INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND MANAGEMENT ISSN: 2583-6129
VOLUME: 03 ISSUE: 04 | APRIL – 2024 DOI: 10.55041/ISJEM0k1493
AN INTERNATIONAL SCHOLARLY | | MULTIDISCIPLINARY | | OPEN ACCESS | | INDEXING IN ALL MAJOR DATABASE & METADATA

SALES PREDICTION BASED ON SARIMAX TIME SERIES ALGORITHM

Ms. S. R Sri Abirami, AP/IT
Information Technology
Nandha Engineering College
Erode
sriabirami@nandhaengg.org

Jeevitha K
Information Technology
Nandha Engineering College
Erode
jeevithait017@gmail.com

Logamithra V
Information Technology
Nandha Engineering College
Erode
vlmithra0403@gmail.com

Sudharshini.V.M
Information Technology
Nandha Engineering College
Erode
sudharshini263@gmail.com

***Abstract*—-Businesses must perform time-series forecasting of seasonal item sales because it enables them to predict future demand and modify their inventory accordingly. This investigation compares the performance of three well-known machine learning methods for time-series forecasting of seasonal item sales: Support Vector Machine (SVM), Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX), and Multi-layer Perception (MLP). A dataset of historical sales data is used to evaluate the algorithms. The data is split into training and testing sets, and measures like Mean Absolute Error (MAE), Relative Absolute Error (RAE), and Root Mean Squared Error (RMSE) are used to assess each algorithm's performance. The analysis's findings support the assertion that SARIMA- X provides greater accuracy than other methods incalculating the seasonal sales of the historical data.**

Keywords—, Time-series forecasting, SARIMAX, ARIMA, Machine Learning

1. INTRODUCTION

Sales forecasting using machine learning and Angular View is a powerful combination that leverages data-driven insights and a dynamic user interface for effective decision-making. There are various methods of forecasting based on historical data or values, which have been refined and optimized through the development of science and technology. In terms of sales forecasting statistical analysis techniques for forecasting sales, including time series, linear regression and other methodologies. Both multiple factorial linear regression models and time series. At the company level, sales forecasting is the major part of the business plan and significant inputs for decision-makingactivities. It is essential for organizations to produce the required quantity at the specified time. For that, sales forecasting will give the idea about how an organization should manage its budgeting, workforce and resources. This forecasting helps the business management to determine how much products should be manufacture, how much revenue can be expected and what could be the requirement of employees, investment and equipment.Machine learning techniques can be used to automatically create precise sales forecasting models using the wealth of sales data and related data. This strategy is significantly easier. It is adaptable, so it can adjust to data changes, and it is not biased by the quirks of a single sales manager Yet, it runs the risk of overestimating the human expert's prediction's accuracy, which is typically faulty. For instance, businesses once produced goods without taking demand or sales volume into account, which resulted in a number of issues. Data on consumer demand for items is crucial for any producer to decide whether to increase or decrease the number of units produced because they don't know how much to sell. Companies will suffer losses if they don't take these guidelines into account when they compete in the market. Various businesses use various metrics to estimate their market and sales.

1.1 ROLE OF MACHINE LEARNING IN SALES FORECASTING

Machine learning algorithms analyze historical sales data, market trends, and various influencing factors to generate accurate predictions. This enables businesses to anticipate demand, optimize inventory, and align their strategies with market dynamics. Machine learning excels in analyzing vast datasets, identifying patterns, and recognizing trends that may not be apparent through traditional methods. By leveraging advanced algorithms, machine learning models can make accurate and data-driven predictions, allowing businesses to anticipate future trends and outcomes with higher precision.Machine learning models can be trained to forecast future sales based on past data. Techniques like regression analysis, time series analysis, and ensemble methods are commonly used to build predictive models that estimate future sales figures with a certain degree of accuracy.

By analyzing customer demographics, purchasing history, and behavior patterns, machine learning algorithms can segment customers into groups with similar characteristics. This segmentation enables sales teams to tailor their strategies and offerings to specific customer segments, improving overall sales performance. Machine learning algorithms can predict future demand for products or services by analyzing various demand drivers such as marketing campaigns, economic indicators, competitor actions, and seasonality. Accurate demand forecasting helps businesses optimize inventory management, production planning, and resource allocation. Machine learning can be used to optimize pricing strategies based on real-time market conditions, demand fluctuations, and competitor pricing. Dynamic pricing algorithms adjust prices dynamically to maximize revenue and profitability, taking into account factors like customer willingness to pay and product demand elasticity.

Sales prediction using machine learning algorithms. Machine learning is transforming every part of life, and this has had a big impact on actual-world events as well. Education, healthcare, engineering, commerce, entertainment, and transportation are just a handful of the many industries where machine learning has made a significant impact. Because it ignores how consumers actually make purchases, the conventional approach to attaining sales and marketing goals no longer benefits firms. This study's goal is to propose a factor for predicting Big Retail Companies' expected sales while accounting for their historical sales. A thorough analysis of sales prediction is conducted using machine learning techniques like linear regression, K-neighbors regression, and XG are used.

1.1.1 AN ADVANCED SALES FORECASTING USING MACHINELEARNING

A retailer can estimate its anticipated future revenues for income earned over a specific period of time with the aid of sales forecasting. As a result, time is crucial in sales forecasting. For sales forecasting in this study, we used data mining approaches including ARIMA models and XG Boost algorithms, which are more effective at manipulating trending sales analyses. The material was once subjected to extensive review in order to identify patterns and outliers that can hinder the prediction system. As it provides crucial information on the visitors a store might anticipate on a particular day, sales forecast plays a crucial role in increasing how successfully stores can operate.

1.1.2. INTELLIGENT SALES PREDICTION USING MACHINE LEARNING TECHNIQUES

Sales forecasting is becoming more and more crucial as e-commerce grows, as accurate and prompt forecasting can assist e-commerce businesses in resolving all the supply and demand-related uncertainty and lowering inventory costs. The majority of commercial businesses rely largely on data sources and demand forecasts of sales patterns. The accuracy of sales projections has a significant effect on business. Data mining techniques are very practical tools for unlocking buried information from a sizable dataset to improve predicting precision and effectiveness. Dealing with huge data and accurate sales forecasting is challenging, though. The use of various data mining techniques could solve these problems. We provided a quick analysis of sales data and sales projection in this report. The use of various data mining techniques could solve these problems. We provided a quick analysis of sales data and sales projection in this report. Every person and every business must be aware of customer demand far in advance of any season to prevent product shortages. As time passes by, the desire of the businesses to be more accurate regarding the predictions will expand tremendously. Improved projections are directly correlated to the company's earnings.

1.1.3 MACHINE LEARNING BASED SALES FORECASTING SYSTEM

Sales forecasting tries to anticipate future demand for sales numbers, reserve the quantity of products, and implement marketing strategies based on the outcomes of the forecasting. In addition to avoiding wasteful overstocking, maintenance expenses, and sales demand patterns, an accurate and reliable forecasting system plays a significant role in decision-making activities in the departments relating to sales, production, purchasing, finance, and accounting. The results of the sales forecasting might be influenced by numerous things. However, during their research,

researchers only collected a small number of information. In this project, both internal and external factors are examined. These include the weather, fuel costs, holidays, CPI, unemployment rate, and discounting tactics, all of which are thought to have a direct impact on consumer demand for sales in supermarkets and their departments. This analysis also takes into account the opinions of subject-matter experts. An experiment process is started to assess the performances of machine learning algorithms after past research papers and publications are used to determine appropriate approaches and algorithms. Additionally, based on those key variables, a novel forecasting approach is suggested. It was created as an accurate machine learning-based sales forecasting system for local supermarkets and the departments in Walmart USA supermarkets to meet the gaps in existing solutions.

II. RELATED WORKS

In this section, discussion and analysis of the research methodologies are done with the goal to forecast the sales.

Previously, there are a lot of sales and demand forecasting work that are performed using Machine Learning. Due to the importance of predictions in various fields, people have tried so many methods to improve the accuracy of the prediction. To handle this work, some of the methods such as auto regressive moving average (ARMA) and auto regressive integrated moving average (ARIMA) will be helpful[4]. Hence, we have performed a lot of trial and error during our process of making this working project. The best performing algorithm in our process was different from other projects that were aiming similar results. After a lot of research, we have tried not to replicate what has been done earlier and try to do something of our own with some traditional algorithms and different parameters and we have surely got better results.

Intelligent sales prediction using machine learning technique: The detailed study and analysis of comprehensible predictive models to improve future sales predictions are carried out in this research. Traditional forecast systems are difficult to deal with the big data and accuracy of sales forecasting. This research focuses on developing an intelligent sales prediction system through the application of machine learning techniques. The goal is to leverage historical sales data to create accurate forecasts, aiding businesses in optimizing inventory management and strategic decision-making. The study explores various machine learning algorithms, including but not limited to regression models, time series analysis, and ensemble methods. The proposed system aims to enhance prediction accuracy by considering factors such as seasonality, promotional events, and market trends.

Forecasting the retail sales of china's catering industry using support vector machines. The forecast of China's catering retail sales was studied in this paper. The seasonal impact was considered in the forecasting. The retail sales were predicted using the seasonal autoregressive integrated moving average (ARIMA) model. Our analysis takes into account the significant seasonal impact on sales, recognizing its pivotal role in the forecasting process. The study extends its predictive capabilities by incorporating the Seasonal Auto-Regressive Integrated Moving Average with exogenous factor (SARIMAX) model, leveraging its prowess in capturing temporal dependencies and trends within time series data. By combining the strength of SVM with the sophistication of seasonal SARIMAX, our approach aims to enhance the accuracy and reliability of retail sales predictions.

Sales prediction using machine learning algorithm. The aim of this paper is to propose a dimension for predicting the future sales of Big Mart Companies keeping in view the sales of previous years. A comprehensive study of sales prediction is done using Machine Learning models. The methodology involves the application of diverse Machine Learning algorithms such as Random Forest, Gradient Boosting, and Support Vector Machines, each contributing to a holistic understanding of sales dynamics. Feature importance analysis is conducted to identify key drivers influencing sales, providing valuable insights for business decision-makers. Additionally, the study delves into the impact of external factors, market trends, and seasonal variations on sales performance. The paper outlines the importance of model interpretability, enabling stakeholders to comprehend the factors influencing sales predictions and fostering trust in the proposed approach. A comparative analysis of different algorithms sheds light on their respective strengths and weaknesses in the context of sales forecasting.

Comparison of different machine learning algorithm for multiple regression on black Friday. This study focuses on the field of prediction models to develop an accurate and efficient algorithm to analyze the customer spending in the past and output the future spending of the customers with same features. This comprehensive study delves into the realm of prediction models with a specific focus on Multiple Regression applied to Black Friday sales data. The

primary objective is to formulate an algorithm that not only accurately analyzes past customer spending patterns but also excels in predicting future spending based on similar features. The research emphasizes the significance of understanding and leveraging historical customer behavior as a foundation for robust predictive models. The study investigates the impact of diverse features on customer spending, recognizing the importance of variables such us products, categories, discounts, and customer demographics.

III. METHODOLOGY

Sales forecasting is the process of predicting future sales. It is the vital part of the financial planning of the business. Most of the companies heavily depend on the future prediction of the sales. Accurate sales forecasting empower the organizations to make informed business decisions and it will help to predict the short-term and long-term performances. A precise forecasting can avoid overestimating or underestimating of the future sales, which may lead to great loss to companies. The past and current sales statistics is used to estimate the future performance. But it is difficult to deal with accuracy of sales forecasting by traditional forecasting. For this purpose, various machine learning techniques have been discovered. In this work, we have taken Black Friday dataset and made a detailed analysis over the dataset. Here, we have implemented the different machine learning techniques with different metrics. By analyzing the performance, we have trying to suggest the suitable predictive algorithm to our problem statement. Sales forecasting, a crucial aspect of strategic planning for businesses, has undergone a transformative evolution with the integration of machine learning and Angular View. This abstract encapsulates the key elements of this innovative approach. Machine learning emerges as a powerful solution, leveraging historical sales data and market trends to generate precise predictions. The synergy between machine learning and Angular View, a dynamic web application framework, introduces a paradigm shift in how businesses interact with and utilize sales forecasting insights. In conclusion, the abstract provides a concise yet comprehensive overview of the transformative impact of combining machine learning with Angular View in the realm of sales forecasting. This innovative approach empowers businesses to navigate the complexities of the market with agility, fostering a data driven decision-making culture and paving the way for sustained growth and competiveness.

Machine Learning algorithms are used to predict the accuracy of the sales which will be used to determine whether to increase or decrease the production. The performance of these algorithms is compared using different accuracy measurement methods.

Data Collection:Identify the relevant sources of historical sales data, such as point-of-sale systems, transaction databases, or online sales platforms. Collect data from as many sources as possible to ensure a comprehensive dataset. Ensure that the collected data is complete and includes all relevant information, such as sales date, item sold, quantity sold, price, and any other relevant variables. Check the quality of the collected data, including the accuracy of the data entry, any missing values, and any anomalies or outliers in the data. Remove any irrelevant data and clean the data before further analysis. Ensure that the data is in a format that is suitable for machine learning algorithms. This may involve converting the data into a structured format, such as a .csv or Excel file, and labelling the data appropriately.

Data Preprocessing: The process of Eliminating any duplicates, missing data, or outliers from the gathered data. This can be accomplished through data cleaning methods including interpolation, imputation, or the complete removal of the problematic data points. To guarantee that all variables have the same range of values, normalize or scale the data. Techniques like min-max scaling, normalization, or log transformation can be used for this. With the gathered data, find and extract pertinent features that can be utilized to train machine learning algorithms. To enhance model performance, this can entail adding new features or combining current features. Dividing the pre-processed data into training and testing sets will allow you to test the machine learning models' accuracy with a subset of the data.

Prediction: The proposed system for forecasting seasonal item sales using machine learning algorithms. SVM is a popular algorithm used for time-series forecasting. It works by finding a hyper plane that separates the data into different classes, with the aim of minimizing the prediction error. Seasonal Autoregressive Integrated Moving Average with exogenous variables (SARIMAX) is a statistical algorithm used for time-series forecasting. It models the time-series data as a combination of seasonal, autoregressive, and moving average components, and incorporates exogenous variables to improve forecast accuracy. Multi-layer perceptron (MLP) is a type of neural network used for time-series forecasting. It works by using multiple layers of interconnected neurons to model complex relationships between the input data and the output forecasts. Ensemble methods, Ensemble methods such as random forests or gradient

boosting can be used to combine the predictions of multiple machine learning algorithms. This can improve the accuracy and stability of the generated forecasts. Evaluation metrics, once the machine learning algorithms have been trained and tested on the pre-processed data, evaluate the accuracy of the generated forecasts using metrics such as Mean Absolute Error, Relative Absolute Error, or Root Mean Squared Error. He suggested method can be utilized to give accurate and dependable forecasts of seasonal item sales by utilizing various prediction approaches and assessing the correctness of the forecasts provided.

SARIMAX The seasonal component of unilabiate time series data is specifically supported by the ARIMA extension known as Seasonal Autoregressive Integrated Moving Average, or Seasonal ARIMA. Together with three new hyper parameters to calculate the auto regression (AR), differencing (I), and moving average (MA) for the seasonal component of the series, it also contains an additional parameter for the seasonality period. A seasonal ARIMA model is produced by adding additional seasonal components to the ARIMA. Although they also incorporate backshifts of the seasonal period, the words that make up the model's seasonal section are pretty comparable to those that make up its non-seasonal portions. SARIMA (p, d, q) x (P, D, Q, s) the parameters for these types of models are as follows: p and seasonal P: indicate number of autoregressive terms

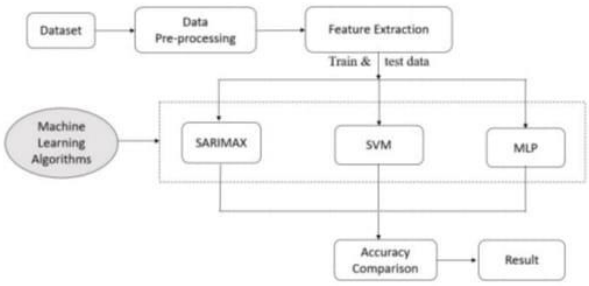

Fig.1 Block Diagram

Forecasting means predicting events of the future, typically based on previous records. For a long time, statistical models were commonly used for the conducting of predictions. The role of generalization in Machine Learning has been considered. Because there isn't much previous data available for a given time series, this impact could be utilized to anticipate sales when a new store or product is introduced [4]. The seasonal autoregressive integrated moving average (SARIMA), Multilayer Perceptron (MLP), and Support Vector Machine (SVM) are supervised learning techniques that are used to predict sales.Support Vector Machine Support Vector Machine or SVM is the common Supervised Learning algorithms used for both Classification and Regression issues. The SVM algorithm aims to build the best line or decision boundary that can divide n-dimensional space into conveniently place the new data point in the right category. The optimal choice boundary is called a hyper plane. SVM chooses extreme points vectors that help to create a hyper plane. Such extreme cases are called help vectors [9]. The equation for Support Vector Regression is: $f(x) = x' \beta + b$.

SARIMAX (Seasonal Autoregressive Integrated Moving Average with exogenous factor) in a sales prediction project can be highly effective, especially when dealing with time-series data that exhibit seasonality and other temporal patterns. Train the SARIMAX model using both the sales data and the exogenous features. Evaluate the model's performance using appropriate validation metrics, taking into account both the accuracy of sales forecasts and the ability of the model to capture the effects of exogenous factors.Choose appropriate SARIMAX orders (p, d, q) based on autocorrelation and partial autocorrelation plots of the sales data. Determine the seasonal orders (P, D, Q) by analyzing seasonal patterns in the data. Decide which exogenous variables to include in the model based on their significance and impact on sales.

Fig.2 Annual vehicle prediction

Data Set Collection:

Data set are collected from Kaggle website and used for sales forecasting, leveraging advanced techniques and models to make accurate predictions and inform business decisions.Choose a dataset that suits your forecasting needs. Look for datasets that include historical sales data along with any additional information you may want to use as exogenous factors (e.g., economic indicators, marketing data). Once the suitable dataset was found, download it from Kaggle. Make sure to read any accompanying documentation or descriptions to understand the dataset's contents and format.

I. RESULTS AND DISCUSSION

By utilizing evaluation criteria like mean absolute error, mean squared error, or root mean square error, you may compare the predicting accuracy of the various machine learning methods. This will make it easier to determine which algorithm is best for the available dataset. To assess the projections' accuracy, visualize the generated forecasts and contrast them with the actual sales data. Line plots, scatter plots, and time series decomposition plots are some of the methods that can be used to do this. By contrasting the anticipated values with the actual values, analyses the forecasting inaccuracies. Recognize any systematic biases or trends in the errors that can point to forecasting model flaws. The integration of machine learning and Angular for sales forecasting presents a transformative approach, combining advanced analytics with a user-friendly interface. This dynamic synergy not only empowers businesses to make informed decisions but also enhances the overall forecasting experience. The combination of machine learning and Angular fosters a data driven decision-making culture, empowering stakeholders with actionable insights derived from comprehensive data analysis. The use of machine learning algorithms significantly improves the precision of sales forecasts, enabling businesses to anticipate market trends with greater accuracy. The Angular interface offers a user-centric design, facilitating seamless interaction and making complex forecasting insights accessible to a wide range of users. The flexibility of Angular allows for agile adaptability, ensuring the platform evolves with changing business needs and technological advancements in the field of sales forecasting. The combination of machine learning and Angular fosters a data-driven decision-making culture, empowering stakeholders with actionable insights derived from comprehensive data analysis.

Sales forecasting is essential to the business sector in all industries. Sales revenue analysis will assist in obtaining the information required to estimate both the revenue and the income with the aid of the sales predictions. The proposed system for forecasting seasonal item sales using machine learning algorithms can provide accurate and reliable forecasts that can help businesses optimize their inventory management and increase sales revenue. Data collection, data pre-processing, feature extraction, prediction, and outcome analysis are only a few of the system's crucial phases. Methods like the multi-layer perceptron, the seasonal autoregressive integrated moving average with exogenous variables, and the support vector machine can be used for forecasting. Performance indicators including correlation, mean absolute error, root mean square error, and relative absolute error are calculated for each of the three machine learning algorithms. According to research, the Multilayer Perceptron is the most effective technique for forecasting sales

II. CONCLUSION

In this sales prediction project, we successfully combined modern web development technologies with machine learning to create a valuable tool for sales forecasting and decision-making. Our project leveraged Angular for the frontend user interface, Flask for the backend API, and machine learning models to predict sales.We trained and deployed machine learning models (e.g., regression, time series forecasting) to predict sales accurately. These models have shown promising results in capturing sales trends and seasonality.The Angular frontend provided an intuitive and user-friendly interface for users to input data and view predictions

REFERENCE

1. Forecasting in financial accounting with artificial intelligence – A systematic literature review and future research agenda Marko Kureljusic, Erik Karger Journal of Applied Accounting ResearchISSN: 0967-5426Open Access. Article publication date: [10 May 2023]

2. Retail Sales Forecasting Using Deep Learning: Systematic Literature ReviewApril 2022Complex Systems Informatics and Modeling QuarterlyDOI:10.7250/csimq.[2022-30.03]

3. A systematic literature review on price forecasting models in construction industryMingxue Ma,Vivian W. Y. Tam,Khoa N. Le &Robert Osei-Kyei Received [25 Nov 2022, Accepted 10 Jun 2023]

4. Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan & Susan Treesa (2018) Intelligent Sales Prediction Using Machine Learning Techniques.

5. Xiangsheng Xie & Gang Hu (2008). Forecasting the Retail Sales of China's Catering Industry.

6. Avinash kumar, Neha Gopal & Jatin Rajput(2020). An Intelligent Model For Predicting the Sales of a Product.

7. Purvika Bajaj, Renesa Ray, Shivani Shedge & Shravani Vidhate(2020). SALES PREDICTION USING MACHINE LEARNING ALGORITHMS.

8. Ching-Seh (Mike) Wu. Pratik Patil & Saravana Gunaseelan(2018). Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data.

9. Nikhil Sunil Elias, Seema Singh(2019).FORECASTING of WALMART SALES using MACHINE LEARNING ALGORITHMS.