA Fitting

arima_model = fit ARIMA (train_data, order=arima_order)

arima_result = arima_model. fit ()

arima_forecast = arima_result. forecast (steps = length of test_data) [: length of test_data]

Step 5: Prepare Data for LSTM Model

Initialize train_sequences and train_targets

For i from 0 to length of train_data – sequence_length:

Append train_data [i: i + sequence_length] to train_sequences

Append train_data [i: i + sequence_length] to train_targets

Convert train_sequences and train_targets to arrays

Step 6: Define and Compile LSTM Model

Model = Sequential ()

Model. Add (LSTM (units=lstm_units, activation = ‘relu’, input_shape = (sequnce_length, 1)))

Model. Add (Dense (units=1))

Model. Compile (optimizer=’adam’, loss = ‘mean_squared_error’)

Step 7: Train the LSTM Model

model.fit(train_sequences,train_targets, epochs=epochs, batch_size=batch_size, verbose = 1, shuffle = False)

Step 8: Make Predictions on Test Set

Initialize test_sequences

For i from 0 to length of test_data – sequence_length:

Append test_data [i: i+sequence_length] to test_sequences

Convert test_sequences to array

lstm_forecast = model. predict (test_sequences). flatten ()

Step 9: Combine ARIMA & LSTM Models Forecasts using Weighted Averaging

Define arima_weight and lstm_weight

min_len = minimum of length of arima_forecast and length of lstm_forecast

hybrid_forecast = (arima_forecast [:min_len] * arima_weight + lstm_forecast [:min_len] * lstm_weight) / (arima_weight + lstm_weight)

Step 10: Plot Results & Calculate Error Metrics

Plot actual data vs hybrid forecast

mse = Mean Square Error of actual and hybrid forecast

rmse = Root Mean Square Error of mse

mae = Mean Absolute Error of actual and hybrid forecast

Print mse, rmse, mae

End Function

#### 3.4.1.3 ARIMA-MLPNNs

##### 3.4.1.3.1 Overview

Combining (ARIMA with MLPNNs) covers the strengths of both linear and non-linear modeling techniques. ARIMA is known at capturing linear trends and seasonality within time-series data and its ability to handle short-term dependencies, enabling it for identifying and modeling predictable patterns. However, it considers weak when dealing with complex, non-linear dependencies. MLPNNs, with their powerful capabilities to learn complex functions, surpass at capturing these non-linear dependencies. By combining these two models, the goal is to handle both linear and non-linear aspect of the data, to improve the overall predictive performance.

**3.4.1.3.2 Methodology**

1. **Data Preparation & Split the data**

Converted the ‘Date’ index of the DataFrame to date time format, then extract the sales values from data frame. Split the data into training and testing sets with an 80-20 ratio.

2. **ARIMA Modeling**

Train the ARIMA model on the training data, using the order set of ARIMA such as (6, 1, 1) which may be adjusted as needed, then forecasting for the length of the test data.

3. **MLPNN Modeling**

At first, prepare the data for multilayer-perceptron neural network model MLPNN by reshaping the training data where the input data points for the MLPNN are the original data points and the target values are the next data points in the sequence. It encompass splitting the training data into input features as (‘X_train’) and target values as (‘y_train’) where each input feature is a data point and the target values are the next data point in the sequence, which has been done the same for the testing data. Secondly, build the MLPNN network by identifying the number of hidden layers and the suitable number of neurons in the single layer, the maximum iterations, the activation function “relu” and the number of random states. Thirdly, train and fitting the MLPNN using the training and target values in the training data. Finally, making predictions using the testing data.

4. **Combine Forecasts**

At first, assign weights to the ARIMA and MLPNN forecasts, as the weights can be adjusted based on model performance or trial and error. Secondly, calculate hybrid forecast by combining the forecasts using the weighted averaging technique.

**3.4.1.3.3 Pseudocode**

The following pseudocode introduces the followed step in developing a hybrid model that combines the ARIMA model and Multilayer Perceptron Neural Network (MLPNN). This approach aims to improve the forecasting performance by capturing both linear and non-linear patterns in the time series data.

Step 1: Prepare the data

Convert "Date" index to date time

Extract sales values from data

Step 2: Split the data into training and testing

train_size = train_ratio * length of sales values

Split sales values into train data and test data

Step 3: ARIMA Fitting

arima_model = fit ARIMA (train_data, order=arima_order)

arima_result = arima_model. fit ()

arima_forecast = arima_result. forecast (steps = length of test_data) [: length of test_data]

Step 4: Reshape Data for MLPNN Model

Function to reshape data for the MLPNN (train_data, test_data, random_state):

# Initialize lists for input features (X) and target values (y)

X_train = [ ]

x_test = [ ]

y_train = [ ]

y_test = [ ]

# Get the split index for training and testing sets splitting using (train_test_split ()) and based on the specified 'test_size' which is 0.2

Split_index = int ((1- test_size) * length of train_data)

# Loop through the train data to create the input features and target values

For i from 0 to split_index – 1:

Append train_data [i] to X_train

Append train_data [i + 1] to y_train

# Loop through the remaining data to create test sets

For i from split_index to length of train_data – 1:

Append train_data [i] to X_test

Append train_data [i + 1] to y_test

Return X_train, y_train, X_test, y_test

End Function

Step 5: Build the MLPNN network

Function for the MLPNN model initialization (hidden_layer_sizes, max_iteration, activation function, random_state):

Create the MLPNN model with specified parameters

mlp_model = MLPRegressor (hidden_layer_sizes = hidden_layer_sizes, max_iteration= max_iteration, activation = activation, random_state = random_state0

Return mlp_model

End Function

Step 6: Training and fitting the MLPNN model

Function for the MLPNN training (mlp_model, X_train, y_train):

Fit the MLPNN model on the training data

mlp_model. fit (X_train, y_train)

Return mlp_model

End Function

Step 7: Making forecasts using the MLPNN model

Function for making forecasts using (mlp_model, Test_data):

# Use the MLPNN model to make prediction on the test data

mlp_forecast = mlp_model. predict (Test_data [:1])

Return mlp_forecast

End Function

Step 8: Combine ARIMA & MLPNN Models Forecasts using Weighted Averaging

Define arima_weight and mlp_weight

min_len = minimum of length of arima_forecast and length of mlp_forecast

hybrid_forecast = (arima_forecast [:min_len] * arima_weight + mlp_forecast [:min_len] * mlp_weight) / (arima_weight + mlp_weight)

Step 9: Plot Results & Calculate Error Metrics

Plot actual data vs hybrid forecast

mse = Mean Square Error of actual and hybrid forecast

rmse = Root Mean Square Error of mse

mae = Mean Absolute Error of actual and hybrid forecast

Print mse, rmse, mae

End Function

#### 3.4.1.4 ARIMA-RBFNN

##### 3.4.1.4.1 Overview

Combining Autoregressive Integrated Moving Average (ARIMA) with Radial Basis Function Neural Networks (RBFNN) takes advantages of both linear and non-linear modeling techniques. ARIMA capabilities to capture linear trends and seasonality, making it powerful for differentiating and modeling predictable patterns. However, it’s less proficient with complex, non-linear dependencies. On the other hand, RBFNN is capable at modeling these non-linear dependencies in the data using radial basis functions (RBFs). The integration of these models, allows to handle both linear and non-linear parts of the data, improving the overall predictive performance.

**3.4.1.4.2 Methodology**

1. **Data Preparation & Split the data**

Converted the ‘Date’ index of the DataFrame to date time format, then extract the sales values from data frame. Split the data into training and testing sets with an 80-20 ratio.

2. **ARIMA Modeling**

Train the ARIMA model on the training data, using the order set of ARIMA such as (6, 1, 1) which may be adjusted as needed, then forecasting for the length of the test data.

3. **RBFNN Modeling**

At first, prepare the data for Radial Basis Function Neural Network (RBFNNs) by reshaping the training data where the input data points for the RBFNN are the original data points and the target values are the next data points in the sequence. It encompass splitting the training data into input features as (‘X_train’) and target values as (‘y_train’) where each input feature is a data point and the target values are the next data point in the sequence, which has been done the same for the testing data. Secondly, define and inherits the RBFNN class from ‘tf.keras.Model’ and represents the Radial Basis Function Neural Network (RBFNN), which contains components of initialization (_init_ method) such as the dimension of input data, the number of RBFs units (hidden layers) in the network, the dimension of output, a variable to store the RBFs centers of the RBF units (initialized as zero), a variable to represent the spread of RBFs (initialized as one) and a dense layer with a ‘relu’ activation function to produce the output from the RBF layer. Thirdly, compile the modes using “Adam” optimizer and mean squared error loss function. Fourthly, fitting and training the RBFNN model using the training sets, epochs. Finally, making predictions using testing sets.

4. **Combine Forecasts**

At first, assign weights to the ARIMA and RBFNN forecasts, as the weights can be adjusted based on model performance or trial and error. Secondly, calculate hybrid forecast by combining the forecasts using the weighted averaging technique.

#### 3.4.1.4.3 Pseudocode

The following pseudocode introduces the followed step in developing a hybrid model that combines the ARIMA model and Radial Basis Function Neural Network (RBFNN). This approach aims to improve the forecasting performance by capturing both linear and non-linear patterns in the time series data.

Step 1: Prepare the data

Convert "Date" index to date time

Extract sales values from data

Step 2: Split the data into training and testing

train_size = train_ratio * length of sales values

# Split sales values into train data and test data

Split sales values into X_train, X_test, y_train, Y_test using train_test_split with test_size = 0.2 and shuffle = False

Step 3: ARIMA Fitting

arima_model = fit ARIMA (train_data = X_train, order=arima_order)

arima_result = arima_model. fit ()

arima_forecast = arima_result. forecast (steps = length of test_data) [: length of Y_test]

Step 4: Define and Compile RBFNN model

The class of RBFNN inherits from (tf.keras.Model)

# Initialize RBFNN model with (input_ dimension, number of RBF units, output_ dimension):

_init_ (input dim, num_rbf_units, output _dim)

# Initialize (centers, betas, dense layer)

self. centers = tf. variable to store the centers of RBF units, initialized as zeros

self.beta = tf. variable to introduce the RBFs spread, initialized as ones.

self.dense = tf. a dense layer with an activation function "relu" and output dimension.

# Define "call" method for forward pass

The input data (x) = tf. Input expansion (expand_ dims) (input, 1)

# Distance calculation

diff = the input data (x) – the centers of RBFs units (self. centers)

# The Squared Euclidian distance (Norm) calculation between the input and the centers

Norm = tf.reduce_sum (diff **2, axis = 1)

# Get the activation values of the RBFs units by applying the radial basis function to the norm

activation values of RBFs (rbf_out) = tf.exp (norm * self. beta)

# Final output of Network

Return self.dense (rbf_out)

End Class

# Compile the RBFNN model using optimizer “adam” and mean squared error loss function:

Rbfnn_model_compile (optimizer = ‘adam’, loss = ‘mean_squred_error’)

Step 5: Training and fitting the RBFNN model using the training set (X_train, y_train)

Rbfnn_mod = rbfnn_model.fit (X_train, y_train, epochs = epochs, batch_size = batch_size)

Step 6: Use the RBFNN model (Rbfnn_model) to make prediction on the test data

Y_pred_rbfnn = Rbfnn_model. predict (X_test)

Step 7: Combine ARIMA & RBFNN Models Forecasts using Weighted Averaging

Define arima_weight and mlp_weight

min_len = minimum of length of arima_forecast and length of mlp_forecast

hybrid_forecast = (arima_forecast [:min_len] * arima_weight + (y_pred_rbfnn * rbfnn_weight)

Step 8**:** Plot Results & Calculate Error Metrics

Plot actual data vs hybrid forecast

mse = Mean Square Error of actual and hybrid forecast

rmse = Root Mean Square Error of mse

mae = Mean Absolute Error of actual and hybrid forecast

Print mse, rmse, mae

#### 3.4.1.5 Weighted Average model for hybrid models

When a single model is created, occasionally the prediction or the model accuracy might not be enough for achieving the best accuracy and the required prediction. As statistical models are able to forecast for short-term horizon and capture linear dependences, while the advanced machine-learning models such as ANNs are able to capture long-term horizons and able to handle more complex patterns with nonlinear dependences. To minimize this problem, and get advantage of both approaches for products sales forecasting based on time series sequential data, multiple models from both approaches (statistical model with ANN model) were combined, to get one with better performance. An ensemble is providing techniques to combine different set of individual models together to enhance the stability and predictive power of the model using diverse ensemble learning techniques such as averaging or weighted average technique. In another meaning, the ensemble learning is a process of using multiple models that are strategically built to solve a particular problem(Bhatnagar, 2023; *Ensemble Averaging (Machine Learning)*, 2021; *Python Code for Weighted Average Ensemble - Google Search*, n.d.; *What Is Ensemble Averaging (Machine Learning)? | Autoblocks Glossary*, n.d.).Weighted average model is an ensemble technique which weight the contribution of each sub-model to the combined prediction through the expected performance of the sub model. This technique enable efficient models to contribute more and less efficient models to contribute less (Bhatnagar, 2023; *Ensemble Averaging (Machine Learning)*, 2021; *What Is Ensemble Averaging (Machine Learning)? | Autoblocks Glossary*, n.d.). Figure 3.19 Illustrates how the ensemble model works:

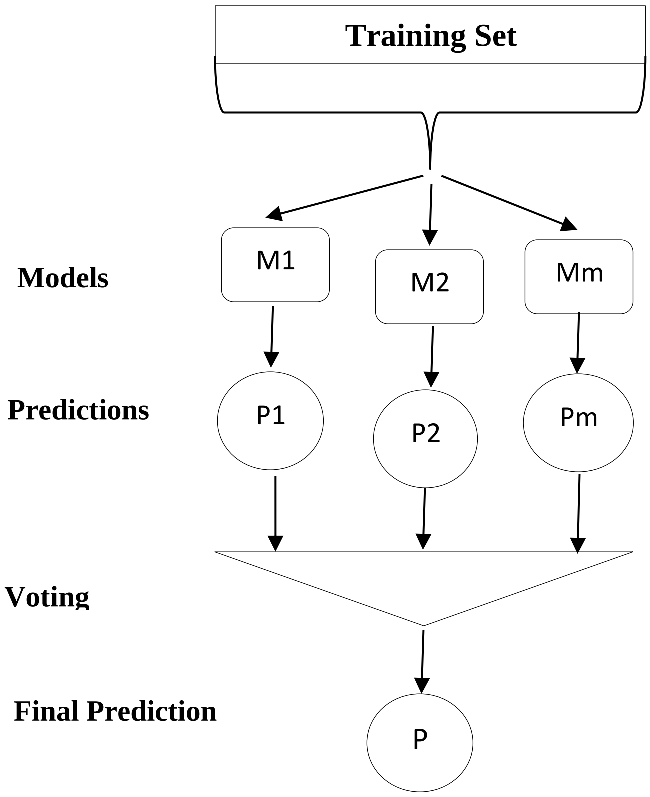

Figure 3. 19 The ensemble technique to combine models predictions into one single model.

The diagram shows that there are many models, where in our case are (statistical models and ANNs models) such as (ARIMA & RNN).All of these models take the same dataset, and each model gives its predictions, then using any of voting techniques to combine the models predictions and gives one model prediction such as weighted averaging by giving each model a weight, and high weight for the model that performs well than the other model. Make predictions using each model. Multiply each model predictions with its assigned weight. Finally, sum up the weighted predictions to get the final predictions. For example, if two models will be trained: ARIMA and RNN models. Then assign weights where ARIMA (weight: .1) and RNN (weight: .9).

The weighted average of predictions from multiple models (one statistical and one ANN model) in our case, can be calculated as follows:

$$\hat{y} = \sum_{i=1}^{n} w_i \hat{y}_i \qquad (3.20)$$

Where $\hat{y}$ the final is predicted value, $n$ is the number of models, $w_i$ is the weight assigned to the $i_{th}$ model's prediction and $\hat{y}_i$ is the predicted value from the $i_{th}$ model.

An example with two models: If there are two models (ARIMA model, RNN model), the weighted average prediction can be expressed as:

$$\hat{y} = w_1 \hat{y}_1 + w_2 \hat{y}_2 \quad (3.21)$$

There are constraints related to the total weight, which typically need to sum up to 1:

$$\sum_{i=1}^{n} w_i = 1 \qquad (3.22)$$

Let's assume the weights and predictions from two models (ARIMA and RNN) models are as follows:

- Weight for Model 1 (ARIMA model)($w_{1)}$) = 0.1, Prediction from Model 1 (ARIMA model) ($\hat{y}_1$) = 10.
- Weight for Model 2 (RNN model)($w_{2)}$) = 0.9, Prediction from Model 2 (RNN model) ($\hat{y}_2$) = 20.

The final prediction ($\hat{y}$) would be calculated as:

$$\hat{y} = (0.1 \ \times 10) + (0.9 \ \times 18) = 1 + 16.2 = 17.2$$

### 3.5 Evaluation Criteria

This work focus on enhancing the prediction process of (Dukkan 11) products sales. The error metrics of mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE), were used for evaluating the models performance as pure models and as hybrid models.

The formula for the mean square error (MSE) is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{N}(y_i - \ p_i)^2 \qquad (3.23)$$

Where $y_i$ is the $i_{th}$ observed value (actual value), $p_i$is the corresponding predicted value for $y_i$ , and $n$ is the number of observations. The $\Sigma$ indicated that a summation is performed over all values of $i$.

The formula for the root mean squared error (RMSE) is:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left\|y_{(i)} - \ \hat{y}(i)\right\|^2}{N}} \qquad (3.24)$$

Where $N$ is the number of data points, $y_{(i)}$ is the observed value (actual value), $\hat{y}(i)$is the corresponding predicted value for $y_{(i)}$ .

The formula for mean absolute error (MAE) is:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \ x| \qquad (3.25)$$

Where $n$ is the number of observations, $x_i$ is the predicted value, and $x$ is the actual value.

And the error between the output and target is calculated as follow:

*Error:*

$$e = \sum (Actual\ value - Predicted\ value) \qquad (3.26)$$

# Chapter Four: Results

# 4. Experiments and Results

### 4.1 Experiments and Results Background

In this chapter, the process of preparing the product sales datasets is introduced. Then, the results of exploratory data analysis using statistical and visualization analysis will be explained and introduced, to study the data behavior and discover the time series patterns. Moreover, the data preparation before applying the statistical models (ARIMA, SARIMA) models will be explained. The main target is to forecast the sales of the product as (combined sales and based on each product level (separately)) based on daily sales records, to be able to forecast future sales for the next time series. For NN models and hybrid models, the data has been split into training and testing sets with (80% for training) and (20% for testing). Our objective is to evaluate the optimum results that occurred using error metrics (MSE, RMSE, and MAE) metrics through the application of pure models individually and as hybrid models. This evaluation will cover scenarios where product sales are combined, as well as scenarios where product sales are kept separate. By analyzing the performance of each approach on these contexts, beneficial insights into the effectiveness of different modeling strategies and their influence on overall sales forecasting accuracy, will be obtained. The experiments and analysis were applied using a computing device with an Intel(R) Core(TM) i7-4510U CPU @ 2.00GHz, 8.00 GB RAM, running on Windows 10 Pro (64-bit). The software environment included Python along with group of libraries: Pandas for data manipulation and analysis, numpy for numerical calculations, tensorflow for neural network modeling and training, scikit-learn for machine learning algorithms, statsmodels for statistical models including ARIMA, SARIMA, and matplotlip for plotting and visualization. Jupiter Notebook (version 6.5.4) was used for coding and introducing the analysis results. The arrangement of components was intended to achieve efficient enforcement of data processing, statistical and machine learning tasks.

## 4.2 Exploratory Data Analysis (EDA)

### 4.2.1 Descriptive Analysis

A comprehensive statistical analysis was applied to obtain insights into the underlying patterns and trends within the product sales dataset. At first, basic statistical measures such as mean, median, standard deviation, min, max, etc were calculated to get a comprehensive understanding of the central tendency and variability of sales data across different categories. Also, to search for relationships between sales performance and different factors such as seasonality, weekends, holidays, and local events. Additionally, it has been used to discover temporal dependencies and long-term trends in sales data to enable a further understanding of product sales dynamics. By extracting weekends for each category, it was clear that some categories were experiencing high sales over the weekend, such as dairies. The main reason is that families prefer purchasing dairies items, a trend that looks to be impacted by the types of meals usually prepared and enjoyed during weekend gatherings. Analysis of consumer behavior uncovers a correlation between the food preferences of households and their shopping habits, particularly on weekends. Table 4.1 shows the descriptive analysis of the “dairies” category based on weekends.

Table 4. 1 Dairies sales based on weekends.

| Year | Is Weekend? | Count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | False | 176 | 120.84 | 51.91 | 7 | 84.37 | 115.75 | 161.97 | 244.17 |
| | True | 130 | 134.36 | 54.41 | 22 | 100.85 | 131 | 166.49 | 326.93 |
| 2022 | False | 208 | 191.11 | 69.16 | 60.9 | 135.68 | 182.45 | 232.95 | 408.03 |
| | True | 157 | 187.66 | 63.19 | 22.5 | 141 | 181.92 | 228.68 | 379.87 |
| 2023 | False | 174 | 322.53 | 106.46 | 80 | 241.19 | 318.29 | 398.79 | 739.00 |
| | True | 130 | 325.1 | 98.41 | 155 | 257.25 | 302.4 | 377.75 | 818.57 |

For the drinks category, it has shown a tendency to buy drinks items during their usual visits to the supermarket or while passing during their usual days. Table 4.2 shows the descriptive analysis of the “drinks” category based on weekends.

Table 4. 2 Drinks sales based on weekends.

| Year | Is Weekend? | Count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | False | 176 | 469.04 | 213.76 | 26.36 | 299.12 | 487.34 | 630.73 | 925 |
| | True | 130 | 461.81 | 194.58 | 42.61 | 295.78 | 487.86 | 595.23 | 857.61 |
| 2022 | False | 208 | 752.34 | 243.01 | 314.49 | 585.18 | 721.43 | 892.6 | 2070.37 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | True | 157 | 717.24 | 218.23 | 296.47 | 526.96 | 703.36 | 890.47 | 1287.12 |
| 2023 | False | 174 | 816.34 | 250.36 | 179 | 645.48 | 786.5 | 970.75 | 1576 |
| | True | 130 | 812.70 | 262.01 | 200.75 | 624.81 | 776.13 | 1002.87 | 1861.7 |

The purchase of cleaning materials witnessed significant increases during the weekdays, because the demand for cleaning supplies is greater during the weekdays than on the weekends, which are a holiday for them than doing household cleaning work. Table 4.3 shows the descriptive analysis of the "cleaning materials" category based on weekends.

Table 4. 3 Cleaning materials sales based on weekends

| Year | Is Weekend? | Count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | False | 176 | 158.26 | 88.13 | 0.00 | 93.75 | 147.00 | 209.87 | 549.00 |
| | True | 130 | 180.91 | 89.48 | 6.00 | 114.19 | 179.52 | 240.8 | 424.00 |
| 2022 | False | 208 | 229.49 | 95.72 | 77.00 | 165.91 | 212.25 | 284.08 | 623.71 |
| | True | 157 | 233.03 | 92.21 | 43.97 | 164 | 222.00 | 298.95 | 507.92 |
| 2023 | False | 174 | 451.90 | 223.33 | 70.50 | 299.67 | 406.00 | 559.95 | 1320.50 |
| | True | 130 | 425.25 | 165.36 | 59.88 | 302.71 | 405.35 | 532.93 | 1041.00 |

The purchase of snacks & chips items is usually from the young age group or school students during the weekdays due to that the supermarket location is surrounded by several student schools, they usually prefer to buy such items as they pass through the supermarket. Table 4.4 shows the snacks & chips category based on weekends

Table 4. 4 Snacks & chips sales based on weekends.

| Year | Is Weekend? | Count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | False | 176 | 727.10 | 276.07 | 95.60 | 525.03 | 716.91 | 923.06 | 1351.03 |
| | True | 130 | 695.64 | 250.39 | 142.40 | 475.36 | 713.66 | 883.11 | 1389.13 |
| 2022 | False | 208 | 1063.21 | 229.93 | 528.13 | 917.46 | 1042.05 | 1213.51 | 2037.09 |
| | True | 157 | 968.87 | 217.66 | 111.61 | 848..03 | 930.70 | 1105.85 | 1547.58 |
| 2023 | False | 174 | 1039.41 | 257.85 | 420.75 | 855.92 | 1023.47 | 1215.82 | 1829.00 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | True | 130 | 998.21 | 203.68 | 400.50 | 844.75 | 996.00 | 1136.52 | 1651.50 |

For ice-cream category items, it has shown significant sales during weekends which might be related to customers' preferences for family outings such as picnics, or visits to parks. Or it might be related to social gatherings at weekends with friends and families where ice cream is often served as a snack or dessert during these gatherings. Table 4.5 shows the ice cream category based on weekends.

Table 4. 5 Ice-cream sales based on weekends.

| Year | Is Weekend? | Count | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | False | 176 | 56.59 | 51.79 | 0.00 | 12.37 | 48.00 | 87.00 | 211.00 |
| | True | 130 | 69.51 | 68.99 | 0.00 | 13.62 | 49.25 | 103.16 | 330.97 |
| 2022 | False | 208 | 94.97 | 78.72 | 0.00 | 21.5 | 83.61 | 158.58 | 324.20 |
| | True | 157 | 99.26 | 85.73 | 0.00 | 25..00 | 84.50 | 159.00 | 381.12 |
| 2023 | False | 174 | 81.92 | 70.49 | 0.00 | 17.00 | 70.00 | 131.75 | 317.00 |
| | True | 130 | 87.62 | 74.46 | 0.00 | 22.75 | 78.00 | 137.32 | 271.00 |

Another descriptive analysis has been applied for monthly sales data for each product category to get more insights into consumer behavior, preferences, and market trends. It appears that cleaning materials experienced high sales in the months between (May to Dec) with an increasing trend from 2021 to 2023. This seasonality behavior refers to an annual tradition which is commonly known as “spring cleaning”, also it refers to seasonal cleaning routines, as the weather becomes warmer and people are motivated to refresh their living spaces. And many other possible reasons such as “back to school season” specifically in August and September. Moreover, as the year comes to a close, some individuals prefer to make end-of-year cleaning routines to start the new year with a fresh and organized home. For detailed tables and further data that support results presented in the exploratory data analysis, please refer to the Appendix section. The appendix contains thorough tables and additional information that provide more insights into monthly sales for cleaning materials category.

For dairies category, it has shown an increased trend over the years from 2021 to 2023, starting from July to December. The increasing trend in dairy consumption particularly in these months, might be due to summer holidays and outdoor activities, as these items are preferable for summer snacks such as outdoor gatherings, barbecues, and picnics which increase consumption of these items. Particular dairy products are more commonly enjoyed during holidays and gatherings in the latter part of the year. For detailed tables and further data that support results presented in the exploratory data analysis, please refer to the Appendix section. The appendix contains thorough tables and additional information that provide more insights into monthly sales for dairies category.

For ice-cream category, it has shown a clear seasonality in summer months especially from May to September over the years. This type of category is available in the summer months and considered as a preferable item because of high temperatures in summer. For detailed tables and further data that support results presented in the exploratory data analysis, please refer to the Appendix section. The appendix contains thorough tables and additional information that provide more insights into monthly sales for ice-cream category.

The drinks category has shown both trend and seasonality patterns. The seasonality pattern in months from May to December with increased trend over the years. Certain types of drinks are preferred during specific seasons for factors such as weather and holidays. In addition, the drinks industry is dynamic, because of the introducing of new flavors and verities, which might lead to drink fluctuations over time. For detailed tables and further data that support results presented in the exploratory data analysis, please refer to the Appendix section. The appendix contains thorough tables and additional information that provide more insights into monthly sales for drinks category.

For snacks & chips category, it has shown an increased trend over the years, and a slight seasonality in the months of (June to December) specifically in September to December. September is the beginning of back-to-school, where families prefer to buy some snacks for their kid's lunchboxes. Also, the supermarket location is near some schools, so students prefer to buy after-school snacks. In the winter and fall seasons from October to December, the weather cools down, and people's consumption of snacks and chips increases. Moreover, some snack manufacturers often supply supermarkets with seasonal flavors that attract consumers and lead to higher sales. For detailed tables and further data that support results presented in the exploratory data analysis, please refer to the Appendix section. The appendix contains thorough tables and additional information that provide more insights into monthly sales for snacks & chips category.

### 4.2.2 Augmented Dicky-Fuller (ADF)

The ADF (Augmented Dicky-Fuller) test is a statistical test used to identify whether a unit root is present in a time series dataset, and helps to evaluate the stationary of a time series which is a crucial step for many time series analysis techniques (*Augmented Dickey-Fuller Test (ADF) in Time Series Analysis - Google Search*, 2020; Wikipedia Contributors, 2019). Table 1. Shows the (ADF) results before and after applying differencing using (d=1) for the combined sales. Table 2. Shows the (ADF) results for each category sales before and after applying differencing using (d=1).

The null hypothesis (H0) of the ADF test: the time series suggests that it is non-stationary if ($p > 0.05$).

The alternative hypothesis (H1) of the ADF test: the time-series indicating that it is stationary if ($p < 0.05$).

Table 4. 6 The (ADF) results before and after differencing for combined sales using order (d=1).

| Differencing(I) | P -value (before differencing) | P- value (after differencing) |
|---|---|---|
| d=1 | 0.12 | 1.44e-10 |

Based on Table 4.6 results, the differenced time series is likely stationary (reject the null hypothesis).

Table 4. 7 The (ADF) results before and after differencing for each category using order (d=1).

| Category | Differencing(I) | P -value (before differencing) | P- value (after differencing) |
|---|---|---|---|
| Drinks | d=1 | 0.11 | 9.3e-27 |
| Dairies | d=1 | 0.38 | 4.6e-26 |
| Cleaning materials | d=1 | 0.33 | 2.9e-25 |
| Ice-cream | d=1 | 0.25 | 2.0e-26 |
| Snacks & Chips | d=1 | 0.05 | 3.3e-14 |

Based on Table 4.7 results, the differenced time series is likely stationary (reject the null hypothesis).

### 4.3 Models Experiments

In this thesis, all experiments were performed using Python, a multi-functional programming language that is widely used in the field of machine learning and data analysis because of its comprehensive libraries and ease of use. Python provides abundant options of ecosystem tools and frameworks that simplify different stages of the research process, from data preprocessing to model evaluation. For more effective use and implementation of Python environment, Anaconda, which is a popular open-source distribution of Python, was used. Anaconda facilitates the setup and configuration of Python environment by providing an inclusive package manager and virtual environment manager. Moreover, Anaconda contains a large array of pre-installed libraries which usually used in data science and machine learning, such as NumPy, Pandas, Sikit-learn, TensorFlow, Matplotlib, and many other libraries, regulate the implementation of algorithms and reducing development time. In addition, Anacondas integrated development environment (IDE), such as Jupiter Notebook, suggests an interactive environment for prototyping code, visualizing data, and documenting research findings. By utilizing Python and the Jupiter Notebook of Anaconda, this thesis has obtained an advantage from a powerful computational environment, allowing for efficient experimentation, accurate analysis, and informative interpretation of results.

The results were obtained for the two scenarios as combined sales for five categories, and based on each product level (separately) based on daily sales records. For both scenarios, two stages were followed. For the first stage, the statistical models (ARIMA and SARIMA) and Neural Networks models (RNNs, LSTM, MLPNNs, and RBFNNs) were applied individually and measured each model performance using the error metrics (MSE, RMSE, and MAE) to get numerical insights into how well each model is performing. Moreover, visualizations were applied to understand how each model is making predictions and where it might be struggling. The dataset was split into 80% for training and 20% for testing. For the second stage, the hybrid models were created using the best statistical model performance with neural network models.

To identify the most appropriate parameter values for ARIMA and SARIMA models, a systematic trial and error approach was employed. This repeated process involved experimenting with different sets of parameters. These parameters included the autoregressive order (p), the order of differencing (d), the moving average order (q), and seasonal parameters for the SARIMA model (P,D,Q, and m). The final parameter values were selected based on the models' performance metrics. To enhance the architecture of the neural network models, a range of experiments was involved to identify the most appropriate number of neurons from (5 to 30) by adding 5 neurons at the time. Furthermore, a decision was made to use a single hidden layer for the MLPNNs.

### 4.3.1 Statistical Models

#### 4.3.1.1 Autoregressive Integrated Moving Average (ARIMA)

The models' experiments were applied under two scenarios, where the first scenario is for products combined sales, and the second one based on each product sales individually (drinks, dairies, ice cream, snacks & chips, and cleaning materials). A systemic trial and error approach was employed to identify sets of parameters which include the autoregressive order (p), the differencing order (d), and the moving average order (q). The error metrics were used to evaluate the model performance and provide numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best parameter combination explored for the ARIMA model was using the set of (6, 1, 1), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), and the moving average order (q=1) respectively, collectively assist the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.1 shows the product's combined sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023)which represents the observed and forecasted behavior for the combination of (6, 1, 1).

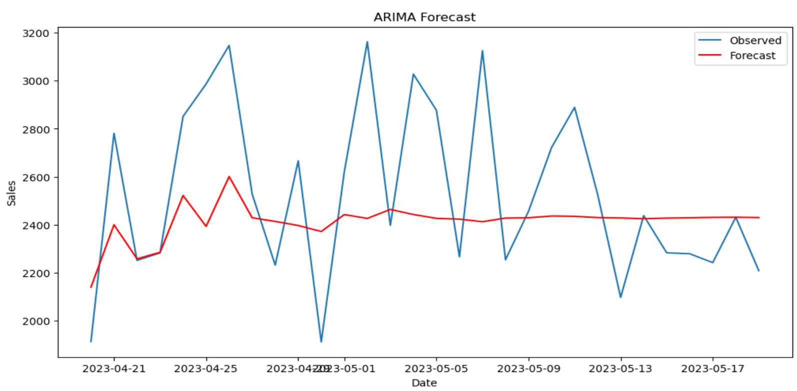


Figure 4. 1 The ARIMA forecasting for products combined sales for one month.

The second scenario shows the experiments for each product sale (Individually). For dairies products sales, the experiment shows the best parameter combination explored for the ARIMA model was using the set of (6, 1, 6), appeared as the most promising, and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), and the moving average order (q=6) respectively, collectively assist the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.2 shows the products' combined sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represents the observed and forecasted behavior for the combination of (6, 1, 6).

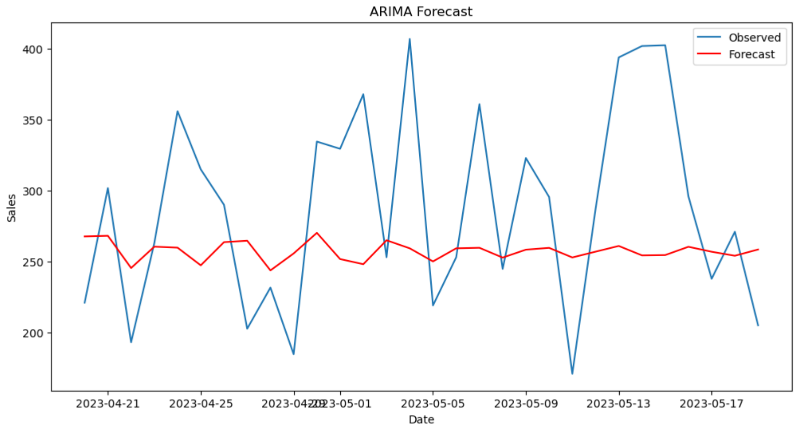


Figure 4. 2 The ARIMA forecasting for dairies products sales for one month.

For drinks products sales, the experiments show the best parameter combination explored for the ARIMA model was using the set of (6, 1, 1), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), and the moving average order (q=1) respectively, collectively assist the model's ability to capture temporal dependencies and fluctuations in the time series data**.** Figure 4.3 shows drinks product sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represents the observed and forecasted behavior for the combination of (6, 1, 1).

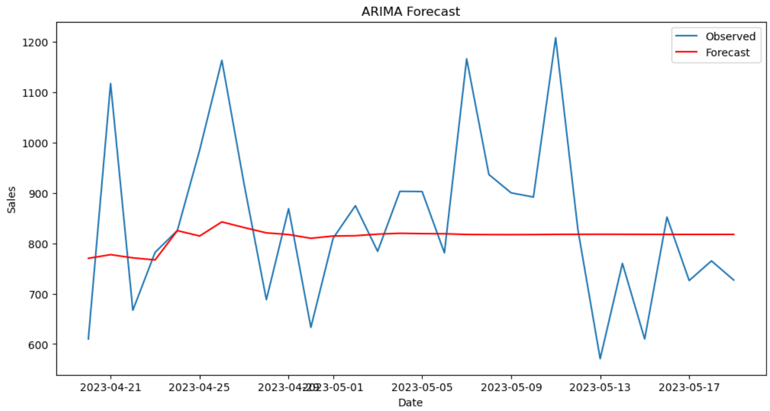

Figure 4. 3 The ARIMA forecasting for drinks products sales for one month.

For ice-cream products sales, the experiments show the best parameter combination explored for the ARIMA model was using the set of (6, 1, 1), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), and the moving average order (q=1) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.4 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (6, 1, 1).

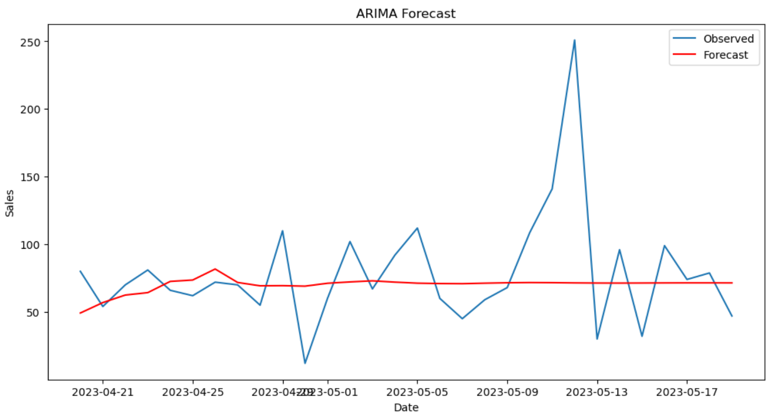


Figure 4. 4 The ARIMA forecasting for ice-cream products sales for one month.

For snacks and chips products sales, the experiments show the best parameter combination explored for the ARIMA model was using the set of (6, 1, 1), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), and the moving average order (q=1) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.5 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (6, 1, 1).

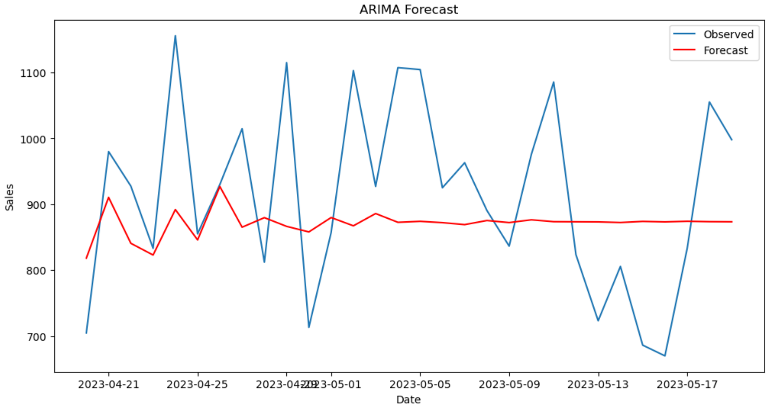


Figure 4. 5 The ARIMA forecasting for snacks & chips products sales for one month.

For cleaning materials products sales, the experiments show the best parameter combination explored for the ARIMA model was using the set of (5, 1, 1), appeared as the most promising, and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 5), the order of differencing (d=1), and the moving average order (q=1) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.6 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (5, 1, 1).

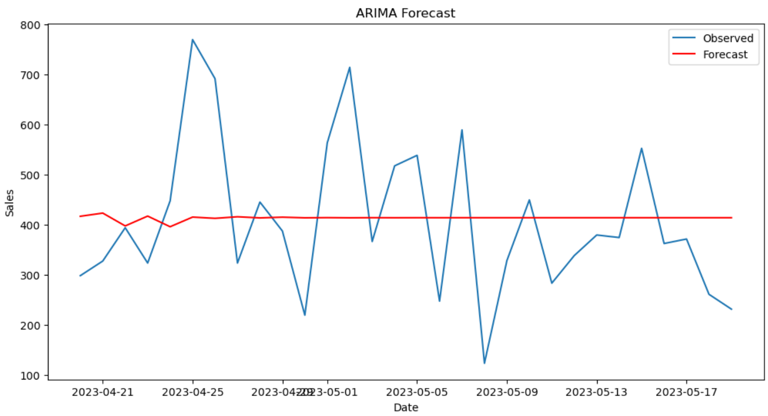


Figure 4. 6 The ARIMA forecasting for cleaning materials products sales for month.

The table below summarizes the best results of under the two scenarios across five categories, as combined sales and for each category sales (individually) based on each error metric for ARIMA model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.8 shows best error metrics of ARIMA model for products combined sales and for each category sales (individually).

Table 4. 8 The error metrics results of ARIMA model for products combined sales and for each category sales (individually).

| | | Error Metrics | | |
|---|---|---|---|---|
| | Parameters | MSE | RMSE | MAE |
| Products Combined Sales | (6,1,1) | 116571.25 | 341.42 | 268.65 |
| Dairies Products Sales | (6,1,6) | 5560.13 | 74.56 | 60.18 |
| Drinks Products Sales | (6,1,1) | 26370.50 | 162.38 | 122.16 |
| Ice-cream Products Sales | (6,1,1) | 1814.51 | 42.59 | 27.09 |
| Snacks & Chips Products Sales | (6,1,1) | 19856.26 | 140.91 | 114.93 |
| Cleaning Materials Products Sales | (5,1,1) | 22803.75 | 151.00 | 122.28 |

#### 4.3.1.2 Seasonal Autoregressive Integrated Moving Average (SARIMA)

The models experiments were applied under two scenarios, where the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials). A systemic trial and error approach was employed to identify sets of parameters which includes the autoregressive order (p), the differencing order (d), ,the moving average order (q) and seasonal parameters for SARIMA model (P,D,Q, and m). The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best parameter combination explored for the SARIMA model was using the set of (6, 1, 1) (1, 1, 0, 12), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), the moving average order (q=1), and seasonal parameters for SARIMA model (P=1, D=1, Q=0, m=12) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.7 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (6, 1, 1) (1, 1, 0, 12).

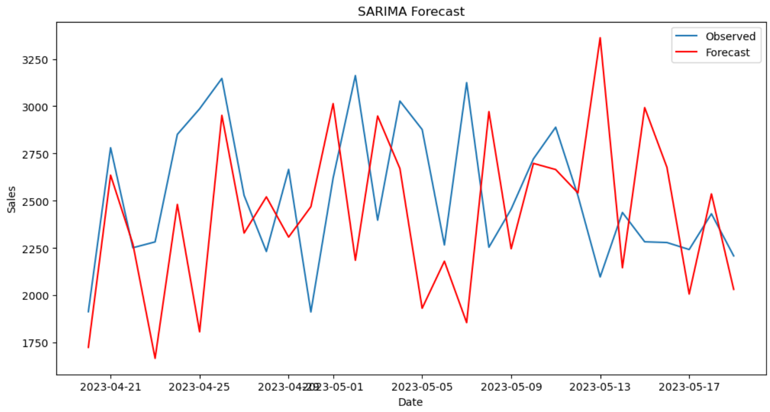


Figure 4. 7 The SARIMA forecasting for products combined sales for one month.

For dairies products sales, the experiments show the best parameter combination explored for the SARIMA model was using the set of (6, 1, 6) (1, 1, 0, 12), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 6), the order of differencing (d=1), the moving average order (q=6), and seasonal parameters for SARIMA model (P=1, D=1, Q=0, m=12) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.8 shows dairies products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (6, 1, 6) (1, 1, 0, 12).

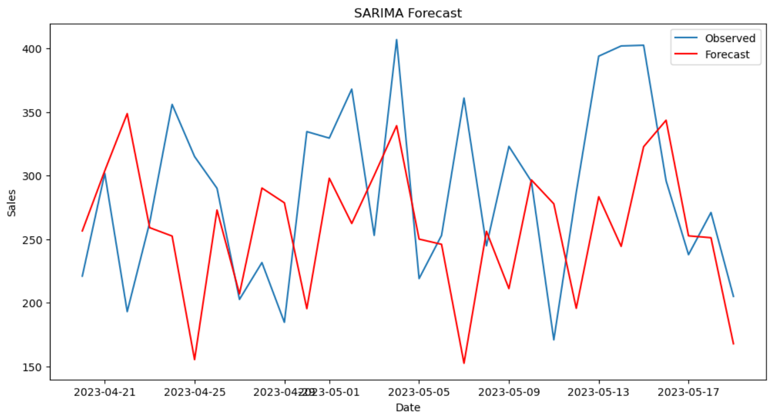


Figure 4. 8 The SARIMA forecasting for dairies products sales for one month.

For drinks products sales, the experiments show the best parameter combination explored for the SARIMA model was using the set of (7, 1, 7) (1, 1, 0, 12), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 7), the order of differencing (d=1), the moving average order (q=7), and seasonal parameters for SARIMA model (P=1, D=1, Q=0, m=12) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.9 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (7, 1, 7) (1, 1, 0, 12).

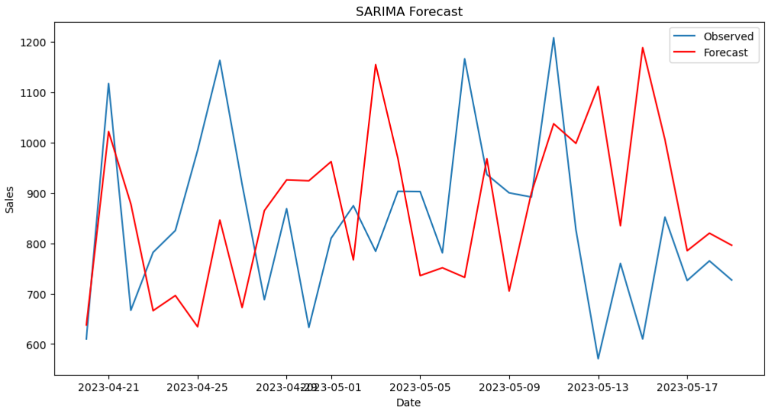


Figure 4. 9 The SARIMA forecasting for drinks products sales for one month.

For ice -cream products sales, the experiments show the best parameter combination explored for the SARIMA model was using the set of (1, 1, 1) (1, 1, 0, 12), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 1), the order of differencing (d=1), the moving average order (q=1), and seasonal parameters for SARIMA model (P=1, D=1, Q=0, m=12) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.10 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (1, 1, 1) (1, 1, 0, 12).

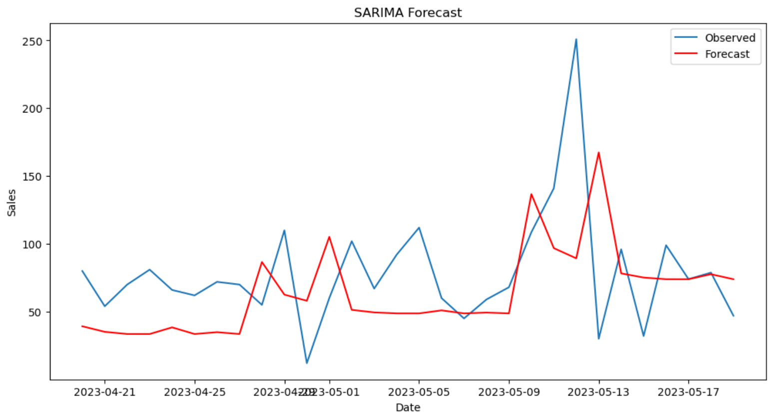

Figure 4. 10 The SARIMA forecasting for ice-cream products sales for one month.

For snacks & chips products sales, the experiments show the best parameter combination explored for the SARIMA model was using the set of (1, 1, 1) (1, 1, 0, 12) appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 1), the order of differencing (d=1), the moving average order (q=1), and seasonal parameters for SARIMA model (P=1, D=1, Q=0, m=12) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.11 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (1, 1, 1) (1, 1, 0, 12).

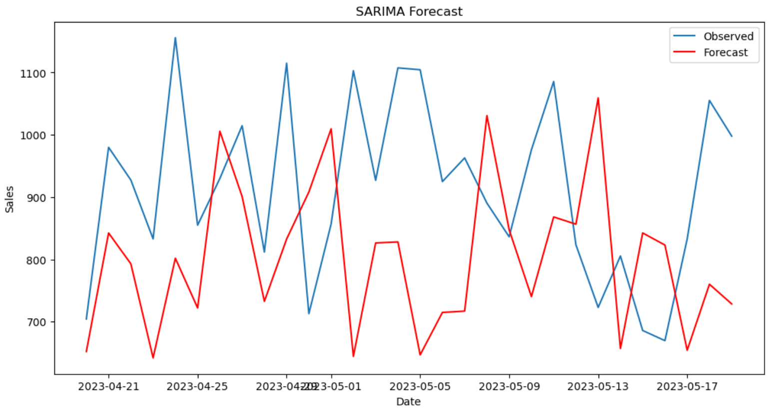


Figure 4. 11 The SARIMA forecasting for snacks & chips products sales for one month.

For cleaning materials products sales, the experiments show the best parameter combination explored for the SARIMA model was using the set of (1, 1, 1) (1, 1, 0, 12), appeared as the most promising and showing the lowest error metrics for the testing set. These parameters, introducing the number of lag observations (p = 1), the order of differencing (d=1), the moving average order (q=1), and seasonal parameters for SARIMA model (P=1, D=1, Q=0, m=12) respectively, collectively assist to the model's ability to capture temporal dependencies and fluctuations in the time series data. Figure 4.12 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) for only one month (20-04-2023) to (20-05-2023) which represent the observed and forecasted behavior for the combination of (1, 1, 1) (1, 1, 0, 12).

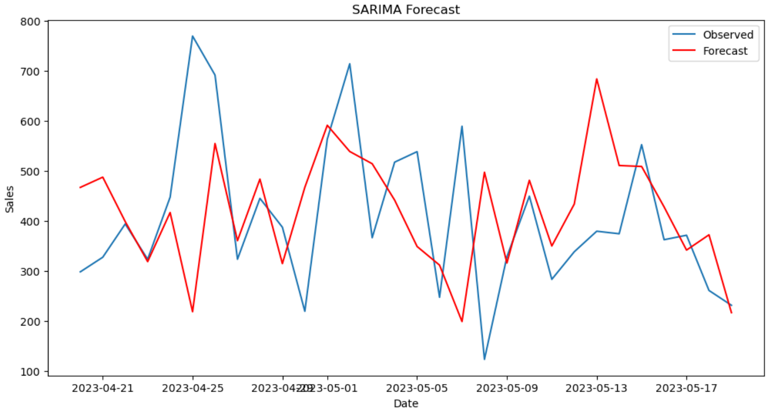


Figure 4. 12 The SARIMA forecasting for cleaning materials products sales for one month.

The table below summarizes the best results under two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for SARIMA model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.9 shows best error metrics of SARIMA model for products combined sales and for each category sales (individually).

Table 4. 9 The error metrics results of SARIMA model for products combined sales and for each category sales (individually).

| | | Error Metrics | | |
|---|---|---|---|---|
| | Parameters | MSE | RMSE | MAE |
| Products Combined Sales | (6,1,1) (1,1,0,12) | 321979.44 | 567.43 | 435.37 |
| Dairies Products Sales | (6,1,6) (1,1,0,12) | 7914.08 | 88.96 | 68.76 |
| Drinks Products Sales | (7,1,7) (1,1,0,12) | 54815.10 | 234.12 | 181.63 |
| Ice-cream Products Sales | (1,1,1) (1,1,0,12) | 2596.91 | 50.95 | 38.18 |
| Snacks & Chips Products Sales | (1,1,1) (1,1,0,12) | 49864.61 | 223.30 | 194.27 |
| Cleaning Materials Products Sales | (1,1,1) (1,1,0,12) | 32784.72 | 181.06 | 126.96 |

#### 4.3.1.3 Summary

Based on error metrics, the ARIMA model surpass the SARIMA model. The ARIMA model has presented lower values across all key error metrics under two scenarios as combined products sales, and based on each product sales (individually), pointing outstanding performance and predictions compared to SARIMA model. The evaluation of models performance detect outstanding advancements when applied based on each products level sales, in contrast to when those products sales were combined with other products sales data. This highlight the importance of analyzing products-specific sales behavior and patterns independently, as such an approach allows models to capture and discover specific patterns and dynamics inherent to each product. The identifiable characteristics and demand fluctuations appeared by individual products sales, focusing the significance of specified analysis for improving forecasting accuracy and marking strategic decision-making.

### 4.3.2 Neural Networks Models

#### 4.3.2.1 Recurrent Neural Networks (RNNs)

The models experiments were applied under two scenarios, the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials). To enhance the architecture of the neural networks model (RNN), a range of experiments were involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function “relu” to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, number of units in Dense layer equal to 1, to specify the value of the output layer prediction (units = 1), the optimizer is “adam”, random seed equal to 42, train-test split ratio equal to 80% for training and 20% for testing. For the model fitting, the sequence length was equal to (10), the epochs equal to (50), batch size to (32). The error metrics were used to evaluate the model performance into how well the model is performing to new data.

For products combined sales, the experiments show the best neurons explored for the RNN model was using (30 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.13 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 30 neurons.

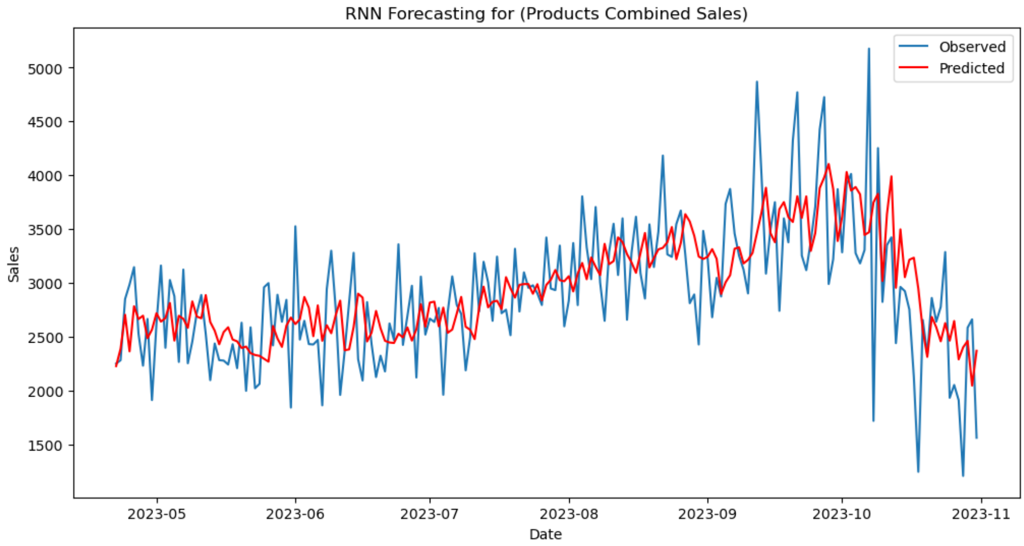


Figure 4. 13 The RNN forecasting for products combined sales using (30 neurons).

The second scenario shows the experiments for each product sales (Individually).For dairies products, the experiments show the best neurons explored for the RNN model was using (20 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.14 shows dairies sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 20 neurons.

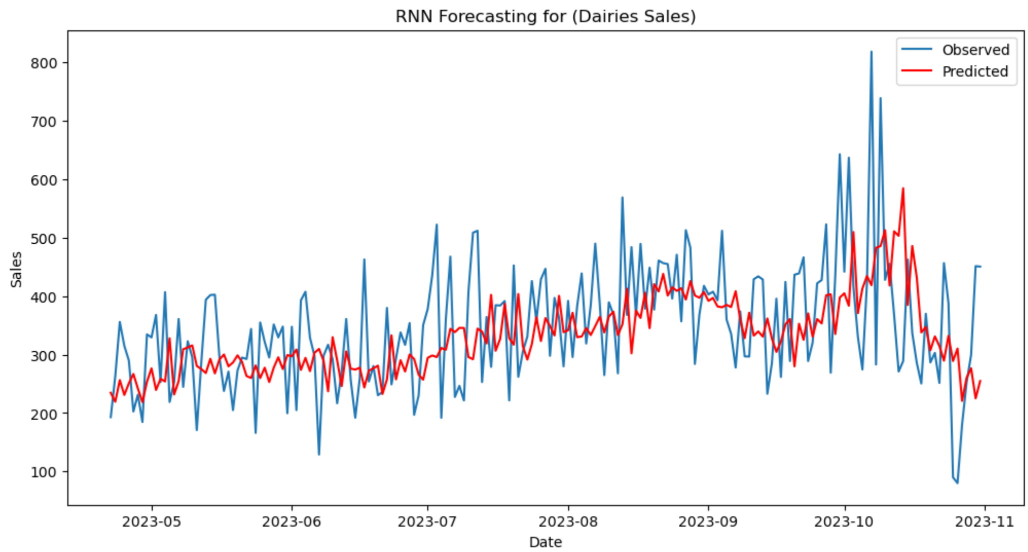


Figure 4. 14 The RNN forecasting for dairies products sales using (20 neurons).

For ice-cream products sales, the experiments show the best neurons explored for the RNN model was using (15 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.15 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 15 neurons.

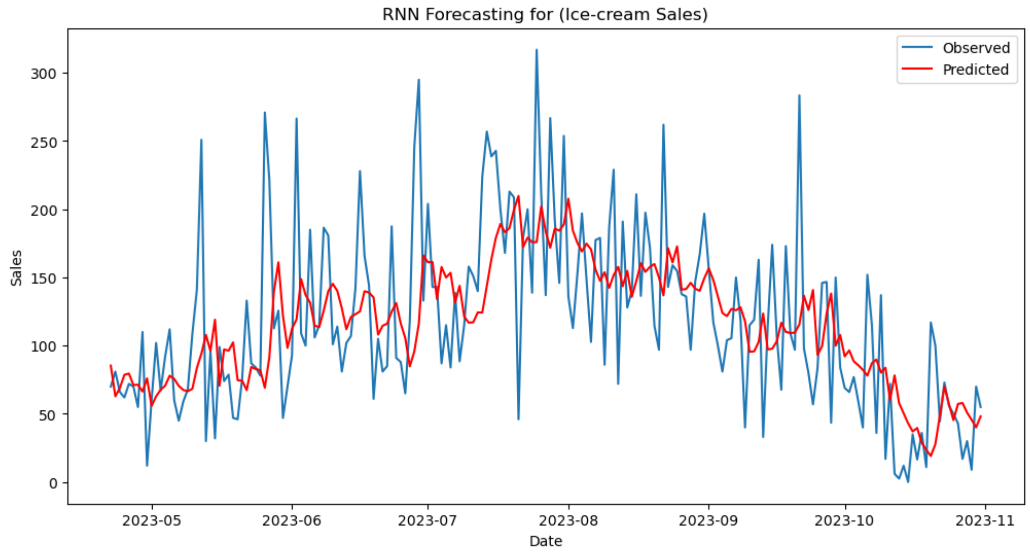


Figure 4. 15 The RNN forecasting for ice-cream products sales using (15 neurons).

For drinks products sales, the experiments show the best neurons explored for the RNN model was using (15 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.16 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 15 neurons.

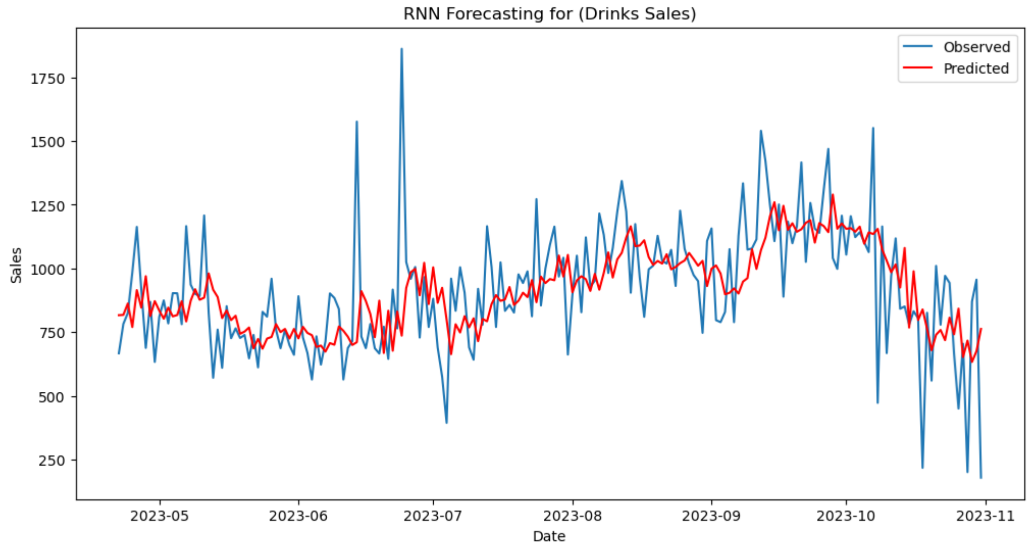


Figure 4. 16 The RNN forecasting for drinks products sales using (15 neurons).

For snacks & chips products sales, the experiments show the best neurons explored for the RNN model was using (20 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.17 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 20 neurons.

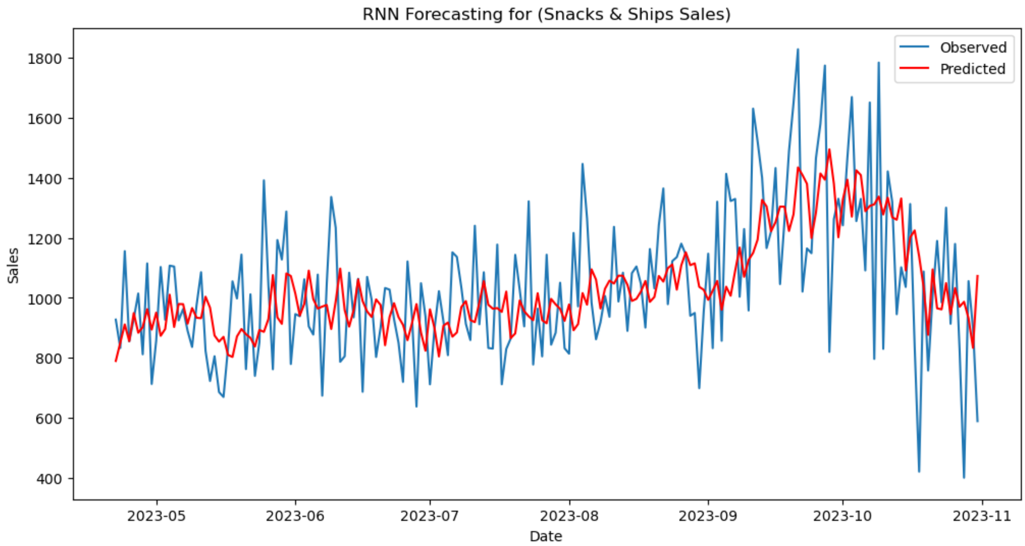


Figure 4. 17 The RNN forecasting for snacks & chips products sales using (20 neurons).

For cleaning materials products sales, the experiments show the best neurons explored for the RNN model was using (15 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.18 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 15 neurons.

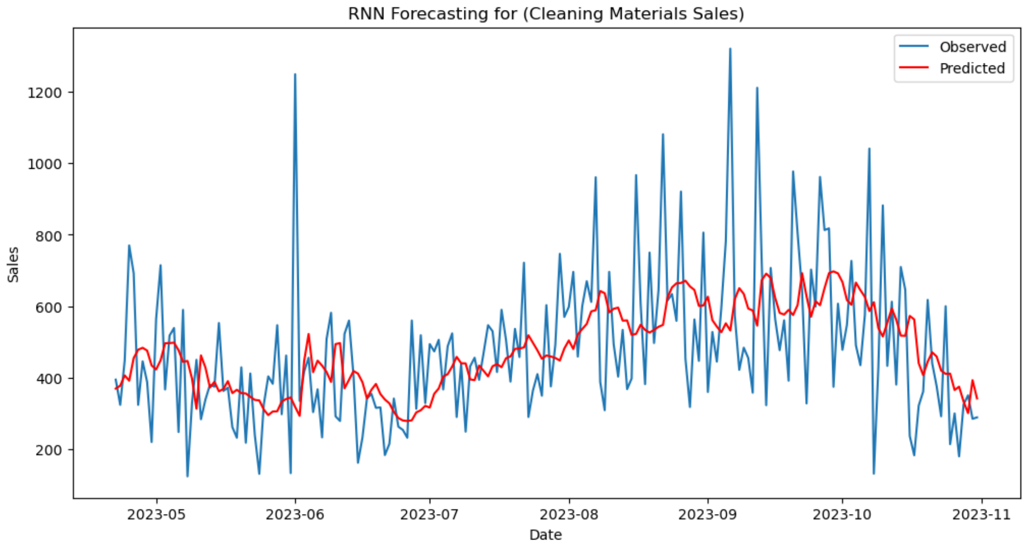


Figure 4. 18 The RNN forecasting for cleaning materials products sales using (15 neurons).

The table below summarizes the best results under two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for RNN model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.10 shows best error metrics of RNN model for products combined sales and for each category sales (individually).

Table 4. 10 The error metrics results of RNN model for products combined sales and for each category sales (individually).

| | | Error Metrics | | |
|---|---|---|---|---|
| | RNN Neurons | MSE | RMSE | MAE |
| Products Combined Sales | 30 | 2559.43 | 505.91 | 390.12 |
| Dairies Products Sales | 20 | 9637.16 | 98.16 | 77.48 |
| Drinks Products Sales | 15 | 4571.72 | 213.80 | 150.58 |
| Ice-cream Products Sales | 15 | 3056.37 | 55.28 | 40.73 |
| Snacks & Chips Products Sales | 20 | 4709.28 | 217.01 | 171.86 |
| Cleaning Materials Products Sales | 15 | 3927.98 | 198.17 | 146.03 |

### 4.3.2.2 Long-Short Term Model (LSTM)

The models experiments were applied under scenarios, the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials). To enhance the architecture of the neural networks model (LSTM), a range of experiments were involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function "relu" to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, number of units in Dense layer equal to 1 to specify the value of the output layer prediction (units = 1), train-test split ratio equal to 80% for training and 20% for testing, and the optimizer "adam". For the model fitting, the sequence length was equal to (10), the epochs equal to (50) and batch size to (32). The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best neurons explored for the LSTM model was using (25 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.19 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 25 neurons.

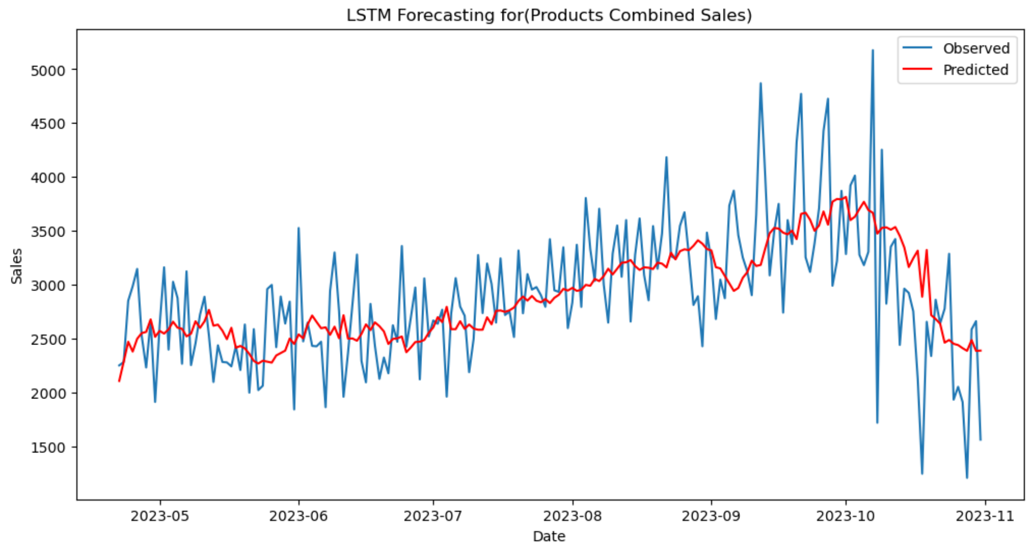


Figure 4. 19 The LSTM forecasting for products combined sales using (25 neurons).

The second scenario shows the experiments for each product sales (Individually).For dairies products, the experiments show the best neurons explored for the LSTM model was using (25 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.20 shows dairies sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 25 neurons.

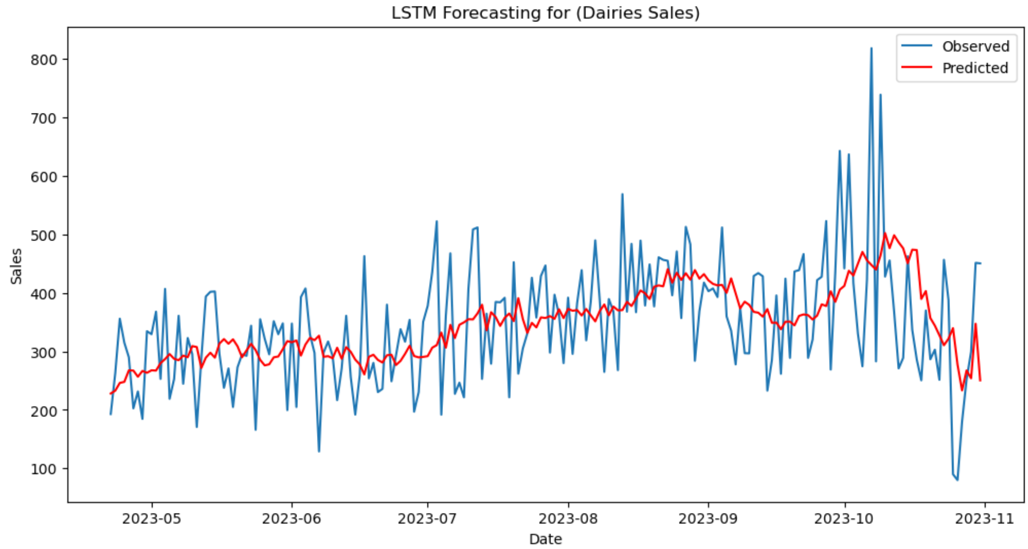


Figure 4. 20 The LSTM forecasting for dairies products sales using (25 neurons).

For ice-cream products sales, the experiments show the best neurons explored for the LSTM model was using (30 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.21 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 30 neurons.

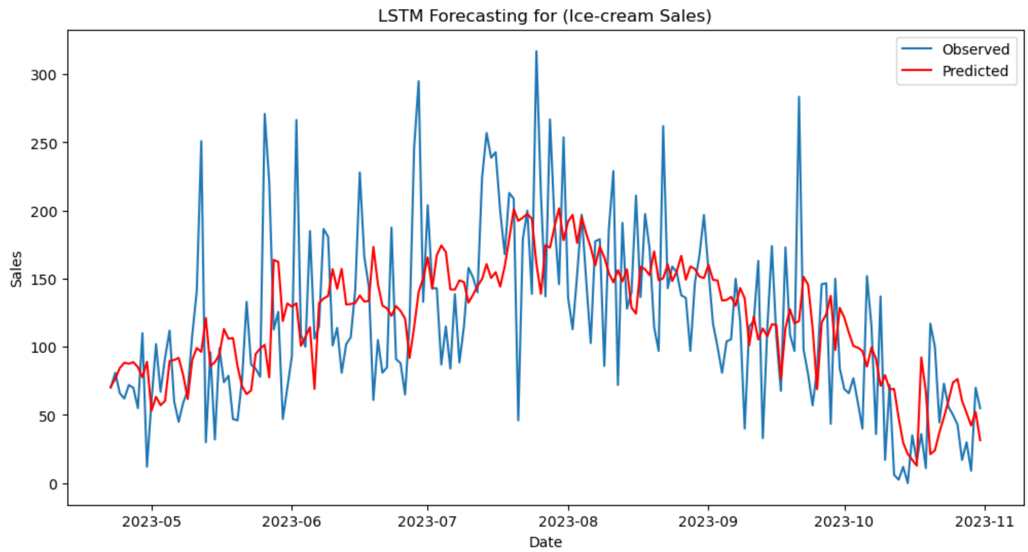


Figure 4. 21 The LSTM forecasting for ice-cream products sales using (30 neurons).

For drinks products sales, the experiments show the best neurons explored for the LSTM model was using (30 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.22 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 30 neurons.

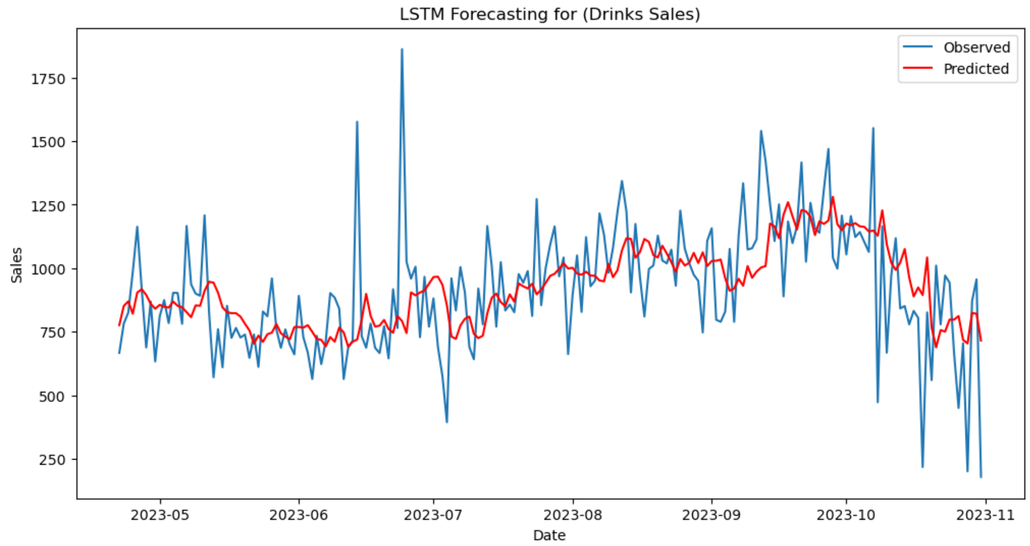


Figure 4. 22 The LSTM forecasting for drinks products sales using (30 neurons).

For snacks & chips products sales, the experiments show the best neurons explored for the LSTM model was using (25 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.23 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 25 neurons.

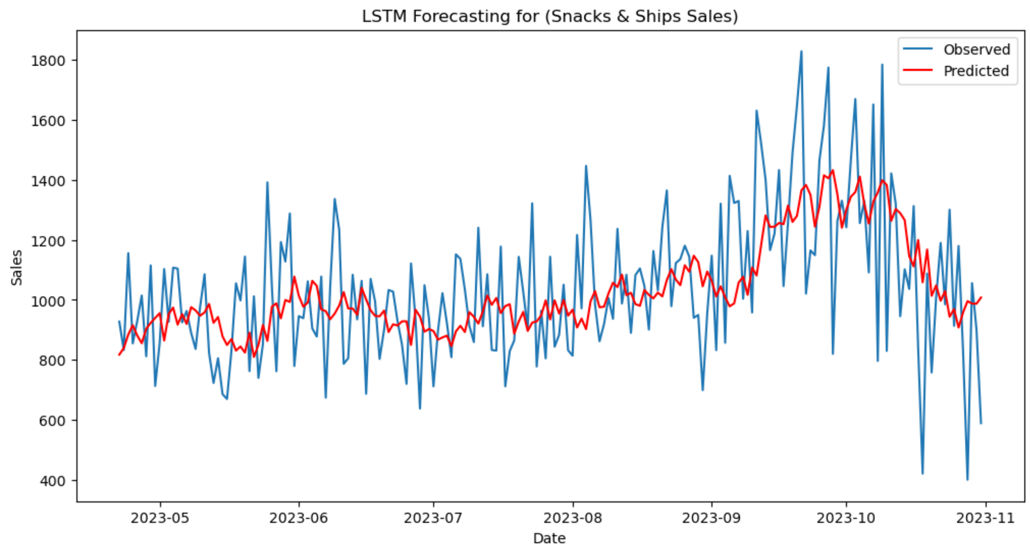

Figure 4. 23 The LSTM forecasting for snacks & chips products sales using (25 neurons).

For cleaning materials products sales, the experiments show the best neurons explored for the LSTM model was using (15 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.24 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 15 neurons.

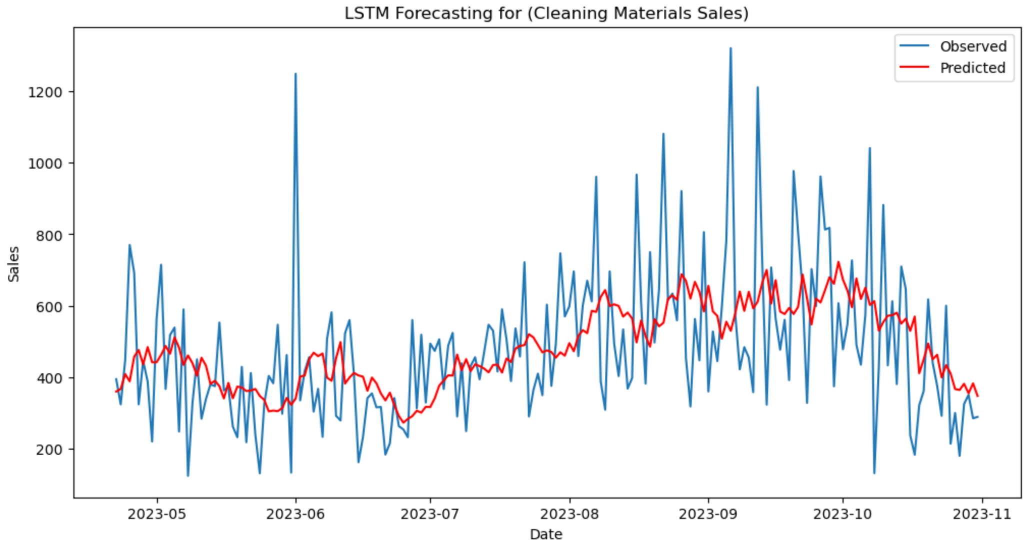


Figure 4. 24 The LSTM forecasting for cleaning materials products sales using (15 neurons).

The table below summarizes the best results under two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for LSTM model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.11 shows best error metrics of LSTM model for products combined sales and for each category sales (individually).

Table 4. 11 The error metrics results of LSTM model for products combined sales and for each category sales (individually).

| | | Error Metrics | | |
|---|---|---|---|---|
| | LSTM Neurons | MSE | RMSE | MAE |
| Products Combined Sales | 25 | 258210.66 | 508.144 | 387.65 |
| Dairies Products Sales | 25 | 9310.37 | 96.49 | 76.25 |
| Drinks Products Sales | 30 | 44571.38 | 211.11 | 149.69 |

| | | | | |
|---|---|---|---|---|
| Ice-cream Products Sales | 30 | 3150.44 | 56.12 | 43.69 |
| Snacks & Chips Products Sales | 25 | 47246.52 | 217.36 | 172.44 |
| Cleaning Materials Products Sales | 15 | 38656.95 | 196.61 | 146.24 |

#### 4.3.2.3 Multilayer Perceptron Neural Networks (MLPNNs)

The models experiments were applied under two scenarios, the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials). To enhance the architecture of multi-layer perceptron neural networks model (MLPNNs), a range of experiments were involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function “relu” to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, and using one hidden layer**,** maximum iterations equal to 1000, learning rate is constant, alpha equal to 0.01 for regularization, random seed equal to 42, train-test split ratio equal to 80% for training and 20% for testing, the optimizer is “adam”, batch size equal to the size of training dataset (batch gradient descent). The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best neurons explored for the MLPNNs model was using (15 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.25 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 15 neurons.

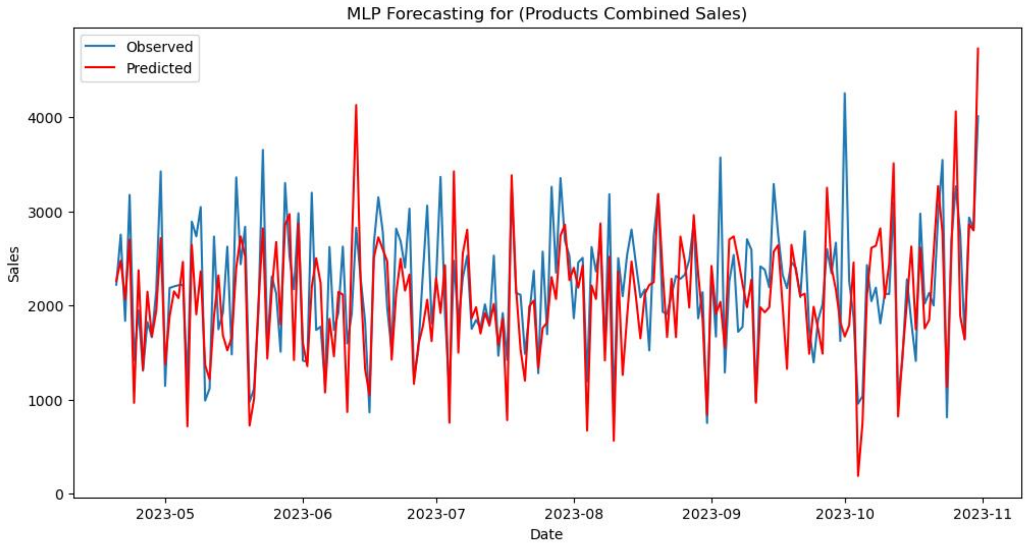


Figure 4. 25 The MLPNNs forecasting for products combined sales using (15 neurons).

The second scenario shows the experiments for each product sales (Individually).For dairies products sales, the experiments show the best neurons explored for the MLPNNs model was using (20 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.26 shows dairies sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 20 neurons.

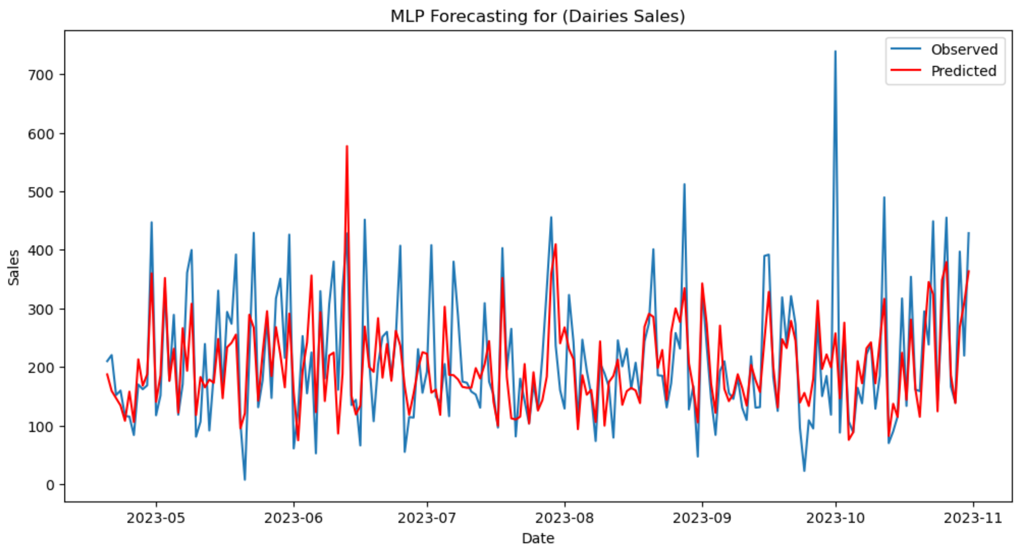


Figure 4. 26 The MLPNNs forecasting for dairies sales using (20 neurons).

For ice-cream products sales, the experiments show the best neurons explored for the MLPNNs model was using (30 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.27 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 30 neurons.

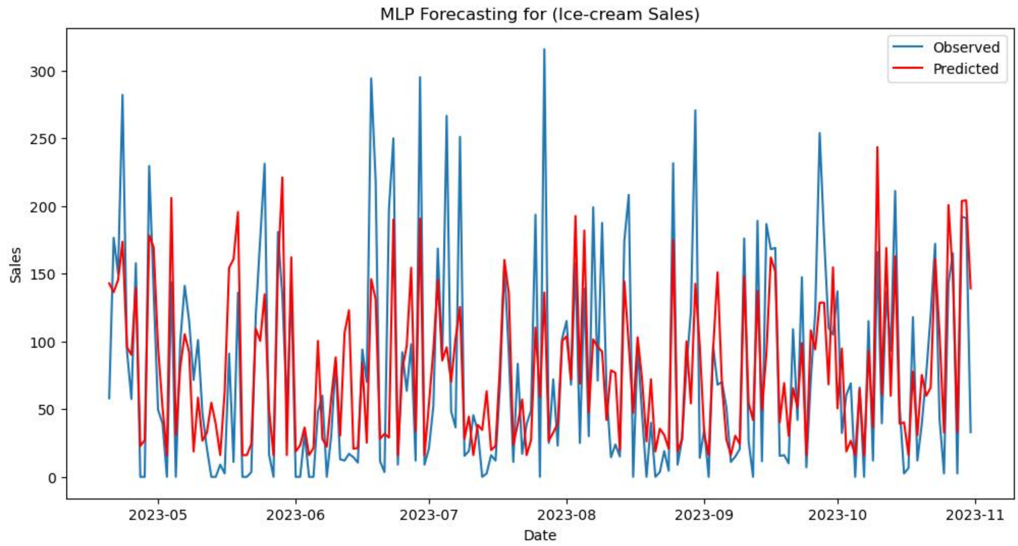


Figure 4. 27 The MLPNNs forecasting for ice-cream sales using (30 neurons).

For drinks products sales, the experiments show the best neurons explored for the MLPNNs model was using (20 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.28 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 20 neurons.

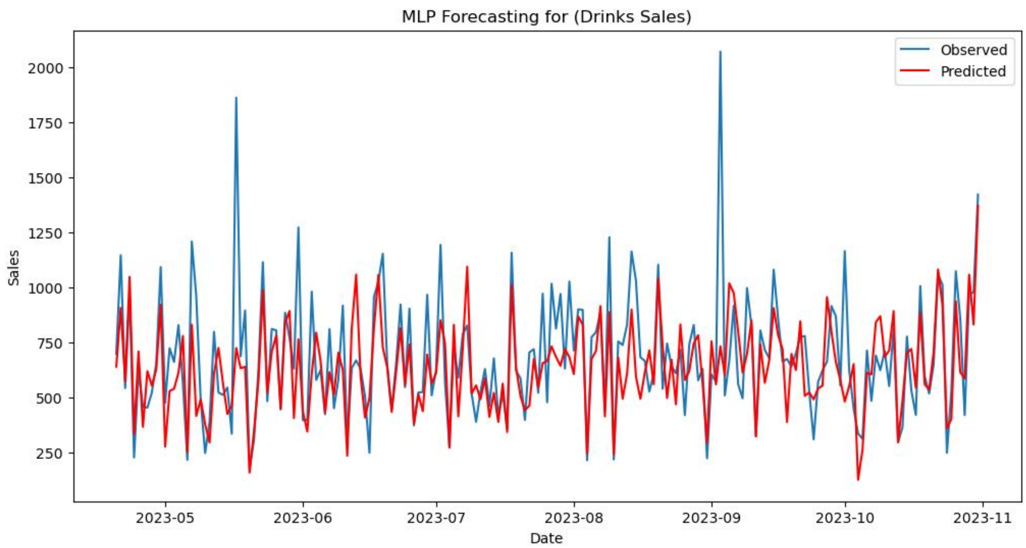


Figure 4. 28 The MLPNNs forecasting for drinks sales using (20 neurons).

For snacks & chips products sales, the experiments show the best neurons explored for the MLPNNs model was using (25 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.29 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 25 neurons.

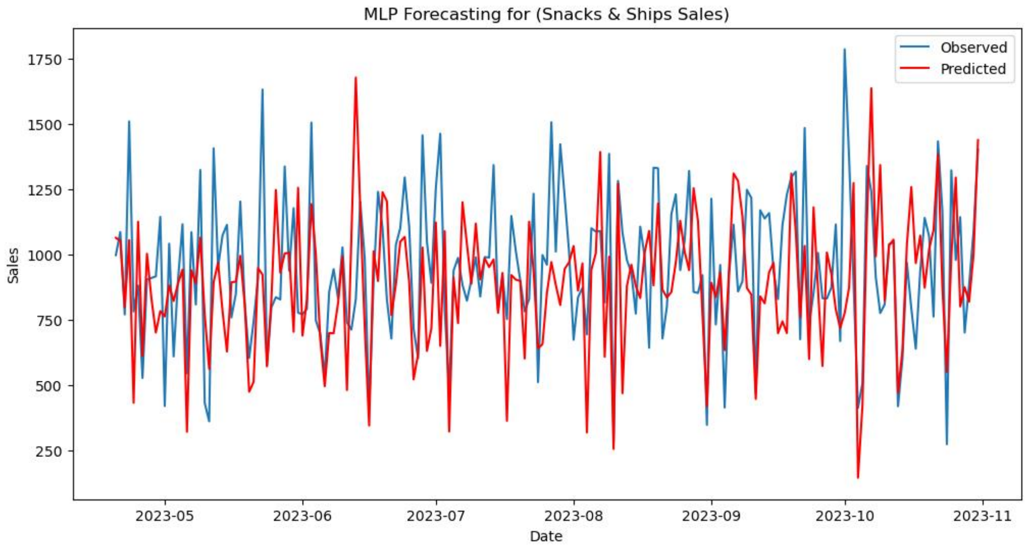


Figure 4. 29 The MLPNNs forecasting for snacks & chips sales using (25 neurons).

For cleaning materials products sales, the experiments show the best neurons explored for the MLPNNs model was using (30 neurons) appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.30 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 30 neurons.

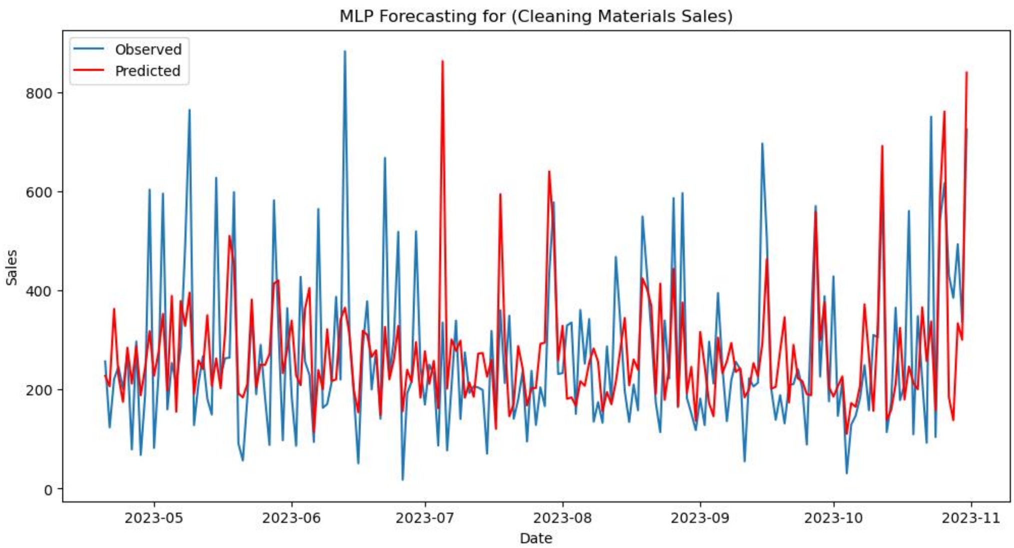

Figure 4. 30 The MLPNNs forecasting for cleaning materials sales using (30 neurons).

The table below summarizes the best results of two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for MLPNNs model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.12 shows best error metrics of MLPNNs model for products combined sales and for each category sales (individually).

Table 4. 12 The error metrics results of MLPNNs model for products combined sales and for each category sales (individually).

| | | Error Metrics | | |
|---|---|---|---|---|
| | MLPNNs Neurons | MSE | RMSE | MAE |
| Products Combined Sales | 15 | 2649.19 | 514.73 | 392.58 |
| Dairies Products Sales | 20 | 6612.22 | 81.31 | 58.29 |
| Drinks Products Sales | 20 | 4500.18 | 212.13 | 140.52 |
| Ice-cream Products Sales | 30 | 2684.08 | 51.80 | 38.59 |
| Snacks & Chips Products Sales | 25 | 7624.03 | 276.12 | 212.80 |
| Cleaning Materials Products Sales | 30 | 1922.36 | 138.66 | 102.07 |

#### 4.3.2.4 Radial Basis Function Neural Network (RBFNNs)

The models experiments were applied under two scenarios, the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials). To enhance the architecture of the neural networks model (RBFNN), a range of experiments was involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function “relu” to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, the input dimension equal to 1 as the number of input feature based on training shape is one which is “date”, the centers of the RBF units are initialized to zero, the beta parameter to control the width of the RBF units was initialized to one, the output dimension equal to one, the RBF is Gaussian function, validation split equal to 0.2, epochs of 50, batch size of 32 and the optimizer “Adam”. The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

The experiments show that the RBFNN model was not able for making forecasting using any of neurons, showing the error metrics for the testing set. Figure 4.31 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 5 neurons- which has provided the same result for the rest of neurons.

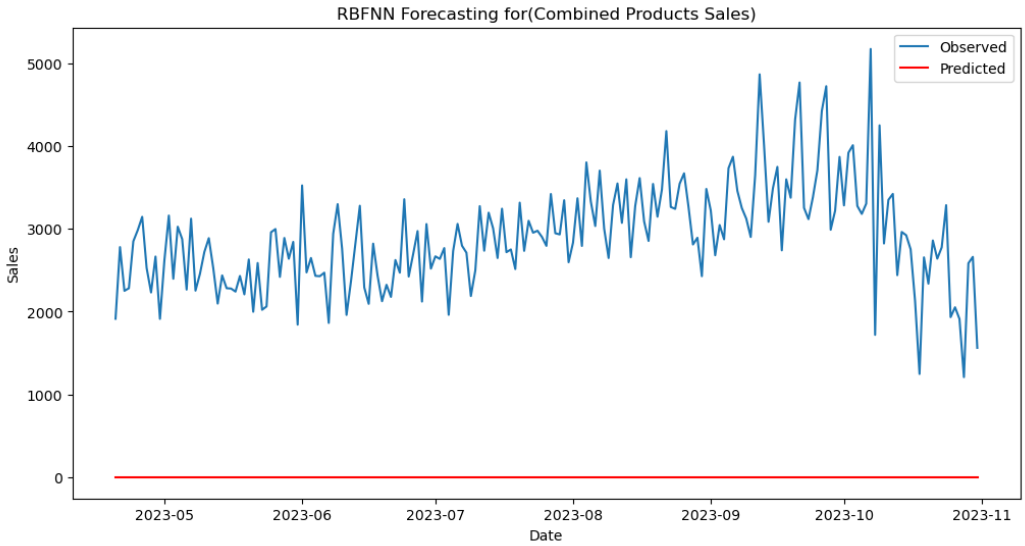


Figure 4. 31 The RBFNN forecasting for products combined sales using (5 neurons).

The second scenario shows the experiments for each product sales (Individually).For dairies products sales, the experiments show that the RBFNN model was not able for making forecasting using any of neurons and showing the error metrics for the testing set. Figure 4.32 shows dairies products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 10 neurons- which has provided the same result for the rest of neurons.

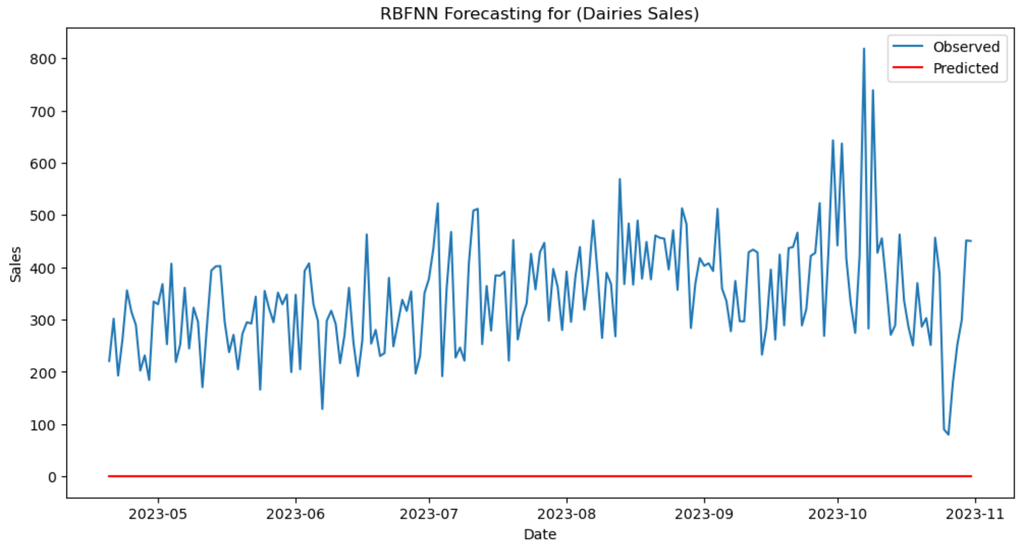


Figure 4. 32 The RBFNN forecasting for dairies sales using (10 neurons).

For ice -cream products sales, the experiment shows that the RBFNN model was not able for making forecasting using any of neurons and showing the error metrics for the testing set. Figure 4.33 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 10 neurons- which has provided the same result for the rest of neurons.

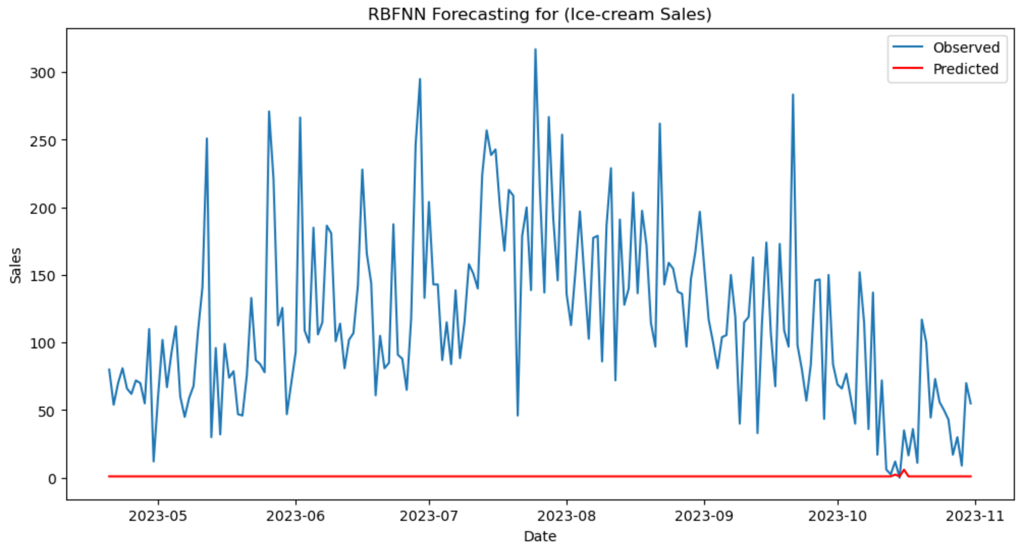


Figure 4. 33 The RBFNN forecasting for ice-cream sales using (10 neurons).

For drinks products sales, the experiments show that the RBFNN model was not able for making forecasting using any of neurons and showing the error metrics for the testing set. Figure 4.34 shows drinks products sales for the testing set starting from (20-04-2023) to (31-

10-2023) which represent the observed and forecasted behavior using 5 neurons- which has provided the same result for the rest of neurons.

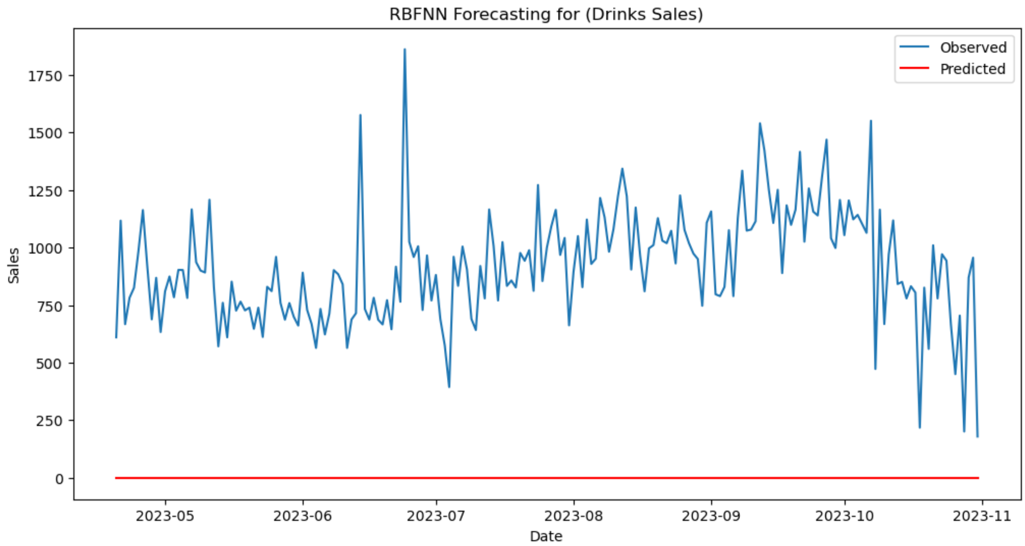


Figure 4. 34 The RBFNN forecasting for drinks sales using (5 neurons).

For snacks & chips products sales, the experiments show that the RBFNN model was not able for making forecasting using any of neurons and showing the error metrics for the testing set. Figure 4.35 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 5 neurons- which has provided the same result for the rest of neurons.

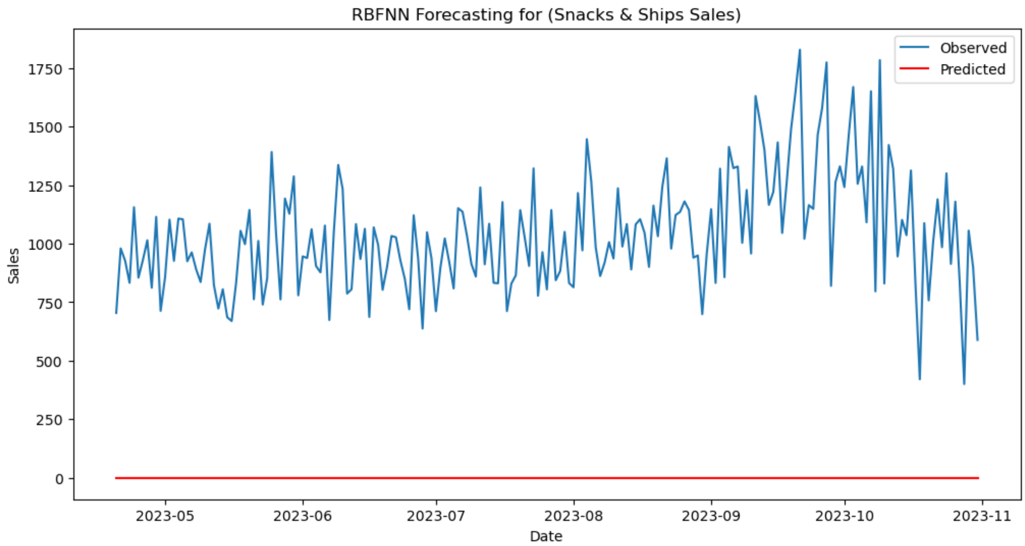


Figure 4. 35 The RBFNN forecasting for snacks & chips sales using (5 neurons).

For cleaning materials products sales, the experiments show that the RBFNN model was not able for making forecasting using any of neurons and showing the error metrics for the testing set. Figure 4.36 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior using 15 neurons- which has provided the same result for the rest of neurons.

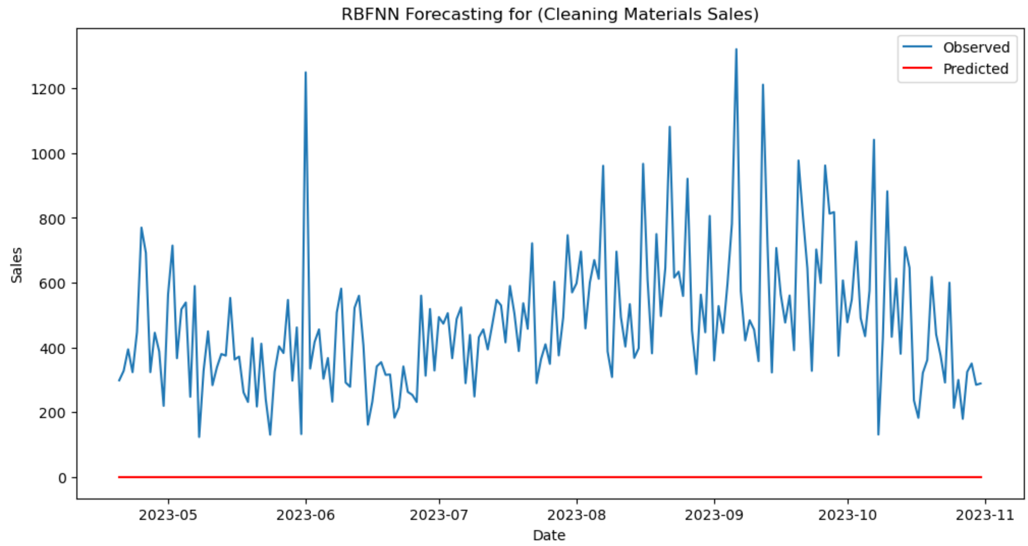


Figure 4. 36 The RBFNN forecasting for cleaning materials sales using (15 neurons).

The table below summarizes the best results of two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for RBFNN model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.13 shows best error metrics of RBFNN model for products combined sales and for each category sales (individually).

Table 4. 13 The error metrics results of RBFNN model for products combined sales and for each category sales (individually).

| | | Error Metrics | | |
|---|---|---|---|---|
| | RBFNN Neurons | MSE | RMSE | MAE |
| Products Combined Sales | 5 | 8808.14 | 296.88 | 289.88 |
| Dairies Products Sales | 10 | 1299.89 | 360.5 | 344.87 |
| Drinks Products Sales | 5 | 9018.48 | 949.67 | 918.05 |
| Ice-cream Products Sales | 10 | 1795.67 | 133.99 | 116.96 |

| | | | | |
|---|---|---|---|---|
| Snacks & Chips Products Sales | 5 | 1139.59 | 1067.66 | 1039.14 |
| Cleaning Materials Products Sales | 15 | 2723.86 | 521.89 | 477.84 |

#### 4.3.2.5 Models Errors Comparisons

For clear comparison of the models forecasting performance, the errors of various models were visualized, showing the differences and accuracy levels of each model for the two scenarios. For the first scenario as combined products sales, the errors of models (LSTM, MLPNNs and RBFNN) compared with the best model performance (RNN), to highlight the discrepancies between the observed and predicted sales values. Figure 4.37, Figure 4.38 and Figure 4.39 shows the error models comparisons between the best models performance (RNN) and other models (LSTM, MLPNNs and RBFNNs) for products combined sales.

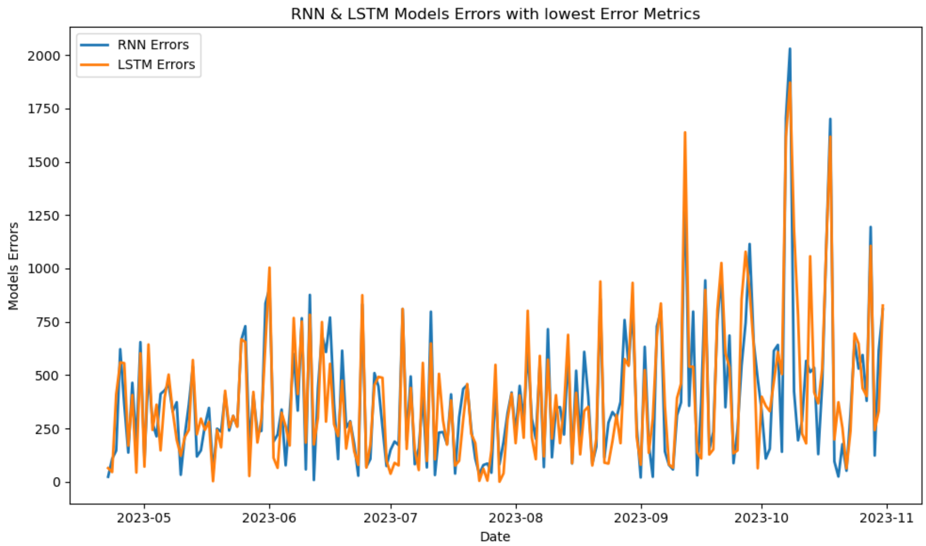


Figure 4. 37 (RNN & LSTM) models errors comparisons with lowest error metrics.

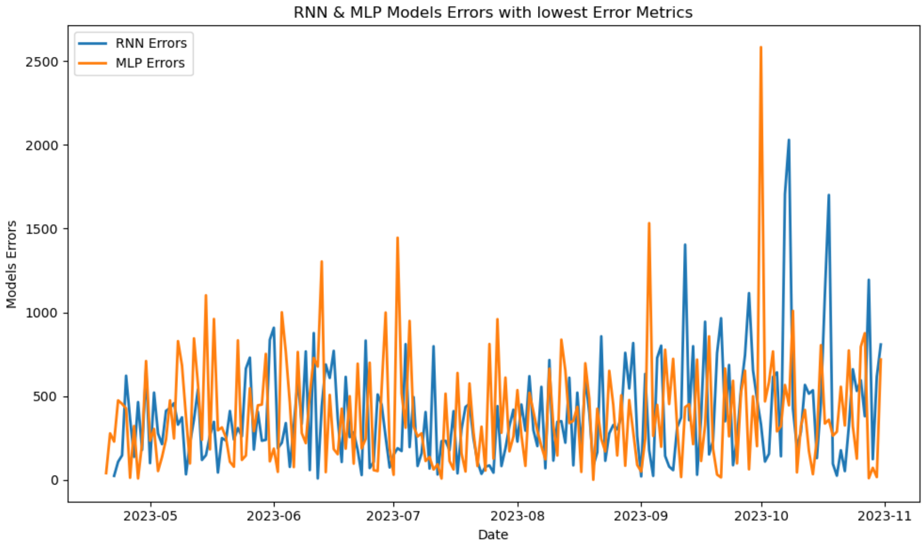


Figure 4. 38 (RNN & MLPNNs) models errors comparisons with lowest error metrics.

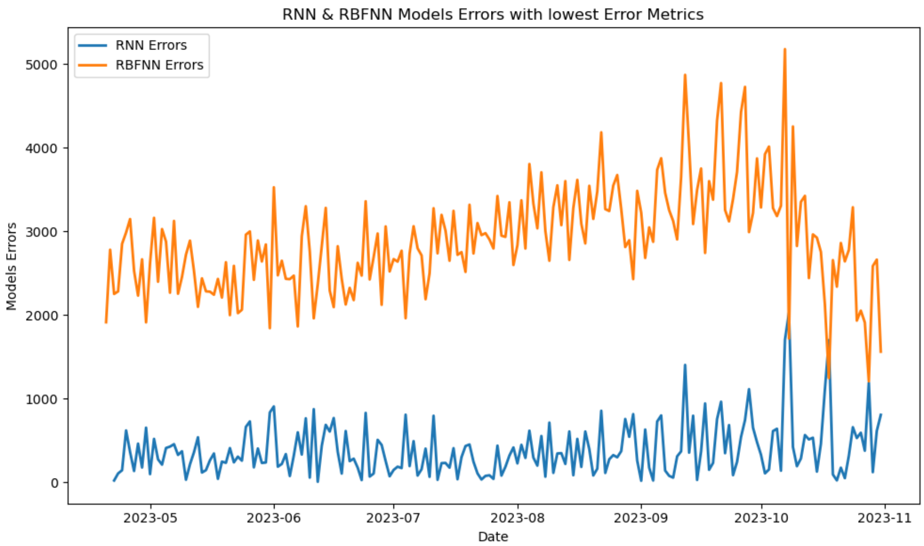


Figure 4. 39 (RNN & RBFNN) models errors comparisons with lowest error metrics.

For each products sales dataset (individually). The models errors visualization for dairies products sales of (LSTM, RNN and RBFNN) models compared with the best model performance (MLPNNs) to highlight the discrepancies between the observed and predicted sales values. Figure 4.40, Figure 4.41 and Figure 4.42 shows the error models comparisons between the best model's performance (MLPNNs) and other models (LSTM, RNN and RBFNN) for dairies products sales.

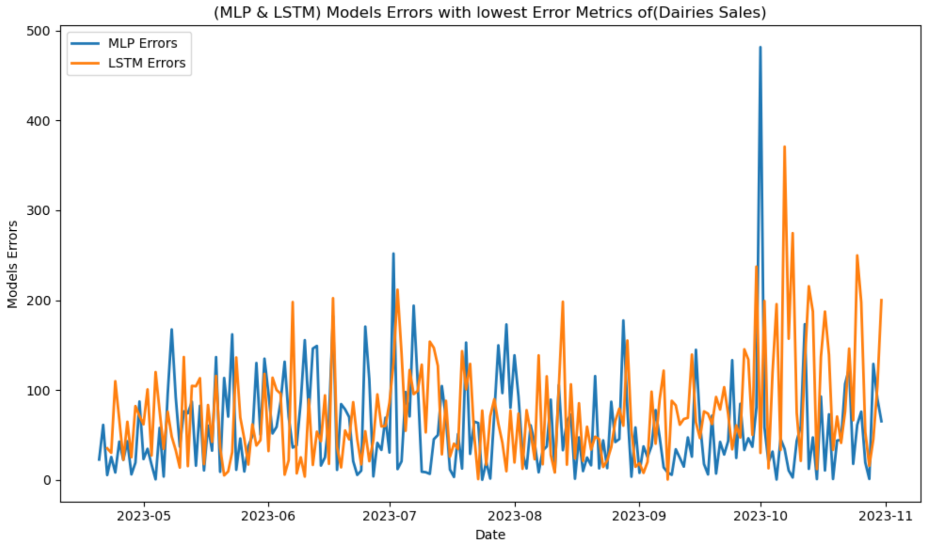


Figure 4. 40 (MLPNNs & LSTM) models errors comparisons with lowest error metrics.

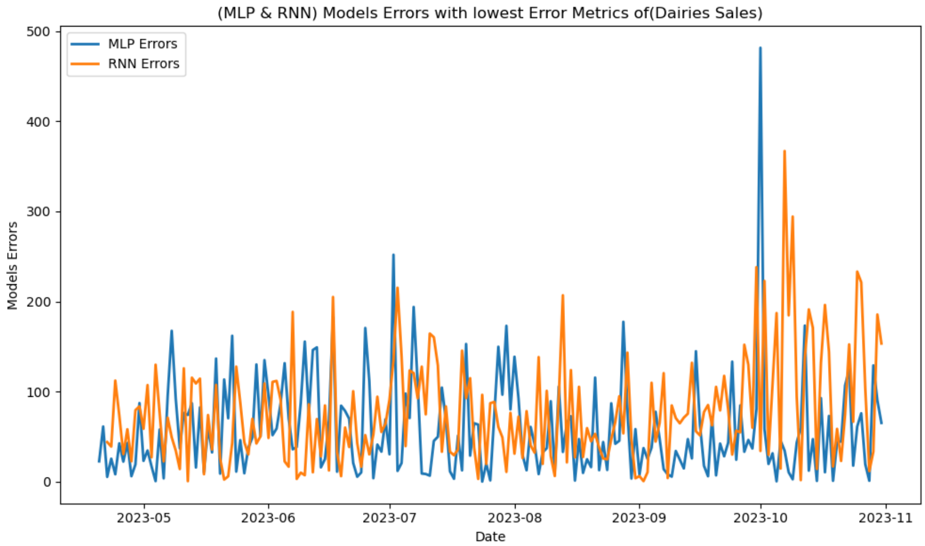


Figure 4. 41 (MLPNNs & RNN) models errors comparisons with lowest error metrics.

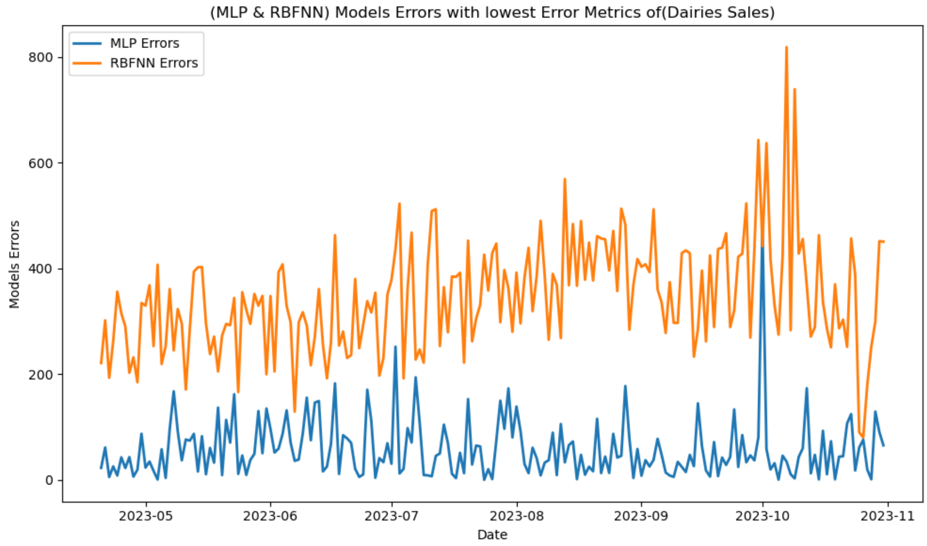


Figure 4. 42 (MLPNNs & RBFNN) models errors comparisons with lowest error metrics.

For ice-cream products sales clarification for the models performance, a visualization of the models errors of (RNN, LSTM and RBFNN) compared with the best model performance (MLPNNs) to highlight the discrepancies between the observed and predicted sales values. Figure 4.43, Figure 4.44 and Figure 4.45 shows the error models comparisons between the best models performance (MLPNNs) and other models (RNN, LSTM and RBFNN) for ice-cream products sales.

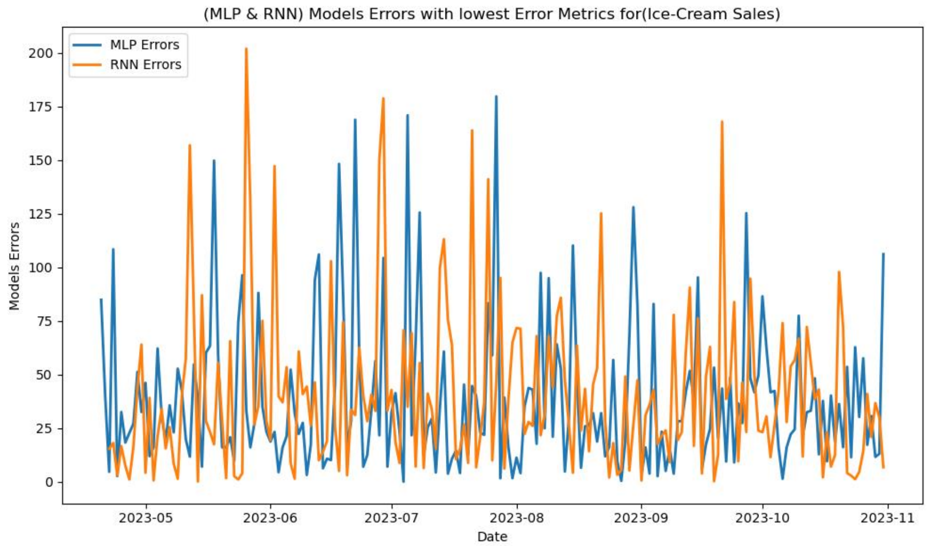


Figure 4. 43 (MLPNNs & RNN) models errors comparisons with lowest error metrics.

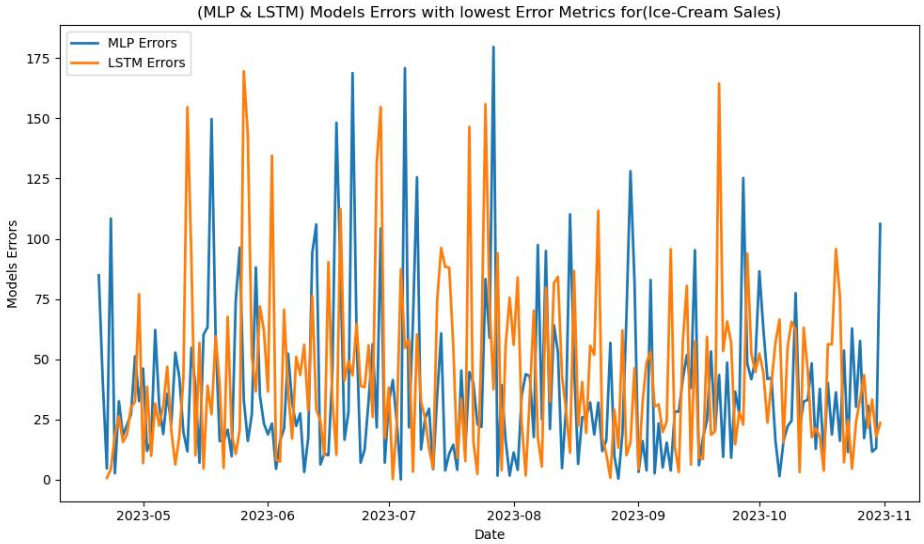


Figure 4. 44 (MLPNNs & LSTM) models errors comparisons with lowest error metrics.

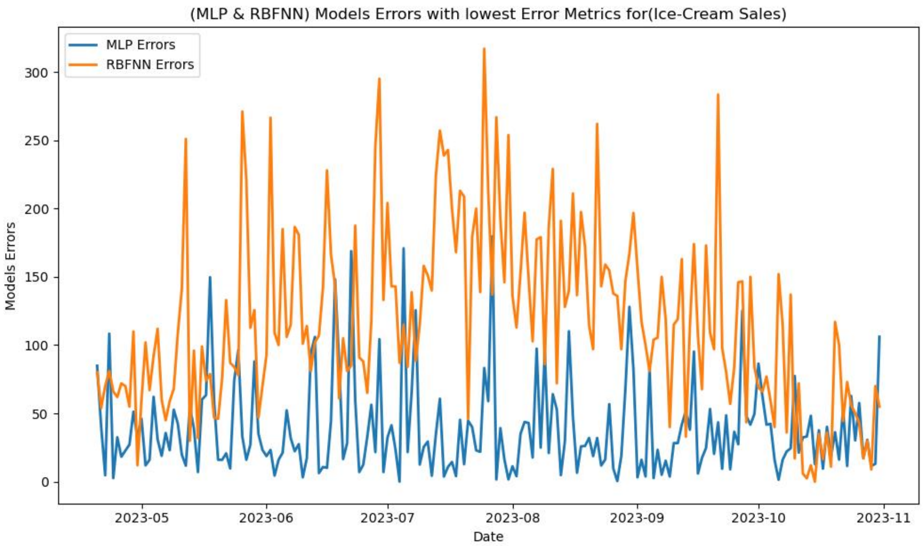


Figure 4. 45 (MLPNNs & RBFNN) models errors comparisons with lowest error metrics.

For drinks products sales clarification for the models performance, a visualization of the models errors of (RNN, MLPNNs and RBFNN) compared with the best model performance (LSTM), to highlight the discrepancies between the observed and predicted sales values. Figure 4.46, Figure 4.47 and Figure 4.48 shows the error models comparisons between the best models performance (LSTM) and other models (RNN, MLPNNs and RBFNN) for drinks products sales.

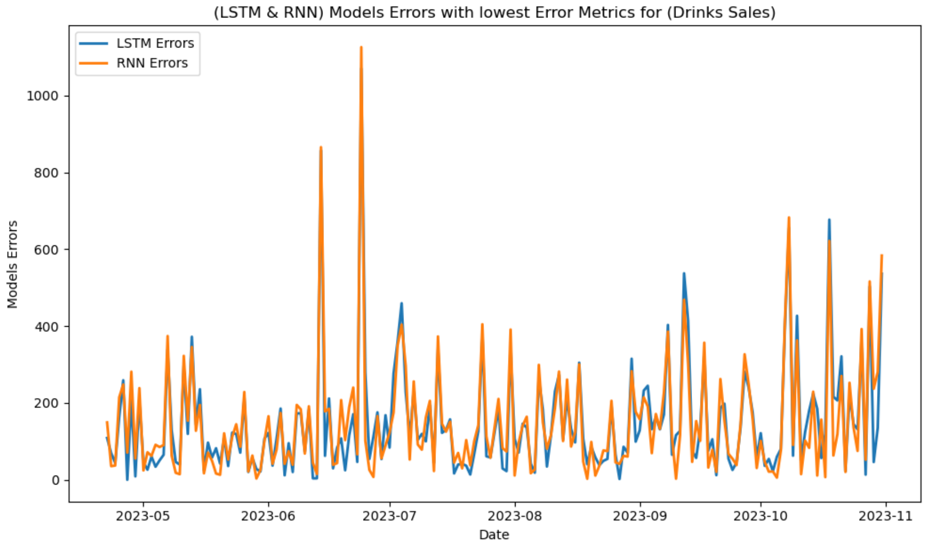


Figure 4. 46 (LSTM & RNN) models errors comparisons with lowest error metrics.

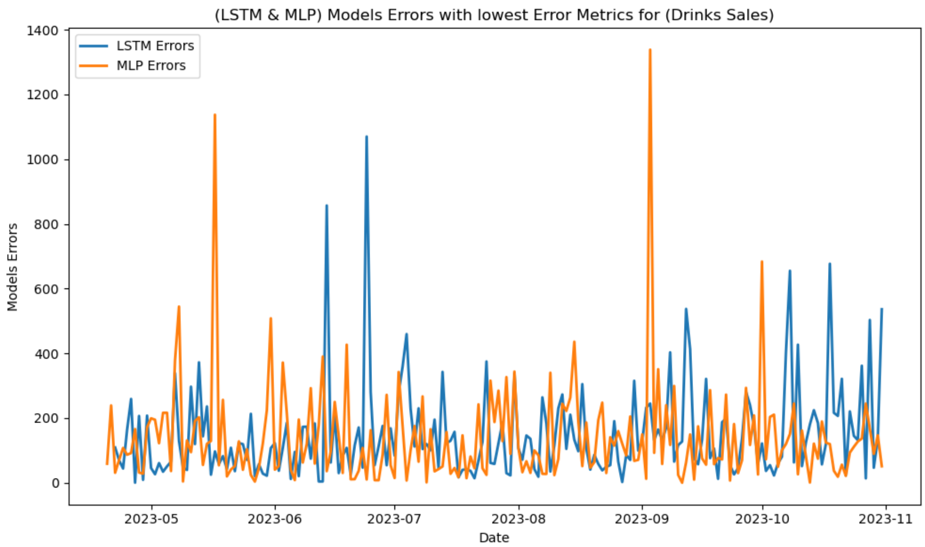


Figure 4. 47 (LSTM & MLPNNs) models errors comparisons with lowest error metrics.

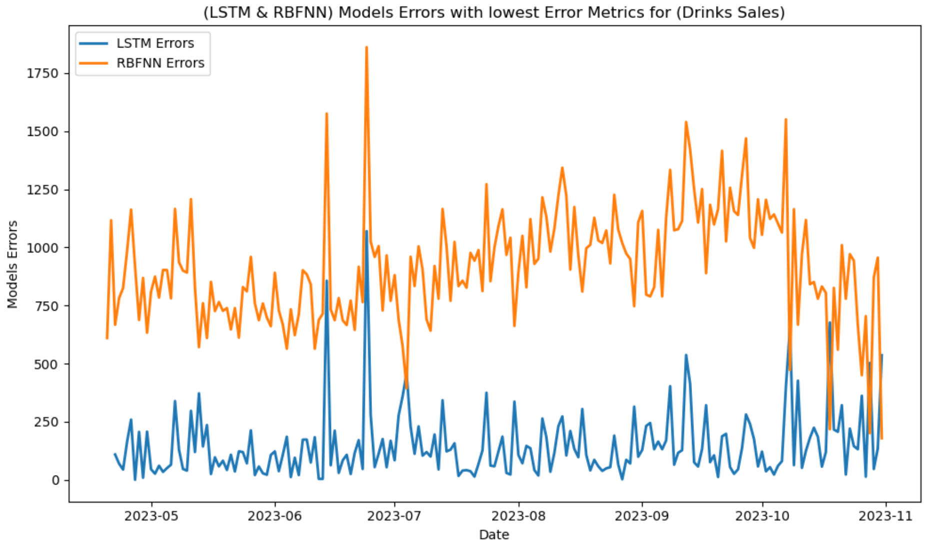


Figure 4. 48 (LSTM & RBFNN) models errors comparisons with lowest error metrics**.**

For snacks & chips products sales clarification for the models performance, a visualization of the models errors of (LSTM, MLPNNs and RBFNN) compared with the best model performance (RNN), to highlight the discrepancies between the observed and predicted sales values. Figure 4.49, Figure 4.50 and Figure 4.51 shows the error models comparisons between the best models performance (RNN) and other models (LSTM, MLPNNs and RBFNN) for snacks & chips products sales.

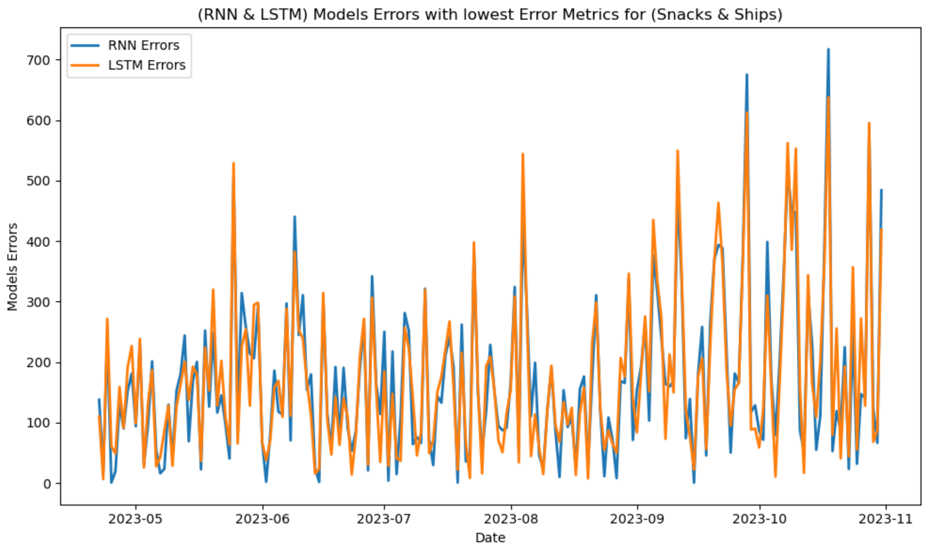


Figure 4. 49 (RNN & LSTM) models errors comparisons with lowest error metrics.

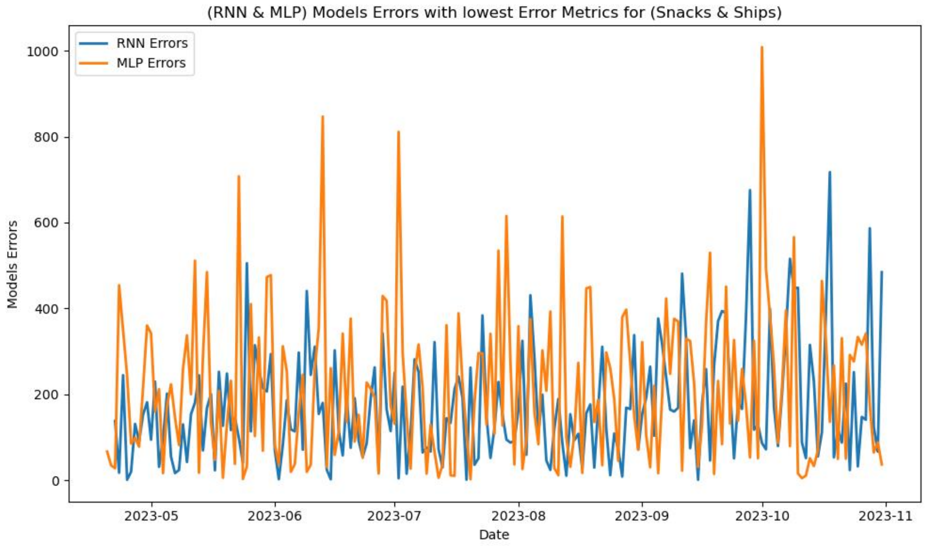


Figure 4. 50 (RNN & MLPNNs) models errors comparisons with lowest error metrics.

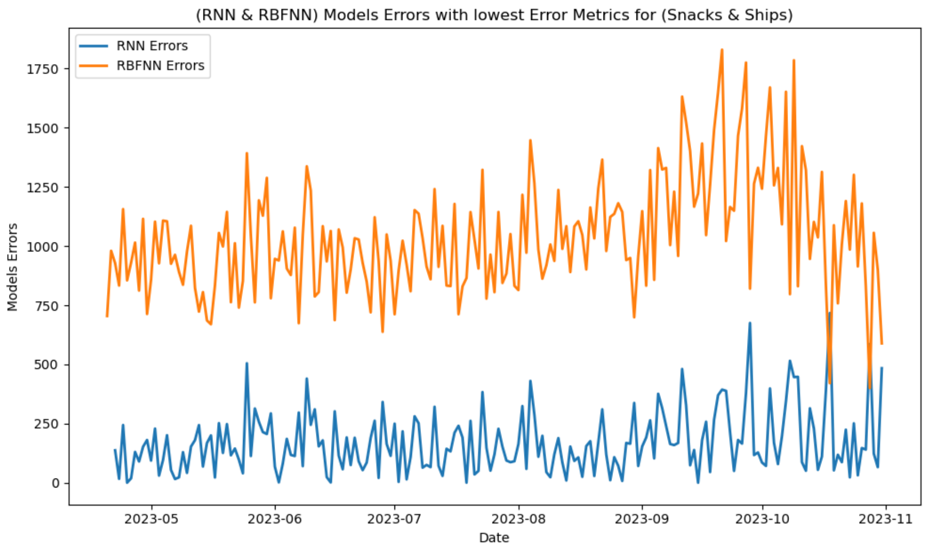


Figure 4. 51 (RNN & RBFNN) models errors comparisons with lowest error metrics.

For cleaning materials products sales clarification for the models performance, a visualization of the models errors of (RNN, LSTM and RBFNN) compared with the best model performance (MLPNNs), to highlight the discrepancies between the observed and predicted sales values. Figure 4.52, Figure 4.53 and Figure 4.54 shows the error models comparisons between the best models performance (MLPNNs) and other models (LSTM, RNN and RBFNN) for cleaning materials products sales.

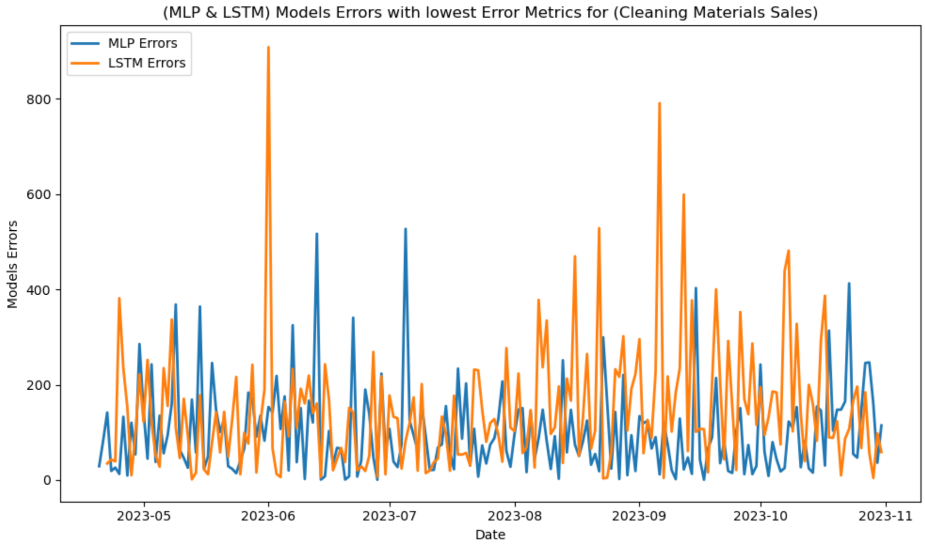


Figure 4. 52 (MLPNNs & LSTM) models errors comparisons with lowest error metrics.

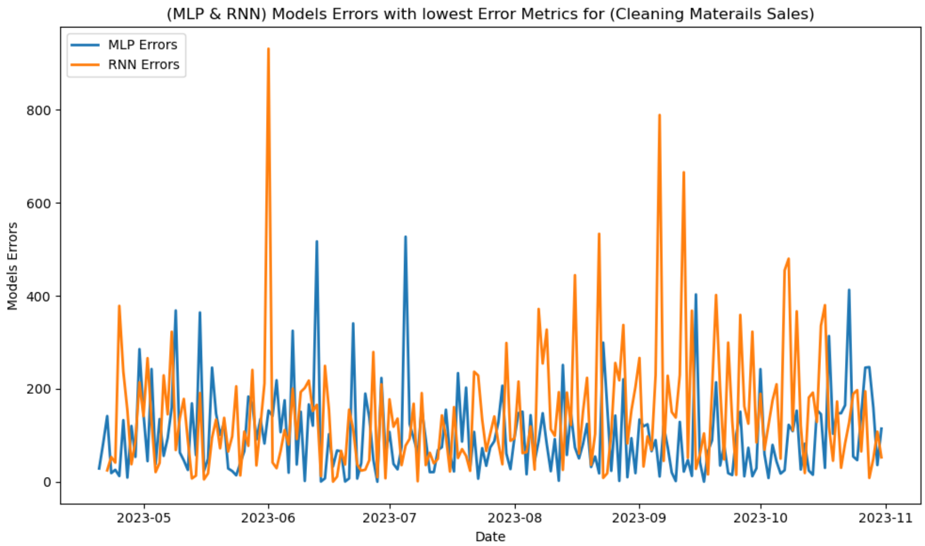


Figure 4. 53 (MLPNNs & RNN) models errors comparisons with lowest error metrics.

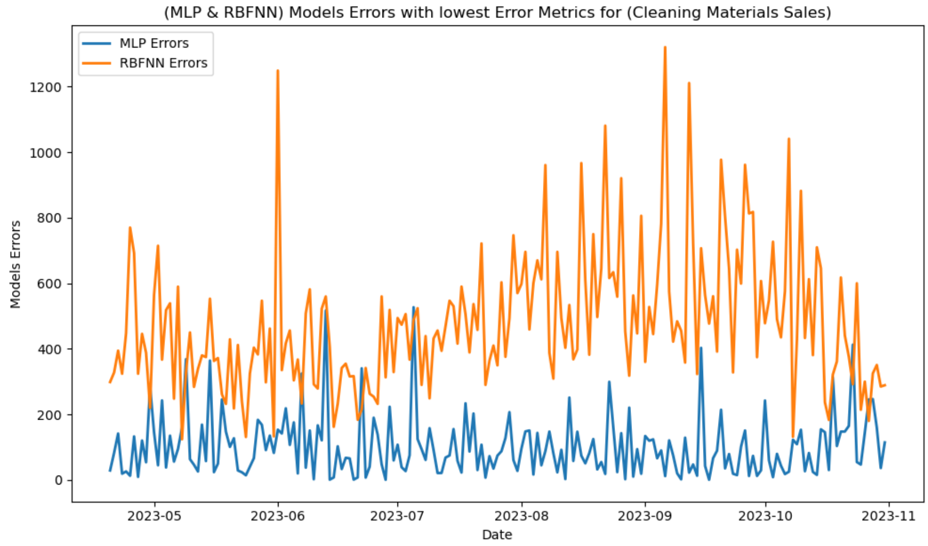


Figure 4. 54 (MLPNNs & RBFNN) models errors comparisons with lowest error metrics.

#### 4.3.2.6 Summary

After applying set of experiments using various forecasting models such as RNN, LSTM, MLPNNs and RBFNN for products combined sales dataset and for each product sales (individually), models performance was evaluated using error metrics including (MSE, RMSE and MAE). For more inclusive comparison of the applied models performance, the following table summarizes the best results gained for each model type. The table includes the model type, number of neurons and key performance metrics. This summary allows for easy comparison. Table 4.2 shows best models under two scenarios based on error metrics.

Table 4. 14 The best model for each scenario based on error metrics

| | | | Error Metrics | | |
|---|---|---|---|---|---|
| | The Model | Neurons | MSE | RMSE | MAE |
| Products Combined Sales | RNN Model | 30 | 2559.43 | 505.91 | 390.12 |
| Snacks & Chips Products Sales | RNN Model | 20 | 4709.28 | 217.01 | 171.86 |
| Dairies Products Sales | MLPNNs Model | 20 | 6612.22 | 81.31 | 58.29 |
| Ice-cream Products Sales | MLPNNs Model | 30 | 2684.08 | 51.80 | 38.59 |
| Cleaning Materials Products Sales | MLPNNs Model | 30 | 1922.36 | 138.66 | 102.07 |
| Drinks Products Sales | LSTM Model | 30 | 4457.38 | 211.11 | 149.69 |

For products combined sales, the results have pointed out that the RNN model using (30 neurons), has achieved the lowest error metrics among all models which provided best forecasting accuracy. It was followed by the LSTM model, achieving a good and close result using (25 neurons).

For each product sales dataset (individually), starting with dairies products sales, results have pointed out that the MLPNNs model using (20 neurons) has achieved the lowest error metrics among other models which provided best forecasting accuracy.it was followed with the LSTM and RNN models, achieving a good and close results using (25 for LSTM) and (20 neurons for RNN).

For ice-cream products sales, results have pointed out that the MLPNNs model using (30 neurons), have achieved the lowest error metrics among all models which–provided best forecasting accuracy.it was followed with the RNN and LSTM models, achieving a very close results using (15 neurons for RNN) and (30 neurons for LSTM).

For drinks products sales, results have pointed out that the three models have achieved a very close results for error metrics values, where the LSTM model error metrics using (30 neurons) followed with MLPNNs model error metrics with very slightly difference using (30 neurons) and the RNN model using (15 neurons).

For snacks & chips products sales, results have pointed out that RNN and LSTM models have achieved a very close results for error metrics values, the RNN model error metrics using (20 neurons) followed with LSTM model error metrics with very slightly difference using (25 neurons).

For cleaning materials products sales, results have pointed out that the MLPNNs model using (30 neurons) has achieved the lowest error metrics among all models which provided best forecasting accuracy, followed with the LSTM and RNN models, achieving close results using (15 neurons) and using (15 neurons) respectively.

### 4.3.3 Hybrid Models

The main goal of employing hybrid models combining both statistical methods of ARIMA and neural networks approaches was to reinforce forecasting accuracy by combining the strengths of each approach. By combining the best performing ARIMA model with different configurations of neural networks, an attempts to take advantage of the strong statistical foundation of ARIMA while employing the flexibility and adaptability of neural networks (RNN, LSTM, MLPNNs and RBFNN) to capture complex patterns in the data.

#### 4.3.3.1 ARIMA-RNN Models

The models experiments were applied under two scenarios, the first scenario for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials).The best statistical model performance (ARIMA) was selected. To enhance the architecture of the neural networks model (RNN), a range of experiments were involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function “relu” to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, and using the optimizer “adam”. For the model fitting, the sequence length was equal to (10), the epochs equal to (50) and batch size to (32).The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best results explored for the hybrid models of (ARIMA & RNN) model were using parameters of ARIMA (6,1,1) with (20 neurons) for RNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.55 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

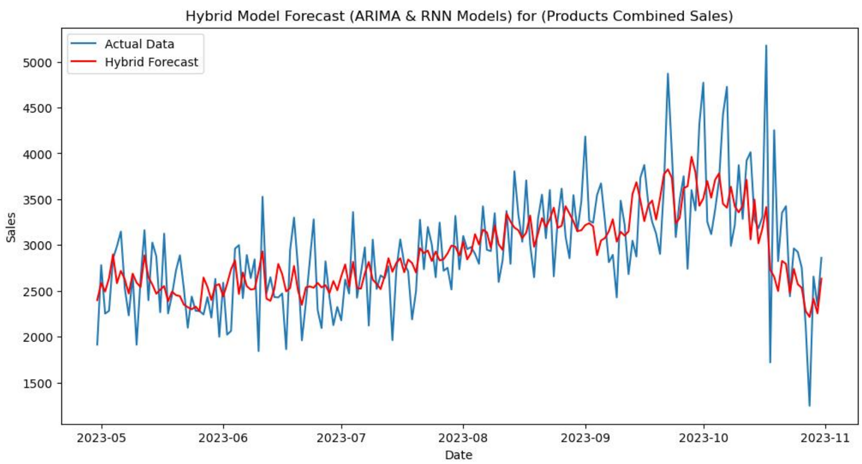


Figure 4. 55 The hybrid model forecasting of (ARIMA-RNN) models for products combined sales.

The second scenario shows the experiments for each product sales (Individually).For dairies products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RNN) model was using parameters of ARIMA (6,1,6) with (25 neurons) for RNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.56 shows dairies products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

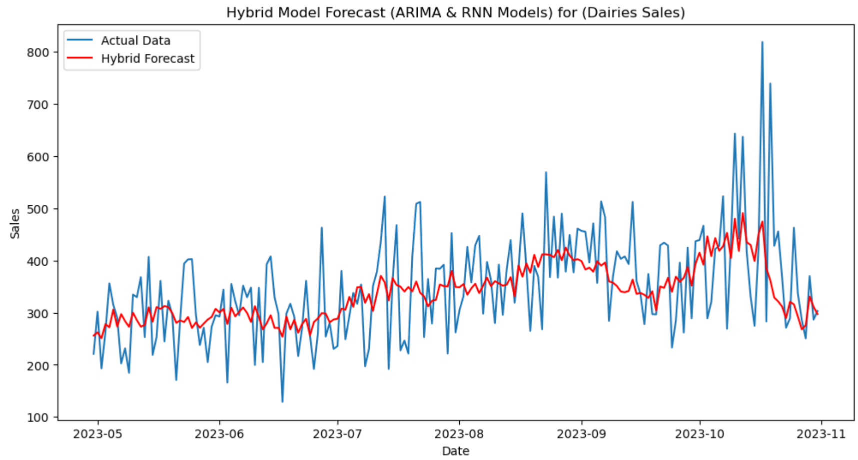


Figure 4. 56 The hybrid model forecasting of (ARIMA-RNN) models for dairies products sales.

For ice-cream products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RNN) model was using parameters of ARIMA (6,1,1) with (20 neurons) for RNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.57 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

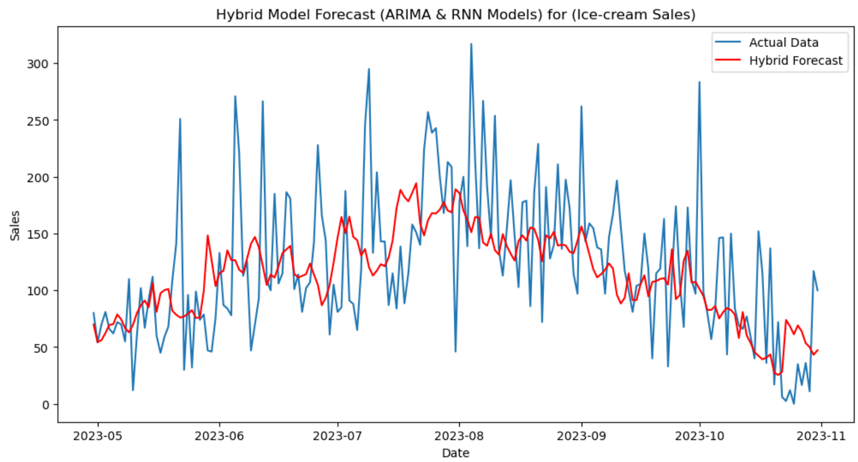

Figure 4. 57 The hybrid model forecasting of (ARIMA-RNN) models for ice-cream products sales.

For drinks products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RNN) model was using parameters of ARIMA (6,1,1) with (20 neurons) for RNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.58 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

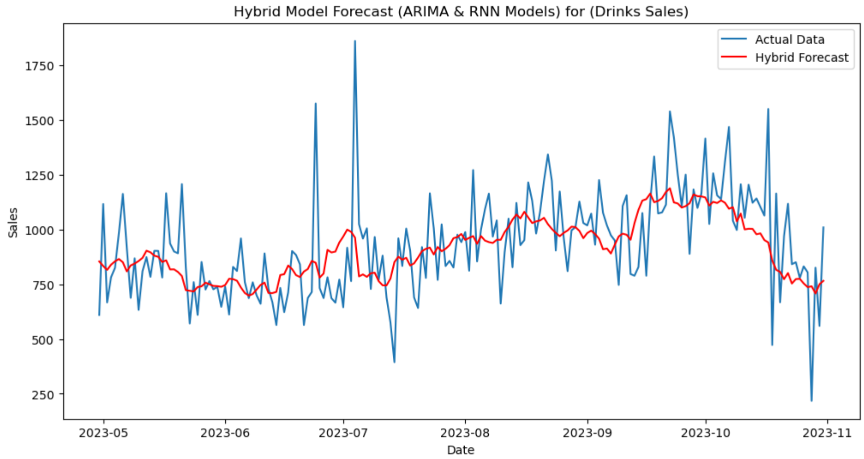


Figure 4. 58 The hybrid model forecasting of (ARIMA-RNN) models for drinks products sales.

For snacks & chips products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RNN) model was using parameters of ARIMA (6,1,1) with (20 neurons) for RNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.59 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

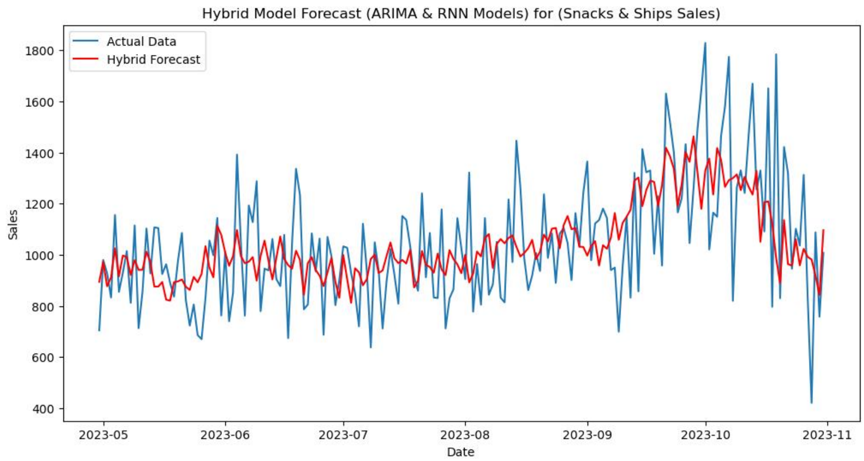

Figure 4. 59 The hybrid model forecasting of (ARIMA-RNN) models for snacks & chips products sales.

For cleaning materials products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RNN) model was using parameters of ARIMA (5,1,1) with (20 neurons) for RNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.60 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

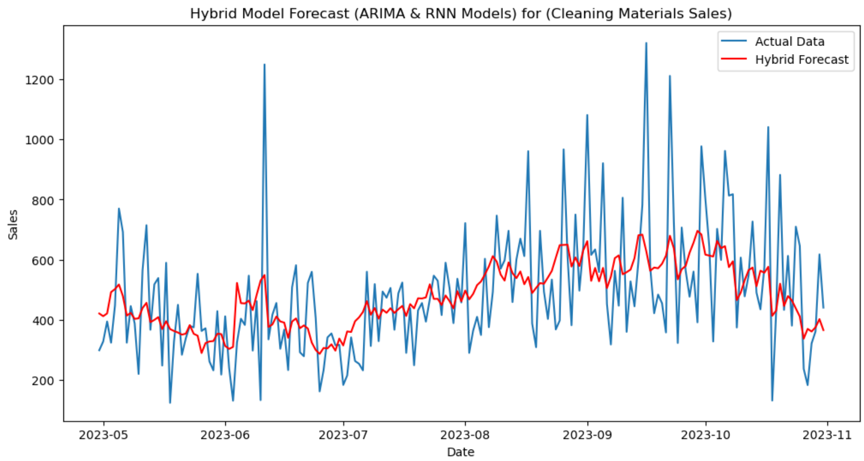


Figure 4. 60 The hybrid model forecasting of (ARIMA-RNN) models for cleaning materials products sales.

The table below summarizes the best results of two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for the hybrid model (ARIMA – RNN) model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.15 shows best error metrics of hybrid model (ARIMA – RNN) model for products combined sales and for each category sales (individually).

Table 4. 15 The best error metrics of the hybrid model (ARIMA-RNN) for the two scenarios.

| | | | Error Metrics | | |
|---|---|---|---|---|---|
| | ARIMA Model | RNN Neurons | MSE | RMSE | MAE |
| Products Combined Sales | (6,1,1) | 20 | 2037.50 | 451.41 | 344.98 |

| | | | | | |
|---|---|---|---|---|---|
| Dairies Products Sales | (6,1,6) | 25 | 7068.51 | 84.07 | 65.15 |
| Drinks Products Sales | (6,1,1) | 20 | 3571.14 | 188.97 | 140.24 |
| Ice-cream Products Sales | (6,1,1) | 20 | 3322.94 | 57.64 | 43.64 |
| Snacks & Chips Products Sales | (6,1,1) | 20 | 4149.57 | 203.71 | 158.78 |
| Cleaning Materials Products Sales | (5,1,1) | 20 | 2899.60 | 170.28 | 127.21 |

#### 4.3.3.2 ARIMA-LSTM Models

The models experiments were applied under two scenarios, the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials).The best statistical model performance (ARIMA) was selected. To enhance the architecture of the neural networks model (LSTM), a range of experiments was involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function “relu” to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, and using the optimizer “adam”. For the model fitting, the sequence length was equal to (10), the epochs equal to (50) and batch size to (32).The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best results explored for the hybrid models of (ARIMA & LSTM) model was using parameters of ARIMA (6,1,1) with (25 neurons) for LSTM appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.61 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

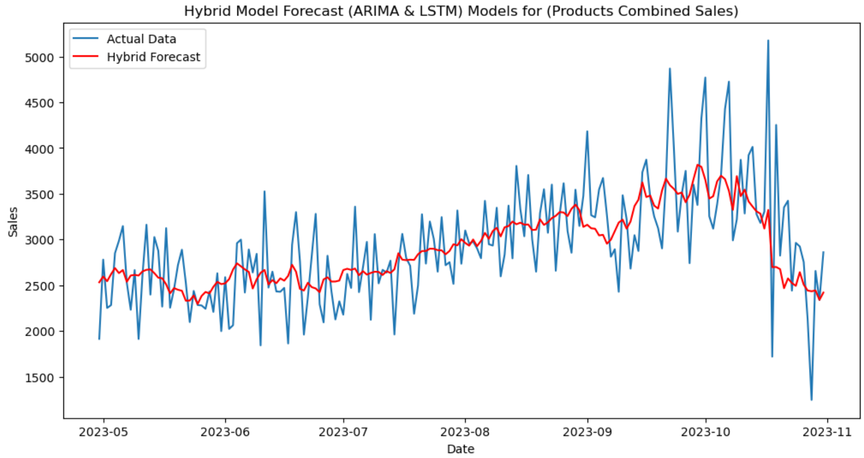


Figure 4. 61 The hybrid model forecasting of (ARIMA-LSTM) models for products combined sales.

The second scenario shows the experiments for each product sales (Individually).For dairies products sales, the experiments show the best results explored for the hybrid models of (ARIMA & LSTM) model was using parameters of ARIMA (6,1,6) with (15 neurons) for LSTM appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.62 shows dairies products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

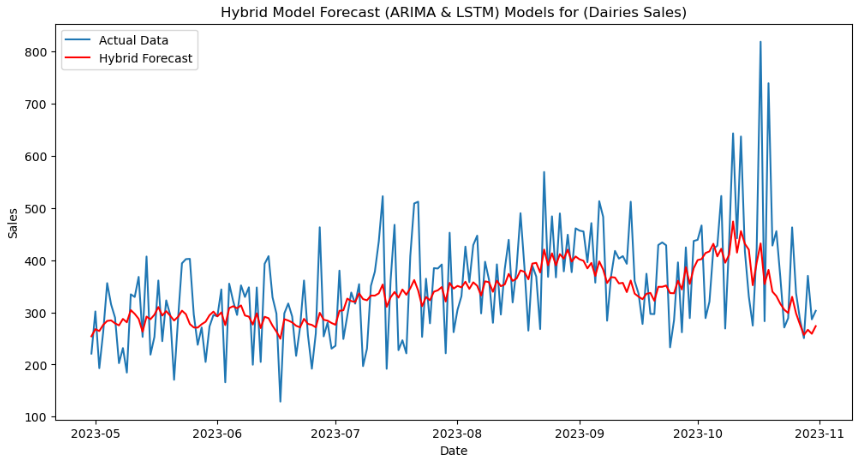


Figure 4. 62 The hybrid model forecasting of (ARIMA & LSTM) models for dairies products sales.

For ice-cream products sales, the experiments show the best results explored for the hybrid models of (ARIMA & LSTM) model was using parameters of ARIMA (6,1,1) with (10 neurons) for LSTM appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.63 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

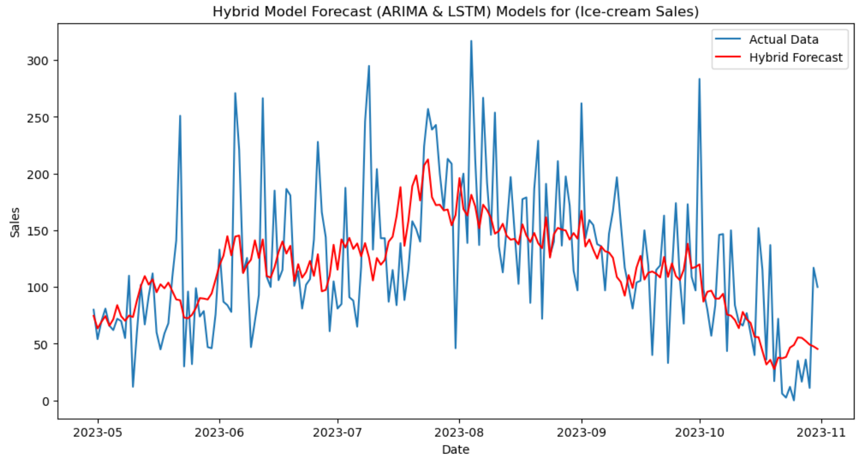


Figure 4. 63 The hybrid model forecasting of (ARIMA-LSTM) models for ice-cream products sales.

For drinks products sales, the experiments show the best results explored for the hybrid models of (ARIMA & LSTM) model was using parameters of ARIMA (6,1,1) with (15 neurons) for LSTM appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.64 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

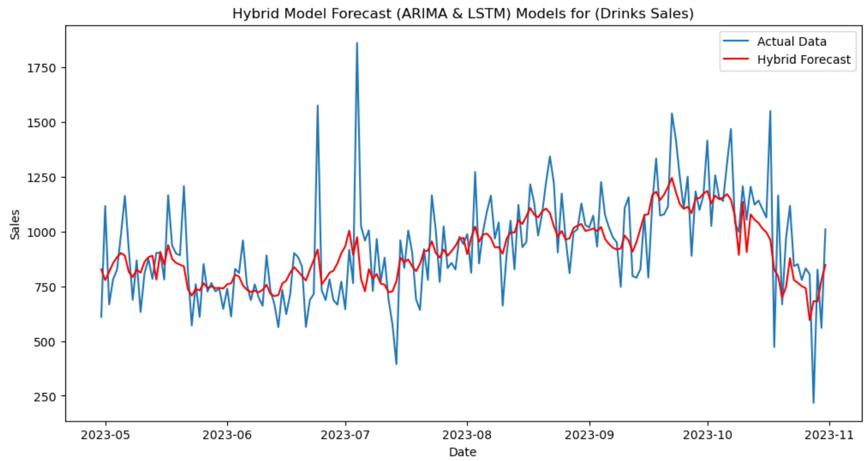


Figure 4. 64 The hybrid model forecasting of (ARIMA-LSTM) models for drinks products sales.

For snacks & chips products sales, the experiments show the best results explored for the hybrid models of (ARIMA & LSTM) model was using parameters of ARIMA (6,1,1) with (25 neurons) for LSTM appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.65 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

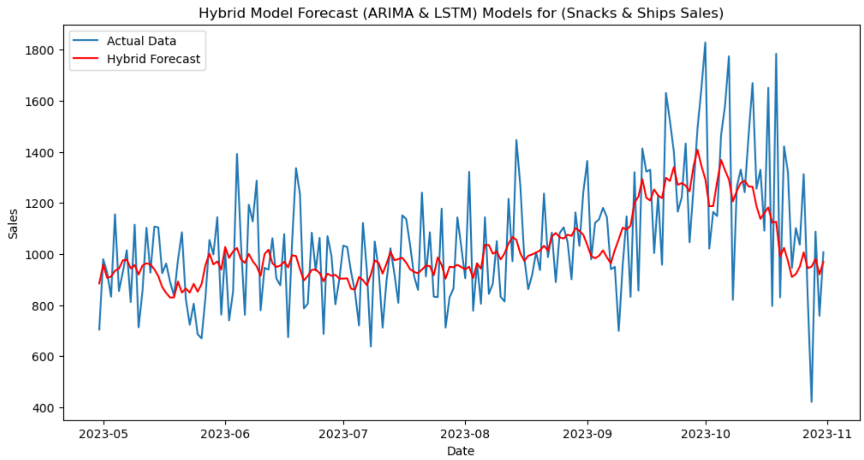


Figure 4. 65 The hybrid model forecasting of (ARIMA-LSTM) models for snacks & chips products sales.

For cleaning materials products sales, the experiments show the best results explored for the hybrid models of (ARIMA & LSTM) model was using parameters of ARIMA (5,1,1) with (30 neurons) for LSTM appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.66 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

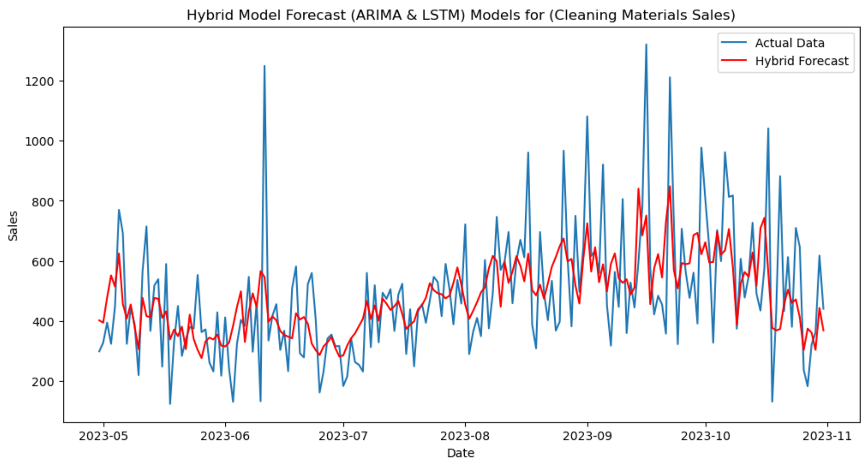


Figure 4. 66 The hybrid model forecasting of (ARIMA-LSTM) models for cleaning materials products sales.

The table below summarizes the best results of two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for the hybrid model (ARIMA – LSTM) model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.16 shows best error metrics of hybrid model (ARIMA – LSTM) model for products combined sales and for each category sales (individually).

Table 4. 16 The best error metrics of the hybrid model (ARIMA-LSTM) under two scenarios.

| | | | Error Metrics | | |
|---|---|---|---|---|---|
| | ARIMA Model | LSTM Neurons | MSE | RMSE | MAE |
| Products Combined Sales | (6,1,1) | 25 | 200417.60 | 447.68 | 337.31 |
| Dairies Products Sales | (6,1,6) | 15 | 6486.78 | 80.54 | 61.19 |
| Drinks Products Sales | (6,1,1) | 15 | 28607.24 | 169.13 | 122.67 |
| Ice-cream Products Sales | (6,1,1) | 10 | 2565.18 | 50.65 | 38.80 |
| Snacks & Chips Products Sales | (6,1,1) | 25 | 35355.39 | 188.03 | 147.47 |
| Cleaning Materials Products Sales | (5,1,1) | 30 | 28201.18 | 167.93 | 125.53 |

#### 4.3.3.3 ARIMA-MLPNNs Models

The models experiments were applied under two scenarios, where the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials).The best statistical model performance (ARIMA) was selected. To enhance the architecture of the multi-layer perceptron neural networks model (MLPNNs), a range of experiments was involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function "relu" to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, using one hidden layer. The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

For products combined sales, the experiments show the best results explored for the hybrid models of (ARIMA-MLPNNs) model was using parameters of ARIMA (6,1,1) with (5 neurons) for MLPNNs appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.67 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

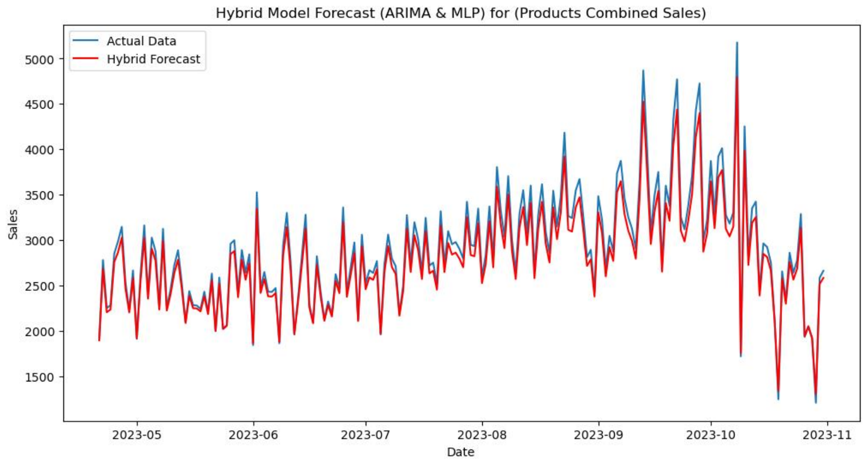


Figure 4. 67 The hybrid model forecasting of (ARIMA-MLPNNs) models for products combined sales.

The second scenario shows the experiments for each product sales (Individually). For dairies products sales, the experiments show the best results explored for the hybrid models of (ARIMA & MLPNNs) model was using parameters of ARIMA (6,1,6) with (5 neurons) for MLPNNs appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.68 shows dairies products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

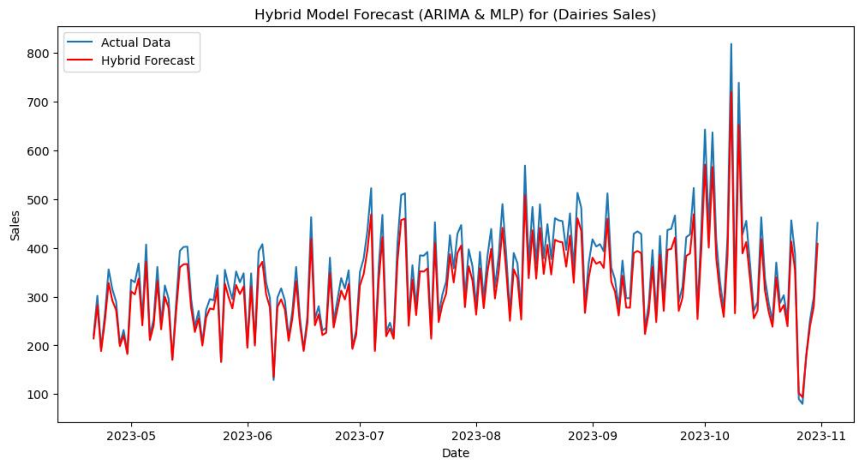


Figure 4. 68 The hybrid model forecasting of (ARIMA-MLPNNs) models for dairies products sales.

For ice-cream products sales, the experiments show the best results explored for the hybrid models of (ARIMA & MLPNNs) model was using parameters of ARIMA (6,1,1) with (25 neurons) for MLPNNs appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.69 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

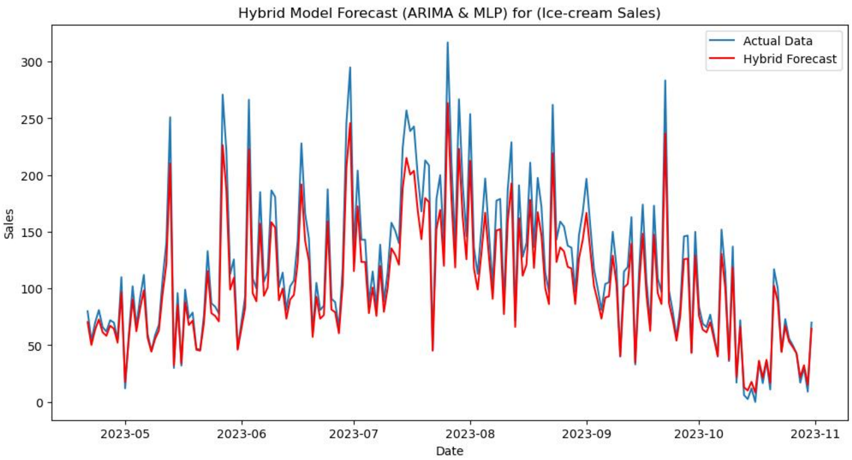


Figure 4. 69 The hybrid model forecasting of (ARIMA-MLPNNs) models for ice-cream products sales.

For drinks products sales, the experiments show the best results explored for the hybrid models of (ARIMA & MLPNNs) model was using parameters of ARIMA (6,1,1) with (25 neurons) for MLPNNs appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.70 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

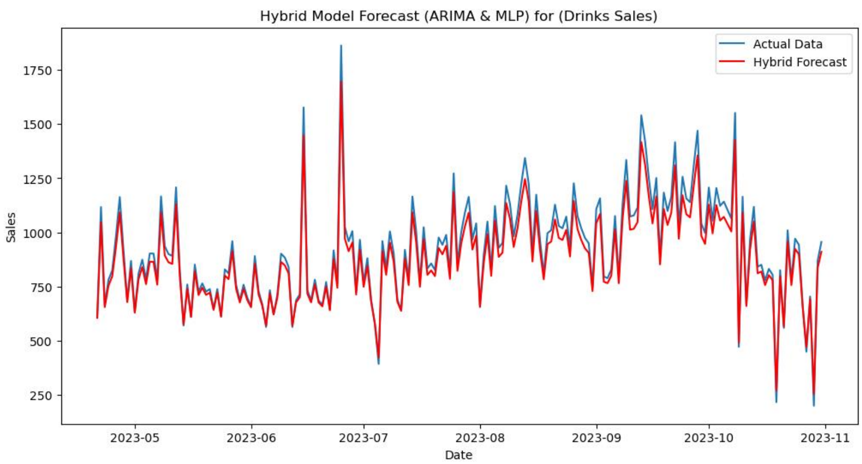


Figure 4. 70 The hybrid model forecasting of (ARIMA-MLPNNs) models for drinks products sales.

For snacks & chips products sales, the experiments show the best results explored for the hybrid models of (ARIMA & MLPNNs) model was using parameters of ARIMA (6,1,1) with (25 neurons) for MLPNNs appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.71 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

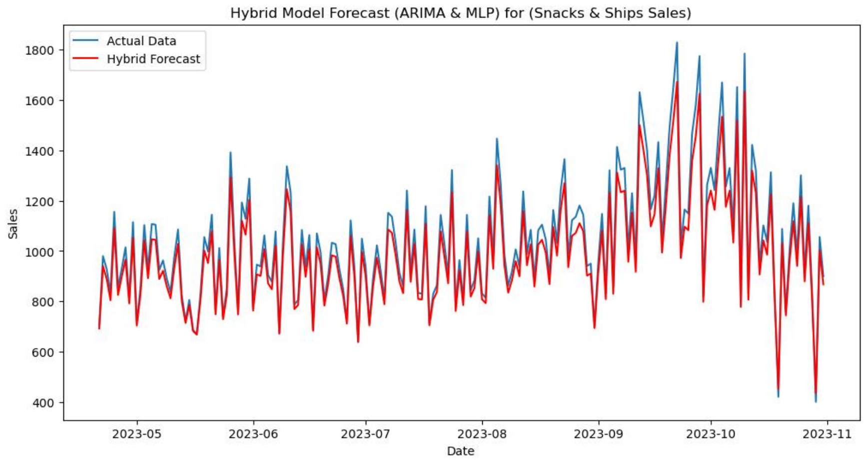


Figure 4. 71 The hybrid model forecasting of (ARIMA-MLPNNs) models for snacks & chips products sales.

For cleaning materials products sales, the experiments show the best results explored for the hybrid models of (ARIMA & MLPNNs) model was using parameters of ARIMA (5,1,1) with (5 neurons) for MLPNNs appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.72 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

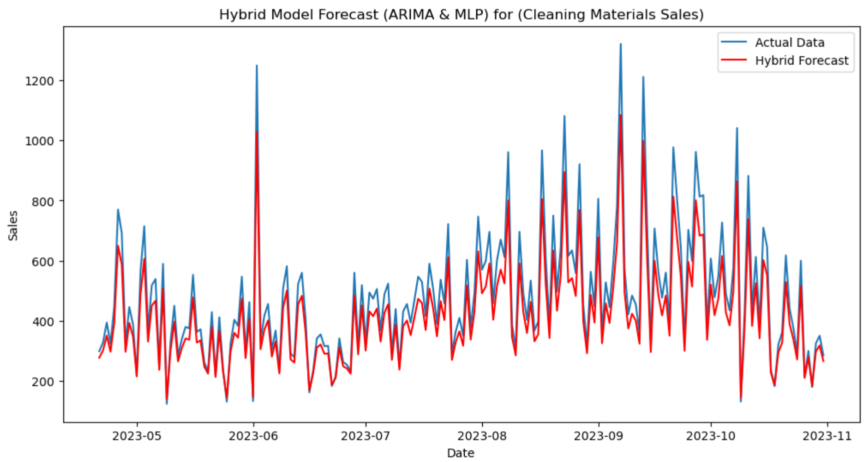

Figure 4. 72 The hybrid model forecasting of (ARIMA-MLPNNs) models for cleaning materials products sales.

The table below summarizes the best results of two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for the hybrid model (ARIMA – MLPNNs) model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.17 shows best error metrics of hybrid model (ARIMA – MLPNNs) model for products combined sales and for each category sales (individually).

Table 4. 17 The best error metrics of the hybrid model (ARIMA-MLPNNs) under two scenarios.

| | | | Error Metrics | | |
|---|---|---|---|---|---|
| | ARIMA Model | MLPNNs Neurons | MSE | RMSE | MAE |
| Products Combined Sales | (6,1,1) | 5 | 1733.03 | 131.64 | 111.32 |
| Dairies Products Sales | (6,1,6) | 5 | 969.22 | 31.13 | 27.06 |
| Drinks Products Sales | (6,1,1) | 25 | 2677.97 | 51.74 | 43.11 |
| Ice-cream Products Sales | (6,1,1) | 25 | 381.99 | 19.54 | 15.62 |
| Snacks & Chips Products Sales | (6,1,1) | 25 | 3720.72 | 60.99 | 52.27 |
| Cleaning Materials Products Sales | (5,1,1) | 5 | 5508.25 | 74.21 | 60.41 |

#### 4.3.3.4 ARIMA-RBFNNs Models

The models experiments were applied under two scenarios, the first scenario is for products combined sales, and the second one based on each products sales individually (drinks, dairies, ice-cream, snacks & chips and cleaning materials).The best statistical model performance (ARIMA) was selected. To enhance the architecture of the neural networks model (RBFNN), a range of experiments was involved to identify the most appropriate number of neurons from (5 to 30) with adding 5 neurons at the time. Furthermore, using the activation function “relu” to introduce the non-linearity and enabling the model to learn complex temporal patterns and relationships in sequential data, using epochs of 50, batch size of 32 and the optimizer “Adam”.

The error metrics were used to evaluate the model performance and provide a numerical insights into how well the model is performing to new data.

The experiment shows the best results explored for the hybrid models of (ARIMA & RBFNN) model was using parameters of ARIMA (6,1,1) with (5 neurons) for RBFNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.73 shows products combined sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

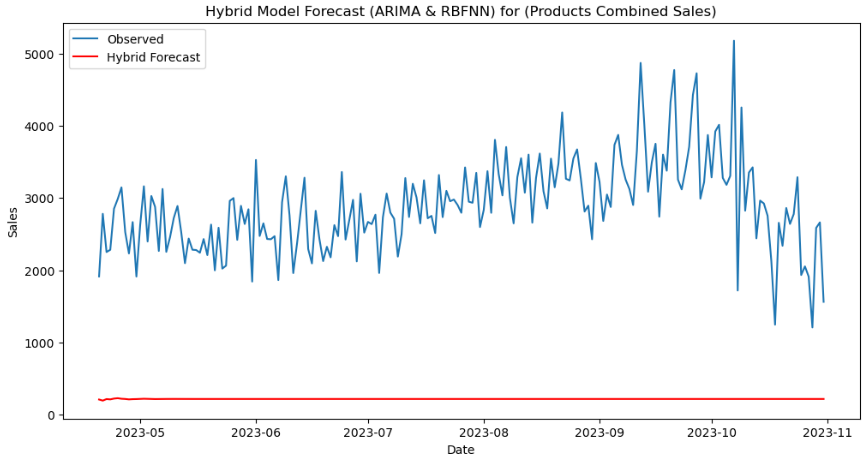


Figure 4. 73 The hybrid model forecasting of (ARIMA-RBFNN) models for products combined sales.

The second scenario shows the experiments for each product sales (Individually). For dairies products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RBFNN) model was using parameters of ARIMA (6,1,6) with (25 neurons) for RBFNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.74 shows dairies products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

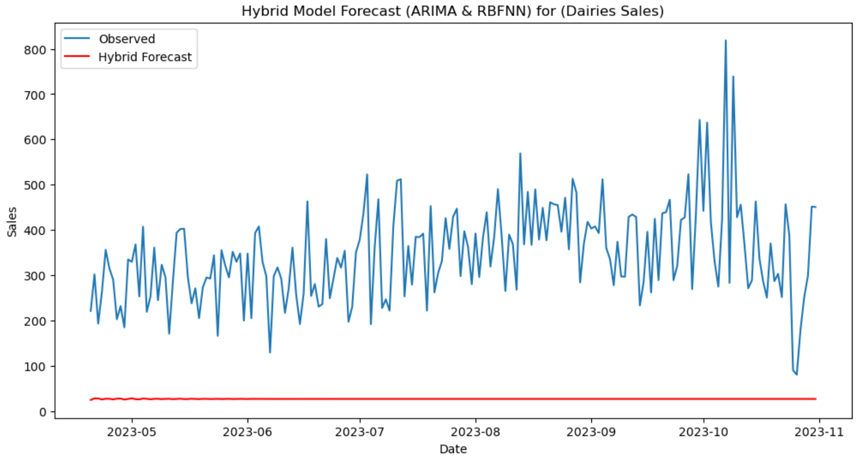


Figure 4. 74 The hybrid model forecasting of (ARIMA-RBFNN) models for dairies products sales.

For ice-cream products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RBFNN) model was using parameters of ARIMA (6,1,1) with (10 neurons) for RBFNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.75 shows ice-cream products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

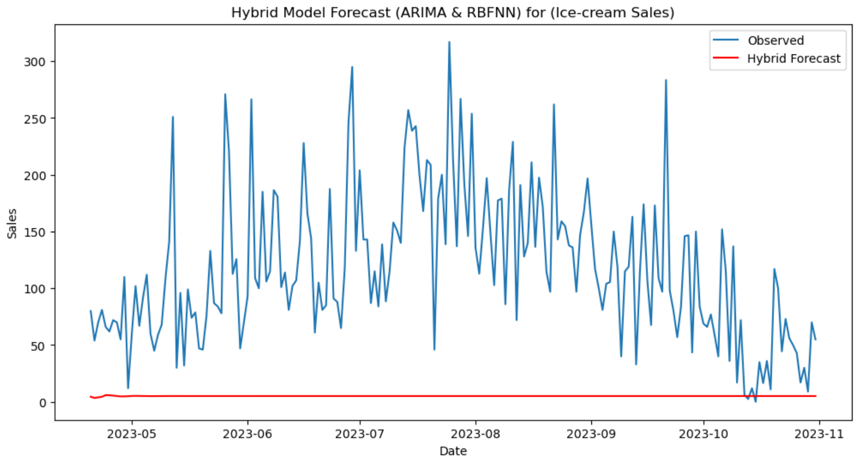


Figure 4. 75 The hybrid model forecasting of (ARIMA-RBFNN) models for ice-cream products sales.

For drinks products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RBFNN) model was using parameters of ARIMA (6,1,1) with (10 neurons) for RBFNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.76 shows drinks products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

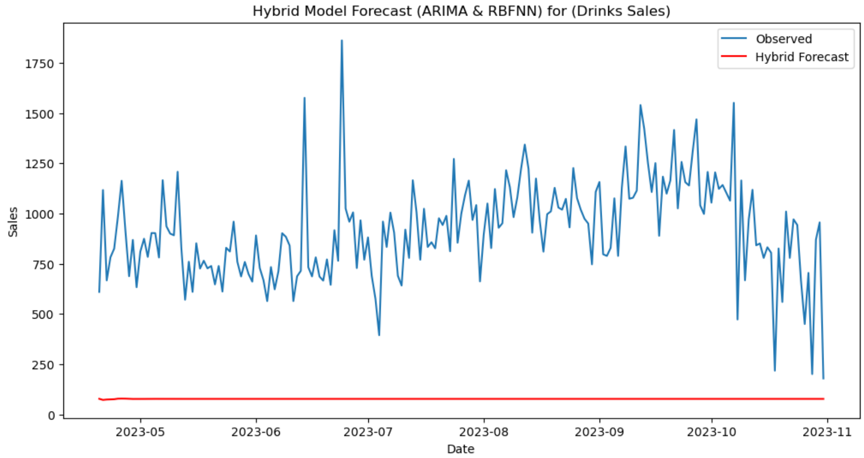


Figure 4. 76 The hybrid model forecasting of (ARIMA-RBFNN) models for drinks products sales.

For snacks & chips products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RBFNN) model was using parameters of ARIMA (6,1,1) with (5 neurons) for RBFNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.77 shows snacks & chips products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

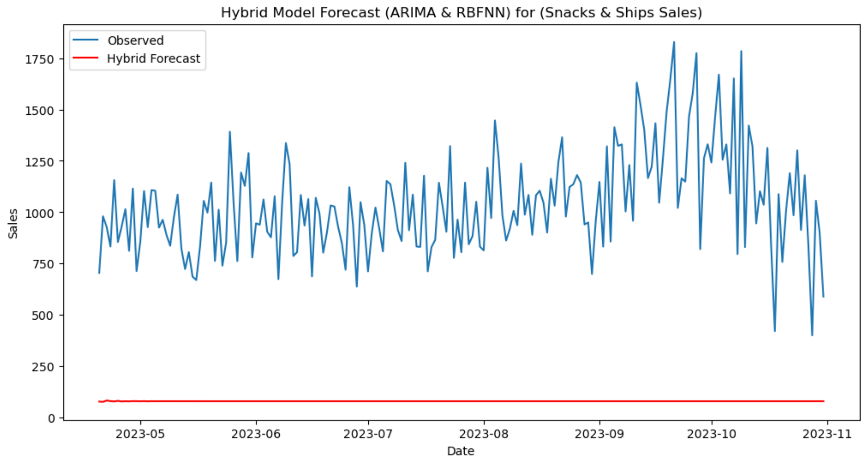


Figure 4. 77 The hybrid model forecasting of (ARIMA-RBFNN) models for snacks & chips products sales.

For cleaning materials products sales, the experiments show the best results explored for the hybrid models of (ARIMA & RBFNN) model was using parameters of ARIMA (5,1,1) with (5 neurons) for RBFNN appeared as the most promising and showing the lowest error metrics for the testing set. Figure 4.78 shows cleaning materials products sales for the testing set starting from (20-04-2023) to (31-10-2023) which represent the observed and forecasted behavior.

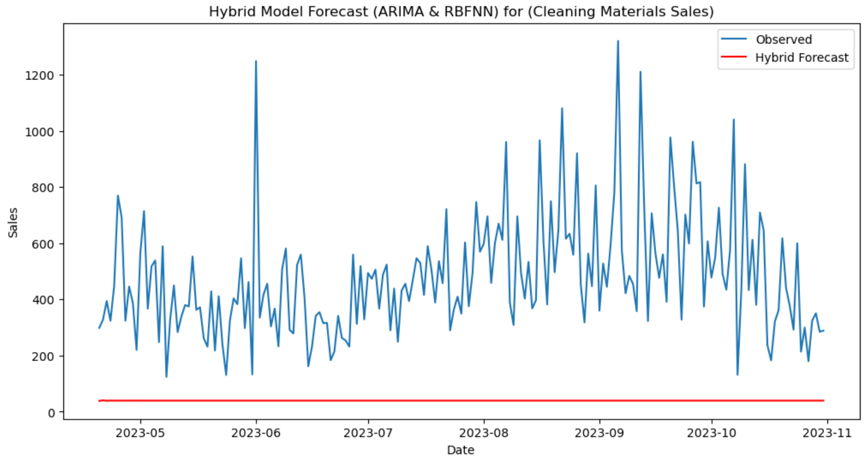


Figure 4. 78 The hybrid model forecasting of (ARIMA-RBFNN) models for cleaning materials products sales.

The table below summarizes the best results of two scenarios across the five categories, as combined sales and for each category sales (individually) based on each error metric for the hybrid model (ARIMA – RBFNN) model. It underlines the finest performance metrics, introducing an obvious overview of the most effective scenario for products sales forecasting. Table 4.18 shows best error metrics of hybrid model (ARIMA – RBFNN) model for products combined sales and for each category sales (individually).

Table 4. 18 The best error metrics of the hybrid model (ARIMA-RBFNN) under two scenarios.

| | | | Error Metrics | | |
|---|---|---|---|---|---|
| | ARIMA Model | RBFNN Neurons | MSE | RMSE | MAE |
| Products Combined Sales | (6,1,1) | 5 | 7591.54 | 2755.35 | 2682.01 |
| Dairies Products Sales | (6,1,6) | 25 | 1129.35 | 336.07 | 319.25 |
| Drinks Products Sales | (6,1,1) | 5 | 7660.41 | 875.23 | 840.82 |
| Ice-cream Products Sales | (6,1,1) | 10 | 1700.78 | 130.41 | 112.94 |
| Snacks & Chips Products Sales | (6,1,1) | 5 | 9807.68 | 990.32 | 959.52 |

| | | | | | |
|---|---|---|---|---|---|
| Cleaning Materials Products Sales | (5,1,1) | 5 | 2369.90 | 486.74 | 439.18 |

#### 4.3.3.5 Hybrid Models Errors Comparisons

For clear comparison of the hybrid models forecasting performance, the errors of various hybrid models were visualized, showing the differences and accuracy levels of each model for the two scenarios. For the first scenario as combined products sales, the errors of models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) compared with the best model performance (ARIMA-MLPNNs) to highlight the discrepancies between the observed and predicted sales values. Figure 4.79, Figure 4.80 and Figure 4.81 shows the error models comparisons between the best models performance (ARIMA-MLPNNs) and other models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) for products combined sales.

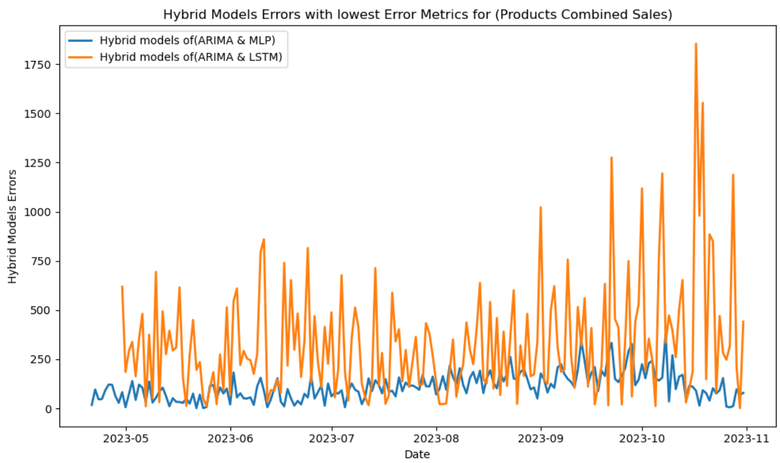


Figure 4. 79 The hybrid models (ARIMA & MLPNNs) and (ARIMA & LSTM) errors comparisons with lowest error metrics.

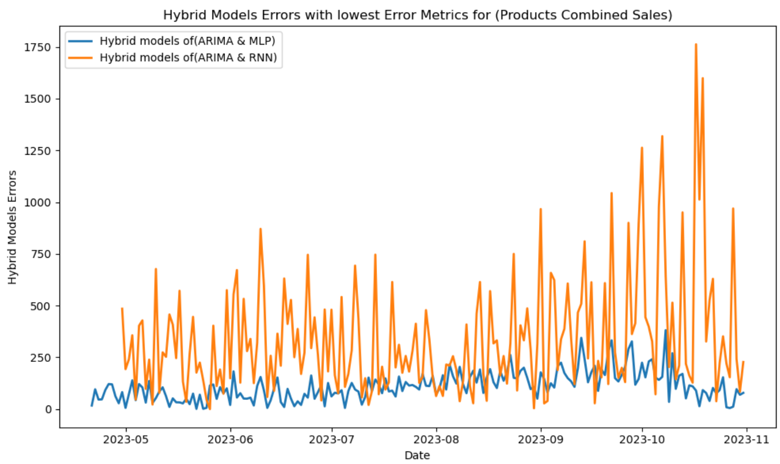

Figure 4. 80 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RNN) errors comparisons with lowest error metrics.

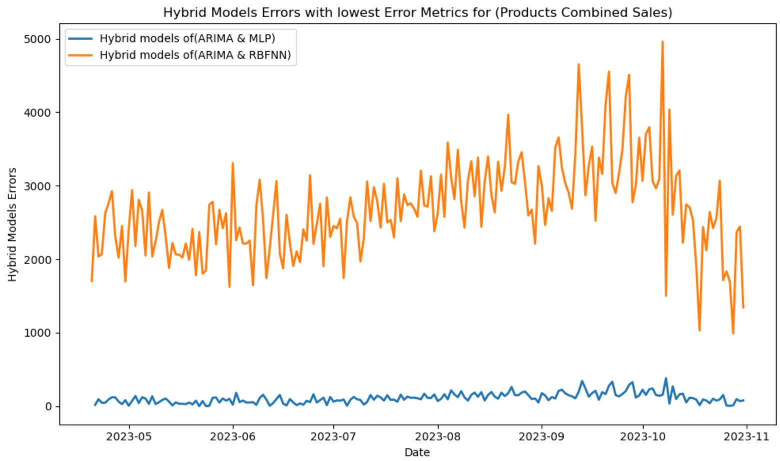


Figure 4. 81 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RBFNN) errors comparisons with lowest error metrics.

For each products sales dataset (individually). The hybrid models errors visualization for dairies products sales of (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) compared with the best model performance (ARIMA-MLPNNs) to highlight the discrepancies between the observed and predicted sales values. Figure 4.82, Figure 4.83 and Figure 4.84 shows the error models comparisons between the best models performance (ARIMA-MLPNNs) and other models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) for dairies products sales.

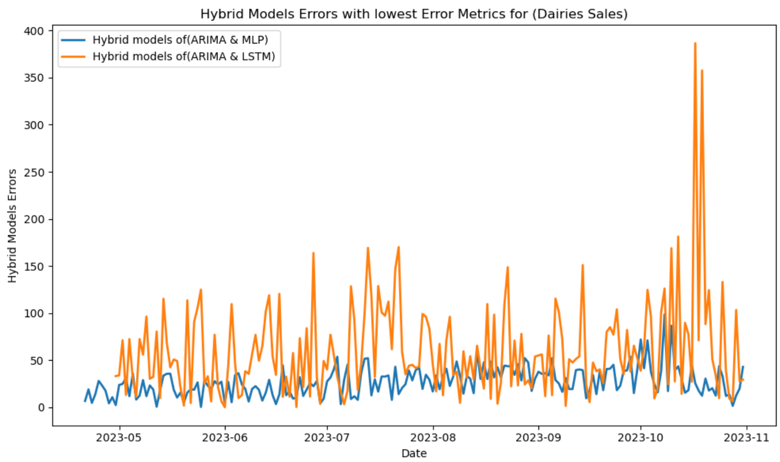


Figure 4. 82 The hybrid models (ARIMA & MLPNNs) and (ARIMA & LSTM) errors comparisons with lowest error metrics.

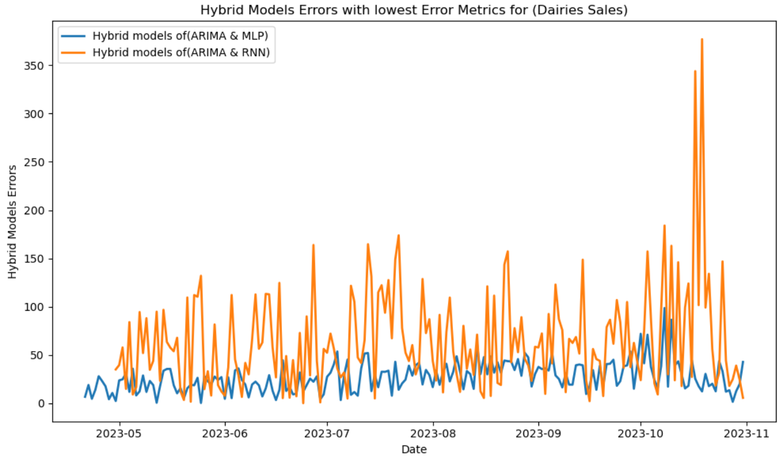


Figure 4. 83 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RNN) errors comparisons with lowest error metrics.

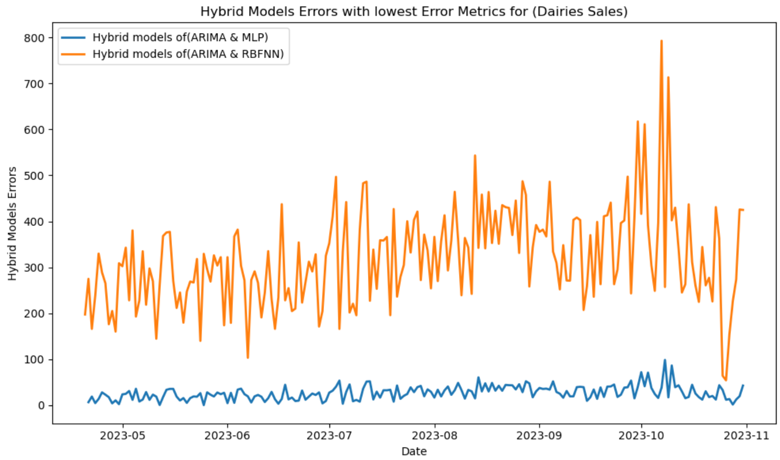


Figure 4. 84 The hybrid models (ARIMA & MLPNNs) and (ARIMA & RBFNN) errors comparisons with lowest error metrics.

For ice-cream products sales clarification for the models performance, a visualization of the models errors of (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) compared with the best model performance (ARIMA-MLPNNs), to highlight the discrepancies between the observed and predicted sales values. Figure 4.85, Figure 4.86 and Figure 4.87 shows the error models comparisons between the best models performance (ARIMA-MLPNNs) and other models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) for ice-cream products sales.

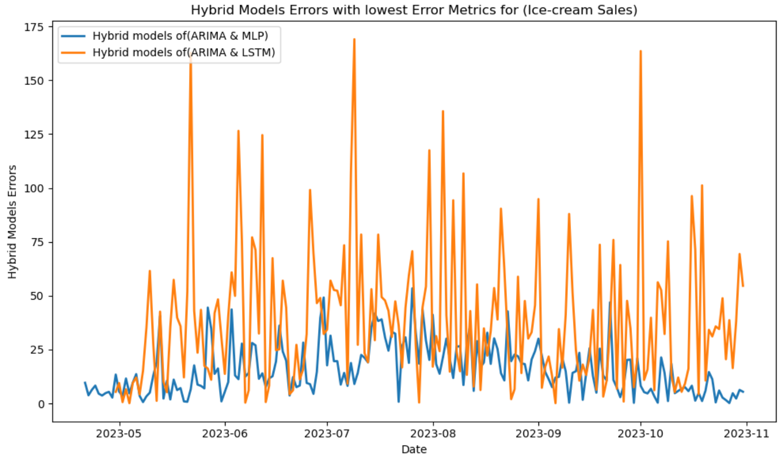


Figure 4. 85 The hybrid models (ARIMA & MLPNNs) and (ARIMA & LSTM) errors comparisons with lowest error metrics.

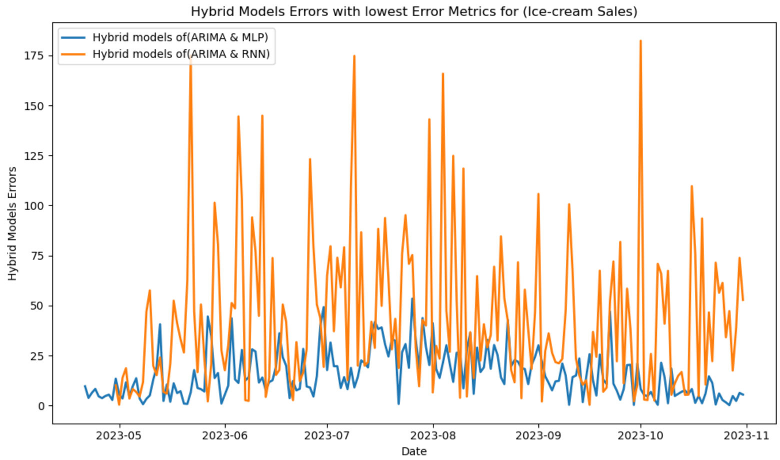


Figure 4. 86 The hybrid models (ARIMA-MLPNNs) and (ARIMA- RNN) errors comparisons with lowest error metrics.

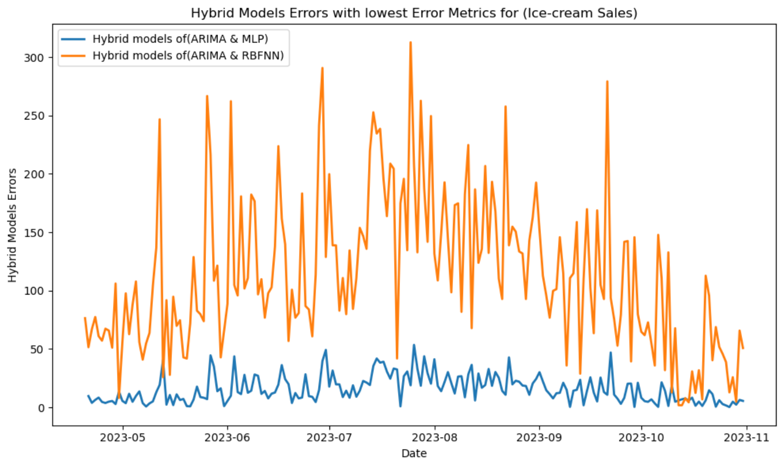

Figure 4. 87 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RBFNN) errors comparisons with lowest error metrics.

For drinks products sales clarification for the models performance, a visualization of the models errors of (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) compared with the best model performance (ARIMA-MLPNNs) to highlight the discrepancies between the observed and predicted sales values. Figure 4.88, Figure 4.89 and Figure 4.90 shows the error models comparisons between the best models performance (ARIMA-MLPNNs) and other models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) for drinks products sales.

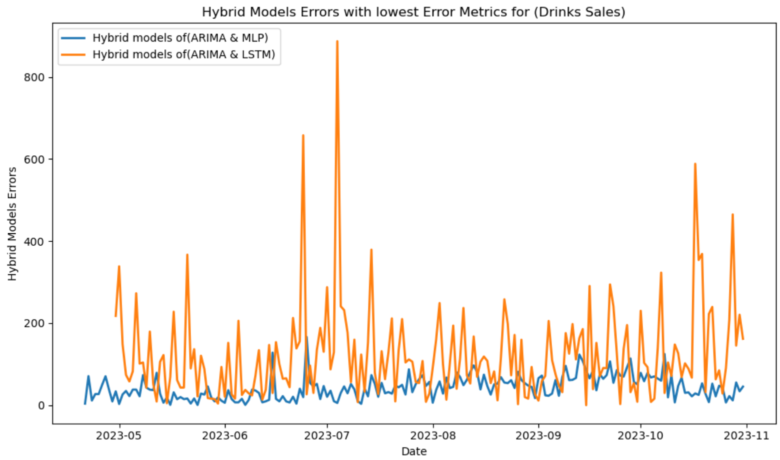


Figure 4. 88 The hybrid models (ARIMA & MLPNNs) and (ARIMA & LSTM) errors comparisons with lowest error metrics.

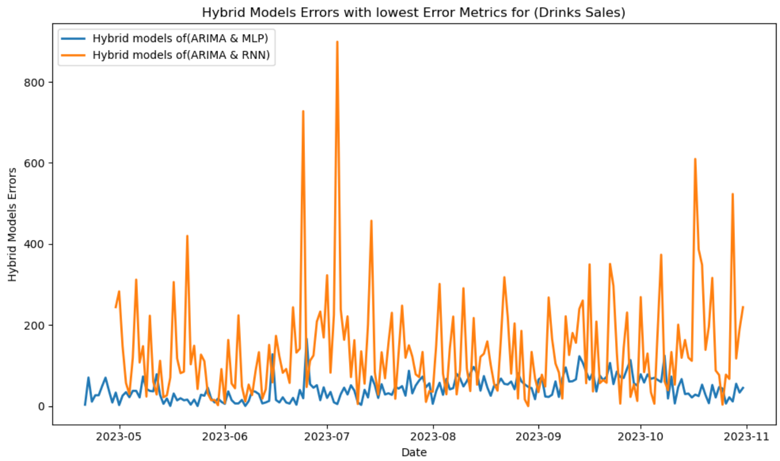


Figure 4. 89 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RNN) errors comparisons with lowest error metrics.

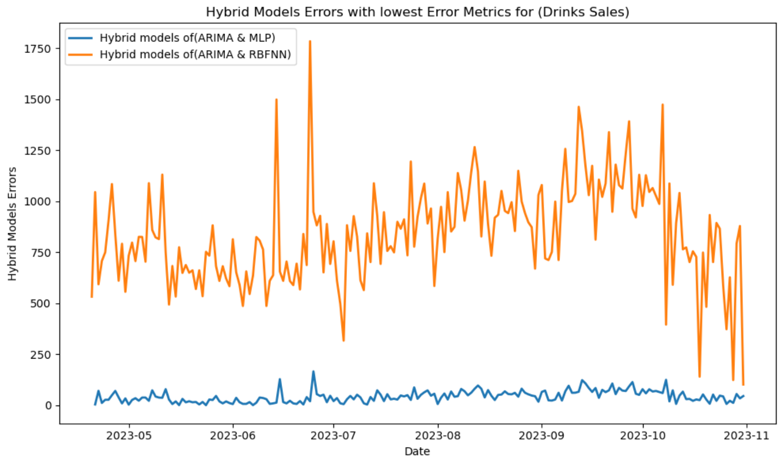


Figure 4. 90 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RBFNN) errors comparisons with lowest error metrics.

For snacks & chips products sales clarification for the models performance, a visualization of the models errors of (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) compared with the best model performance (ARIMA-MLPNNs), to highlight the discrepancies between the observed and predicted sales values. Figure 4.91, Figure 4.92 and Figure 4.93 shows the error models comparisons between the best models performance (ARIMA-MLPNNs) and other models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) for snacks & chips products sales.

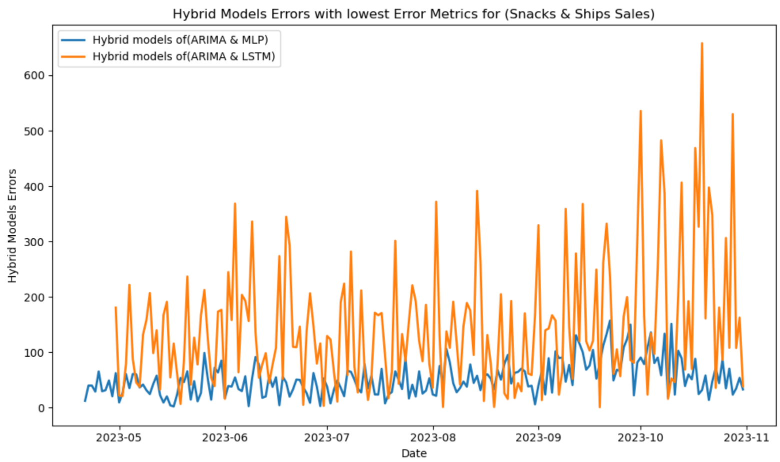


Figure 4. 91 The hybrid models (ARIMA & MLPNNs) and (ARIMA & LSTM) errors comparisons with lowest error metrics.

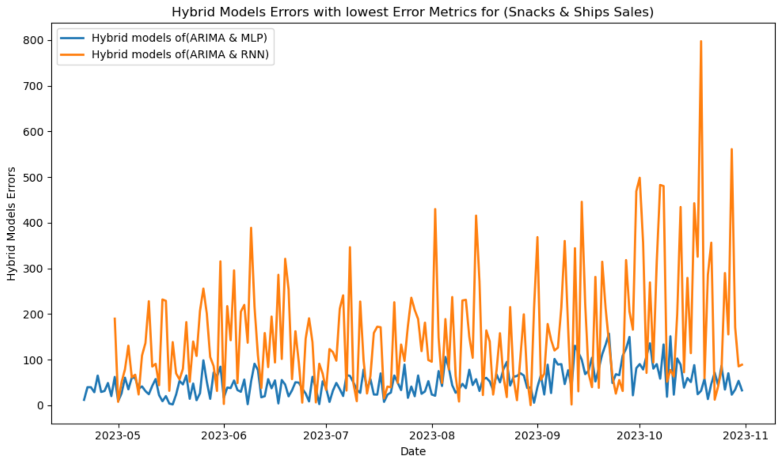


Figure 4. 92 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RNN) errors comparisons with lowest error metrics.

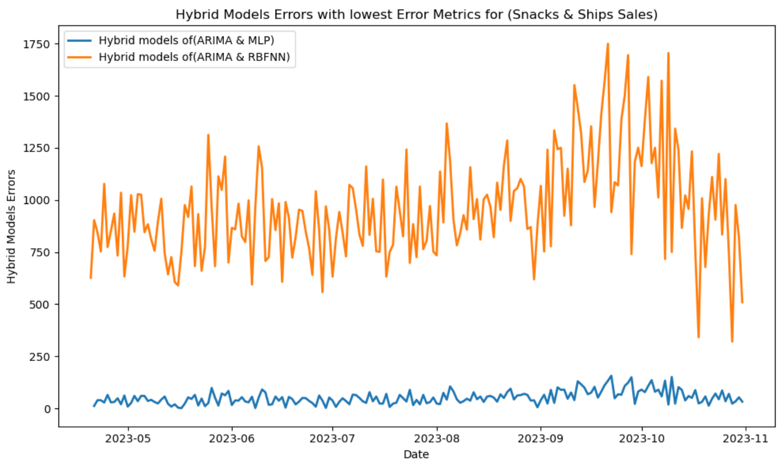


Figure 4. 93 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RBFNN) errors comparisons with lowest error metrics.

For cleaning materials products sales clarification for the models performance, a visualization of the models errors (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) compared with the best model performance (ARIMA-MLPNNs), to highlight the discrepancies between the observed and predicted sales values. Figure 4.94, Figure 4.95 and Figure 4.96 shows the error models comparisons between the best models performance (ARIMA-MLPNNs) and other models (ARIMA-LSTM), (ARIMA-RNN) and (ARIMA-RBFNN) for cleaning materials products sales.

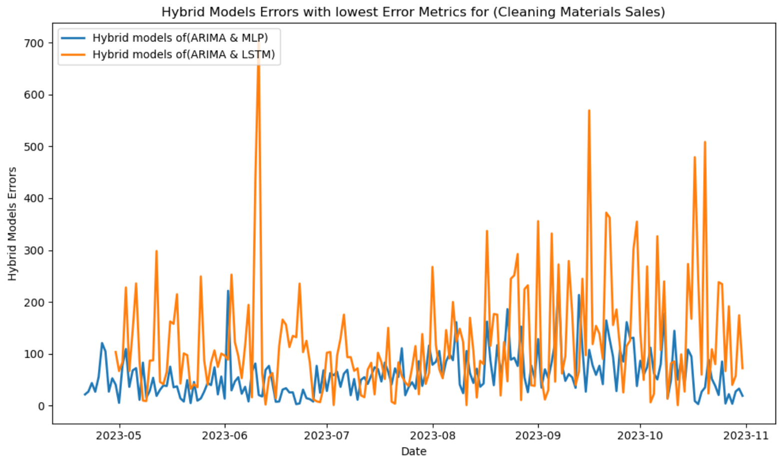


Figure 4. 94 The hybrid models (ARIMA & MLPNNs) and (ARIMA & LSTM) errors comparisons with lowest error metrics.

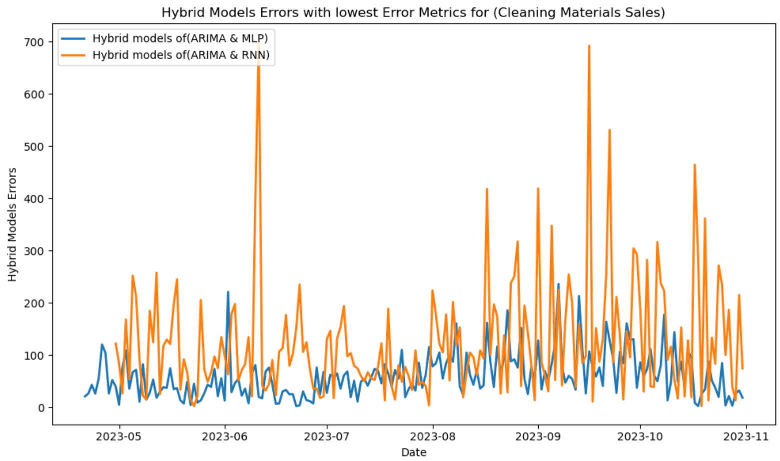


Figure 4. 95 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RNN) errors comparisons with lowest error metrics.

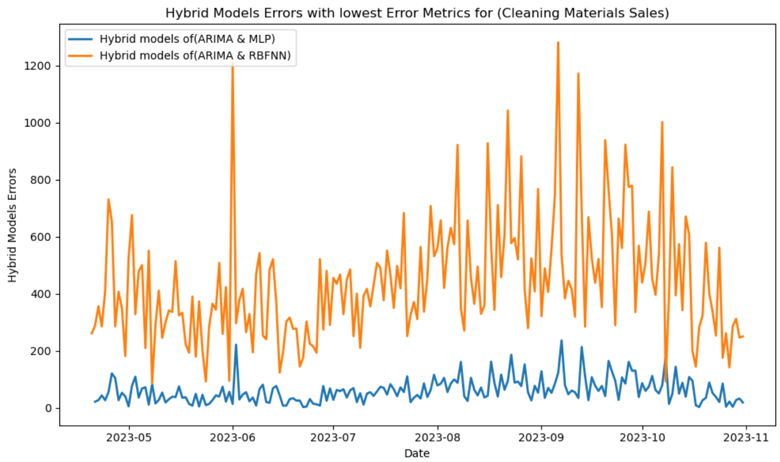

Figure 4. 96 The hybrid models (ARIMA-MLPNNs) and (ARIMA-RBFNN) errors comparisons with lowest error metrics.

#### 4.3.3.6 Summary

After evaluating the performance of hybrid models combining ARIMA and neural networks(RNN, LSTM, MLPNNs and RBFNN), their accuracy was examined using key error metrics including (MSE, RMSE and MAE).The results mark that most hybrid models achieved promising levels of forecasting accuracy. In the evaluation of hybrid models for products combined sales, our analysis detected that the hybrid model of (ARIMA-MLPNNs) surpassed all other hybrid models, including all standalone statistical and neural networks models (NNs), when evaluated as individual models and in hybrid configurations. While other hybrid models (ARIMA-LSTM and ARIMA-RNN) excelled the performance of all single statistical and NNs models, except for ARIMA model, which presented comparable results. The hybrid model (ARIMA-RBFNN) excelled the RBFNN model only but its performance still remained poor and unable to predict. The following Table 4.19 presents the summarized metrics the best error metrics of the hybrid models for products combined sales.

Table 4. 19 The best error metrics of the hybrid models for products combined sales.

| | The Hybrid Models | Error Metrics | | |
|---|---|---|---|---|
| | | MSE | RMSE | MAE |
| Products Combined Sales | ARIMA-MLPNN | 1733.03 | 131.64 | 111.32 |
| | ARIMA-LSTM | 20041.60 | 447.68 | 337.31 |
| | ARIMA-RNN | 2037.50 | 451.41 | 344.98 |
| | ARIMA-RBFNN | 7591.54 | 275.35 | 268.01 |

In the evaluation of hybrid models for products sales (individually). For dairies products sales, our analysis detected that the hybrid model of (ARIMA-MLPNNs) surpassed all other hybrid models, including all standalone statistical and neural networks models (NNs), when evaluated as individual models and in hybrid configurations. Followed with hybrid models of (ARIMA-LSTM) excelled the performance of all single statistical and NNs models, except the ARIMA model, while (ARIMA-RNN) provide good performance but not better than individual ARIMA and MLPNNs models. The hybrid model (ARIMA-RBFNN) excelled the RBFNN model only but its performance still remained poor and unable to predict. Table 4.20 summarizes the best error metrics of the hybrid models for dairies products sales.

Table 4. 20 The best error metrics of the hybrid models for dairies products sales.

| | The Hybrid Models | Error Metrics | | |
|---|---|---|---|---|
| | | MSE | RMSE | MAE |
| Dairies Products Sales | ARIMA-MLPNN | 969.22 | 31.13 | 27.06 |
| | ARIMA-LSTM | 6486.78 | 80.54 | 61.19 |
| | ARIMA-RNN | 7068.51 | 84.07 | 65.15 |
| | ARIMA-RBFNN | 1129.35 | 336.07 | 319.25 |

In the evaluation of hybrid models for ice-cream products sales, our analysis detected that the hybrid model of (ARIMA-MLPNNs) significantly surpassed all other hybrid models, including all standalone statistical and neural networks models (NNs), when evaluated as individual models and in hybrid configurations. Followed with hybrid models (ARIMA-LSTM) and (ARIMA-RNN) where the differences between them were relatively small, providing similar levels of effectiveness among standalone and hybrid models, except with ARIMA model. The hybrid model (ARIMA-RBFNN) performance still remained poor and unable to predict. Table 4.21 summarizes the best error metrics of the hybrid models for ice-cream products sales.

Table 4. 21 The best error metrics of the hybrid models for dairies products sales.

| | The Hybrid Models | Error Metrics | | |
|---|---|---|---|---|
| | | MSE | RMSE | MAE |
| Ice-cream Products Sales | ARIMA-MLPNN | 381.99 | 19.54 | 15.62 |
| | ARIMA-LSTM | 2565.18 | 50.65 | 38.80 |
| | ARIMA-RNN | 3322.94 | 57.64 | 43.64 |
| | ARIMA-RBFNN | 1700.78 | 130.41 | 112.94 |

In the evaluation of hybrid models for drinks products sales, our analysis detected that the hybrid model of (ARIMA-MLPNNs) surpassed all other hybrid models, including all standalone statistical and neural networks models (NNs), when evaluated as individual models and in hybrid configurations. Followed with hybrid models (ARIMA-LSTM) and (ARIMA-

RNN) which excelled the performance of all single statistical and NNs models, except for ARIMA model, with comparable results. The hybrid model (ARIMA-RBFNN) excelled the RBFNN model only but its performance still remained poor and unable to predict. Table 4.22 summarizes the best error metrics of the hybrid models for ice-cream products sales.

Table 4. 22 The best error metrics of the hybrid models for drinks products sales.

| | The Hybrid Models | Error Metrics | | |
|---|---|---|---|---|
| | | MSE | RMSE | MAE |
| Drinks Products Sales | ARIMA-MLPNN | 2677.97 | 51.74 | 43.11 |
| | ARIMA-LSTM | 2860.24 | 169.13 | 122.67 |
| | ARIMA-RNN | 3571.14 | 188.97 | 140.24 |
| | ARIMA-RBFNN | 7660.41 | 875.23 | 840.82 |

In the evaluation of hybrid models for snacks & chips products sales, our analysis detected that the hybrid model of (ARIMA-MLPNNs) surpassed all other hybrid models, including all standalone statistical and neural networks models (NNs), when evaluated as individual models and in hybrid configurations. Followed with hybrid models (ARIMA-LSTM) and (ARIMA-RNN) which excelled the performance of all single statistical and NNs models, except for ARIMA model, with comparable results. The hybrid model (ARIMA-RBFNN) excelled the RBFNN model only but its performance still remained poor and unable to predict. Table 4.23 summarizes the best error metrics of the hybrid models for snacks & chips products sales.

Table 4. 23 The best error metrics of the hybrid models for snacks & chips products sales.

| | The Hybrid Models | Error Metrics | | |
|---|---|---|---|---|
| | | MSE | RMSE | MAE |
| Snacks & Chips Products Sales | ARIMA-MLPNN | 3720.72 | 60.99 | 52.27 |
| | ARIMA-LSTM | 3535.39 | 188.03 | 147.47 |
| | ARIMA-RNN | 4149.57 | 203.71 | 158.78 |
| | ARIMA-RBFNN | 9807.68 | 990.32 | 959.52 |

In the evaluation of hybrid models for cleaning materials products sales, our analysis detected that the hybrid model of (ARIMA-MLPNNs) surpassed all other hybrid models, including all standalone statistical and neural networks models (NNs), when evaluated as individual models and in hybrid configurations. Followed with hybrid models (ARIMA-LSTM) and (ARIMA-RNN) where the differences between them are relatively small, but excelled the performance of all single statistical and NNs models, except for standalone ARIMA and MLPNNs models , with comparable results. The hybrid model (ARIMA-RBFNN) excelled the RBFNN model only but its performance still remained poor and unable to predict. Table 4.24 summarizes the best error metrics of the hybrid models for cleaning materials products sales.

Table 4. 24 The best error metrics of the hybrid models for cleaning materials products sales.

| | The Hybrid Models | Error Metrics | | |
|---|---|---|---|---|
| | | MSE | RMSE | MAE |
| Cleaning Materials Products Sales | ARIMA-MLPNN | 5508.25 | 74.21 | 60.41 |
| | ARIMA-LSTM | 2820.18 | 167.93 | 125.53 |
| | ARIMA-RNN | 2899.60 | 170.28 | 127.21 |
| | ARIMA-RBFNN | 2369.90 | 486.74 | 439.18 |

### 4.3.4 Future Forecasting

Future forecasting results detect insights into hybrid models forecasting capabilities for next month in (November). For both scenarios, the best hybrid models results were applied which are (ARIMA-MLPNNs) and (ARIMA-LSTM) models. Starting with products combined sales, Figure 4.97 and Figure 4.98 show the future forecasting results for products combined sales for November using the two hybrid models.

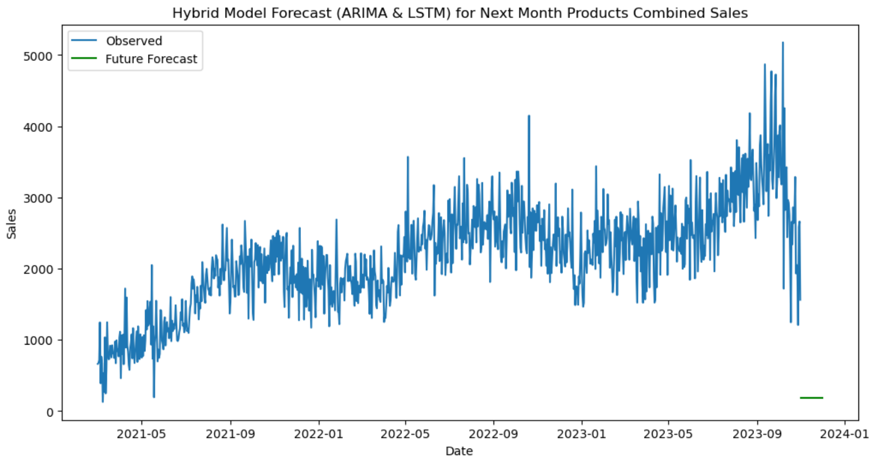


Figure 4. 97 The (ARIMA-LSTM) model forecasting for next month products combined sales.

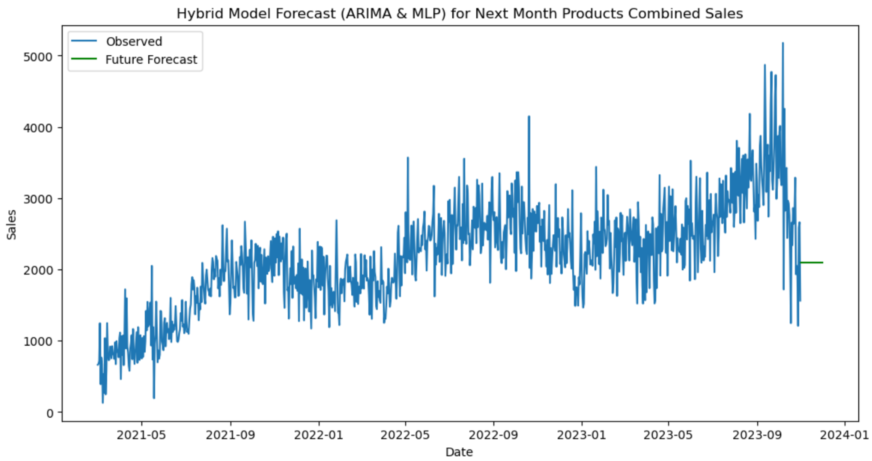


Figure 4. 98 The (ARIMA-MLPNNs) model forecasting for next month products combined sales.

Results have shown that the hybrid model (ARIMA-LSTM) was unable to forecast for the next month, while the hybrid model (ARIMA-MLPNNs) was able to make fixed sales forecasting equal to nearly 2000 and a little more for next month. The forecasting results in this scenario were not good compared to the forecasting results for each product individually.

The second scenario shows future forecasting for each product sales (Individually). Figure 4.99 and Figure 4.100 show future forecasting for next month dairies sales using (ARIMA-LSTM) and (ARIMA-MLPNNs). Figure 4.101 and Figure 4.102 show closer view of the models movements in November forecasting using (ARIMA-LSTM) and (ARIMA-MLPNNs).

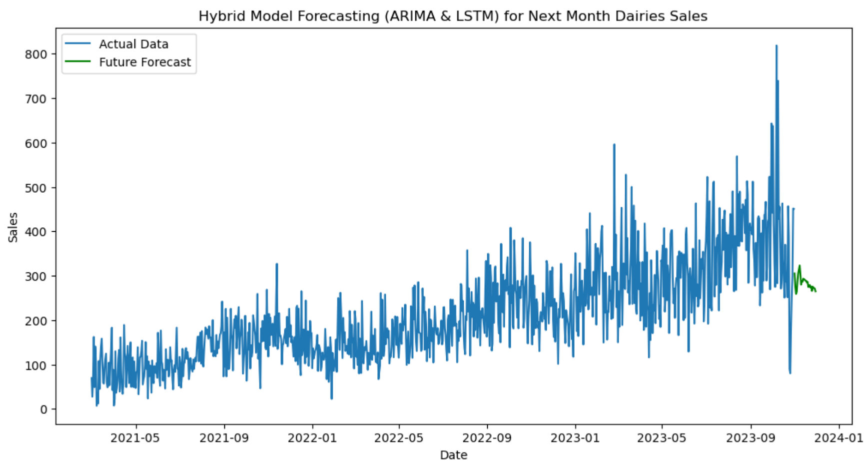


Figure 4. 99 The (ARIMA-LSTM) model forecasting for next month dairies products sales.

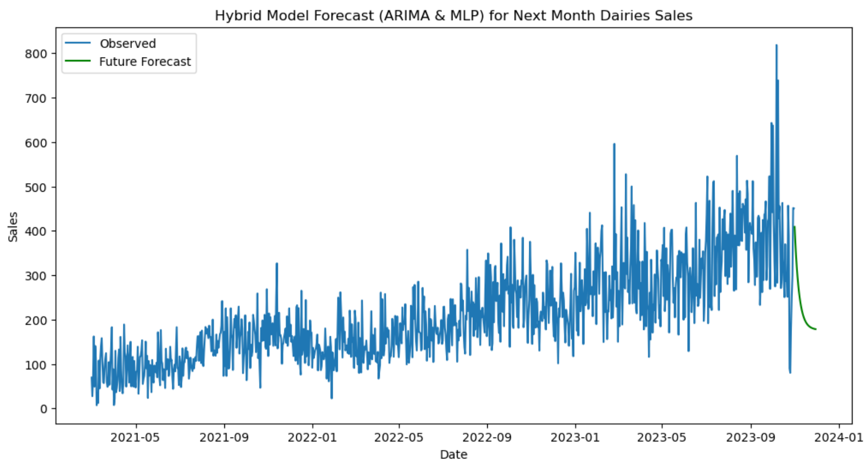


Figure 4. 100 The (ARIMA-MLPNNs) model forecasting for next month dairies products sales.

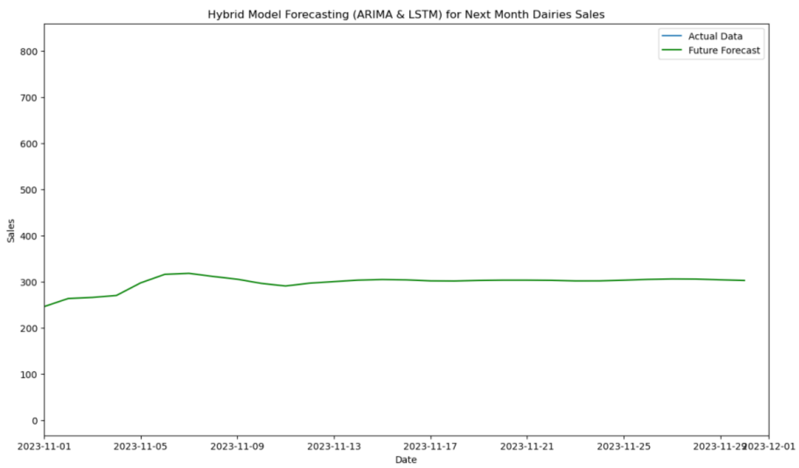

Figure 4. 101 The (ARIMA-LSTM) model forecasting for November dairies products sales.

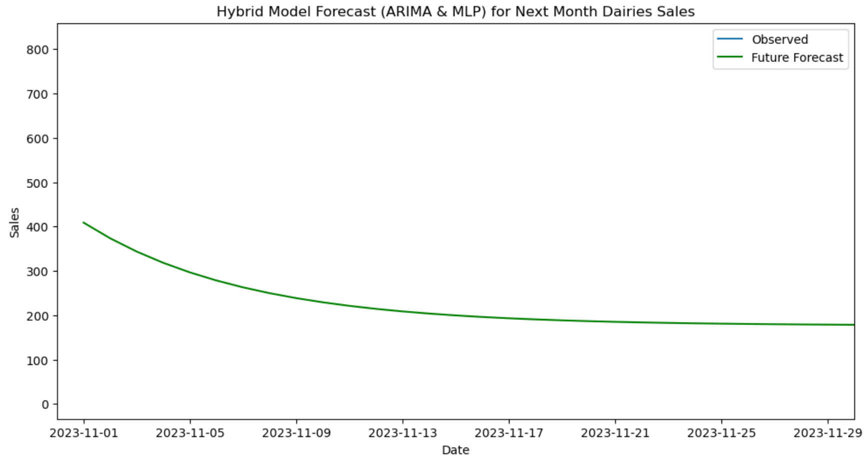


Figure 4. 102 The (ARIMA-MLPNNs) model forecasting for November dairies products sales.

Results have shown that hybrid model of (ARIMA & LSTM) was able to make forecasting to the middle of November, while hybrid model of (ARIMA-MLPNNs) model forecasting was appeared to be a decreasing trend. The hybrid model of (ARIMA-MLPNNs) ability to forecast decreasing trend might be depending on the information available in data and the model capacity to capture relevant patterns such as trends, while (ARIMA-LSTM) seems to capture more complex temporal dependencies and patterns in data, leading to more accurate predictions in terms of capturing trends and seasonality. Referring to this period of expectation, the recent difficult war events were witnessed which actually caused a shortage on sales within the supermarket.

For ice-cream products sales, Figure 4.115 and Figure 4.116 show future forecasting for next month using (ARIMA-LSTM) and (ARIMA-MLPNNs). Figure 4.103 and Figure 4.104 show closer view of the models movements in November forecasting using (ARIMA-LSTM) and (ARIMA-MLPNNs).

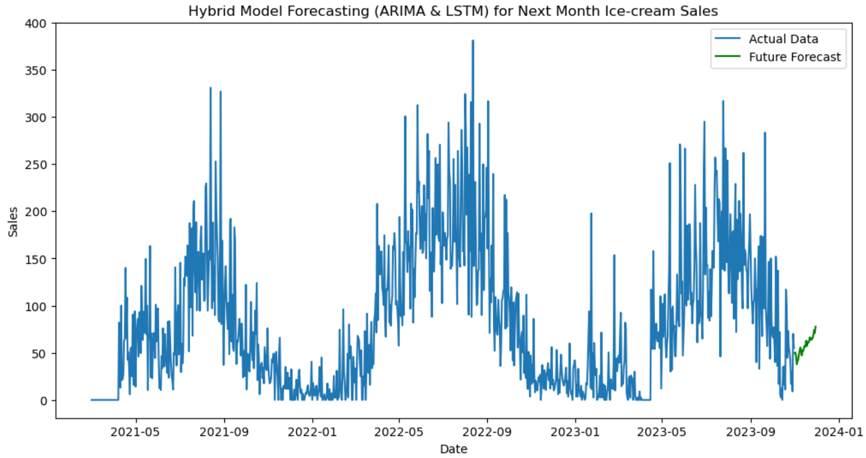


Figure 4. 103 The (ARIMA-LSTM) model forecasting for next month ice-cream products sales.

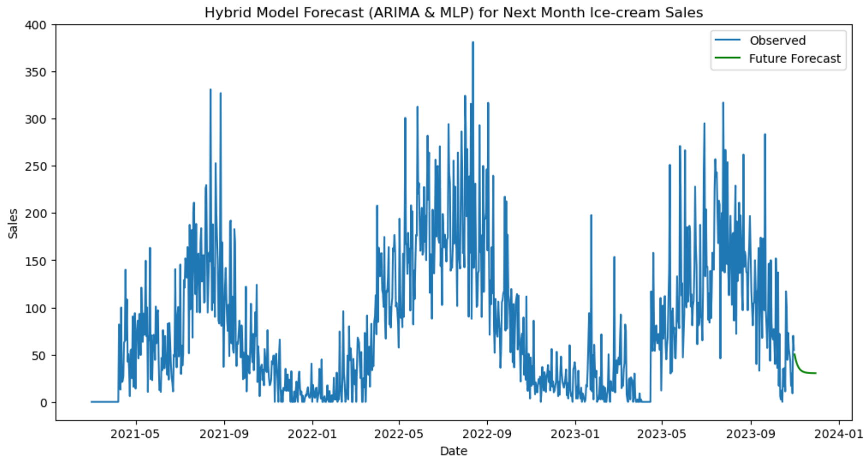


Figure 4. 104 The (ARIMA-MLPNNs) model forecasting for next month ice-cream products sales.

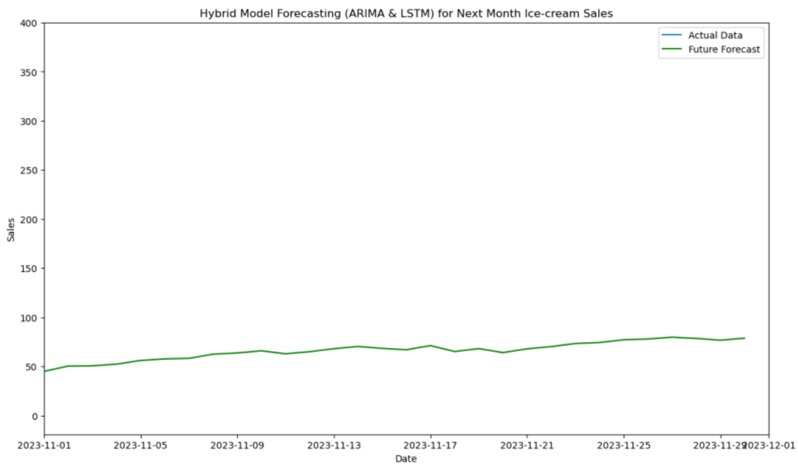

Figure 4. 105 The (ARIMA-LSTM) model forecasting for November ice-cream products sales.

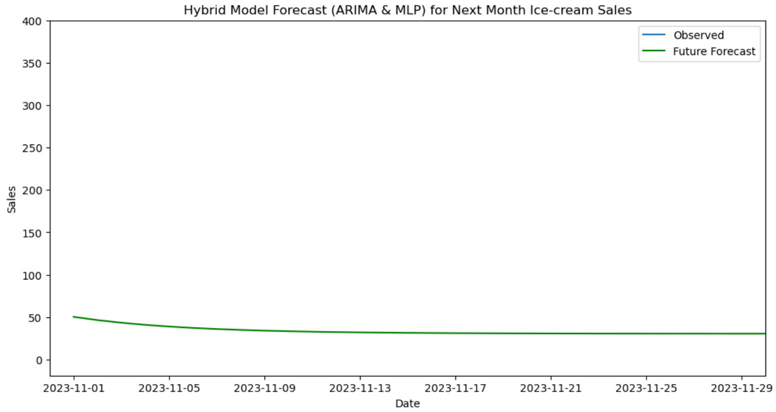


Figure 4. 106 The (ARIMA-MLPNNs) model forecasting for November ice-cream products sales.

Results have shown that hybrid model of (ARIMA-LSTM) was able to make forecasting appear like stable with very lower sales, while the hybrid model of (ARIMA-MLPNNs) forecasting was appeared to be a slightly decreasing trend and closely to be stable. Referring to the statistical analysis of such a period for selling ice-cream products, from the month of November, these products witness a high stagnation and very low sales that are closer to stagnation, given that ice-cream products witness high sales in summer seasons from May to September.

For drinks products sales, Figure 4.119 and Figure 4.120 show future forecasting for next month drinks sales using (ARIMA-LSTM) and (ARIMA-MLPNNs). Figure 4.107 and Figure 4.108 show closer view of the models movements in November forecasting using (ARIMA-LSTM) and (ARIMA-MLPNNs).

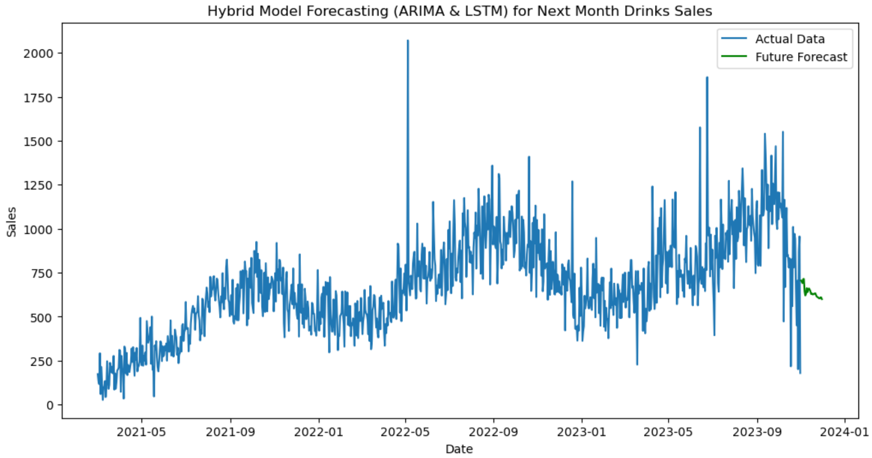


Figure 4. 107 The (ARIMA-LSTM) model forecasting for next month drinks products sales.

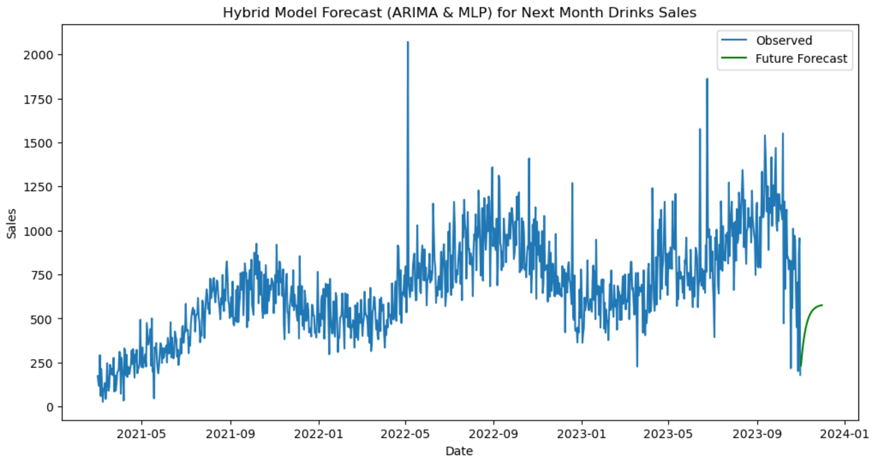


Figure 4. 108 The (ARIMA-MLPNNs) model forecasting for next month drinks products sales.

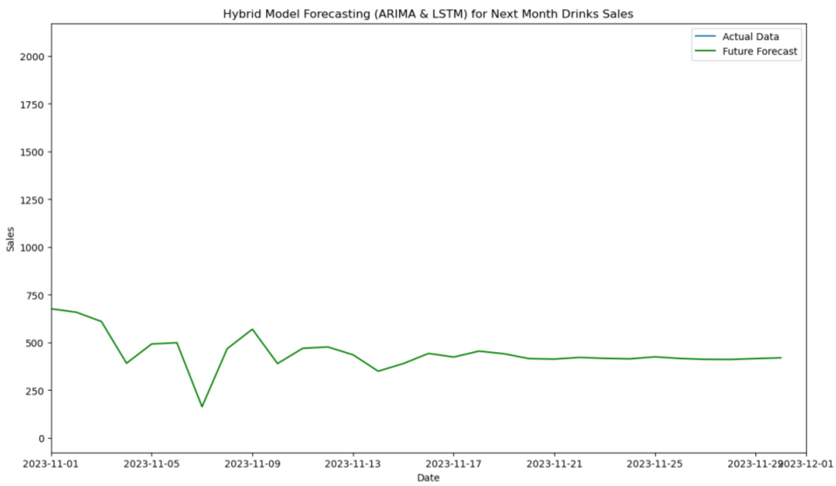

Figure 4. 109 The (ARIMA-LSTM) model forecasting for November drinks products sales.

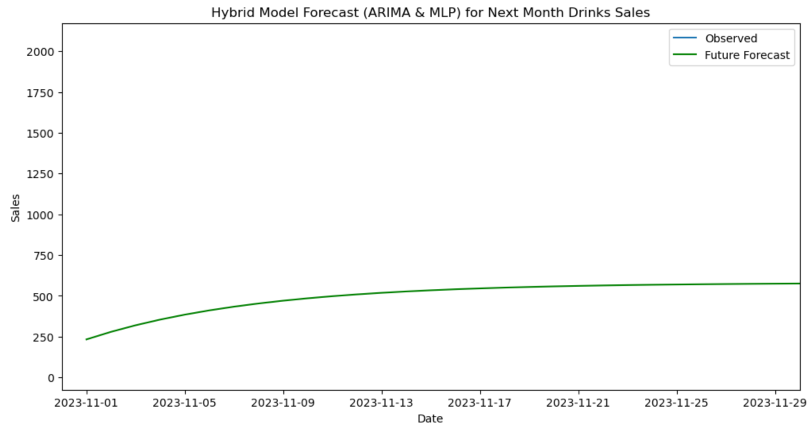


Figure 4. 110 The (ARIMA-MLPNNs) model forecasting for November drinks products sales.

Results have shown that hybrid model of (ARIMA-LSTM) was able to make forecasting very well, as it seems able to capture more complex and temporal dependencies and patterns in data such as trend and seasonality together. The drinks sales witness trend and seasonality together. For the forecasting of the hybrid model (ARIMA-MLPNNs), it was appeared to be as increasing trend, MLPNNs models may struggle to effectively capture more complex patterns such as the simultaneous appearance of trend and seasonality in drinks data.

For snacks & chips products sales, Figure 4.111 and Figure 4.112 show future forecasting for next month snacks & chips sales using (ARIMA-LSTM) and (ARIMA-MLPNNs). Figure 4.113 and Figure 4.114 show closer view of the models movements in November forecasting using (ARIMA-LSTM) and (ARIMA-MLPNNs).

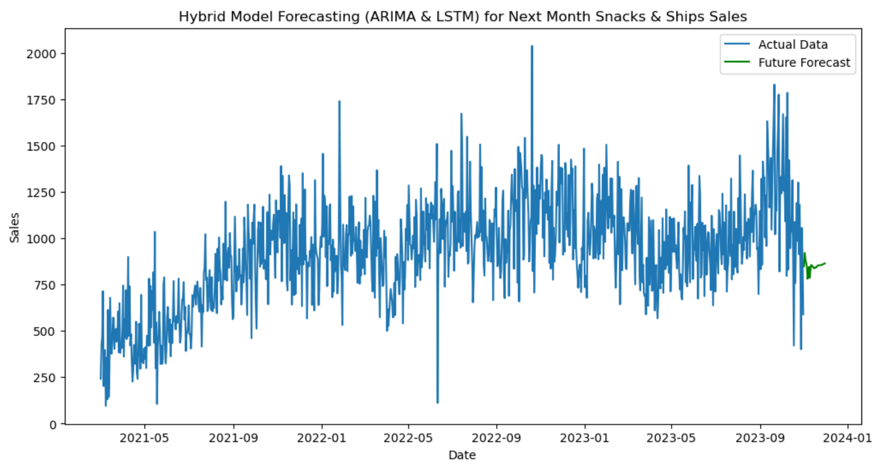

Figure 4. 111 The (ARIMA-LSTM) model forecasting for next month snacks & chips products sales.

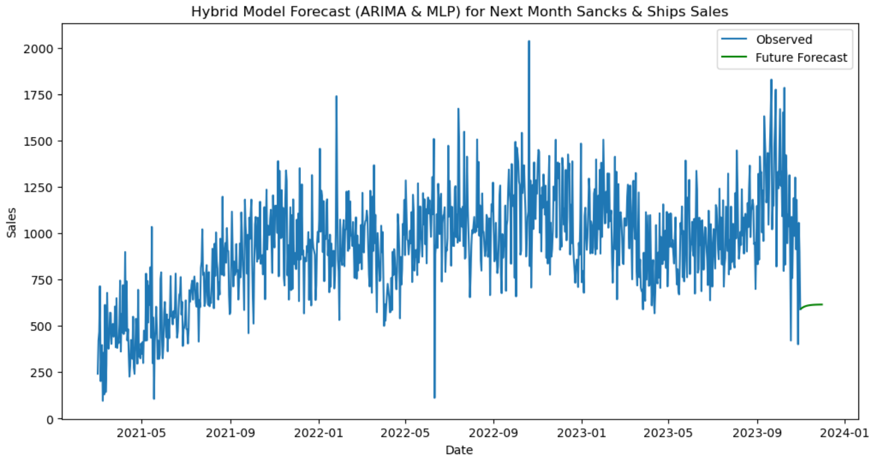


Figure 4. 112 The (ARIMA-MLPNNs) model forecasting for next month snacks & chips products sales.

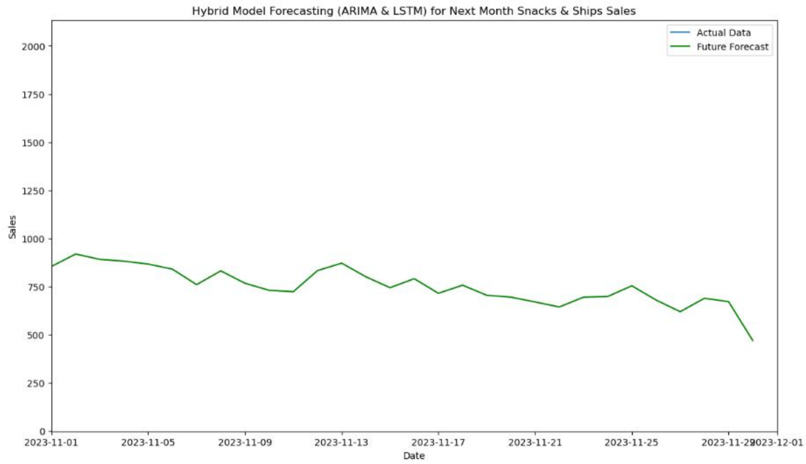


Figure 4. 113 The (ARIMA-LSTM) model forecasting for November snacks & chips products sales.

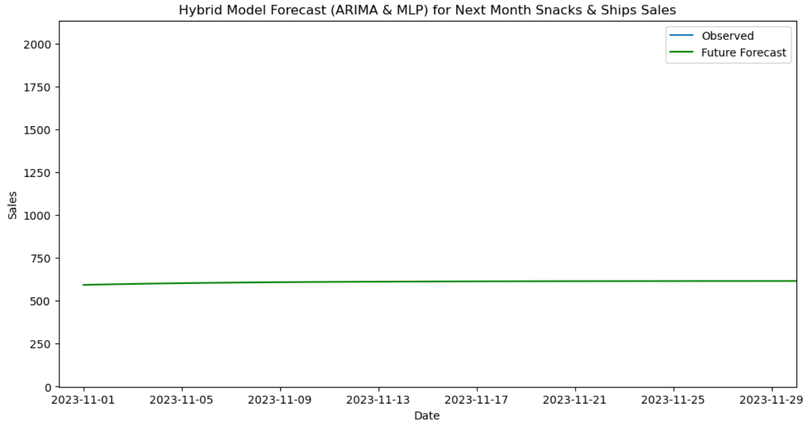


Figure 4. 114 The (ARIMA-MLPNNs) model forecasting for November snacks & chips products sales.

Results have shown that the hybrid model of (ARIMA-LSTM) was able to make forecasting very well, as it seems able to capture more complex and temporal dependencies and patterns in data such as trend and seasonality together. The snacks & chips sales witness trend and seasonality together. For the hybrid model forecasting of (ARIMA-MLPNNs), it was stable and struggled to effectively capture more complex patterns such as the simultaneous appearance of trend and seasonality in snacks & chips data.

For cleaning materials products sales, Figure 4.115 and Figure 4.116 show future forecasting for next month's cleaning materials sales using (ARIMA-LSTM) and (ARIMA-MLPNNs). Figure 4.117 and Figure 4.118 show a closer view of the models' movement in the November forecasting using (ARIMA-LSTM) and (ARIMA-MLPNNs).

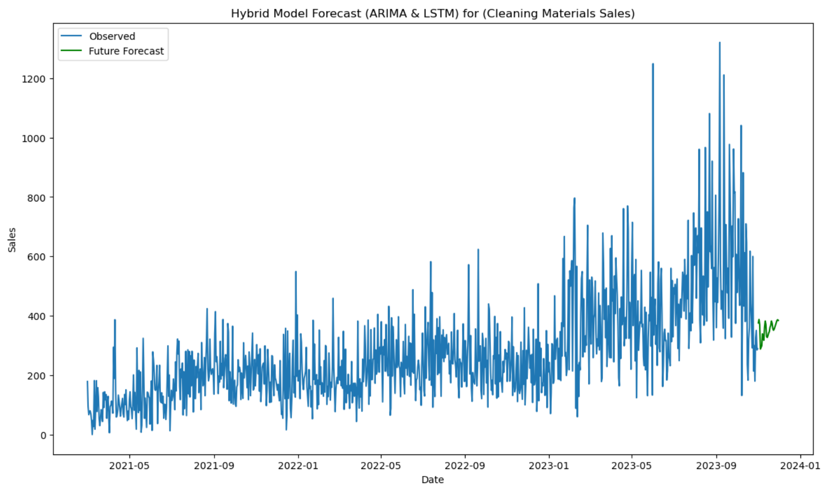


Figure 4. 115 The (ARIMA-LSTM) model forecasting for next month cleaning materials products sales.

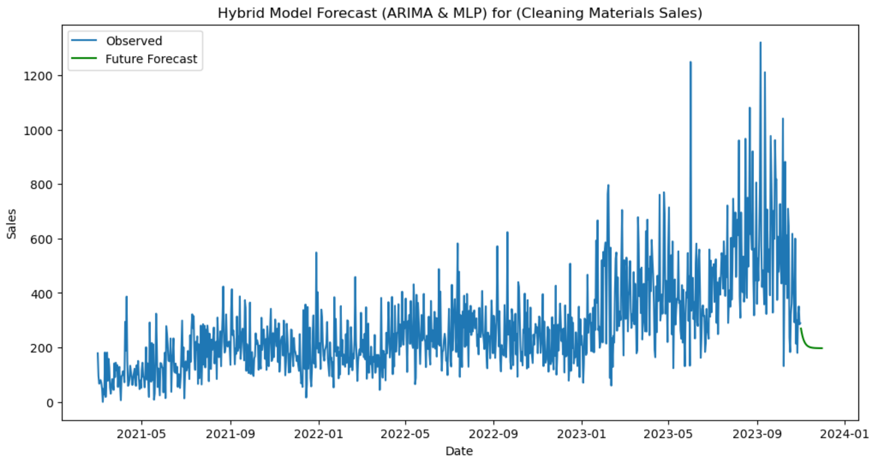


Figure 4. 116 The (ARIMA-MLPNNs) model forecasting for next month cleaning materials products sales.

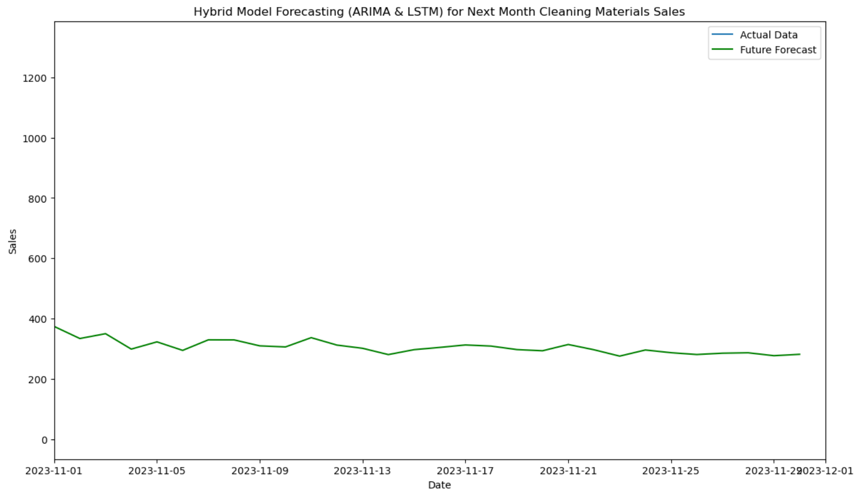


Figure 4. 117 The (ARIMA-LSTM) model forecasting for November cleaning materials products sales.

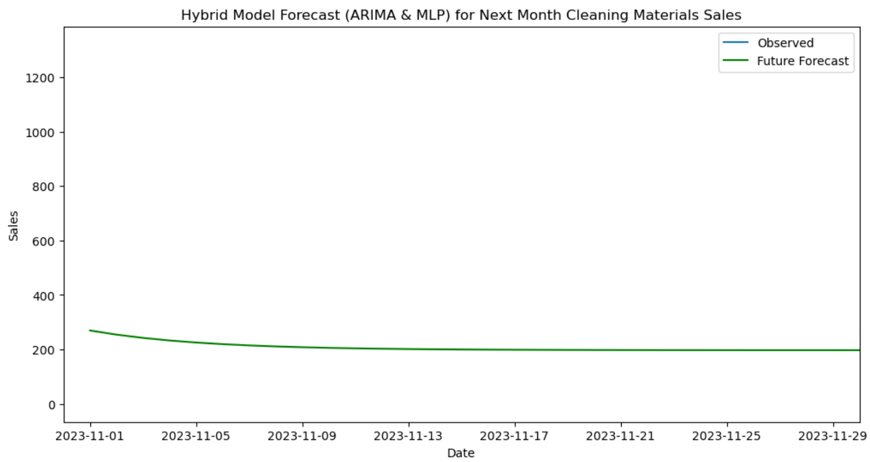

Figure 4. 118 The (ARIMA-MLPNNs) model forecasting for November cleaning materials products sales.

Results have shown that the hybrid model of (ARIMA-LSTM) was able to make forecasting very well, as it seems able to capture more complex and temporal dependencies and patterns in data such as trend and seasonality together. The cleaning materials sales witness trend and seasonality together. For the hybrid model forecasting of (ARIMA-MLPNNs), it has slightly decreased trend and then was stable and struggled to effectively capture more complex patterns such as the simultaneous appearance of trend and seasonality in cleaning materials data.

# Chapter Five: Discussion

# 5. Conclusion & Future Works

### 5.1 Conclusion

In the light of rapid development that the business sector is constantly witnessing, especially in the retail sector and among small to medium businesses. The ability to accurately forecast sales is essential for business sustainability across various industries. Nowadays, with the coming of artificial intelligence (AI) and machine learning technologies, alongside traditional statistical methods, businesses have proficient mechanisms to derive key insights from large amounts of data and make informed decisions. Sales forecasting using AI and statistical techniques has appeared as a revolutionary approach, allowing organizations to predict future sales trends with unprecedented accuracy. Employing the predictive capabilities of AI models affects inventory management optimization, customer satisfaction, the ability to meet customer needs and preferences, and overall business performance. This thesis applies two statistical models (ARIMA and SARIMA) and four advanced machine learning models (LSTM, RNNs, MLPNNs, and RBFNNs) both individually and as hybrid models to gain an advantage of both statistical and neural network characteristics. These models have been applied to local supermarket product sales located in Ramallah (Dukkan 11) to predict the supermarket sales for next month under two scenarios: the first scenario combines the five categories of sales, and the other one, based on each product level (individually). This approach allows us to know which is better for forecasting sales, would it be better to combine the sales of the product, or to forecast sales of each product separately. The model performance was evaluated depending on the best error metrics (MSE, RMSE, and MAE). It is worth noting that hybrid models, which combines statistical models and neural networks models, were also applied. These compared to decide whether applying models as an individual model would provide better prediction performance than in the form of a hybrid model. The weighted average technique was used to combine individual models, which weighed the contribution of each sub-model to the combined prediction through the expected performance of the sub-model. Datasets were gathered from the point of sale (POS) of a local supermarket in Ramallah (Dukkan 11) covering five product sales from (01-03-2021) to (31-10-2023). For training purposes, 80% of the original data was used, and 20% were reserved for testing. The predicted supermarket sales for both scenarios as combined product sales and based on each product level were given for the next month. After conducting experiments using statistical models (ARIMA & SARIMA) and neural networks models (RNNs, LSTM, MLPNNs and RBFNNs) individually and as hybrid models (ARIMA & RNNs), (ARIMA & LSTM), (ARIMA & MLPNNs) and (ARIMA & RBFNNs), the analysis of sales forecasting performance shows that modeling individual products achieve better results compared to forecasting combined products sales. Several key factors may account for these results. Firstly, each product demonstrates unique sales patterns impacted by factors such as seasonality, consumer preferences, supermarket location, local events, and holidays. By applying models to individual products, forecasting models are able to capture these behaviors and factors better, resulting in more accurate predictions. Secondly, products often show changing levels of demand, trend behavior, and seasonality. However, modeling products separately allows forecasting models to adjust to these dynamics more efficiently, and increases the forecasting accuracy. In addition, preserving data accuracy by modeling individual products allows the model to take advantage of accurate information, which might be fogged when combining sales data for five products. Finally, focusing on individual products enhances forecasting accuracy by preventing the need to fence possible conflicting trends or seasonality through different products. However, by focusing the forecasting approach on the special characteristics of each product, models can provide insightful predictions, which affect decision-making and business outcomes. Inspecting the performance of hybrid forecasting models under two scenarios as combined sales for five products (dairies, ice cream, drinks, snacks & chips, and cleaning materials), and based on each level of products (individually), it grants valuable insights into their efficiency in predicting future sales. In scenario one, as combined products sales, the hybrid models, combining (ARIMA & MLPNNs) presented better forecasting results compared with ARIMA & MLPNNs individually, as well as all other models. This was followed by the (ARIMA & LSTM), which presented better forecasting results compared with LSTM individually and all other models but not for ARIMA. Despite the benefits of combining both models, the forecasts didn't live up to expectations, pointing to difficulties in capturing complex patterns in sales data, which appeared in future sales forecasting in November 2023.On the contrary, scenario two analysis of hybrid forecasting models (ARIMA & MLPNNs) followed by (ARIMA & LSTM) detects interesting insights into the dynamics of future sales prediction, specifically for the next month's sales in November 2023. Our analysis reveals a significant variation in the performance of these two hybrid models, presenting their distinguished strengths and weaknesses across five products for making future forecasting sales for the next month (November 2023). Notably, the hybrid model of (ARIMA & LSTM) showed marked ability in capturing complex sales patterns observed in specific products such as drinks, snacks & chips, and cleaning materials products. Moreover, an intriguing observation appeared, focusing on the relative success of the hybrid model in predicting ice cream sales which has consistent purchasing behavior in November. The second hybrid model (ARIMA & MLPNNs) faced notable challenges, struggling to produce accurate forecasts for certain products within November. The hybrid model was not able to make forecasting for more complex patterns that appeared in some product sales such as snacks & chips and cleaning materials products, and appeared as stable forecasting for specific sales value, while for some products such as drinks and dairies was able to make future forecasting as increasing and decreasing trends. This suggests that while the hybrid models struggled with future forecasting for some products, they presented the potential for capturing specific trends and seasonality effectively.

## 5.2 Challenges and Limitations

During this study, several challenges and limitations were raised. One considerable issue was the Point of Sale (POS) device, which was not able to provide an expanded period of sales data because of software limitations. This limited the amount of historical data for analysis. Moreover, different programming challenges were raised, including difficulties in implementing and optimizing some machine learning models. Additionally, the limited computational power of my laptop device caused constraints on running complex models, which required more time. These challenges spot areas for more enhancements and considerations in future research efforts.

## 5.3 Future Works

For future work, it is significant to expand the size of the dataset with extended period and integrating more categories, which would provide more thorough analysis and enhance the models performance. More discovering of the models should include applying them within real-world applications, guaranteeing their practical relevance and utility. Moreover, experimenting more hyper parameters for neural networks, and parameters for statistical models to improve models forecasting performance, experimenting the SARIMA model with other neural networks as part of hybrid models component, and other broader range of hybrid models and configurations could reveal new forecasting abilities between various methodologies such as genetic algorithms with neural networks and deep learning hybrids such as (CNN-LSTM).These points will help to a profound understanding of the models power and limitations, to enhance predictive performance.

## References


*A Guide to Time Series Analysis in Python | Built In*. (n.d.). Builtin.com. https://builtin.com/data-science/time-series-python.

Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing*, *7*(1), 136–144. https://doi.org/10.1016/j.asoc.2005.06.001.

Alon, I., Qi, M., & Sadowski, R. J. (2001). Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, *8*(3), 147–156. https://doi.org/10.1016/s0969-6989(00)00011-4.

Alwan, L. C., & Roberts, H. V. (1988). Time-Series Modeling for Statistical Process Control. *Journal of Business & Economic Statistics*, *6*(1), 87–95. https://doi.org/10.1080/07350015.1988.10509640.

Arunraj, N. S., & Ahrens, D. (2015). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. *International Journal of Production Economics*, *170*, 321–335. https://doi.org/10.1016/j.ijpe.2015.09.039.

*Augmented Dickey-Fuller Test (ADF) in Time Series Analysis - Google Search*. (2020). Google.com.https://www.google.com/search?q=Augmented+Dickey-Fuller+Test+(ADF)+in+Time+Series+Analysis&oq=Augmented+Dickey-Fuller+Test+(ADF)+in+Time+Series+Analysis&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIKCAEQABiABBiiBNIBBzc4NWowajeoAgCwAgA&sourceid=chrome&ie=UTF-8#fpstate=ive&vld=cid:a81ad49a.

Babu, C. N., & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. *Applied Soft Computing*, *23*, 27–38. https://doi.org/10.1016/j.asoc.2014.05.028.

Bhatnagar, S. (2023, October 16). *Ensemble Methods in Machine Learning*. Medium. https://medium.com/@shashank25.it/ensemble-methods-in-machine-learning-2d4cc7513c77.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*. Springer.

Broomhead, D. S., & Lowe, D. (1988). Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems*, *2*, 321–355.

Chase, C. W. (2013). *Demand-Driven Forecasting: A Structured Approach to Forecasting, 2nd Edition.* John Wiley & Sons.

Chase, W, C., & Jr. (2014). Innovations in Business Forecasting: Predictive Analytics. *The Journal of Business Forecasting; Flushing*, *33*(2), 26–30, 32.

Chopra, S., & Meindl, P. (2019). *Supply Chain Management: Strategy, Planning and Operation* (7th ed.). Pearson Education.

De Almeida, F. (2021). *Sales forecasting for a supermarket chain in Natal, Brazil: an empirical assessment*. Universidade Federal do Rio Grande do Norte – UFRN.

Denny. (2015, September 17). Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs [Review of *Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs*]. *Denny's Blog*.https://dennybritz.com/posts/wildml/recurrent-neural-networks-tutorial-part-1/.

Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., & Moncada-Herrera, J. A. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*, *42*(35), 8331–8340. https://doi.org/10.1016/j.atmosenv.2008.07.020.

Dorne, G. (n.d.). *Neural Networks for Time Series Processing*. 26. Retrieved May 11, 2023, from https://machine-learning.martinsewell.com/ann/Dorf96.pdf.

Dr. Eng. Yousef Abuzir, & Mr. Abdul Rahman M.Baraka. (2019). Financial Stock Market Forecast Using Data Mining in Palestine. Palestinian Journal of Technology and Applied Sciences (PJTA), 2(1), 38–48. https://doi.org/10.5281/zenodo.2576331.

Dugan, M. T., Shriver, K. A., & Silhan, P. A. (1994). How to Forecast Income Statement Items for Auditing Purposes. *The Journal of Business Forecasting Methods & Systems*, *13*(2), 22.

E., B., & K., N. (2023). Sales Forecasting and Organizational Performance in Bakery Industry in Nigeria. *British Journal of Management and Marketing Studies*, *6*(1), 97–112. https://doi.org/10.52589/bjmms-mcg2fuin.

E., B., & K., N. (2023). Sales Forecasting and Organizational Performance in Bakery Industry in Nigeria. *British Journal of Management and Marketing Studies*, *6*(1), 97–112. https://doi.org/10.52589/bjmms-mcg2fuin.

Eglite, L., & Birzniece, I. (2022). Retail Sales Forecasting Using Deep Learning: Systematic Literature Review. *Complex Systems Informatics and Modeling Quarterly*, *30*, 53–62. https://doi.org/10.7250/csimq.2022-30.03.

Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. *International Journal of Information Management Data Insights*, *2*(1), 100058. Sciencedirect. https://doi.org/10.1016/j.jjimei.2022.100058.

*Ensemble averaging (machine learning)*. (2021, November 30). Wikipedia. https://en.wikipedia.org/wiki/Ensemble_averaging_ (machine-learning).

Florance, M, M., Sawicz, & S, M. (1993). Positioning sales forecasting for better results .*The Journal of Business Forecasting Methods & Systems; Flushing*, *12*(4), 20.

Franses, P. H., & Draisma, G. (1997). Recognizing changing seasonal patterns using artificial neural networks. *Journal of Econometrics*, *81*(1), 273–280. https://doi.org/10.1016/s0304-4076(97)00047-x.

Garcia-Pedrajas, N., Hervas-Martinez, C., & Munoz-Perez, J. (2003). COVNET: a cooperative coevolutionary model for evolving artificial neural networks. *IEEE Transactions on Neural Networks*, *14*(3), 575–596. https://doi.org/10.1109/tnn.2003.810618.

George & Jenkins, G. M. (1976). *Time Series Analysis*.
Georgoff, D. M., & Murdick, R. G. (1986). *Manager's Guide to Forecasting*.

Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, *23*(5), 499–510. https://doi.org/10.1016/s0261-5177(02)00009-2.

Granger, J., & Newbold, P. (2014). *Forecasting Economic Time Series*. Academic Press.

Haataja,T.(2016). *Sales Forecasting in Small and Medium-Sized Enterprises*. https://www.theseus.fi/bitstream/handle/10024/106191/Haataja_Timo.pdf?sequence=1&isAllowed=y.

Hamzacebi, C. (2008). Improving artificial neural networks' performance in seasonal time series forecasting. *Information Sciences*, *178*(23), 4550–4559. https://doi.org/10.1016/j.ins.2008.07.024.

Hibon, M., & Evgeniou, T. (2005). To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting*, *21*(1), 15–24. https://doi.org/10.1016/j.ijforecast.2004.05.002.

Hill, T., O'Connor, M., & Remus, W. (1996). Neural Network Models for Time Series Forecasts. *Management Science*, *42*(7), 1082–1092. https://doi.org/10.1287/mnsc.42.7.1082.

Huber, J., Gossmann, A., & Stuckenschmidt, H. (2017). Cluster-based hierarchical demand forecasting for perishable goods. *Expert Systems with Applications*, *76*, 140–151. https://doi.org/10.1016/j.eswa.2017.01.022.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). Otexts.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). Otexts.

Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition. *Procedia Computer Science*, *48*, 173–179. https://doi.org/10.1016/j.procs.2015.04.167.

Kleiner, B. (1977). Time Series Analysis: Forecasting and Control. *Technimetrics*, *19*(3), 343–344. https://doi.org/10.1080/00401706.1977.10489562.

Lawless, & Mark. (2014). Predictive Analytics: An Opportunity for Better Demand Planning and Forecasting. *Journal of Business Forecasting*, *33*(4), 44.

Leung, C. K.-S., MacKinnon, R. K., & Wang, Y. (2014). A machine learning approach for stock price prediction. *Proceedings of the 18th International Database Engineering & Applications Symposium on - IDEAS '14*. https://doi.org/10.1145/2628194.2628211.

Loureiro, A. L. D., Miguéis, V. L., & da Silva, L. F. M. (2018). Exploring the use of deep neural networks for sales forecasting in fashion retail. *Decision Support Systems*, *114*, 81–93. https://doi.org/10.1016/j.dss.2018.08.010.

Luxhøj, J. T., Riis, J. O., & Stensballe, B. (1996). A hybrid econometric—neural network modeling approach for sales forecasting. *International Journal of Production Economics*, *43*(2-3), 175–192. https://doi.org/10.1016/0925-5273(96)00039-4.

Marie-Aude Aufaure, & ZimányiE. (2013). Business Intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July 15-21, 2012, Tutorial Lectures. Springer Berlin Heidelberg.

Mondal, P., Shit, L., & Goswami, S. (2014). Study of Effectiveness of Time Series Modeling (Arima) in Forecasting Stock Prices. *International Journal of Computer Science, Engineering and Applications*, *4*(2), 13–29. https://doi.org/10.5121/ijcsea.2014.4202.

Moon, M. A. (2018). *Demand and supply integration the key to world-class demand forecasting*. Boston, Mass. Berlin Walter De Gruyter Inc.

Nelson, Hill, Remus, & O'Connor. (1994). Can neural networks applied to time series forecasting learn seasonal patterns: an empirical investigation. *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences HICSS-94*. https://doi.org/10.1109/hicss.1994.323316.

Nilsson, N. J. (2003). *Artificial intelligence: a new synthesis*. Kaufmann.

O'Donovan, T. M. (1983). *Short Term Forecasting*. John Wiley & Sons.

Polanski, A., & Stoja, E. (2017). Forecasting multidimensional tail risk at short and long horizons. *International Journal of Forecasting*, *33*(4), 958–969. https://doi.org/10.1016/j.ijforecast.2017.05.005.

*Python code for weighted average ensemble - Google Search*. (n.d.). Www.google.com. Retrieved May 11,2024,fromhttps://www.google.com/search?q=python+code+for+weighted+average+ensemble&oq=python+&gs_lcrp=EgZjaHJvbWUqBggCECMYJzIHCAAQABiPAjIGCAEQRRhAMgYIAhAjGCcyBwgDEAAYgAQyBwgEEAAYgAQyBggFEEUYPDIGCAYQRRg8MgYIBxBFGDzSAQg0MjQyajBqN6gCALACAA&sourceid=chrome&ie=UTF-8#fpstate=ive&vld=cid:023ec7a7.

Ramos, P., Santos, N., & Rebelo, R. (2015). Performance of state space and ARIMA models for consumer retail sales forecasting. *Robotics and Computer-Integrated Manufacturing*, *34*, 151–163. https://doi.org/10.1016/j.rcim.2014.12.015.

Rivas, V. M., Merelo, J. J., Castillo, P. A., Arenas, M. G., & Castellano, J. G. (2004). Evolving RBF neural networks for time-series forecasting with EvRBF. *Information Sciences*, *165*(3-4), 207–220. https://doi.org/10.1016/j.ins.2003.09.025.

*Sales Forecasting Management: A Demand Management Approach*. (2005). https://doi.org/10.4135/9781452204444.
*Sales Forecasting*. (n.d.). DealHub. https://dealhub.io/glossary/sales-forecasting/.

Silvia Priscila, S., Suman Rajest, S., Tadiboina, S. N., Regin, R., & András, S. (2023). Analysis of Machine Learning and Deep Learning Methods for Superstore Sales Prediction. *FMDB Transactions on Sustainable Computer Letters*, *1*(1), 1–11.

Singh, S. (2016). Forecasting: Academia versus Business. *Foresight: The International Journal of Applied Forecasting, International Institute of Forecasters*, *41*, 46-47.

Staff, S. (2024, May 6). *Demand Forecasting: A Complete Guide*. Salesforce.

Tandon, C., Revankar, S., Palivela, H., & Parihar, S. S. (2021). How can we predict the impact of the social media messages on the value of cryptocurrency? Insights from big data analytics. *International Journal of Information Management Data Insights*, *1*(2), 100035. https://doi.org/10.1016/j.jjimei.2021.100035.

Tkáč, M., & Verner, R. (2016). Artificial neural networks in business: Two decades of research. *Applied Soft Computing*, *38*, 788–804. https://doi.org/10.1016/j.asoc.2015.09.040.

*What is ensemble averaging (machine learning)? | Autoblocks Glossary*. (n.d.). Www.autoblocks.ai. https://www.autoblocks.ai/glossary/ensemble-averaging.

Wikipedia Contributors. (2019, May 7). *Augmented Dickey–Fuller test*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Augmented_Dickey%E2%80%93Fuller_test.

Yan, X.-B., Wang, Z., Yu, S.-H., & Li, Y.-J. (2005). *Time series forecasting with RBF neural network*. https://doi.org/10.1109/icmlc.2005.1527764.

Yang, Y., Fan, C., & Xiong, H. (2021). A novel general-purpose hybrid model for time series forecasting. *Applied Intelligence*, *52*(2), 2212–2223. https://doi.org/10.1007/s10489-021-02442-y.

Yin, P., Dou, G., Lin, X., & Liu, L. (2020). A hybrid method for forecasting new product sales based on fuzzy clustering and deep learning. *Kybernetes*, *49*(12), 3099–3118. https://doi.org/10.1108/k-10-2019-0688.

Zhang, G. P., & Kline, D. M. (2007). Quarterly Time-Series Forecasting With Neural Networks. *IEEE Transactions on Neural Networks*, *18*(6), 1800–1814. https://doi.org/10.1109/tnn.2007.896859.

Zhang, G. Peter. & Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, *160*(2), 501–514. https://doi.org/10.1016/j.ejor.2003.08.037.

Zhang, G. Peter. (2003). Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*, *50*, 159–175. https://doi.org/10.1016/s0925-2312(01)00702-0.

Zhao, K., & Wang, C. (2017, August 26). Sales Forecast in E-commerce using Convolutional Neural Network. *ArXiv.org*. https://doi.org/10.48550/arXiv.1708.07946.

## Appendices

Appendix 1 The descriptive analysis for cleaning materials category monthly sales.

| Year | Month | Count | Mean | Std | Min | 25% | 50% | 75% | Max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | Mar | 31.0 | 89.54 | 49.42 | 0.00 | 53.90 | 80.00 | 127.98 | 181.96 |
| | Apr | 30.0 | 108.53 | 74.03 | 6.00 | 64.75 | 92.06 | 119.75 | 386.94 |
| | May | 31.0 | 119.17 | 75.39 | 8.00 | 73.50 | 118.00 | 143.50 | 324.21 |
| | Jun | 30.0 | 142.60 | 73.81 | 13.00 | 101.25 | 136.50 | 190.95 | 298.89 |
| | Jul | 31.0 | 190.63 | 76.45 | 64.00 | 132.50 | 174.00 | 268.25 | 321.87 |
| | Aug | 31.0 | 213.01 | 71.97 | 76.00 | 178.50 | 221.00 | 243.60 | 424.00 |
| | Sep | 30.0 | 227.85 | 88.70 | 102.50 | 161.55 | 206.44 | 267.59 | 413.99 |
| | Oct | 31.0 | 197.29 | 62.05 | 95.00 | 152.50 | 206.50 | 230.72 | 342.00 |
| | Nov | 30.0 | 203.41 | 61.03 | 97.50 | 161.37 | 206.91 | 245.50 | 302.92 |
| | Dec | 31.0 | 187.13 | 120.46 | 16.00 | 117.46 | 149.50 | 251.55 | 549.00 |
| 2022 | Jan | 31.0 | 187.09 | 76.00 | 53.00 | 149.21 | 181.00 | 214.79 | 384.84 |
| | Feb | 28.0 | 195.97 | 82.15 | 77.00 | 146.75 | 173.05 | 226.61 | 459.00 |
| | Mar | 31.0 | 176.84 | 78.10 | 43.97 | 137.47 | 164.00 | 198.46 | 382.16 |
| | Apr | 30.0 | 232.11 | 83.44 | 79.00 | 179.65 | 213.40 | 277.77 | 405.08 |
| | May | 31.0 | 256.45 | 93.43 | 65.00 | 212.75 | 243.38 | 319.75 | 431.93 |
| | Jun | 30.0 | 233.98 | 85.94 | 98.50 | 185.00 | 216.25 | 260.50 | 487.86 |
| | Jul | 31.0 | 281.44 | 114.71 | 92.00 | 194.91 | 273.00 | 356.45 | 582.38 |
| | Aug | 31.0 | 259.48 | 65.94 | 123.50 | 234.35 | 252.50 | 298.70 | 407.71 |
| | Sep | 30.0 | 258.63 | 117.45 | 112.00 | 171.15 | 233.00 | 316.66 | 623.71 |
| | Oct | 31.0 | 243.87 | 92.46 | 92.44 | 172.28 | 247.00 | 304.34 | 440.00 |
| | Nov | 30.0 | 217.69 | 83.76 | 100.00 | 156.49 | 184.22 | 280.95 | 427.11 |
| | Dec | 31.0 | 225.79 | 96.36 | 78.00 | 165.50 | 201.95 | 289.92 | 507.92 |
| 2023 | Jan | 31.0 | 300.27 | 137.19 | 70.50 | 198.44 | 277.00 | 335.70 | 667.00 |
| | Feb | 28.0 | 397.80 | 200.81 | 59.88 | 240.68 | 380.81 | 543.19 | 796.63 |
| | Mar | 31.0 | 389.41 | 126.78 | 178.50 | 298.78 | 372.74 | 490.67 | 678.60 |
| | Apr | 30.0 | 421.82 | 158.99 | 164.00 | 324.00 | 400.75 | 483.76 | 770.00 |
| | May | 31.0 | 373.66 | 143.51 | 124.00 | 272.73 | 372.00 | 456.00 | 714.60 |
| | Jun | 30.0 | 381.45 | 201.47 | 162.00 | 266.96 | 332.00 | 446.25 | 1249.00 |
| | Jul | 31.0 | 466.97 | 112.99 | 249.00 | 391.50 | 469.00 | 527.00 | 746.60 |
| | Aug | 31.0 | 594.62 | 197.86 | 309.00 | 450.00 | 598.00 | 683.00 | 1081.00 |
| | Sep | 30.0 | 630.40 | 250.50 | 323.00 | 447.50 | 585.28 | 767.62 | 1320.50 |
| | Oct | 31.0 | 448.13 | 210.69 | 131.50 | 296.00 | 428.00 | 588.00 | 1041.00 |

Appendix 2 The descriptive analysis for dairies category monthly sales.

| Year | Month | Count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | Mar | 31.0 | 85.58 | 45.79 | 7.00 | 50.33 | 88.24 | 111.95 | 182.90 |
| | Apr | 30.0 | 87.17 | 45.95 | 7.50 | 49.00 | 89.60 | 119.58 | 189.09 |
| | May | 31.0 | 87.92 | 32.72 | 23.50 | 68.50 | 84.50 | 108.00 | 153.50 |
| | Jun | 30.0 | 99.22 | 36.63 | 44.00 | 75.09 | 90.68 | 110.51 | 183.00 |
| | Jul | 31.0 | 111.23 | 29.72 | 48.00 | 93.00 | 107.00 | 130.10 | 172.00 |
| | Aug | 31.0 | 152.59 | 34.25 | 73.28 | 129.11 | 159.00 | 172.94 | 241.89 |
| | Sep | 30.0 | 158.16 | 47.67 | 73.20 | 125.93 | 164.77 | 201.63 | 229.37 |
| | Oct | 31.0 | 152.45 | 49.80 | 46.50 | 126.98 | 151.50 | 180.22 | 268.48 |
| | Nov | 30.0 | 175.05 | 44.32 | 100.48 | 144.47 | 165.96 | 202.05 | 326.93 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Dec | 31.0 | 156.91 | 47.41 | 75.97 | 119.39 | 155.17 | 177.00 | 265.14 |
| 2022 | Jan | 31.0 | 116.99 | 32.93 | 22.50 | 93.21 | 121.00 | 137.18 | 180.24 |
| | Feb | 28.0 | 158.11 | 46.02 | 83.42 | 124.31 | 153.47 | 183.82 | 261.68 |
| | Mar | 31.0 | 137.71 | 42.93 | 79.71 | 106.68 | 126.50 | 156.10 | 243.20 |
| | Apr | 30.0 | 154.66 | 50.09 | 67.09 | 117.73 | 144.82 | 181.55 | 260.97 |
| | May | 31.0 | 175.98 | 49.30 | 99.00 | 144.16 | 172.22 | 195.73 | 285.52 |
| | Jun | 30.0 | 174.18 | 41.65 | 104.44 | 145.34 | 169.17 | 199.12 | 271.44 |
| | Jul | 31.0 | 189.04 | 45.70 | 105.03 | 157.83 | 187.17 | 224.32 | 281.55 |
| | Aug | 31.0 | 218.05 | 56.19 | 120.15 | 182.88 | 209.39 | 248.92 | 357.43 |
| | Sep | 30.0 | 249.92 | 68.08 | 131.29 | 193.21 | 256.53 | 300.76 | 371.81 |
| | Oct | 31.0 | 260.61 | 68.89 | 138.50 | 204.07 | 266.50 | 288.47 | 408.03 |
| | Nov | 30.0 | 229.00 | 52.97 | 130.11 | 192.99 | 223.41 | 258.93 | 333.42 |
| | Dec | 31.0 | 209.85 | 52.72 | 101.41 | 178.00 | 205.83 | 245.20 | 326.97 |
| 2023 | Jan | 31.0 | 270.02 | 67.80 | 145.00 | 235.25 | 256.00 | 311.36 | 440.84 |
| | Feb | 28.0 | 292.11 | 91.65 | 149.43 | 226.73 | 289.00 | 339.40 | 596.00 |
| | Mar | 31.0 | 311.81 | 93.52 | 150.00 | 244.75 | 289.00 | 361.23 | 527.60 |
| | Apr | 30.0 | 253.87 | 71.22 | 116.00 | 199.90 | 237.13 | 311.70 | 417.77 |
| | May | 31.0 | 299.70 | 67.26 | 166.00 | 253.00 | 295.50 | 349.85 | 407.00 |
| | Jun | 30.0 | 291.56 | 73.39 | 129.00 | 239.25 | 293.00 | 345.27 | 463.00 |
| | Jul | 31.0 | 358.31 | 92.94 | 191.85 | 279.50 | 364.50 | 427.50 | 522.55 |
| | Aug | 31.0 | 404.07 | 74.02 | 265.00 | 368.30 | 389.60 | 458.80 | 569.00 |
| | Sep | 30.0 | 383.62 | 91.52 | 233.00 | 296.77 | 399.38 | 432.75 | 643.00 |
| | Oct | 31.0 | 366.81 | 159.03 | 80.00 | 278.80 | 337.00 | 446.30 | 818.57 |

Appendix 3 The descriptive analysis for ice-cream category monthly sales.

| Year | month | count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | Mar | 31.0 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.000 | 0.00 |
| | Apr | 30.0 | 42.07 | 39.07 | 0.00 | 7.74 | 30.44 | 65.12 | 140.00 |
| | May | 31.0 | 74.20 | 34.51 | 10.50 | 57.25 | 70.50 | 93.84 | 163.18 |
| | Jun | 30.0 | 50.97 | 30.82 | 10.50 | 29.75 | 47.00 | 67.68 | 140.50 |
| | Jul | 31.0 | 127.28 | 50.97 | 29.50 | 94.75 | 139.00 | 161.75 | 211.00 |
| | Aug | 31.0 | 152.70 | 67.18 | 37.00 | 105.00 | 148.90 | 170.94 | 330.97 |
| | Sep | 30.0 | 85.80 | 47.70 | 23.90 | 53.00 | 72.00 | 107.46 | 192.00 |
| | Oct | 31.0 | 50.50 | 29.43 | 6.00 | 33.00 | 45.00 | 58.41 | 124.00 |
| | Nov | 30.0 | 25.75 | 17.56 | 0.00 | 12.50 | 23.25 | 38.75 | 66.00 |
| | Dec | 31.0 | 10.09 | 10.94 | 0.00 | 0.00 | 6.50 | 13.50 | 40.50 |
| 2022 | Jan | 31.0 | 8.37 | 10.82 | 0.00 | 0.00 | 4.50 | 11.75 | 45.00 |
| | Feb | 28.0 | 24.69 | 27.62 | 0.00 | 2.50 | 16.00 | 44.37 | 96.00 |
| | Mar | 31.0 | 44.41 | 30.30 | 0.00 | 24.71 | 38.50 | 66.95 | 113.00 |
| | Apr | 30.0 | 125.59 | 36.31 | 66.92 | 100.39 | 116.99 | 158.68 | 207.87 |
| | May | 31.0 | 159.03 | 59.26 | 57.50 | 116.19 | 162.33 | 182.16 | 312.59 |
| | Jun | 30.0 | 184.93 | 50.46 | 88.00 | 154.43 | 173.81 | 217.36 | 281.98 |
| | Jul | 31.0 | 185.72 | 48.18 | 101.53 | 153.56 | 167.71 | 223.70 | 294.14 |
| | Aug | 31.0 | 192.24 | 75.97 | 87.93 | 140.16 | 187.45 | 242.55 | 381.12 |
| | Sep | 30.0 | 120.49 | 62.54 | 36.00 | 77.12 | 107.03 | 152.41 | 316.80 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Oct | 31.0 | 64.67 | 34.53 | 3.50 | 39.25 | 56.34 | 96.44 | 119.50 |
| | Nov | 30.0 | 24.61 | 16.16 | 0.00 | 14.87 | 21.00 | 30.15 | 86.00 |
| | Dec | 31.0 | 22.28 | 14.83 | 0.00 | 10.75 | 20.00 | 31.00 | 60.00 |
| 2023 | Jan | 31.0 | 27.86 | 39.06 | 0.00 | 3.50 | 15.00 | 34.84 | 197.78 |
| | Feb | 28.0 | 22.22 | 31.13 | 0.00 | 2.50 | 12.50 | 27.12 | 153.50 |
| | Mar | 31.0 | 24.56 | 26.20 | 0.00 | 2.50 | 18.00 | 37.00 | 92.50 |
| | Apr | 30.0 | 37.57 | 44.56 | 0.00 | 0.00 | 7.25 | 69.00 | 157.87 |
| | May | 31.0 | 95.92 | 58.20 | 30.00 | 60.00 | 78.80 | 110.28 | 271.00 |
| | Jun | 30.0 | 136.07 | 60.83 | 61.00 | 94.75 | 111.50 | 177.10 | 295.00 |
| | Jul | 31.0 | 174.44 | 62.01 | 46.00 | 138.74 | 168.00 | 212.50 | 317.00 |
| | Aug | 31.0 | 153.81 | 43.29 | 72.00 | 132.00 | 148.70 | 182.80 | 262.00 |
| | Sep | 30.0 | 113.88 | 49.88 | 33.00 | 83.74 | 108.00 | 146.52 | 283.50 |
| | Oct | 31.0 | 52.37 | 39.89 | 0.00 | 17.00 | 44.50 | 71.00 | 152.00 |

Appendix 4 The descriptive analysis for drinks category monthly sales.

| Year | month | count | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | Mar | 31.0 | 158.12 | 73.06 | 26.36 | 101.10 | 147.74 | 211.52 | 310.92 |
| | Apr | 30.0 | 240.35 | 84.83 | 33.70 | 191.87 | 240.90 | 295.49 | 493.00 |
| | May | 31.0 | 298.35 | 94.50 | 46.00 | 229.00 | 296.50 | 355.50 | 500.51 |
| | Jun | 30.0 | 343.91 | 65.36 | 235.50 | 295.50 | 327.85 | 390.62 | 478.73 |
| | Jul | 31.0 | 496.12 | 91.20 | 302.15 | 427.33 | 512.50 | 554.00 | 665.50 |
| | Aug | 31.0 | 634.79 | 82.89 | 495.39 | 582.64 | 615.50 | 713.45 | 824.31 |
| | Sep | 30.0 | 629.28 | 121.03 | 450.11 | 505.43 | 662.69 | 732.43 | 834.21 |
| | Oct | 31.0 | 678.46 | 116.62 | 502.17 | 578.88 | 687.92 | 752.52 | 925.00 |
| | Nov | 30.0 | 648.65 | 122.73 | 382.36 | 570.03 | 633.53 | 739.92 | 918.90 |
| | Dec | 31.0 | 531.55 | 111.05 | 386.35 | 440.63 | 516.44 | 576.67 | 853.87 |
| 2022 | Jan | 31.0 | 516.98 | 118.29 | 296.47 | 419.53 | 503.84 | 622.81 | 724.99 |
| | Feb | 28.0 | 497.62 | 88.08 | 328.97 | 451.68 | 498.43 | 559.80 | 643.36 |
| | Mar | 31.0 | 507.89 | 95.00 | 314.49 | 432.95 | 513.88 | 588.18 | 674.02 |
| | Apr | 30.0 | 581.51 | 142.13 | 334.57 | 483.08 | 522.10 | 672.62 | 915.15 |
| | May | 31.0 | 780.65 | 264.18 | 534.27 | 654.85 | 728.71 | 820.39 | 2070.37 |
| | Jun | 30.0 | 766.87 | 130.24 | 569.91 | 682.96 | 752.73 | 827.92 | 1152.32 |
| | Jul | 31.0 | 877.46 | 154.74 | 603.40 | 740.10 | 874.63 | 988.75 | 1174.48 |
| | Aug | 31.0 | 974.18 | 173.80 | 626.36 | 868.83 | 977.78 | 1105.76 | 1358.88 |
| | Sep | 30.0 | 967.83 | 138.10 | 688.27 | 882.54 | 969.67 | 1032.87 | 1311.28 |
| | Oct | 31.0 | 930.74 | 186.78 | 611.24 | 777.81 | 915.77 | 1055.35 | 1409.25 |
| | Nov | 30.0 | 784.83 | 130.01 | 595.63 | 667.50 | 778.70 | 863.09 | 1082.37 |
| | Dec | 31.0 | 642.12 | 179.66 | 363.23 | 497.63 | 664.46 | 751.63 | 1269.00 |
| 2023 | Jan | 31.0 | 619.72 | 125.90 | 361.81 | 546.46 | 619.00 | 693.00 | 947.62 |
| | Feb | 28.0 | 571.96 | 99.38 | 376.92 | 491.92 | 568.84 | 647.60 | 773.27 |
| | Mar | 31.0 | 623.47 | 128.54 | 226.87 | 547.85 | 635.73 | 698.86 | 821.91 |
| | Apr | 30.0 | 779.22 | 206.24 | 473.57 | 610.79 | 780.00 | 904.92 | 1239.98 |
| | May | 31.0 | 803.09 | 142.42 | 570.80 | 726.50 | 781.00 | 883.10 | 1207.82 |
| | Jun | 30.0 | 835.27 | 272.71 | 564.00 | 687.12 | 749.00 | 899.25 | 1861.70 |
| | Jul | 31.0 | 888.06 | 185.65 | 394.00 | 795.50 | 904.68 | 1002.25 | 1272.00 |
| | Aug | 31.0 | 1035.98 | 135.78 | 747.00 | 950.75 | 1019.00 | 1124.94 | 1343.00 |
| | Sep | 30.0 | 1136.35 | 196.42 | 789.00 | 1048.85 | 1133.75 | 1251.00 | 1540.00 |
| | Oct | 31.0 | 841.10 | 312.59 | 179.00 | 686.05 | 851.00 | 1059.00 | 1551.00 |

Appendix 5 The descriptive analysis for snacks & chips category monthly sales.

| year | month | count | mean | Std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | Mar | 31.0 | 425.32 | 156.03 | 95.60 | 365.87 | 429.42 | 509.36 | 714.30 |
| | Apr | 30.0 | 457.36 | 166.22 | 226.35 | 336.53 | 422.81 | 535.76 | 898.80 |
| | May | 31.0 | 500.96 | 198.17 | 106.00 | 371.15 | 435.50 | 608.25 | 1034.08 |
| | Jun | 30.0 | 558.06 | 103.70 | 362.48 | 494.12 | 553.09 | 601.55 | 782.38 |
| | Jul | 31.0 | 670.85 | 132.71 | 404.62 | 615.00 | 665.33 | 755.40 | 1021.00 |
| | Aug | 31.0 | 802.12 | 150.12 | 563.00 | 667.25 | 809.45 | 908.35 | 1197.00 |
| | Sep | 30.0 | 874.39 | 185.44 | 460.36 | 772.72 | 840.28 | 996.08 | 1181.51 |
| | Oct | 31.0 | 942.12 | 169.01 | 512.00 | 849.64 | 924.00 | 1062.71 | 1235.23 |
| | Nov | 30.0 | 1005.01 | 214.03 | 641.50 | 835.91 | 989.91 | 1148.70 | 1389.13 |
| | Dec | 31.0 | 902.43 | 188.90 | 567.32 | 838.28 | 901.14 | 965.41 | 1351.03 |
| 2022 | Jan | 31.0 | 982.40 | 243.14 | 531.92 | 827.08 | 960.77 | 1115.96 | 1739.51 |
| | Feb | 28.0 | 962.89 | 126.39 | 754.17 | 885.00 | 955.73 | 1029.25 | 1227.50 |
| | Mar | 31.0 | 951.02 | 180.68 | 573.41 | 817.64 | 917.77 | 1067.24 | 1367.08 |
| | Apr | 30.0 | 771.99 | 180.19 | 499.97 | 627.09 | 751.13 | 893.32 | 1179.80 |
| | May | 31.0 | 993.29 | 150.03 | 736.46 | 883.40 | 960.09 | 1128.08 | 1284.96 |
| | Jun | 30.0 | 1056.07 | 247.49 | 111.61 | 948.71 | 1101.71 | 1167.78 | 1508.86 |
| | Jul | 31.0 | 1081.99 | 229.81 | 654.33 | 931.31 | 1055.42 | 1144.68 | 1671.67 |
| | Aug | 31.0 | 1034.08 | 164.80 | 666.01 | 934.09 | 1010.48 | 1093.40 | 1505.76 |
| | Sep | 30.0 | 1010.53 | 196.62 | 676.28 | 852.90 | 1002.97 | 1182.32 | 1373.50 |
| | Oct | 31.0 | 1185.59 | 278.22 | 658.85 | 995.70 | 1216.53 | 1337.19 | 2037.09 |
| | Nov | 30.0 | 1167.20 | 203.86 | 776.21 | 1005.06 | 1223.60 | 1311.89 | 1504.40 |
| | Dec | 31.0 | 1065.99 | 232.75 | 709.49 | 884.00 | 1007.14 | 1336.16 | 1483.74 |
| 2023 | Jan | 31.0 | 1054.83 | 205.04 | 678.83 | 901.31 | 1036.28 | 1208.17 | 1504.48 |
| | Feb | 28.0 | 1061.38 | 194.45 | 643.25 | 907.57 | 1063.98 | 1219.88 | 1412.65 |
| | Mar | 31.0 | 967.85 | 221.35 | 588.99 | 756.50 | 982.00 | 1127.28 | 1325.00 |
| | Apr | 30.0 | 853.71 | 165.08 | 567.80 | 721.82 | 840.00 | 967.50 | 1156.00 |
| | May | 31.0 | 950.53 | 181.20 | 669.70 | 814.53 | 927.00 | 1094.31 | 1392.00 |
| | Jun | 30.0 | 949.04 | 161.17 | 637.66 | 856.92 | 938.50 | 1058.75 | 1336.67 |
| | Jul | 31.0 | 956.21 | 158.21 | 711.80 | 832.50 | 912.00 | 1068.08 | 1322.00 |
| | Aug | 31.0 | 1053.80 | 163.64 | 699.00 | 945.10 | 1032.00 | 1153.50 | 1446.70 |
| | Sep | 30.0 | 1287.72 | 266.23 | 820.00 | 1147.87 | 1291.88 | 1457.41 | 1829.00 |
| | Oct | 31.0 | 1087.50 | 333.58 | 400.50 | 868.500 | 1088.00 | 1306.89 | 1784.56 |

**التنبؤ بمبيعات المنتجات باستخدام نماذج الإحصاء والتعلم الالي – دراسة حالة**

**إسراء عبد الإله عبد الحميد خنفر**

**لجنة الإشراف : البروفيسور محمد عوض ، الدكتور أحمد عويس ، البروفيسور يوسف عوّاد ضراغمة.**

## ملخص

يعبتر التنبؤ بالمبيعات أداة محورية لإدارة الآعمال من مختلف التخصصات ، وأساس لبناء عملية تخطيط فعالة في مؤسسة الآعمال. تظهر أولوية أصحاب العمل بشكل أساسي في متابعتهم الدقيقة لتقديرات المبيعات للحد من مشاكل سوء تقدير المبيعات بالزيادة او النقصان وتأثيرها على العمليات والتكاليف.في هذه الأطروحة قمنا بتطبيق اثنين من النماذج الاحصائية هما المتوسط المتحرك الانحدار التلقائي (ARIMA ) و المتوسط المتحرك الموسمي ذاتي الانحدار الذاتي (SARIMA) بالاضافة الى أربع شبكات عصبية (NNs) وهي الشبكات العصبية المتكررة (RNNs) ، ذاكرة طويلة قصيرة المدى (LSTM) ، الشبكات العصبية متعددة الطبقات (MLPNNs) بالاضافة الى الشبكات العصبية ذات الآساس الشعاعي (RBFNNs) ، ثم تم دمج نموذج إحصائي مع كل نموذج من نماذج الشبكة العصبية لبناء أربعة نماذج هجينة.تبحث هذه الدراسة عن مدى كفاءة التنبؤ بالمبيعات والقدرة على التقاط أنماط مختلفة لخمسة منتجات ، كنماذج فردية ونماذج هجينة من الشبكات الاحصائية والعصبية .خلال الدراسة تم تطبيق سيناريوهين لتطبيق هذه النماذج ، السيناريو الآول عبارة عن دمج مبيعات المنتجات الخمسة باستخدام نموذجين احصائيين، بالاضافة الى أربعة نماذج للشبكات العصبية بشكل منفرد ، ثم مزيج من نموذج إحصائي مع الشبكات العصبية الاربعه للخروج بنموذج هجين. أما السيناريو الثاني كان بالاعتماد على مبيعات كل منتج على حدة. تم تقييم أداء النماذج للسيناريوهين باستخدام مقاييس الخطأ (MAE, RMSE & MAE).أظهرت النتائج أن النموذج الهجين (ARIMA–MLPNNs) ذات قدرة على التنبؤ بشكل أفضل مقارنة بالنماذج الاحصائية الفردية والنماذج الشبكية العصبية ، و النماذج الهجينة الاخرى ل (ARIMA–RNN, ARIMA–LSTM & ARIMA–RBFNN) لمبيعات المنتجات مجتمعه، حيث حقق النموذج الهجين بقرأته الخاصة بال RMSE 131.64 متبوعا بالنموذج الهجين ARIMA–LSTM الذي أظهر RMSE 447.68 متفوقا الأخير على النموذج الاحصائي الفردي SARIMA والنماذج الشبيكة العصبية الاربعة بالإضافة الى بقية النماذج الهجينة. أما بالنسبة للسيناريو الثاني، فقد حقق النموذج الهجين

ARIMA–MLPNNs بقراءة ال RMSE 31.13 لمنتجات الالبان ، 19.54 للمثلجات ، 51.74 للمشروبات ، 60.99 للتسالي والشيبس و 74.21 لمنتجات مواد التنظيف، بينما أظهر النموذج الهجين ARIMA–LSTM أداء أفضل من النموذج الاحصائي الفردي SARIMA وبقية الشبكات العصبية الفردية والنماذج الهجينة الاخرى بقراءة ال RMSE 80.54 لمنتجات الآلبان ، 50.64 للمثلجات ، 169.13 للمشروبات ، 188.03 للتسالي والشيبس و 167.93 لمواد التنظيف . تشير النتائج أن النماذج الهجينة يمكن أن توفر تنبؤات أكثر دقة وموثوقية للتنبؤ بمبيعات المنتجات.

كلمات مفتاحية:
التنبؤ بالمبيعات ، النماذج الاحصائية ، شبكات عصبية ، النماذج الهجينة ، مقاييس الخطأ.