Implementation of models for Demand forecasting for e-commerce using time series forecasting

Astha Sharma
CSE Department
Chandigarh University
Mohali (140413), India
astha27sharma@gmail.com

Bharti Sahu
Assistant Professor CSE
Chandigarh University
Mohali (140413), India
bhartisahu8001@gmail.com

Mayank Kishore
CSE Department
Chandigarh University
Mohali (140413), India
mayankkishorekarn@gmail.com

Kriti Kumari
CSE Department
Chandigarh University
Mohali (140413), India
21bcs5159@cuchd.in

Singh Surajkant Shubhnath
CSE Department
Chandigarh University
Mohali (140413), India
suraj1621singh@cuchd.in

***Abstract*—The study delves into the significance of demand forecasting within the e-commerce realm, crucial for businesses to anticipate future customer needs and optimize inventory management and resource allocation strategies. Focusing on employing time series forecasting methodologies, specifically XGBOOST and SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous factors) models, the research aims to predict demand accurately in e-commerce scenarios. By analyzing historical sales data and integrating pertinent external variables like promotional events and seasonal trends, these models endeavor to forecast forthcoming demand trends for a variety of products or services offered by e-commerce platforms. The investigation scrutinizes the intricacies of XGBOOST and SARIMAX models, highlighting their merits and limitations in the context of e-commerce demand forecasting. Moreover, the study assesses model performance using metrics like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to gauge accuracy and reliability in demand prediction. The insights gleaned from this study contribute to advancing demand forecasting techniques in e-commerce, offering valuable guidance for businesses aiming to refine operational efficiency and elevate customer satisfaction through precise demand projections.The SARIMA model deployed in this work has a MAPE value of 0.027 which is better than any previously formed SARIMA models on this particular dataset. The XGBOOST model implemented in this work has RMSE value of 2852.098 which is lesser than any of the work done on this dataset before. Also the calculated accuracy comes out to be 98.348% for XGBOOST which is remarkable and is best on this data.**

Keywords—Demand Forecasting, Customers, Time Series Data, Customer behaviour.

I. Introduction

In the realm of e-commerce, demand forecasting has become an indispensable aspect of business operations. It enables companies to anticipate forthcoming customer demands and optimize supply chain logistics accordingly. This introduction provides an overview of the historical evolution of demand forecasting in e-commerce, highlighting existing solutions employed by businesses to address this critical operational need. The trajectory of demand forecasting in e-commerce dates back to the nascent stages of online retailing when businesses primarily relied on basic methods and intuition to anticipate customer demands. However, as e-commerce platforms burgeoned and transaction volumes surged, the need for more sophisticated forecasting methodologies became apparent.

In the early 2000s [1], with the advent of data analytics and statistical modeling, e-commerce businesses began embracing time series forecasting techniques. Methods such as moving averages, exponential smoothing, and simple linear regression were employed to project demand for products and services. While these techniques laid the groundwork for demand forecasting, they often struggled to capture the intricacies of e-commerce markets characterized by rapidly shifting consumer preferences and market dynamics. [2] As technology progressed, so did demand forecasting in e-commerce. Advanced statistical techniques and machine learning algorithms revolutionized forecasting approaches. Models like XGBOOST and SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous factors) gained prominence for their ability to analyze complex time series data and incorporate external variables influencing demand.

In a retail context, the demand forecasting problem involves accurately predicting the future demand for various products based on historical sales data and external factors. For instance, a grocery store chain aims to forecast the demand for fresh produce such as fruits and vegetables. By analyzing historical sales data, including seasonal variations and trends, as well as external factors like weather patterns, holidays, and local events, the store can anticipate fluctuations in demand. [3] In the automotive industry, a car manufacturer aims to forecast the demand for specific vehicle models and configurations. By analyzing historical sales data, market trends, economic indicators, and consumer preferences, the manufacturer can anticipate demand for different types of vehicles such as sedans, SUVs, and electric cars. This helps the manufacturer optimize production planning, allocate resources efficiently across production lines, and adjust inventory levels at dealerships to meet customer demand while minimizing excess inventory and production bottlenecks.

For achieving the desired results we have taken “Walmart Sale Data" from Kaggle which consists of 4 datasets. These are - **Features, Stores, Train and Test.** The dataset comes from an American retail organization, Walmart Inc. The data includes information from 45 Walmart division stores, tracking their activities from 2010 to 2012, especially highlighting their weekly sales promotions.

The following figure shows the different attributes present in the **“features”** dataset .

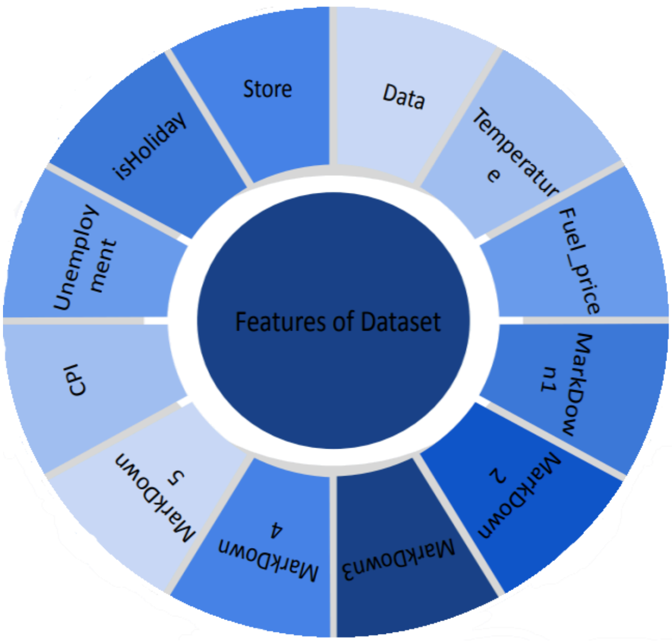

Figure 1: Features dataset

The following figure shows the different attributes present in the **"stores"** dataset .

Figure 2: Stores dataset

II. Literature review

This part pertains to the existing study conducted by many scholars in the field of Demand Forecasting. The purpose of this section is to highlight and reinforce this information via the use of the provided Literature Review.

In the study, A. Abbaspour et al. [1] used five machine learning models, i.e., ANFIS,MLPNN,RBFNN, DWTNN, and GMDH, and the parameters to identify the best model were RMSE, MAE, and correlation coefficient. This leads to the conclusion that the RBFNN model had the best performance with a score of 0.99176 in the interaction between targets and outputs. The study conducted by Arnab Mitra et al. [2] has compared the performance of five regression techniques, i.e., random forest (RF), extreme gradient boosting (XGBoost), gradient boosting,adaptive boosting (AdaBoost), and artificial neural network (ANN), with a hybrid (RF-XGBoost-LR) model, and for prediction, the ARIMA model has been used. The result parameters in this study were mean absolute error(MAE), mean squared error(MSE), and R^2. This leads to the conclusion that RF-XGboost-LR has performed the best among them, with a score of 4.79e-05 in MSE, 0.0024 in MAE, and 0.9551 in R^2.

The study conducted by Koussaila Hamiche et al. [3] has discussed time series analysis, which is robust, model-free, and does not require any prior patterns to predict demand. This leads to the conclusion that for no trend, no seasonality (NT and NS) reaches 25.1% and 26.57%, respectively, and for other trends, the range is between 1.5% and 15%. The study conducted by Zeynep Hilal Kimichi et al. [4] has discussed an improved demand forecasting approach. This approach has used a deep learning approach with time series models and support vector regression models. The parameter used to calculate the efficiency of this approach was MAPE (mean absolute percentage error). This leads to the result of improvement of 2.69% success rate in the MAPE.

In their research, M. Nasseri et al.[5] conducted a comparison between the demand prediction accuracy of tree-based ensembles and deep learning models based on long-short-term memory (LSTM). The parameters for this calculation were MAPE,MAE,RMSE,R^2. This leads to the conclusion that the ETR (extra tree regressors) models perform better than the DL models. The work done by Aamer et al. [6] reviewed the use of machine learning applications in demand forecasting. A total of 1870 papers were retrieved, and the study was conducted on the different models. This leads to the result that neural networks,artificial neural networks, support vector regression, and support vector machines were the best and most commonly used models.

The study conducted by Batuhan Cocaoglu et al. [7] discussed models that are up-to-date and also discussed the impact of models in business. This leads to the conclusion that %5 impact leads to improvement to 1 million and %15 impact leads to 2.65 million. In the study, Majid Rafiee et al.[8] machine learning is one of the most important parts for predicting demands, understanding customers and collecting their feedbacks.It is far better than the traditional ways we used to enhance the performance. In conclusion a study is identified that helps in targeting the customers and in the enhancement of the sales and customer engagements.By knowing the customer needs and purchasing behavior retailers can improve their strategies to complete the customer needs and can increase sales.

In the study, Gerald Reiner et al.[9] focuses on evaluating improvements in supply chain processes. It discusses the impact of forecast errors on bullwhip effect and other performance measures. It also describes the evaluation model used from the reference dataset.In conclusion time series models show the best performance in comparison to regression model and optimal performance across most models is seen with 20 observations. The study conducted by Carla Freitas Silveira Netto et al.[10] examines demand forecasting in marketing, focusing on common models, challenges with big data, types of data used, and areas for future research.It suggests exploring forecasting for durable goods, integrating diverse data sources,and maximizing the utility of location data without hindering implementation.

In the study, Ing. Andrea Kolková et al.[11] found that both deep learning and statistical methods were effective in predicting demand from e-commerce data. However, deep learning models were deemed impractical for business use due to their high costs and complexity. It leads to the conclusion that Facebook Prophest is one of the most useful model for small and medium-sized enterprises and then TBATS and SARIMA as it showed limitations in capturing multi-seasonality and daily data.

Research conducted by R. Rathipriya1 et al.[12]confirms that both shallow and deep neural networks are effective for forecasting demand in pharmaceutical companies. Shallow networks can accurately predict future demand for items, aiding decision-making in production and supply chain operation According to Loubna Terrada et al.[13] This study aims to improve demand forecasting using advanced Deep Learning methods like ARIMA and LSTM. By analyzing historical transaction data, the goal is to enhance accuracy and transition towards Smart Supply Chain Management.

In the study, Akash Singh et al.[14] discuss Demand forecasting as it plays the most important part in demand planning with supply chain management and also predict the future product performance based on past data. There were two types of data, Historical and forecast proportions, which were examined to enhance the accuracy. The study conducted by Tugay et al. [15] implemented stacked generalization, also referred to as stacking ensemble learning, to forecast demand. Subsequently, they assessed all methodologies using authentic data obtained from the e-commerce enterprise. The null hypothesis was rejected in the ANOVA test at a significance level of 5%, indicating that the predictions made by Random Forest (RF) and Linear Regression (LR) are notably superior to those of other methods in the first and second levels, respectively.

According to the research conducted by Smirnov et al. [16] they used information regarding the products demand from Ozon online marketplace. The algorithm's inputs consist of product attributes including price, name, category, and textual description. Regression challenges were addressed using different iterations of the gradient boosting algorithm, including XGBoost, LightGBM, and CatBoost. Currently, the forecasting accuracy stands at approximately 4.00. This system can operate autonomously or integrate seamlessly within a larger, more intricate framework.In their work, Alqatawna et al.[17] utilized time-series analysis techniques for forecasting the resource requirements of logistics delivery firms. This initiative facilitates the fulfillment of objectives and fosters growth. The SARIMAX model was identified as the top performer among the various methods evaluated. Across the UAE, KSA, and KWT regions, the SARIMAX model exhibited exceptional precision in forecasting order volumes and trends, with MAPE scores of 0.097, 0.158, and 0.137 respectively, along with corresponding RMSE values of 0.134, 0.199, and 0.215.

According to the work done by Moroff et al.[18] they implemented six distinct forecasting models from both statistics and machine learning and underwent evaluation concerning forecast accuracy and implementation complexity. The findings underscore the promise of novel forecasting models while emphasizing the importance of thorough, application-specific assessments to discern the merits and drawbacks of the available methodologies. During this study, the deep learning model Multi-Layer Perceptron (MLP) achieved the highest overall score of 81.9%. Following closely, the triple exponential smoothing model by Holt and Winters obtained the second-best value at 80.8%.

The research conducted by Feizabad et al.[19] introduces hybrid forecasting techniques, specifically Neural Network models and ARIMAX. The approach integrates both time series data and explanatory variables. The approach was put into practice and evaluated in the context of functional products within a steel manufacturing environment. According to the research conducted by Ho et al. [20] employs machine learning methods to forecast retail demand at 1C Company across numerous products. The model utilized in this study integrates traditional statistical approaches with machine learning techniques, particularly a hybrid Support Vector Machine. The study conducted by Farzana et al. [21] examines a range of cutting-edge methods for demand forecasting, with a particular emphasis on machine learning. Future research areas include exploring regression-based methods, hybrid models, and ensemble models. This study provides insight into demand forecasting within the realm of machine learning

III. Methodology

In this section of the study, the concentration is on the methodology employed for conducting Demand Forecasting.

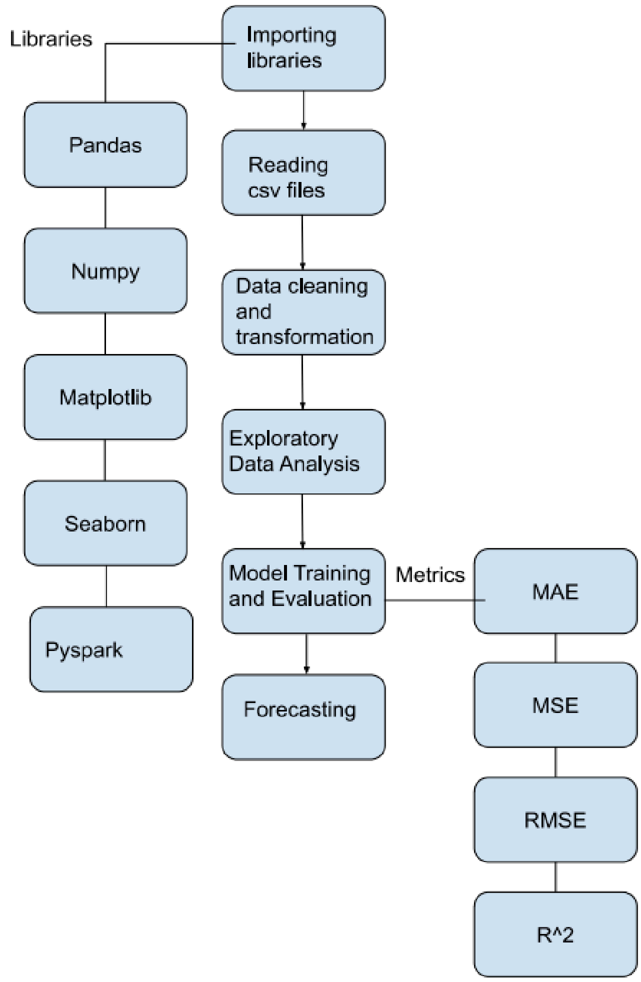

Figure 3: Methodology

In order to target the problem specified related to segmentation of customer, 4 dataset has been taken from Kaggle which were given by Walmart for sales forecasting and these were titled as “Features”, “Store”, “Train”, “Test”.

“Features.csv” consists of 421570 entries and contains the following information:

- **Store:** This column identifies each store uniquely within the dataset.
- **Date:** This column records the date of data collection, indicating when the measurements or observations were taken.
- **Temperature:** This column shows the temperature recorded at the time of data collection, typically in Celsius or Fahrenheit.
- **Fuel_Price:** This column indicates the fuel price at the time of data collection, which is useful for analyzing economic trends.
- **MarkDown1, MarkDown2, MarkDown3, MarkDown4, MarkDown5:** These columns represent different types of discounts or markdowns offered by the stores. Each markdown column may correspond to a specific promotional event.
- **CPI:** CPI stands for Consumer Price Index, and this column contains the CPI value at the time of data collection.
- **Unemployment:** This column shows the unemployment rate at the time of data collection, indicating the percentage of unemployed individuals in the labor force.
- **IsHoliday:** This column is a binary indicator (0 or 1) that indicates whether the date corresponds to a holiday. It can be used to analyze sales and consumer behavior during holiday periods.

“Stores.csv” contains these information:

- **Store:** This column serves as a unique identifier for each store in the dataset, distinguishing one store from another.
- **Type:** In this column, the type or category of each store is specified. This categorization can include designations like "A," "B," or "C," which may denote different store formats, market segments, or strategies.
- **Size:** This column provides information about the physical size of each store, typically measured in terms of floor area or square footage. Store size is an important factor for analyzing store performance, resource allocation, and operational decision-making.

In the first phase outliers, identified as weekly sales exceeding $200,000, were removed from the dataset to prevent skewing the model's predictions. Categorical columns like store type were converted to numerical values to facilitate numerical computations and model training.
These constraints were essential steps in preparing the data for analysis and building reliable predictive models.

Handling Missing Values in Markdown Columns:

- In the code, missing values in markdown columns were handled by converting 'NA' strings to null values using PySpark's withColumn() and when() functions.
- After converting 'NA' strings to null, the null values were filled with 0 using PySpark's otherwise() function.

Treatment of Outliers:

- Outliers in the data were identified based on the weekly sales exceeding $200,000.
- These outliers were removed from the dataset before model training using the query() method in Pandas, which allows filtering rows based on specified conditions.

Conversion of Categorical Columns to Numerical Values:

- Categorical columns like store type ('A', 'B', 'C') were converted to numerical values ('0', '1', '2') for model training.
- This conversion was done using Pandas' replace() method to map categorical values to numerical equivalents.

In the second phase **exploratory data analysis** has been done on the data. This involves the following explorations:

- Stores are categorized into three types: A, B, and C based on their sizes. Nearly half of the stores, which are larger than 150,000, fall into the A category. Sales trends vary according to store type.

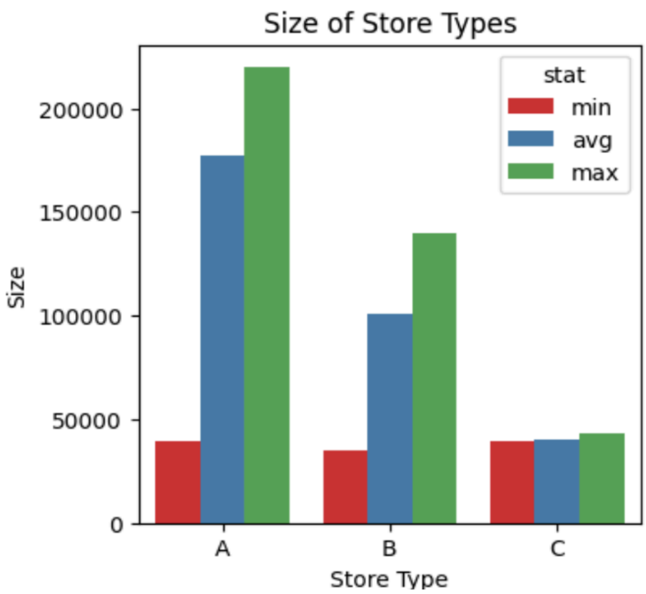

Figure 4: Size of Store Types

- As anticipated, sales during holidays surpass those on regular days. The Christmas holiday marks the end of the year, yet shoppers typically flock to stores during the 51st week. Among holidays, Thanksgiving stands out with higher sales, a distinction acknowledged by Walmart.
- January consistently records lower sales compared to other months, attributed to the heightened sales during November and December. Following two months of

robust sales, consumers tend to prioritize cost-saving measures in January.

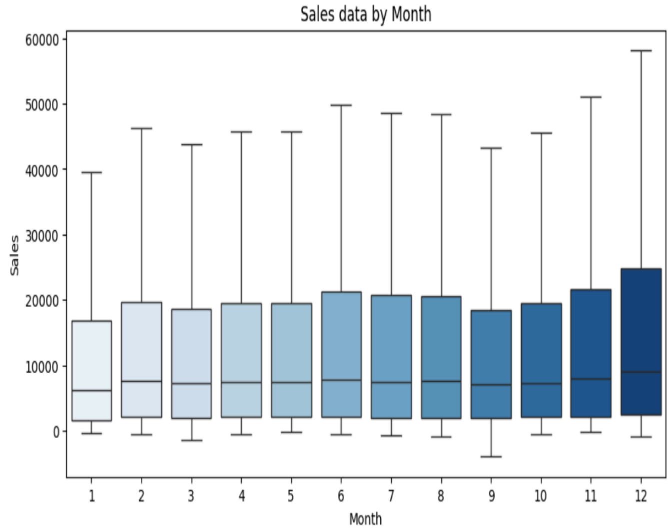

Figure 5: Sales data by Month

- There is no discernible pattern between CPI, temperature, unemployment rate, and fuel price concerning weekly sales.
- Sales in 2010 surpassed those in 2011 and 2012. However, the absence of November and December sales data for 2012 does not significantly reduce its overall sales compared to 2010. Therefore, with the inclusion of the last two months' data, 2012 could potentially take the lead.

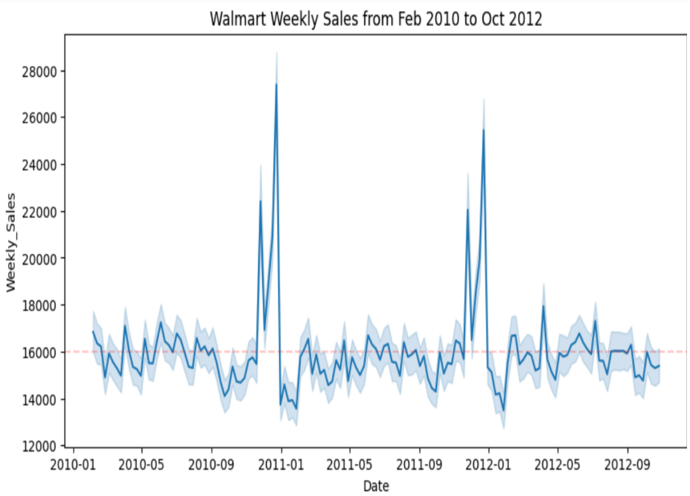

Figure 6: Weekly sales

In the **third phase**, we have implemented the **correlation matrix** for identifying the specific features which can be proved useful for our prediction.
The correlation matrix provides correlation coefficients between pairs of numerical columns.

Most Correlated Columns:

MarkDown4 and MarkDown1: These columns have a high positive correlation, suggesting that they are strongly related. This could indicate similar patterns or effects in their respective data.
MarkDown1 and MarkDown5: These columns also show a positive correlation, though slightly lower than MarkDown4 and MarkDown1. They may share some commonality in their influence on the dataset.
MarkDown2 and Temperature: These columns exhibit a negative correlation, implying an inverse connection. This indicates that as one variable (MarkDown2) increases, the other variable (Temperature) decreases, and conversely.

Least Correlated Columns:
Fuel_Price and Unemployment: These columns show a low correlation close to 0, indicating little to no linear relationship between them. Changes in fuel prices do not seem to be strongly related to changes in unemployment rates in the dataset.
CPI and Weekly_Sales: The Consumer Price Index (CPI) and Weekly Sales columns also exhibit a low correlation, implying that fluctuations in CPI do not have a strong linear impact on weekly sales in this dataset.
MarkDown3 and MarkDown1: These columns show a relatively low correlation compared to other markdown columns, suggesting that they may have different patterns or influences on the dataset.

In the **fourth phase**, the two time series **XGBoost** and **SARIMA** models are implemented.

XGBOOST:
XGBoost is an ensemble learning technique rooted in decision trees, celebrated for its superior performance across diverse machine learning tasks. Its mechanism involves the iterative creation of decision trees, with each tree focused on rectifying errors from the preceding ones. optimizing a predefined objective function. XGBoost handles complex relationships in the data, captures nonlinear patterns, and is robust to overfitting due to regularization techniques.In the given work, XGBoost was used to predict weekly sales based on features like store details, department, date, holiday status, and additional factors like temperature, fuel price, markdowns, CPI, and unemployment.

The results obtained from the XGBoost model were evaluated using metrics like accuracy, mean absolute error (MAE), mean squared error **(MSE)**, root mean squared error **(RMSE)**, and explained variance score (R2). The high accuracy suggests that the model effectively predicted weekly sales..

SARIMA Model:

SARIMA is a time series forecasting model that integrates autoregressive (AR), differencing (I), and moving average (MA) components along with seasonal patterns (S). This combination enables it to capture both trends and seasonal fluctuations present in the data. It works by identifying and modeling the autocorrelation and seasonality in time series data, making it useful and efficient for forecasting tasks with clear time-dependent patterns. SARIMA models are particularly useful for analyzing and forecasting sales data, where seasonality and trends play a significant role.

In the given work, SARIMA was used for overall sales forecasting, aggregating sales data across all stores and departments to predict future sales trends.

The results obtained from the SARIMA model were evaluated based on metrics like mean absolute percentage error **(MAPE)**, mean absolute error **(MAE)**, and root mean squared error **(RMSE)**. The model's performance was assessed in terms of its ability to accurately forecast overall sales trends over time.

IV. Results And Discussions

The following are the results obtained by performing the study:

Accuracy	98.348
MAE	1570.744
MSE	8.13E+06
RMSE	2852.098
R2	0.98348

Table 1: XGB Regressor Evaluation Metrics

MAPE	0.027
MAE	1.28E+06
RMSE	1584689.36
MSE	8.43E+06

Table 2: SARIMA Evaluation Metrics:

1. The XGBoost model showed high accuracy in predicting weekly sales, with a low MAE, MSE, and RMSE, indicating minimal error in predictions.
2. SARIMA demonstrated good performance in overall sales forecasting, as evidenced by low MAPE, MAE, and RMSE values, signifying accurate predictions of sales trends.
3. It's important to note that while XGBoost excelled in capturing complex relationships and nonlinear patterns in weekly sales data, SARIMA proved effective in modeling and forecasting seasonal variations and trends in overall sales.

The following figure shows the actual vs predicted values of sales by implementing the **XGBOOST** model.

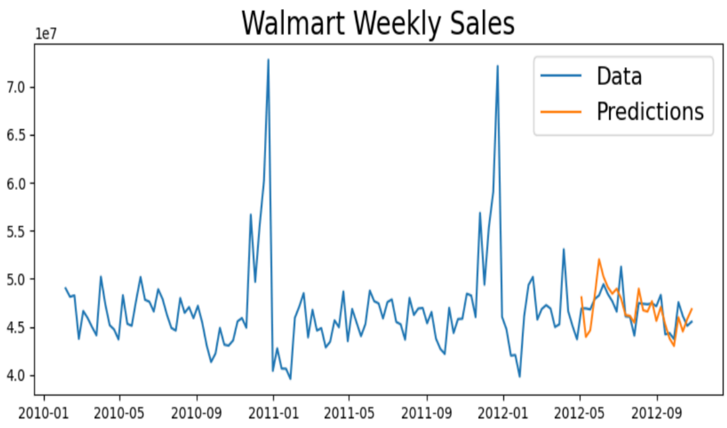

Figure 9: Actual vs predicted data

The following figure shows the RMSE values produced by using the **XGBOOST** model on the train and test data respectively.

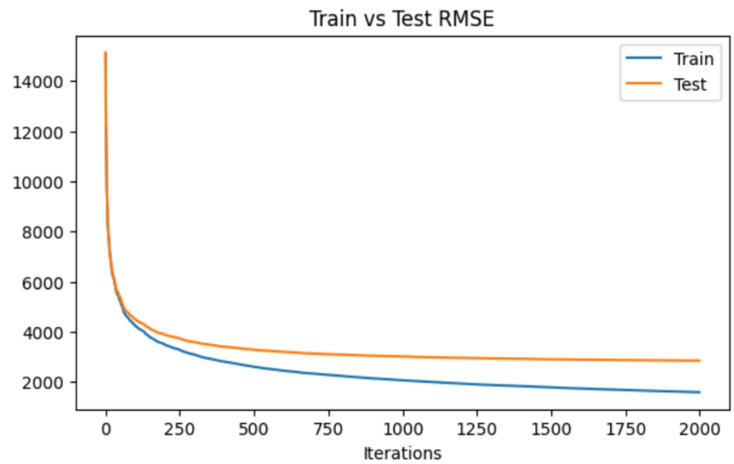

Figure 10: RMSE of train vs test

The following Table shows the different evaluation metrics of the model in comparison to the already existing ones.

Evaluation Metrics	ARIMA	SARIMA	SARIMA (Our model)	XGBOOST (Existing model)	XGBOOST(Our model)
MAE	317530.9	3.64	1.28E+06	2.49E+03	1.57E+03
MAPE	0.143	0.087	0.027	0.056	0.268
MSE	1.17E+07	2.42E+07	8.43E+06	9.82E+06	8.13E+06
RMSE	0.287	2213147	1584689	4001.29	2852.098

Table 3: Comparison of evaluation metrics

The combination of these models can offer a comprehensive understanding of sales dynamics, supporting strategic planning and optimization in retail operations.

V. Conclusion And Future Scope

The SARIMA (Seasonal Autoregressive Integrated Moving Average) is tailored for forecasting time series data, analyzing historical sales data to detect trends, seasonal variations, and anomalies. Through techniques like lagged observations and differencing, SARIMA can capture autocorrelation and seasonal effects in the data, providing precise forecasts for upcoming time intervals. In contrast, the strength of XGBoost lies in capturing intricate data relationships and nonlinearities, making it adept at predicting time-dependent variables like sales demand over time. Trained on past sales data alongside pertinent features such as holiday status and economic factors, XGBoost learns influential patterns and trends in sales behavior, empowering it to forecast future sales accurately.

The XGBoost model and SARIMA model exhibited strong performance in predicting weekly sales and overall sales forecasting, respectively. XGBoost achieved high accuracy with minimal prediction errors, effectively capturing complex relationships and nonlinear patterns in the weekly sales data. Conversely, SARIMA accurately modeled seasonal variations and trends in overall sales, delivering precise forecasts. The integration of these models provides a comprehensive insight into sales dynamics, facilitating strategic decision-making and resource allocation in retail operations.

Moving forward, there are several avenues for enhancing predictive capabilities in retail operations. One avenue involves exploring ensemble methods such as stacking or blending XGBoost with complementary models to improve predictive performance and capture a wider array of patterns and trends. Additionally, conducting rigorous feature engineering can uncover new relevant features, enhancing model accuracy and capturing nuanced data relationships. Hyperparameter tuning can optimize model performance for both XGBoost and SARIMA models. Exploring advanced time series techniques such as Prophet, LSTM, or attention-based models can further improve forecasting accuracy, particularly for long-term sales predictions. Integrating these models with real-time data streams and business processes enables dynamic decision-making based on the latest sales insights. Extending predictive analytics to include inventory management aspects like demand forecasting, stock replenishment strategies, and inventory optimization can streamline operations while meeting customer demands efficiently. Lastly, implementing market basket analysis techniques can uncover product associations and customer purchase patterns, leading to targeted marketing efforts and personalized recommendations.

References

[1] A. Abbaspour Ghadim Bonab, “A comparative study of demand forecasting based on machine learning methods with time series approach,” J. Appl. Res. Ind. Eng, no. Online First, Mar. 2021, doi: 10.22105/jarie.2021.246283.1192.

[2] A. Mitra, A. Jain, A. Kishore, and P. Kumar, “A Comparative Study of Demand Forecasting Models for a Multi-Channel Retail Company: A Novel Hybrid Machine Learning Approach,” Oper. Res. Forum, vol. 3, no. 4, p. 58, Sep. 2022, doi: 10.1007/s43069-022-00166-4.

[3] K. Hamiche, H. Abouaïssa, G. Goncalves, and T. Hsu, “A Robust and Easy Approach for Demand Forecasting in Supply Chains,” IFAC-PapersOnLine, vol. 51, no. 11, pp. 1732–1737, Jan. 2018, doi: 10.1016/j.ifacol.2018.08.206.

[4] Z. H. Kilimci et al., “An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain,” Complexity, vol. 2019, pp. 1–15, Mar. 2019, doi: 10.1155/2019/9067367.

[5] M. Nasseri, T. Falatouri, P. Brandtner, and F. Darbanian, “Applying Machine Learning in Retail Demand Prediction—A Comparison of Tree-Based Ensembles and Long Short-Term Memory-Based Deep Learning,” Applied Sciences, vol. 13, no. 19, Art. no. 19, Jan. 2023, doi: 10.3390/app131911112.

[6] A. Aamer, L. P. E. Yani, and I. M. Priyatna, “Data Analytics in the Supply Chain Management: Review of Machine Learning Applications in Demand Forecasting,” Operations and Supply Chain Management An International Journal, vol. 14, pp. 1–13, Jan. 2021, doi: 10.31387/oscm0440281.

[7] A. Z. Acar, B. Kocao?lu, and B. Y?lmaz, “Demand Forecast, Up-to-date Models, and Suggestions for Improvement an Example of a Business,” Journal of Global Strategic Management, vol. 8, pp. 26–37, Jun. 2014, doi: 10.20460/JGSM.2014815650.

[8] M. Zohdi, M. Rafiee, V. Kayvanfar, and A. Salamiraad, “Demand forecasting based machine learning algorithms on customer information: an applied approach,” International Journal of Information Technology, vol. 14, Feb. 2022, doi: 10.1007/s41870-022-00875-3.

[9] G. Reiner and J. Fichtinger, “Demand forecasting for supply processes in consideration of pricing and market information,” International Journal of Production Economics, vol. 118, no. 1, pp. 55–62, Mar. 2009, doi: 10.1016/j.ijpe.2008.08.009.

[10] A. Kolková and M. Navrátil, “Demand Forecasting in Python: Deep Learning Model Based on LSTM Architecture versus Statistical Models,” Acta Polytechnica Hungarica, vol. 18, pp. 123–141, Jan. 2021, doi: 10.12700/APH.18.8.2021.8.7.

[11] R. Rathipriya, A. A. Abdul Rahman, S. Dhamodharavadhani, A. Meero, and G. Yoganandan, “Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model,” Neural Comput & Applic, vol. 35, no. 2, pp. 1945–1957, Jan. 2023, doi: 10.1007/s00521-022-07889-9.

[12] L. Terrada, M. El Khaili, and O. Hassan, “Demand Forecasting Model using Deep Learning Methods for Supply Chain Management 4.0,” International Journal of Advanced Computer Science and Applications, vol. 13, May 2022, doi: 10.14569/IJACSA.2022.0130581.

[13] P. S. Smirnov and V. A. Sudakov, “Forecasting new product demand using machine learning,” J. Phys.: Conf. Ser., vol. 1925, no. 1, p. 012033, May 2021, doi: 10.1088/1742-6596/1925/1/012033.

[14] A. Alqatawna, B. Abu-Salih, N. Obeid, and M. Almiani, “Incorporating Time-Series Forecasting Techniques to Predict Logistics Companies’ Staffing Needs and Order Volume,” Computation, vol. 11, no. 7, Art. no. 7, Jul. 2023, doi: 10.3390/computation11070141.

[15] N. U. Moroff, E. Kurt, and J. Kamphues, “Machine Learning and Statistics: A Study for assessing innovative Demand Forecasting Models,” Procedia Computer Science, vol. 180, pp. 40–49, Jan. 2021, doi: 10.1016/j.procs.2021.01.127.

[16] J. Feizabadi, “Machine learning demand forecasting and supply chain performance,” International Journal of Logistics Research and Applications, vol. 25, no. 2, pp. 119–142, Feb. 2022, doi: 10.1080/13675567.2020.1803246.

[17] S. F. G and P. N, “Machine Learning in Demand Forecasting - A Review.” Rochester, NY, Nov. 19, 2020. doi: 10.2139/ssrn.3733548.

[18] T. Ho, L. Tran, H. Tran, and S. Dao, “Machine Learning in Demand Forecasting,” Jan. 2022.

[19] E. Tarallo, G. K. Akabane, C. I. Shimabukuro, J. Mello, and D. Amancio, “Machine Learning in Predicting Demand for Fast-Moving Consumer Goods: An Exploratory Research,” IFAC-PapersOnLine, vol. 52, no. 13, pp. 737–742, Jan. 2019, doi: 10.1016/j.ifacol.2019.11.203.

[20] R. Carbonneau, R. Vahidov, and K. Laframboise, “Machine Learning-Based Demand Forecasting in Supply Chains,” IJIIT, vol. 3, pp. 40–57, Oct. 2007, doi: 10.4018/jiit.2007100103.

[21] D. T. Wiyanti, I. Kharisudin, A. B. Setiawan, and A. K. Nugroho, “Machine-learning algorithm for demand forecasting problem,” J. Phys.: Conf. Ser., vol. 1918, no. 4, p. 042012, Jun. 2021, doi: 10.1088/1742-6596/1918/4/042012.