Forecasting SKU Demand Using Time Series and Machine Learning Models

submitted in partial fulfillment of the requirements
for the degree of

BACHELOR OF TECHNOLOGY

in

CHEMICAL ENGINEERING

by

SWAPNIL SHARMA CH20B024

Supervisor(s)

Dr. M. Nabil

DEPARTMENT OF CHEMICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY TIRUPATI

DECEMBER 2023

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission to the best of my knowledge. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Place: Tirupati
Date: 03-12-2023

Signature
SWAPNIL SHARMA
CH20B024

BONA FIDE CERTIFICATE

This is to certify that the report titled **FORECASTING SKU DEMAND USING TIME SERIES AND MACHINE LEARNING MODELLING**, submitted by **Mr. Swapnil Sharma**, to the Indian Institute of Technology, Tirupati, for the award of the degree of **Bachelor of Technology**, is a bona fide record of the project work done by him under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Place: Tirupati
Date: 03-12-2023

Dr. M. Nabil
Guide
Assistant Professor
Department of Chemical Engineering
IIT Tirupati - 517619

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. M. Nabil for his invaluable guidance and unwavering support throughout the course of my B.Tech project. His expertise, encouragement, and insightful feedback have been instrumental in shaping the trajectory of my research.

I am particularly thankful for Dr. Nabil's mentor ship in the realm of forecasting SKU demand using time series and machine learning modelling. His deep understanding of the subject matter, along with his patience and dedication, has significantly enriched my learning experience and contributed to the success of this project.

I extend my heartfelt thanks to Dr. M. Nabil for being an exemplary guide, providing valuable insights, and fostering an environment of intellectual curiosity. His commitment to academic excellence has been a source of inspiration, and I am truly grateful for the opportunity to work under his guidance.

ABSTRACT

KEYWORDS: Time Series modelling, SKU Demand Forecasting, Time Series Regression, ARMA, AR, MA, ARIMA

In the dynamic landscape of modern business, accurate demand forecasting for Stock Keeping Units (SKUs) has become paramount for effective inventory management and overall operational efficiency. This report delves into the pivotal role of SKU demand forecasting, elucidating its significance in the context of contemporary business operations and shedding light on the intricate complications associated with the forecasting process.The first phase of our project concentrates on leveraging time series models, namely Auto-regressive (AR), Moving Average (MA), and their combination, Auto-regressive Moving Average (ARMA), to enhance the accuracy of SKU demand forecasting. Time series models have proven to be robust tools in capturing the temporal patterns inherent in SKU demand data. However, the forecasting process is not without its challenges, and this report meticulously outlines the complexities involved.To address these challenges and elevate the performance of the time series models, we employ sophisticated diagnostic tools such as Partial Auto-correlation Function (PACF) and Auto-correlation Function (ACF) plots. These tools play a crucial role in identifying and mitigating the impact of auto-correlation, enabling a more refined and precise forecasting model.By providing a comprehensive understanding of the importance of SKU demand forecasting in the broader business context and elucidating the intricacies of time series modeling with AR, MA, and ARMA, this report aims to contribute valuable insights for businesses seeking to optimize their inventory management processes and improve overall supply chain efficiency.

TABLE OF CONTENTS

ABBREVIATIONS

SKU	Stock Keeping Unit
AR	Auto Regressor
MA	Moving Average
ARIMA	Auto-regressive Integrated Moving Average
ARMA	Auto-regressive Moving Average
PACF	Partial Auto-correlation Function
ACF	Auto-correlation Function
ML	Machine Learning
DL	Deep Learning
RNN	Recurrent Neural Networks
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MSE	Mean Square Error

CHAPTER 1

INTRODUCTION

1.1 Objectives and Scope

The overarching goal of this B.Tech project is to comprehensively explore and harness the potential of Time Series Forecasting within the business domain. The specific objectives are outlined as follows:

1. Scope of Time Series Forecasting:
   (a) Investigate and analyze the extensive scope of Time Series Forecasting in the context of business operations.
   (b) Explore how Time Series Models can effectively incorporate and account for intricate patterns and external factors such as store closures, holidays, stockouts, advertising campaigns, pricing fluctuations, and sales trends.
   (c) Evaluate the significance of these factors in influencing demand patterns and their integration into forecasting models to enhance accuracy and relevance.
2. Performance of Time Series Modelling:
   (a) Assess the capabilities of Time Series Models, including Autoregressive (AR), Moving Average (MA), and their combination (ARMA), in generating accurate and reliable forecasts.
   (b) Explore the suitability of Time Series Models in capturing temporal dependencies and trends within historical data, leading to improved forecasting precision.
   (c) Investigate the adaptability and robustness of these models across diverse business scenarios, emphasizing their potential to enhance decision-making processes related to inventory management and supply chain optimization.
3. Role of PACF and ACF in Model Enhancement:
   (a) Investigate the utilization of diagnostic tools such as Partial Auto-correlation Function (PACF) and Auto-correlation Function (ACF) in evaluating and refining Time Series Models.
   (b) Explore how PACF and ACF plots can identify and address issues related to auto-correlation, thereby contributing to the optimization of model performance.
   (c) Emphasize the practical implementation of PACF and ACF to enhance the accuracy and reliability of Time Series Models, ensuring their effectiveness in real-world forecasting scenarios.
4. Significance of Historical Data:
   (a) Highlight the paramount importance of historical data in the Time Series Forecasting process.
   (b) Explore the relationships and dependencies within historical data, emphasizing their role in informing forecasting models and improving the understanding of demand patterns over time.

(c) Evaluate the critical role of historical data in building robust, adaptable models capable of handling diverse business conditions and contributing to strategic decision-making.

By achieving these objectives, the project aims to provide valuable insights and contribute to the advancement of Time Series Forecasting methodologies, particularly within the dynamic landscape of business operations.

CHAPTER 2

Literature Review

The literature survey presented here encapsulates key findings from diverse research efforts aimed at advancing SKU retail sales forecasting methodologies within distinct industry frameworks. In oneAndrade et al. (2022) study focusing on the grocery sector, the authors delve into the complexities of sales forecasting by integrating factors such as public holidays and item perishability. Employing the XG Boost machine learning algorithm, this research underscores the significance of advanced techniques in addressing the intricate challenges posed by perishable goods. Shifting the focus to computer productsElalem et al. (2023) characterized by short life cycles, another study adopts a hybrid approach, combining traditional time series models like ARIMAX with cutting-edge Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs). This hybrid modeling strategy recognizes the need for adaptability when forecasting sales for products with transient life spans. In a third research effort Taghiyeb et al. (2023), the emphasis is placed on the forecasting of SKU sales, particularly considering the impact of promotions and holidays. Employing a diverse set of machine learning algorithms, including Gradient Boosting, Random Forest, and XG Boost, alongside Artificial Neural Networks (ANNs), this study provides a holistic exploration of forecasting techniques. The inclusion of comparative analyses underscores the importance of accounting for promotional dynamics and holidays, while the integration of ANNs reflects a commitment to capturing intricate patterns in sales data. Collectively, these studies contribute to a nuanced understanding of SKU retail sales forecasting, showcasing the evolving landscape of methodologies tailored to industry-specific challenges.

Extending the literature survey to encompass an additional study, a research effortSeyedan et al. (2023) within the online retail industry has been investigated. This paper places emphasis on the multifaceted dynamics of online retail sales forecasting by integrating factors such as prices, discounts, and past demand. Moreover, the study delves into the critical aspect of safety stock prediction in conjunction with demand forecasting. Employing a sophisticated arsenal of machine learning and neural network methodologies, including Artificial Neural Networks (ANNs), Long Short-Term Memory networks

(LSTMs), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs), the research underscores the need for comprehensive modeling approaches in the online retail landscape. By exploring the interplay between pricing strategies, promotional discounts, and historical demand patterns, this study contributes to a nuanced understanding of the intricate dynamics inherent in online retail sales forecasting.

Through the comprehensive literature survey, several common variables emerge as pivotal elements in SKU forecasting, each playing a distinctive role in enhancing the accuracy and robustness of predictive models. Holidays, as highlighted in multiple studies, consistently prove to be a significant factor influencing SKU forecasting. The temporal disruptions introduced by holidays impact consumer behavior and purchasing patterns, necessitating the incorporation of these events into forecasting models. The literature underscores the importance of understanding how holidays can cause fluctuations in demand, enabling more precise predictions during such periods. PromotionsKourentzes et al. (2016) emerge as another crucial variable influencing SKU forecasting, as evidenced by the examined research efforts. Promotional activities introduce unique demand patterns, often characterized by sudden spikes or fluctuations. Accurate forecasting demands the incorporation of promotion-related variables, allowing models to adapt to the irregularities introduced by marketing campaigns and special events. Stock outs, observed across various studies, present a substantial challenge in SKU forecasting. Instances where products go out of stock can lead to missed sales opportunities and customer dissatisfaction. Literature emphasizes the necessity of accounting for stock out events, employing strategies such as safety stock prediction, to mitigate the impact on forecasting accuracy and maintain optimal inventory levels. Weekends, identified as a recurring variable in the literature, introduce temporal considerations into SKU forecasting. Consumer behavior often exhibits variations between weekdays and weekends, and models need to adapt to these patterns. The literature survey underscores the significance of incorporating day-of-week effects to capture the fluctuations associated with weekends, ensuring more precise predictions and effective inventory management.

CHAPTER 3

Time Series Modelling

3.1 AR, MA, ARMA and ARIMA

3.1.1 Auto Regressor (AR) model

Autoregressive models are regression models applied on lag series generated using the original time series. Recall in multiple linear regression, the output is a linear combination of multiple input variables. In the case of autoregression models, the output is the future data point and it can be expressed as a linear combination for past p data points. p is the lag window. The auto regressive model can be denoted as the equation:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \ldots + \alpha_p y_{t-p} + \varepsilon \tag{3.1}$$

$$\text{or} \qquad y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon \tag{3.2}$$

where ε is the noise. (α_{i}) are the coefficients that need to be learned from the data. This can be referred to as an auto-regressive model with p lags or an AR(p) model. In an AR(p) model, lag series is a new predictor used to fit the dependent variable, which is still the original series value,y_t. Auto-regressive models are remarkably flexible at handling a wide range of different time series patterns. To get the values of the coefficient there are various methods like Ordinary Least Squares, Gradient Descent and other types of the Gradient Descent algorithms for example Stochastic Batch Gradient Descent Algorithm.

3.1.2 Moving Average (MA) model

The moving average model (MA model)is a statistical approach for modeling univariate time series data. It specifies that the output variable is cross-correlated with a non-identical to itself random-variable. The MA model is a special case and key component of the more general ARMA and ARIMA models of time series. The finite MA model is always stationary. The moving-average model should not be confused with the moving average, a distinct concept despite some similarities.

The notation ‘MA(q)‘ refers to the moving average model of order ‘q‘. A moving-average model is conceptually a linear regression of the current value of the series against current and previous (observed) white noise error terms or random shocks. The random shocks at each point are assumed to be mutually independent and to come from the same distribution, typically a normal distribution, with location at zero and constant scale.

Fitting a moving-average model is generally more complicated than fitting an auto-regressive model. A moving-average model can be fit in the context of time-series analysis by smoothing the time series curve by computing the average of all data points in a fixed-length window. This technique is known as Moving Average Smoothing and can be used for data preparation, feature engineering, and forecasting.

$$y_t = \mu + \theta_0\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q} \tag{3.3}$$

$$\text{or} \qquad y_t = \sum_{i=0}^{q} \theta_i\varepsilon_{t-i} + \mu \tag{3.4}$$

Here, y_t is the time series at time t. μ is the mean of the time series. ($\varepsilon_{\text{t-i}}$) is the error in the previous forecasts at t-i. (θ_{q}) are the parameters of the model, representing the weights on past error terms. q is the order of the MA model.(μ) is the mean of the time series.

3.1.3 Auto-regressive Moving Average (ARMA) model

The Auto-regressive Moving Average (ARMA) model is a combination of Auto-regressive (AR) and Moving Average (MA) models. It’s a time-series model that includes both auto-regressive terms, representing the dependence on past values of the series, and moving average terms, representing the dependence on past forecast errors.
The general form of an ARMA(p, q) model is given by:

$$y_t = c + \theta_0\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \ldots + \theta_q\varepsilon_{t-q} + \alpha_p y_{t-p} \tag{3.5}$$

$$\text{or} \qquad y_t = \sum_{i=0}^{p} \theta_i\varepsilon_{t-i} + \sum_{i=1}^{p} \alpha_i y_{t-i} + c. \tag{3.6}$$

where all notations depicts the same as the AR and MA model. c is the only new term that is a constant.

3.1.4 Auto-regressive Integrated Moving Average (ARIMA) model

The Auto-regressive Integrated Moving Average (ARIMA) model is a popular time-series forecasting model that combines auto-regressive (AR) and moving average (MA) components with differencing to handle non-stationary time series. The notation for an ARIMA model is often denoted as ARIMA(p, d, q), where:

$$(1-\alpha_1 B-\alpha_2 B^2-\ldots-\alpha_p B^p)(1-B)^d Y_t=(1+\theta_1 B+\theta_2 B^2+\ldots+\theta_q B^q)\varepsilon_t \tag{3.7}$$

$$\text{or} \qquad (1-\sum_{i=1}^{p}\alpha_i B^i)(1-B)^d Y_t=(1+\sum_{j=1}^{q}\theta_j B^j)\varepsilon_t \tag{3.8}$$

where, all parametes as have same notations as mentioned in the AR, AR, ARMA and ARIMA. Instead B which is the back shift operator :

$$BY_t=Y_{t-1} \tag{3.9}$$

and d is the differencing order.

3.1.5 PACF

PACF stands for Partial Auto-Correlation Function. The PACF is a statistical measure used in time series analysis to describe the correlation between a certain observation and its lag, after accounting for the effects of other lags in between.n other words, the PACF at lag k is the correlation between the series and its own lag k values,controlling for the influence of all the lags from 1 to k1. It helps identify the direct relationship between the current observation and its (k_{th}) lag, excluding the indirect influences of the intermediate lags. The mathematical representation of PACF is :

$$\phi_{kk}=\frac{\gamma_{t,t-k}-(\sum_{i=1}^{k-1}\phi_{k-1,i})\gamma_{t-i,t-k}}{\sqrt{(1-\sum_{i=1}^{k-1}\phi_{k-1,i}^2)(1-\sum_{i=1}^{k-1}\phi_{k-1,i}^2)}} \tag{3.10}$$

where $\gamma_{t,t-k}$ is the auto-covariance or ACF value between the time series values at time t and $t-k$, and $\phi_{k-1,i}$ represents the PACF at lags 1 through $k-1$.

3.1.6 ACF

The auto-correlation function (ACF) is a statistical tool used to measure the correlation between a time series and its lagged values. It is a plot of the correlation coefficient between the observations at different lags. The ACF is used to identify the presence of seasonality in a time series. The mathematical representation of ACF is :

$$\rho_k = \frac{\mathrm{Cov}(y_t, y_{t-k})}{\sqrt{\mathrm{Var}(y_t) \cdot \mathrm{Var}(y_{t-k})}} \tag{3.11}$$

where the all notations are same as mentioned above. (ρ_{k}) is the the ACF at lag k.

3.2 Case Study

3.2.1 Description of Dataset

The dataset is taken from the Kaggle UNIVERSITY OF NICOSIA & Kaggle (2020) and is the classic M5 forecasting dataset of Walmart stores. The datasets contains the 1914 days of SKU data as the training set and the 28 days of the data for the evaluation. Dataset focuses on the 10 stores of the three states. The states are Texas, California and Wisconsin. Texas contains of 3 stores, Wisconsin 3 stores and California 4 stores. In total there are total 3049 SKUs. These SKUs are divided into three categories Hobbies, Household and Foods. And each category is also divide into subcategories. There are 2 sub categories in Hobbies, 2 in Households and 3 in Foods. Please refer the image for clear information.[Figure 3.1].

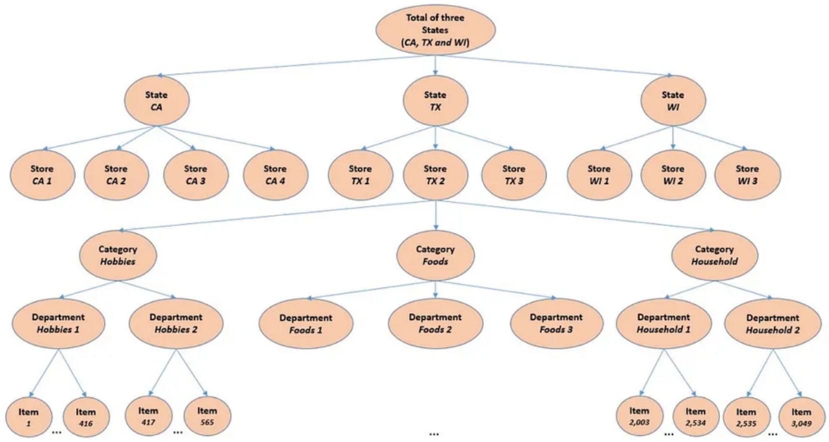

Figure 3.1: Flow Diagram of the dataset

3.3 Methodology

Now, let's talk about the methodology. Firstly, we have checked whether the time series that we want to forecast is stationary or not. In the context of time series analysis, a time series is considered stationary if its statistical properties, such as mean and variance, remain constant over time. On the other hand, a non-stationary time series is one whose statistical properties change over time. We have done this using Augmented Dickey-Fuller (ADF) test. The null hypothesis of the ADF test is that the time series possesses a unit root (i.e., it is non-stationary). The alternative hypothesis is that the time series is stationary. The test statistic is compared to critical values to make a decision regarding the null hypothesis. After that we have created the PACF plots for max lags 30 [Figure 3.3]. PACF helps you to find the order and the significant lags u should consider while creating the model. We have considered the most significant lags based on the P-value>0.05. We have chosen only the two most significant lags so to avoid the overfitting. We have got the parameters of the AR or coefficients of Y_t using the method of ordinary least squares. After predicting the forecast we move towards the MA model. In MA model the most important step is to get the previous errors in the forecast. So we have calculated the residuals now we can get the order of the MA model from the ACF plot [Figure 3.2] to get the same lags u have considered while PACF also but we have chosen the different approach. Once getting the errors, we have considered the model to be same as the AR model and have got the the significant lags for that model from ACF. We have chosen only two of the most significant lags and have got the coefficients of the model via the OLS (Ordinary Least Squares)method. After the predicting forecast on

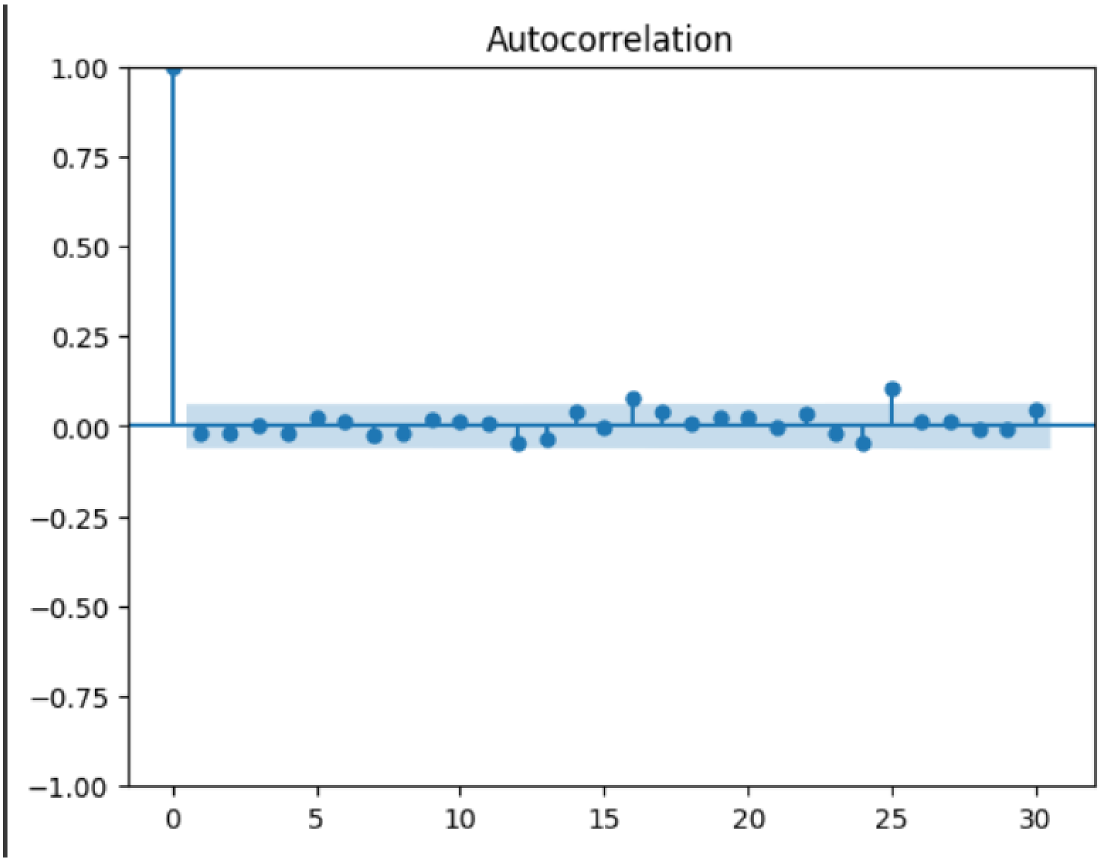

Figure 3.2: ACF Plot for max lags 30

the testing set we have calculated the error functions.

3.3.1 Time Series Models for SKU Demand Forecasting

The models we considered for forecasting are AR, MA and ARMA. The order of the AR model is 2 where I got the first two significant lags that are t-16 and t-25. The respective model can be written as.

$$y_t = \alpha_1 y_{t-16} + \alpha_2 y_{t-25} + \varepsilon \tag{3.12}$$

here the (ε) is considered to be zero.

Now, coming to the MA model the order is 2 same and here we have not considered the (μ) while modelling. The order of MA model is same as AR i.e. 2. The model will look like.

$$y_t = \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-17} + \mu \tag{3.13}$$

And for the final ARMA model we have considered the orders of (2,2) means including the same lags and errors from AR and MA respectively. Therefore the model will look

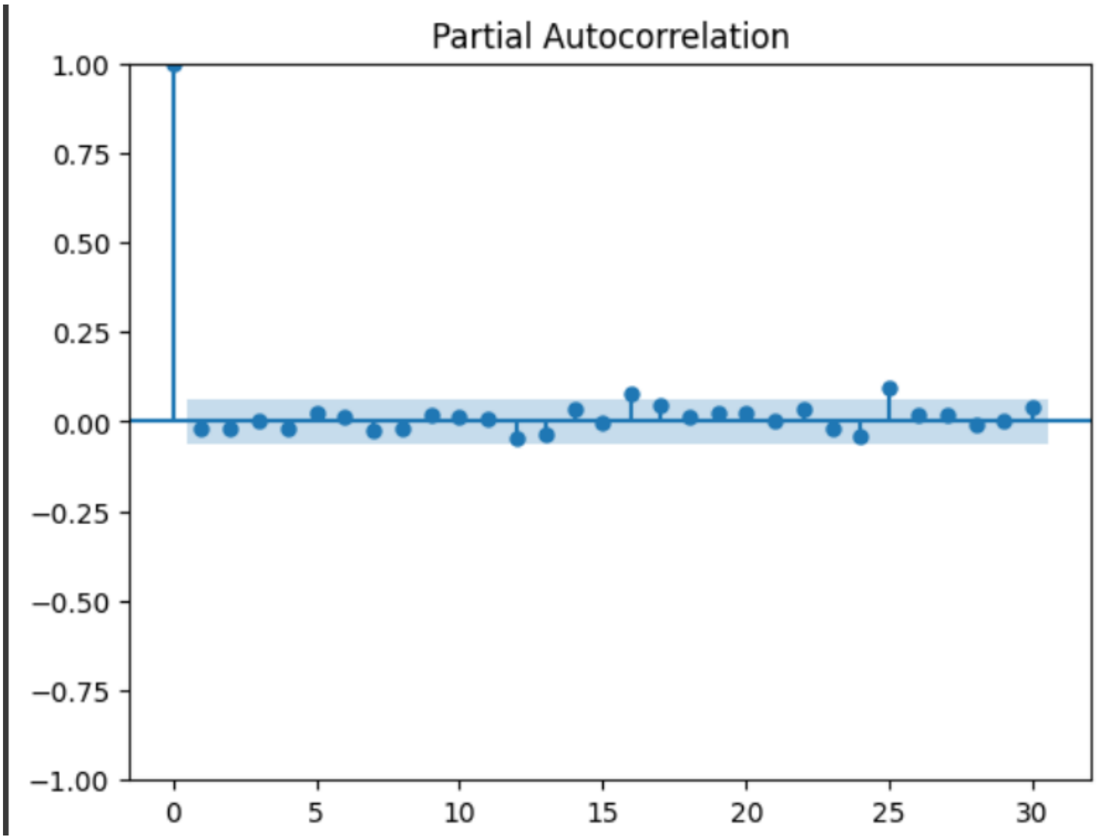

Figure 3.3: PACF Plot for max lags 30

like:

$$y_t = \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-17} + \alpha_1 y_{t-16} + \alpha_2 y_{t-25} + c \tag{3.14}$$

here, c is the constant term that is considered to be zero.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Results

The table of the performance of the respective time series models on the training set is created.The forecasting is done on the store id = "CA_2" and category id = "HOBBIES_1_001".[Table 4.2]
Same has been done for the testing set.[Table 4.1].
From the ACF and PACF we got the significant lags to be t-16 i.e. of 16 days and t-25 i.e. of 25 days for the AR and for MA model errors to be considered at day t and t-17.

Table 4.1: Performance Table of testing

Model Name	RMSE	MSE	MAE
AR	0.845	0.714	0.7477
MA	0.12808	0.0164	0.094

Table 4.2: Performance Table of training

Model Name	RMSE	MSE	MAE
AR	0.835	0.6975	0.665
MA	0.1044	0.0109	0.077

4.1.1 Performance Plots of AR model on each of the sets.

1. On training set.[Figure 4.1].

2. On testing set.[Figure 4.2].

4.1.2 Performance Plots of MA model on each of the sets.

1. On training set.[Figure 4.3]

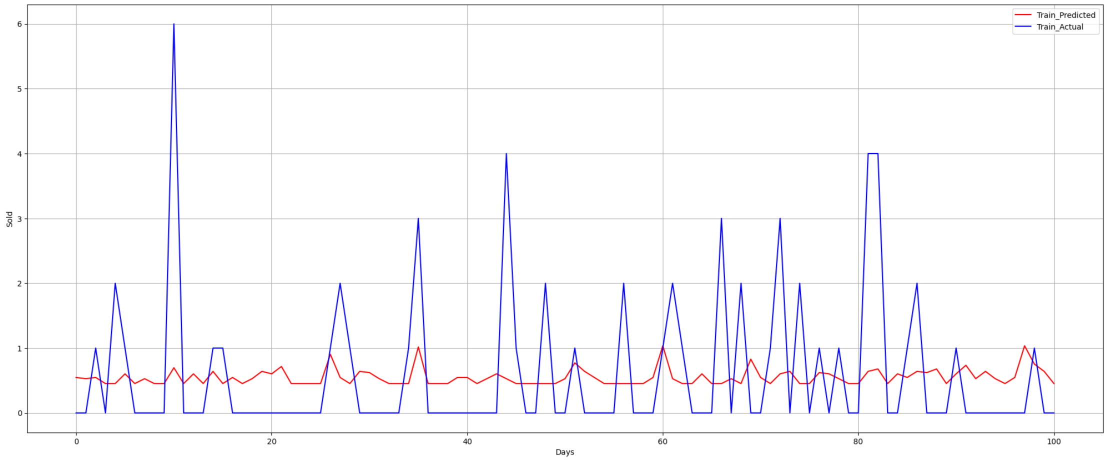

Figure 4.1: Actual vs predicted on training set of AR.

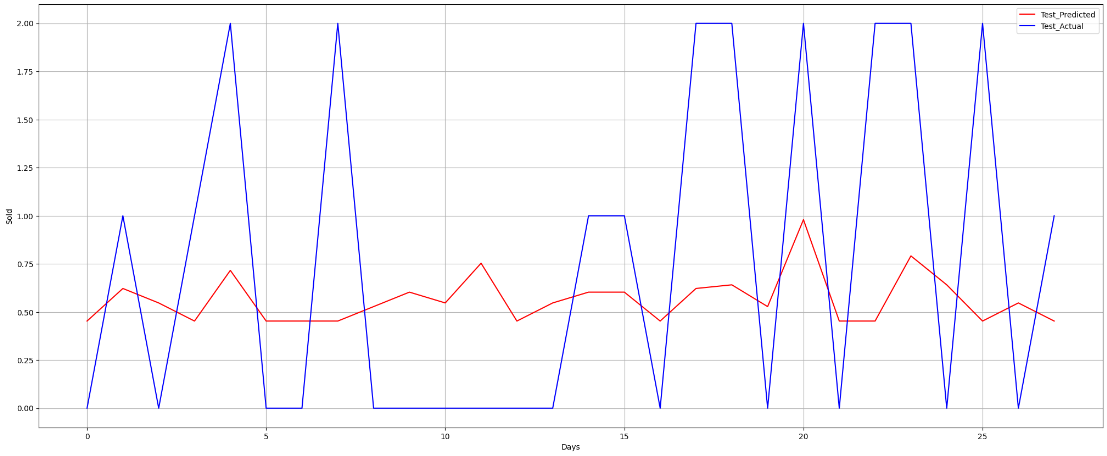

Figure 4.2: Actual VS Predicted on testing set of AR.

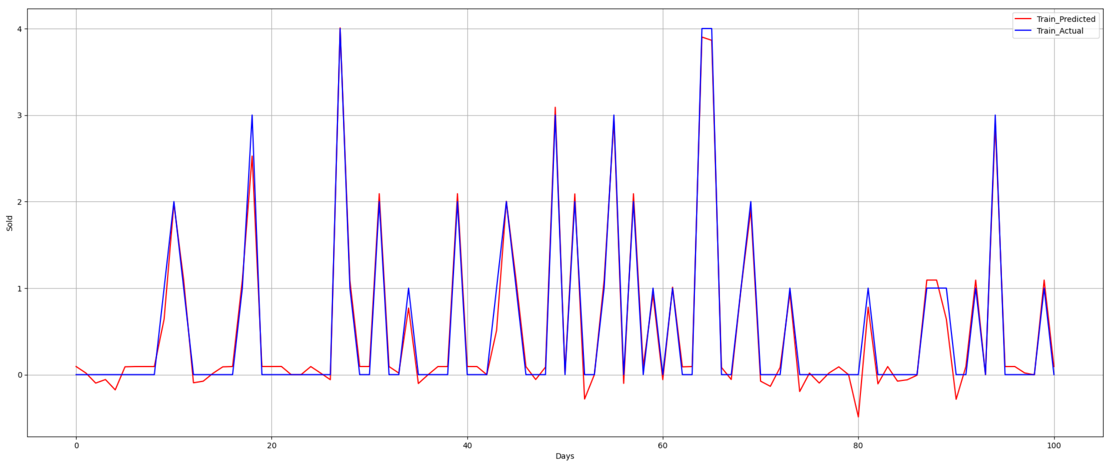

Figure 4.3: Actual vs predicted on training set of MA.

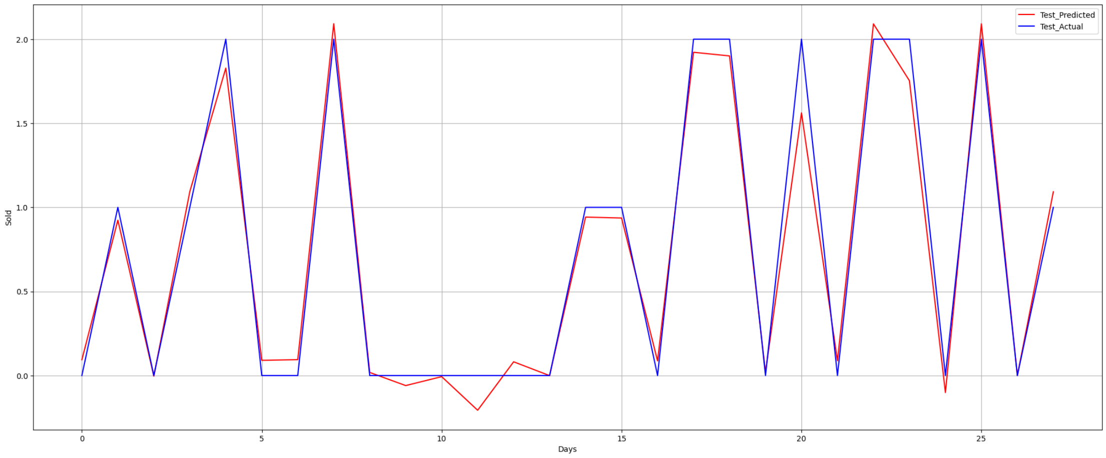

Figure 4.4: Actual vs predicted on testing set of MA.

2. On testing set.[Figure 4.4]

4.2 Discussions

There were the other significant lags having P-value>0.05 but we have considered the two most significant. Using more number of lags may result in overfitting. We have tried to implement the ARMA model also but the case of overfitting was arising also with 4 number of variables.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

In the course of this project, various time series models, including Auto-Regressive (AR), Moving Average (MA), and Auto-Regressive Moving Average (ARMA) models, were applied to the dataset for analysis. The primary objective was to assess their efficacy in extracting essential information from the data. Upon experimentation, it was observed that the Auto-Regressive (AR) model, with an order of 2, did not perform adequately on the dataset. It struggled to capture crucial patterns and trends within the data. However, by considering the errors of the AR model and constructing a Moving Average (MA) model with the same order (2), the performance significantly improved. The MA model demonstrated its capability to extract vital information from the dataset. Notably, as the time series was found to be stationary, there was no imperative need to develop an Auto-Regressive Integrated Moving Average (ARIMA) model. Nevertheless, the ARMA model, with consistent orders for AR and MA, resulted in overfitting and the results were not satisfactory.
In conclusion, the MA model emerged as the most suitable choice, successfully addressing various factors, including stockouts, changes in pricing, store closures, and the evolution of pricing and sales trends. The model's robust performance underscores its versatility and effectiveness in dealing with the complexities of the dataset.

5.2 Future work

The current study opens avenues for further exploration and refinement of time series analysis methodologies. The following areas of future work are identified to enhance the scope and depth of the research:

1. **Exploration of More Complex Datasets:** To broaden the applicability and robustness of the models developed, future efforts will involve working with more

complex datasets. The inclusion of diverse factors, such as market dynamics, economic indicators, and external events, will be considered. In particular, we plan to apply the developed models to the Theta Direct's dataset, which presents a more intricate set of challenges and opportunities.

2. **Comparison with Advanced Machine Learning Techniques**: The current study focused on traditional time series models, namely AR, MA, and ARMA. To benchmark our results and explore alternative approaches, future work will involve the application of advanced machine learning techniques. Recurrent Neural Networks (RNNs) and other sophisticated algorithms will be implemented and compared against the performance of our existing models. This comparative analysis aims to identify the strengths and weaknesses of different approaches in handling time series data.

3. **Optimization and Fine-Tuning:** Continuous refinement of the models will be pursued through optimization and fine-tuning. This involves adjusting hyperparameters, exploring different lag orders, and implementing feature engineering techniques to achieve better model performance.

4. **Improve the ARMA model:** To reduce the overfitting in the case of ARMA model.

CHAPTER 6

REFERENCES

1. Taghiyeh, S., Lengacher, D. C., Sadeghi, A. H., Sahebi-Fakhrabad, A., & Handfield, R. B. (2023, September). A novel multi-phase hierarchical forecasting approach with machine learning in supply chain management. Supply Chain Analytics, 3, 100032. https://doi.org/10.1016/j.sca.2023.100032.

2. Andrade, L., & Cunha, C. B. (2022). Disaggregated Retail Forecasting: A Gradient Boosting Approach. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4129889.

3. Elalem, Y. K., Maier, S., & Seifert, R. W. (2023, October). A machine learning-based framework for forecasting sales of new products with short life cycles using deep neural networks. International Journal of Forecasting, 39(4), 1874–1894. https://doi.org/10.1016/j.ijforecast.2022.09.005

4. Seyedan, M., Mafakheri, F., & Wang, C. (2023, September). Order-up-to-level inventory optimization model using time-series demand forecasting with ensemble deep learning. Supply Chain Analytics, 3, 100024. https://doi.org/10.1016/j.sca.2023.100024

5. Ilic, I., Görgülü, B., Cevik, M., & Baydoğan, M. G. (2021, December). Explainable boosted linear regression for time series forecasting. Pattern Recognition, 120, 108144. https://doi.org/10.1016/j.patcog.2021.108144

6. van Steenbergen, R., & Mes, M. (2020, December). Forecasting demand profiles of new products. Decision Support Systems, 139, 113401. https://doi.org/10.1016/j.dss.2020.113401.

7. Kourentzes, N., & Petropoulos, F. (2016, November). Forecasting with multivariate temporal aggregation: The case of promotional modelling. International Journal of Production Economics, 181, 145–153. https://doi.org/10.1016/j.ijpe.2015.09.011

8. Tadayonrad, Y., & Ndiaye, A. B. (2023, September). A new key performance indicator model for demand forecasting in inventory management considering supply chain reliability and seasonality. Supply Chain Analytics, 3, 100026. https://doi.org/10.1016/j.sca.2023.100026

9. Jahnke, H., Ulrich, M., Pesch, R., Senge, R., & Langrock, R. (2019). Distributional Regression for Demand Forecasting in e-grocery. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.

10. Li, C., & Lim, A. (2018, September). A greedy aggregation–decomposition method for intermittent demand forecasting in fashion retailing. European Journal of Operational Research, 269(3), 860–869. https://doi.org/10.1016/j.ejor.2018.02.029

11. Ma, S., & Fildes, R. (2021, January). Retail sales forecasting with meta-learning. European Journal of Operational Research, 288(1), 111–128. https://doi.org/10.1016/j.ejor.2020.05.038

12. Kaggle Dataset link:- https://www.kaggle.com/competitions/m5-forecasting data