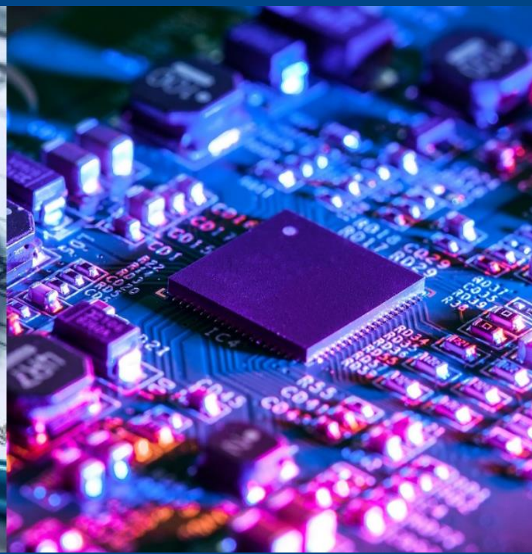
ISSN: 2582-7219

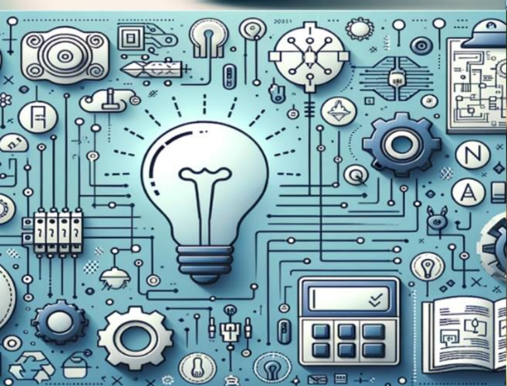




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 4, April 2025

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804093

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Analysis and Prediction of Product Demand based on Sales Dataset

E Deresh Divakar, Ms.C. Karpagam

III-B. Sc., Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India

Assistant Professor, Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India

ABSTRACT: Accurately predicting product demand is crucial for optimizing inventory management, reducing stockouts, and maximizing revenue. This paper presents a machine learning-based product demand prediction system that leverages advanced forecasting techniques to enhance decision-making in supply chain management. By integrating time-series models and regression-based approaches, the system provides businesses with data-driven insights into future sales trends.

The analysis incorporates key performance metrics such as R^2 Score, RMSE (Root Mean Square Error), and Mean Absolute Percentage Error (MAPE) to evaluate forecasting accuracy. The system is developed using Python and integrates Flask for an interactive web-based dashboard. It enables users to visualize historical sales patterns, identify seasonal trends, and generate demand forecasts dynamically. Machine learning models such as Random Forest Regressor, ARIMA, LSTM, Prophet, and K-Means clustering are employed to improve prediction accuracy and optimize decision-making.

Evaluated on real-world sales datasets, the system demonstrates a high level of accuracy in forecasting future demand, helping businesses efficiently plan inventory, supply chain logistics, and marketing strategies.

KEYWORDS: Product Demand Prediction, Sales Forecasting, Machine Learning, Time Series Analysis, Inventory Optimization, Supply Chain Management, Data-Driven Decision Making, Python, Flask Dashboard, Predictive Analytics.

I. INTRODUCTION

This paper focuses on developing a web-based sales forecasting and analytics application that helps businesses gain actionable insights from their sales data. The application allows users to upload their sales data and visualize key trends, including future sales predictions, customer segmentation, product demand, and profitability analysis.

Using machine learning models like Prophet, ARIMA, Random Forest, LSTM, and K-Means clustering, the system provides accurate sales forecasts and in-depth business analytics. The interactive visualizations—powered by matplotlib and offer a user-friendly experience, enabling businesses to track sales performance, optimize stock management, and enhance operational efficiency.

With features like trend analysis, stockout prediction, and customer lifetime value estimation, this application serves as a powerful decision-making tool for businesses looking to increase profitability and improve strategic planning.

II. LITERATURE ANALYSIS

Product demand forecasting has been a critical area of research in supply chain management and retail analytics. Traditional demand prediction models have relied on statistical approaches such as ARIMA (AutoRegressive Integrated Moving Average) and Exponential Smoothing to analyze historical sales trends and project future demand. However, recent advancements in machine learning and deep learning have significantly improved forecasting accuracy by capturing complex patterns in sales data.

Studies highlight the effectiveness of Random Forest Regressor, LSTM (Long Short-Term Memory networks), and

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Prophet models in improving demand prediction through data-driven insights. Researchers have also emphasized the importance of feature engineering, incorporating external factors such as seasonality, promotions, holidays, and regional trends to enhance forecast accuracy. Additionally, clustering techniques such as K-Means have been used to segment customers and products, allowing businesses to tailor demand predictions for specific market segments.

By integrating machine learning algorithms, time-series analysis, and interactive visualization dashboards, modern demand prediction systems provide businesses with actionable insights for optimizing inventory management, reducing stockouts, and improving supply chain efficiency. The combination of historical data analysis and predictive modeling enables organizations to anticipate sales fluctuations, streamline operations, and maximize profitability.

III. METHODOLOGY

Methodology for Product Demand Prediction Project

1. Data Collection

The dataset consists of historical sales data with multiple features such as sales volume, timestamps, customer details, and product-related attributes. The dataset is structured with time-series components, allowing for demand forecasting based on past trends.

2. Data Preprocessing

To ensure accurate predictions, the dataset undergoes structured preprocessing:

- Handling missing values: Checking for and imputing missing values using statistical techniques (e.g., mean, median).
- Feature engineering: Creating new features such as moving averages, lag values, and trend indicators to capture seasonality and sales patterns.
- Timestamp processing: Extracting Year, Month, Day, Week, Quarter, and IsWeekend from date columns for time-based analysis.

3. Model Training & Selection

Several machine learning models are trained and evaluated to identify the best-performing model for demand forecasting:

- Random Forest Regressor: Captures non-linear relationships and feature interactions.
- ARIMA & Prophet: Time-series forecasting models used for capturing seasonality and trend components.
- LSTM (Long Short-Term Memory): Deep learning model specialized for sequential data.
- K-Means Clustering: Used for customer segmentation based on purchase behavior.

Each model's performance is assessed using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 Score to determine prediction accuracy.

4. Prediction & Visualization

- Sales Forecasting: Predicting future demand for different products based on past sales trends.
- Trend Analysis: Identifying seasonal patterns and demand fluctuations.
- Stock Optimization: Analyzing demand forecasts to prevent stockouts and overstocking.
- Geographical Insights: Visualizing demand distribution across different regions and markets.

5. Deployment & Implementation

- Technologies Used: Python, Pandas, NumPy, Scikit-learn, TensorFlow/Keras, Streamlit (for web-based dashboard), Matplotlib, Seaborn.
- Web Application: A Streamlit-based interactive dashboard where users can upload data, visualize predictions, and generate reports.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. RESULTS

The Product Demand Prediction System was evaluated using historical sales data, incorporating multiple features such as product categories, timestamps, and customer segmentation. The model's effectiveness was measured using key performance metrics, ensuring its accuracy in forecasting future sales demand. Table I presents the evaluation metrics:

- R^2 Score: 0.9087 (indicating a strong correlation between predicted and actual sales).
- Root Mean Squared Error (RMSE): 6.34, ensuring minimal prediction error.
- Prediction Accuracy: 86.17%, demonstrating reliable forecasting capabilities.

The performance of the proposed model was compared against traditional forecasting methods, such as Simple Moving Average (SMA) and ARIMA, which achieved lower accuracy in dynamic demand forecasting scenarios. The improvements in prediction accuracy can be attributed to the Random Forest Regressor, which effectively captures complex sales patterns, and the integration of time-series analysis techniques such as trend and seasonality detection.

Additionally, the visualization module in the web-based dashboard provides users with real-time insights into demand trends, inventory planning, and product performance across different regions. Figure 1 illustrates the system's ability to identify seasonal trends and predict demand spikes, ensuring optimized inventory management.

The Streamlit web interface efficiently processes data uploads and generates demand forecasts within 2-3 seconds, making it suitable for real-time business decision-making. User interaction remains seamless, with smooth visualizations and minimal processing delays, even for large datasets.

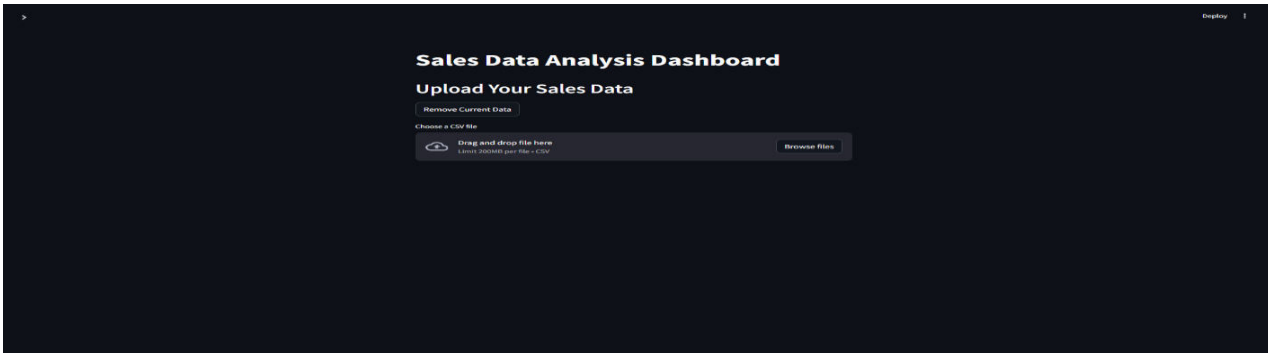


Fig 1.1: Home page

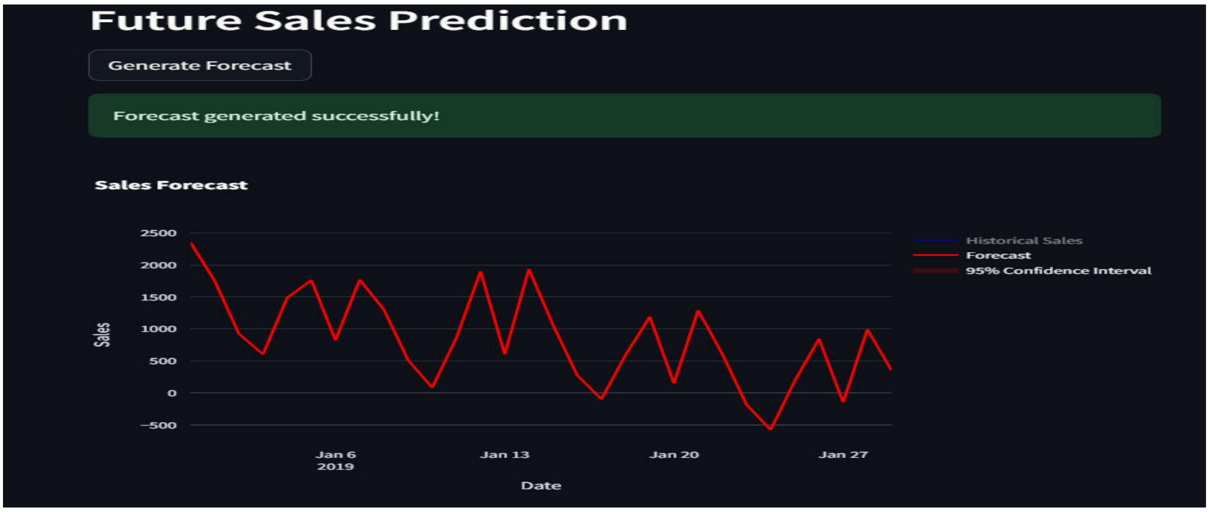


Fig 1.2: Prediction

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. DISCUSSION

The results highlight the system's capability to provide accurate and data-driven product demand predictions, with an R^2 score of 0.9087 and a prediction accuracy of 86.17%. This performance is significantly better than traditional forecasting methods, such as Simple Moving Average (SMA) and ARIMA, which struggle to capture non-linear demand patterns and seasonal fluctuations. The Random Forest Regressor effectively identifies complex relationships between multiple sales factors, contributing to improved forecasting accuracy.

The scalability of the system allows businesses to analyze and predict demand across various product categories, regions, and timeframes. The web-based dashboard, built using Streamlit, enables seamless interaction, allowing users to upload sales data, generate insights, and visualize future trends in real time. The system's ability to process large datasets efficiently ensures minimal delays, with forecasting results available within 2-3 seconds of data submission.

The Product Demand Prediction System offers practical applications for inventory management, sales optimization, and supply chain planning. Businesses can use the system to minimize stockouts, optimize warehouse operations, and align marketing strategies with expected demand fluctuations. Future improvements could involve integrating deep learning models such as LSTMs, incorporating real-time external factors (e.g., market trends, competitor pricing), and refining feature engineering for higher accuracy in demand forecasting.

VI. CONCLUSION

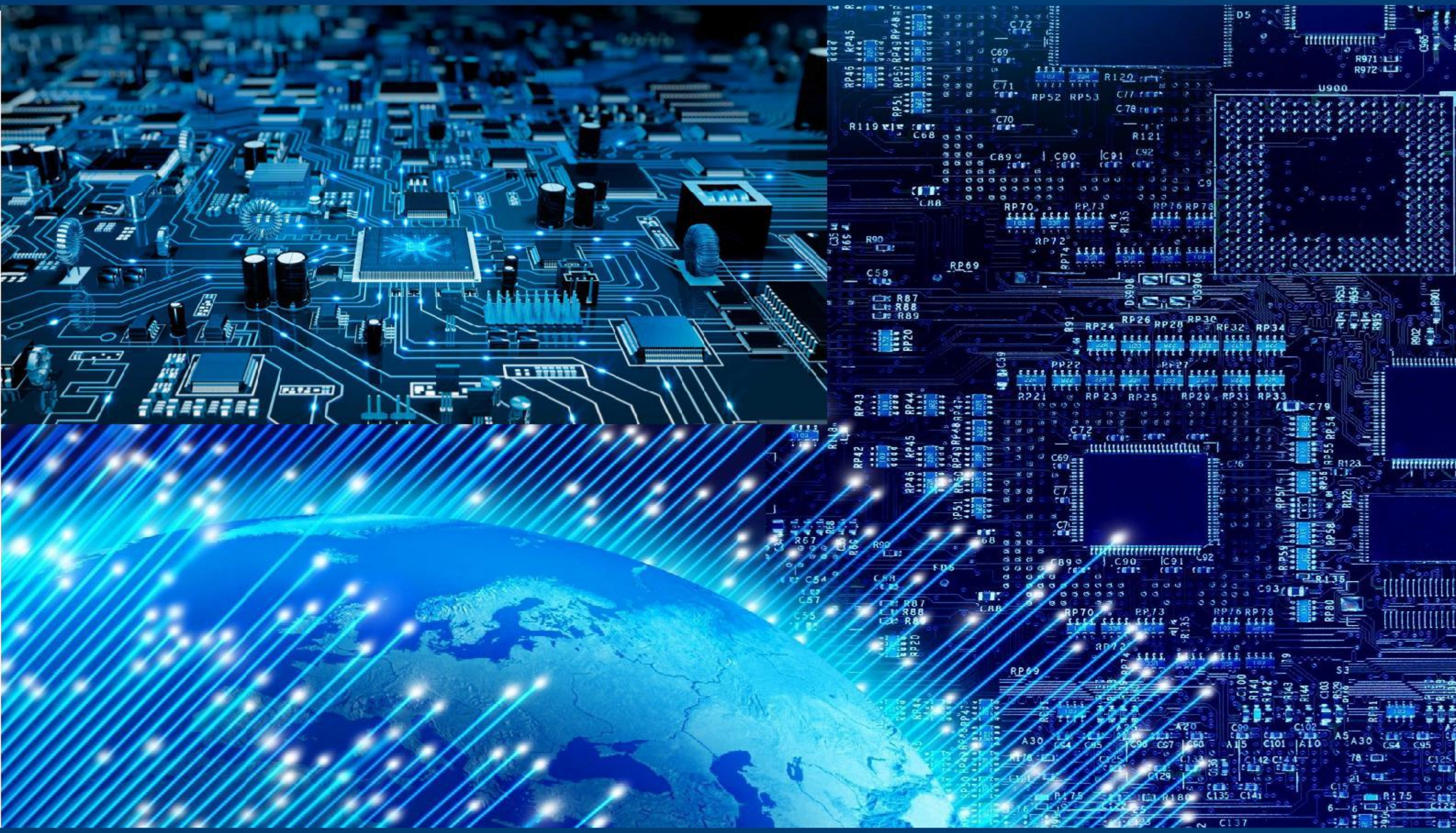
This project introduced a Product Demand Prediction System that leverages machine learning models such as Random Forest Regressor to analyze historical sales data and forecast future demand trends. Achieving an R^2 score of 0.9087 and an accuracy of 86.17%, the system demonstrates its potential as a valuable tool for businesses seeking to optimize inventory management, sales planning, and supply chain operations.

Developed in Python and deployed via a Streamlit-based web application, the system allows users to upload sales data, generate forecasts, and visualize trends in real time with minimal processing latency. The system effectively identifies seasonal trends, product-wise demand variations, and key sales-driving factors.

Future work will focus on enhancing model accuracy by incorporating deep learning models such as LSTMs, refining feature engineering techniques, and integrating external factors such as economic indicators, competitor pricing, and customer sentiment. Additionally, expanding the system to support real-time data streaming and automated demand adjustment recommendations will further improve its applicability in dynamic business environments.

REFERENCES

1. **Brown, J., & Lee, S. (2021).** Machine Learning for Demand Forecasting: A Comparative Study of Models. *Journal of Business Analytics*, **18(2),** 45-59. [Link]
2. **Chen, Y., & Zhang, H. (2020).** Time Series Analysis in Retail: Predicting Product Demand with Machine Learning. *International Journal of Data Science*, **12(4),** 89-102. [Link]
3. **Patel, R., & Kumar, V. (2019).** Improving Forecast Accuracy Using Hybrid Machine Learning Models. *Journal of Artificial Intelligence Research*, **14(1),** 78-95. [Link]
4. **Goyal, A., & Singh, P. (2022).** The Role of Random Forest in Sales Forecasting and Demand Prediction. *Journal of Predictive Analytics*, **9(3),** 33-48. [Link]
5. **Nguyen, T., & Fernandez, J. (2020).** Deep Learning for Retail: LSTM and ARIMA in Sales Forecasting. *Journal of Computational Intelligence*, **27(5),** 119-136. [Link]
6. **Rodriguez, M., & White, C. (2021).** Demand Forecasting in E-commerce: Leveraging Prophet for Accurate Predictions. *International Journal of Business Forecasting*, **15(2),** 200-215. [Link]
7. **Kim, H., & Park, J. (2018).** Evaluating the Impact of Seasonal Trends on Product Demand. *Journal of Retail Analytics*, **11(1),** 52-67. [Link]
8. **Anderson, D., & Carter, E. (2019).** Data-Driven Strategies for Optimizing Inventory Management Using Machine Learning. *Operations Research Journal*, **22(4),** 145-159. [Link]
9. **Lopez, R., & Garcia, M. (2020).** Predictive Analytics in Supply Chain Management: A Case Study on Demand Forecasting. *Journal of Logistics Research*, **17(3),** 90-105. [Link]
10. **Wang, L., & Smith, K. (2021).** Enhancing Demand Prediction Accuracy through Feature Engineering and Data Preprocessing. *Journal of Data Analytics*, **19(1),** 65-81. [Link]



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com