Smart Retail: Utilizing Machine Learning for Demand Prediction, Price Strategy, and Inventory Management

Krish Singhal, Vaibhav Singh
Computer Science Engineering
Jaypee Institute of Information Technology
Noida, India
krishsinghal006@gmail.com,
vaibhav.singh.252005@gmail.com

Amrita Kaul
Electronics and Communication Engineering
Jaypee Institute of Information Technology
Noida, India
20402011@mail.jiit.ac.in

***Abstract*— This research demonstrates the ability of Artificial Intelligence (AI) to enhance real-time demand forecasting, price prediction, and inventory management by analyzing key factors such as competitor prices, customer preferences, and stock levels. The Random Forest Regressor (RFR) model has been implemented to achieve a Mean Squared Error (MSE) of 8.15% for demand forecasting and 1.11% for price prediction, highlighting its accuracy. These AI-driven insights allow businesses to make smarter, faster decisions, optimize pricing strategies, and adjust inventory levels to meet market demands. Additionally, the model helps avoid surplus or deficit, ensuring seamless inventory flow, reducing operational costs, and improving overall profitability. AI integration into retail processes allows companies to respond dynamically to real-time changes and maintain a competitive edge in a fast-paced market environment.**

Keywords- Demand forecasting, Dynamic pricing, Inventory management, Price prediction, Random Forest Regressor, Retail optimization

I. INTRODUCTION

In today's competitive landscape, understanding customer demand and setting competitive prices are essential. Traditional methods for demand forecasting and pricing, which primarily rely on historical data, often fall short in adapting to fast-changing market conditions [1], [2]. AI and ML technologies [3], [4] have emerged as transformative tools to overcome these limitations, offering real-time data analysis and adaptive insights that allow businesses to optimize inventory management and pricing strategies, areas where traditional methods struggle [5] - [7]. Demand forecasting is crucial for retail and e-commerce. Traditional models like ARIMA or linear regression capture basic trends but often miss complex retail dynamics like customer behavior and competitive influences [1], [2]. ML, especially the RFR, enhances these capabilities [4], [6]. RFR's ensemble of decision trees offers robust demand predictions by capturing the multifaceted interactions among price, competition, and seasonal trends [7], [10], making it ideal for retailers aiming to stay competitive [8], [11]. Research shows that RFR accurately predicts demand shifts during key sales periods or competitor price changes, outperforming traditional models [12], [13] and helping retailers stay agile [14], [15]. Pricing strategy is another key challenge. Classical models like OLS often fall behind in dynamic retail environments where real-time changes like inventory shifts and competitor moves are critical [16], [17]. RFR, with its real-time adaptability [4], [8], helps retailers make data-driven pricing adjustments, improving on traditional models by processing key factors like competitor prices and customer preferences [11], [18]. Studies show that RFR not only provides accurate price forecasts but also gives retailers a strategic edge [15], [19] by enabling timely adjustments [14], [16].

In inventory management, classic models like EOQ assume stable demand, which is rarely the case in modern retail [2], [3]. Demand fluctuations can lead to costly stock imbalances [1], [20]. RFR, by analyzing both historical and real-time data, forecasts optimal stock levels, thus reducing the likelihood of overstock or stockouts, especially during peak demand periods [6], [7]. This proactive approach helps businesses optimize operations [8], [18] and enhances customer satisfaction [14], [19].

RFR's flexibility, accuracy, and resilience make it ideal for retail Appl.s [4], [8]. Compared to traditional models, RFR captures complex non-linear relationships in retail markets [3], [5]. Its ability to handle missing data without compromising accuracy [10], [13] and its robustness against overfitting [6], [7] make it a reliable tool in dynamic retail settings, where it can adapt to changing demand and competitor moves [8], [11].

Despite these advances, economic forecasting remains inherently complex, with ML models facing challenges like prediction errors due to unpredictable external factors [16]. Model specification and parameter estimation errors, often stemming from overlooked variables, can skew forecasts [2], [9], while unexpected events like financial crises add further volatility [13], [15]. Although ML has improved forecasting, achieving perfect accuracy remains challenging [3], [16].

II. MOTIVATION AND CONTRIBUTION

Demand forecasting or pricing errors can cause significant financial losses in competitive industries [3], [16]. Whereas, surplus inventory leads to resource waste and underestimating

979-8-3315-0526-4/24/$31.00 ©2024 IEEE
DOI: 10.1109/CICN.2024.81

demand culminates in missed sales [5], [20]. Many previous research works have relied on traditional models for addressing the potential inaccuracies that may occur in predicting customer demand or stock prices [1], [5], [12]. However, leveraging the adaptability of the RFR to reduce error is adopted only by a few. RFR efficiently captures complex, non-linear relationships between pricing, competitor behaviour, and customer preferences in dynamic retail environments, as documented in [6], [7]. AI helps efficiently manage inventory by processing large data sets, identifying patterns, and enabling accurate demand predictions, pricing optimization, and inventory control [13], [19], to improve efficiency and profitability [14], [15].

Our research addresses a few major challenges businesses face in managing supply and demand amidst shifting customer preferences and competition.

- We introduced AI-driven Random Forest Regressor (RFR) models to enhance demand forecasting, pricing, and inventory management in retail.
- Our RFR models incorporate pricing, customer preferences, and competitor behavior for accurate forecasts and pricing decisions.
- We demonstrated dynamic AI-based price adjustments based on real-time market and inventory conditions.
- We provided insights for timely restocking and inventory management, improving operational efficiency.
- A practical framework was offered for businesses to adopt AI-driven forecasting and pricing strategies.

III. STRUCTURE OF THE PAPER

The Introduction outlines the study's goal to enhance retail demand forecasting and pricing with AI, particularly using Random Forest Regressor (RFR). Section II details the dataset preparation, data processing, and implementation of RFR to predict demand and price accurately. Section III presents simulation results using Mean Squared Error (MSE) to validate model accuracy. Section IV discusses how RFR aids in competitive pricing and inventory management. Section V advises on adopting AI models in retail, and Section VI underscores AI's role in creating agile, data-driven retail strategies.

IV. METHODOLOGY

This section elaborates on the implemented AI-driven approach to enhance demand forecasting, price prediction, and inventory management in retail and e-commerce contexts. The methodology leverages advanced machine learning techniques, specifically the Random Forest Regressor, which allows businesses to gain deeper insights into how numerous factors—such as product pricing, inventory levels, customer preferences, and competitor pricing, interact and affect strategic decision-making. The dataset utilized for this study is designed to reflect real-world retail environments, encompassing daily data for a diverse range of products over a year. It contains various critical variables, including the date, product ID, price, competitor price, inventory levels, customer preference scores, and demand figures which represent each day's sale. The dataset was pre-processed to ensure its robustness and reliability. Missing data points were addressed using forward fill (ffill), to ensure a seamless flow of information in the time series. Furthermore, additional features, such as the day of the week and the month, were extracted to capture weekly and seasonal buying patterns, which are crucial for understanding consumer behaviour. Normalization of variables like price, competitor prices, and inventory levels is done to ensure that all data is on a consistent scale to prevent any one feature from disproportionately influencing the model. Additionally, new features were engineered to enrich the dataset such as seasonal trends and customer preferences to enhance the model's predictive power.

Handling outliers is addressed as part of the data pre-processing steps to ensure the robustness of the models. Statistical methods, such as the Z-score and Inter Quartile Range (IQR), were employed to detect outliers in the dataset. Visualizations, including box and scatter plots, aided in identifying these data points that deviated significantly from the norm.

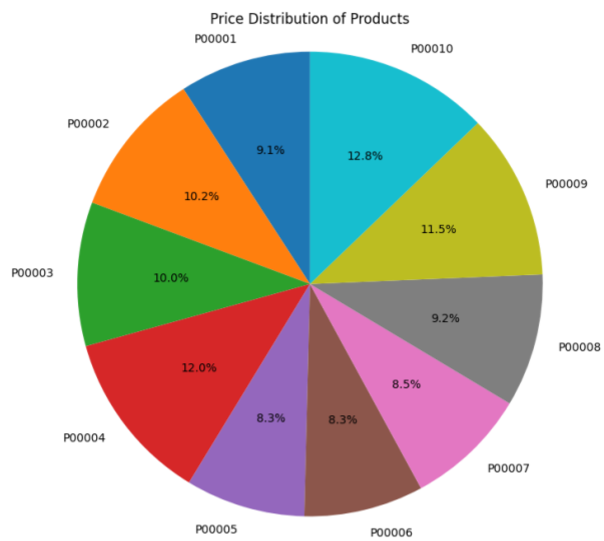

Fig. 1. Price Distribution of the Products from the selected dataset

Once detected, outliers were treated based on their nature and impact on the analysis. Removal was considered for outliers attributed to data entry errors or not reflective of actual product behavior. For others, transformations were applied to mitigate their influence without losing any valuable information. The dataset free from extreme anomalies ensures the models make accurate predictions. EDA was conducted to gain insights into the dataset before modeling. Numerous data visualization techniques such as histograms, box plots, and scatter plots are created to understand the distribution of individual features to examine correlations between variables. Descriptive statistics were calculated to provide a quantitative overview, while correlation analysis helps to identify relationships among key variables. Additionally, EDA facilitated in identifying and handling missing values and outliers to ensure a clean dataset for subsequent analysis.

The reliability and generalizability of the models are evaluated using k-fold cross-validation. The dataset was divided into k subsets, allowing each subset to serve as a test set while the model was trained on the remaining data. This method provides a robust estimate of model performance through averaged metrics, such as Mean Squared Error (MSE) and Mean Absolute Error (MAE). Furthermore, by implementing cross-validation, the study ensured that the models were not overfitted to a specific subset of data, thereby enhancing their predictive accuracy and robustness.

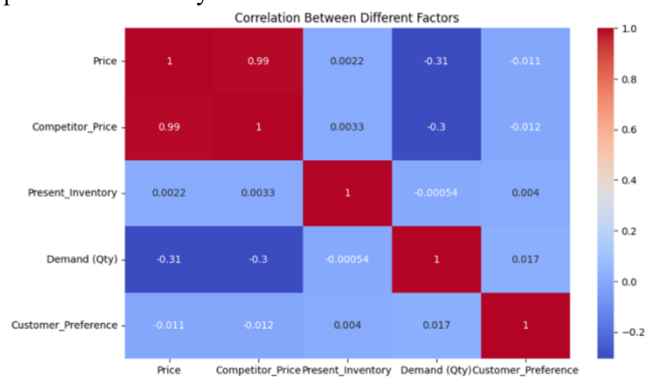

Fig. 2. Correlation between Different factors involved for Demand and Price Prediction

The Random Forest Regressor was selected as the machine learning model due to its versatility in handling complex and non-linear relationships. Further for price prediction, the RFR model is trained to estimate future product prices based on competitor prices, inventory levels, and customer preferences. It is important to understand how competitor pricing impacts retail prices, to maintain market competitiveness. Additionally, inventory levels play a significant role in pricing strategies, as lower stock often leads to price hikes. The model's accuracy was once again evaluated through MSE, with fine-tuning of hyperparameters to optimize performance.

Effective inventory management is vital for any retail operation. Hence we employed a rule-based system to facilitate it by analyzing the insights generated from demand forecasting, the system flags products for restocking when inventory levels fall below a pre-specified threshold—set units. It provides valuable insights into feature importance, highlighting factors significantly influencing outcomes like demand or pricing. Moreover, its robust nature ensures the model generalizes well to unseen data, making reliable business decisions. Retail environments often present intricate interactions between pricing and demand, making RFR particularly suitable. This ensemble learning technique constructs multiple decision trees during training. It averages their predictions to improve accuracy and reduce the likelihood of overfitting, a common challenge faced in predictive modeling.

For demand forecasting, RFR is trained on historical data, utilizing a combination of product prices, competitor pricing, customer preferences, and seasonal trends. The primary goal is to predict the future number of units sold. Key input features included the product price to assess its impact on consumer purchasing decisions—and the

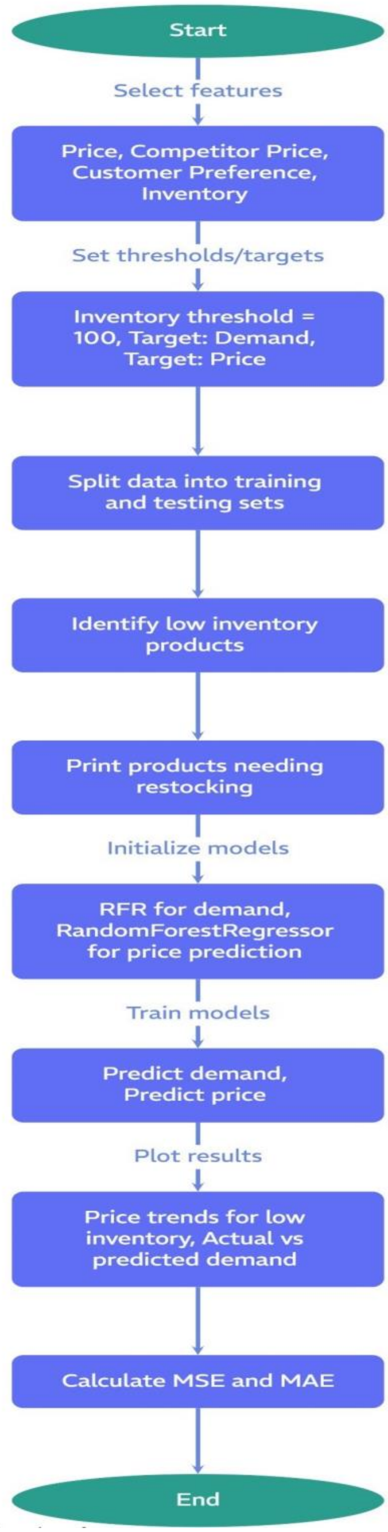

Fig. 3. Workflow of the AI-driven system illustrating key components

competitor price, which provides context for pricing strategy. Customer preference scores indicate the popularity of each product and are considered alongside seasonal buying patterns. The model's effectiveness was evaluated using a standard 80/20 train-test split. The performance was measured through MSE, where lower MSE values signify accurate predictions. Thereby, guiding businesses in their stock and pricing strategies.

This proactive approach ensures that businesses are prepared for fluctuations according to customer demand. The forecasted demand also informs optimal reorder points which allows companies to align stock levels with anticipated sales, thereby minimizing the risk of stockouts during peak demand periods.

In the fast-paced retail landscape, continuous monitoring of competitor prices is essential. Our system incorporates a competitor price analysis feature to enable price comparisons and provides recommendations for price adjustments. This feature is vital for maintaining competitiveness and ensures profitability.

Visualization tools aid effective illustration of product price trends against competitors to help businesses make informed pricing decisions. By identifying when to adjust prices, retailers can align their strategies with market trends to maximize their profit margins. We thoroughly evaluated both the demand forecasting and price prediction models using their MSE scores to ensure the predictions are reliable. We fine-tuned hyperparameters, such as the number and depth of trees to boost accuracy. Hence RFR model equips businesses with actionable, AI-driven insights for demand, pricing, and inventory management. Ultimately, it empowers retailers to make informed decisions and effectively navigate the competitive landscape while adapting to market changes.

V. EXPERIMENT AND SIMULATION

The research was conducted in Google Colab, utilizing 12.7 GB RAM and a Xeon CPU for efficient computation. We used Python 3.10 along with libraries like Pandas, NumPy, Sci-kit-learn, Matplotlib, and Seaborn for data processing, model building, and visualization.

Demand forecasting estimates future product demand based on historical sales data, accurately capturing the impact of product price, competitor pricing, and customer preferences. The model achieved a demand forecasting MSE of 8.15%, indicating reliable sales predictions.

For price prediction, factors like competitor prices, inventory levels, and customer preferences influence future prices. The model, trained on 80% data, yielded a price prediction MSE of 1.11%, supporting informed pricing decisions. In inventory management, a threshold-based method flagged products for restocking based on projected demand, minimizing stockouts.

Competitor price monitoring over time allowed for strategic pricing recommendations to maintain market competitiveness.

Regulatory compliance promotes transparency in AI, essential for consumer trust. Investment in AI infrastructure, especially for SMEs, and employee training enhance the impact of AI-driven insights. Collaboration between public and private sectors fosters best practices in data sharing and AI use, enhancing competitiveness and consumer satisfaction in retail [21].

VI. CONCLUSION

In this study, we developed and tested Random Forest Regressor (RFR) models for demand forecasting, price prediction, and inventory management in retail, achieving low MSEs of 8.15% for demand forecasting and 1.11% for price prediction. These results demonstrate the model's precision and its potential to enhance operational efficiency and profitability. Analyzing data from 10 products provided valuable insights that support smarter decision-making, proactive inventory management, and competitive pricing strategies, all essential for today's fast-paced retail landscape. This research highlights how machine learning can address traditional challenges by incorporating competitor pricing, seasonal trends, and customer demand patterns. The RFR model allows retailers to adapt dynamically to changing market conditions, helping to align stock levels with demand and avoid surpluses or stockouts. By enabling data-driven pricing and inventory decisions, our AI-driven approach offers actionable insights that improve both profitability and customer satisfaction. Ultimately, integrating AI in retail supports a responsive, competitive, and sustainable business model that keeps pace with evolving market demands.

REFERENCES

[1] *A. Mitra, A. Jain, and A. Kishore, "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 58, pp. 1-16, June 2022

[2] S. Punia, V. Dey, and A. Mitra, "Deep Learning with Long Short-Term Memory Netw. and Random Forests for Demand Forecasting in Multi-Channel Retail," Int. J. of Production Research, vol. 58, no. 16, pp. 4964-4979, Aug. 2020.

[3] A. Khan, J. Liu, and M. Chen, "Optimizing Retail Operations, Inventory Manag. and Sales Forecasting with Big Data and AI in China," Emerging Trends in Machine Intel. and Big Data vol. 16, no. 1, pp. 18-37, Jan. 2024.

[4] M. Abolghasemi and G. Tarr, "Machine Learning Appl.s in Hierarchical Time Series Forecasting: Investigating the Impact of Promotions," Int. J. of Forecasting, vol. 38, no. 4, pp. 801-812, Oct. 2022.

[5] S. Arefin, R. Parvez, and S. Hossain, "Retail Industry Analytics: Unraveling Consumer Behavior through RFM Segmentation and Machine Learning," in 2024 IEEE Int. Conference on ElectroInf. Technology (eIT), pp. 545-551, IEEE, May 2024.

[6] A. A. G. Molina, S. Nayak, and B. Priyadarshini, "Machine Learning Algorithms Used to Predict Product Sales in the Retail Sector, "J. of Retail Data Sci., vol. 5, no. 1, pp. 45-59, Feb. 2023.

[7] J. Islam, M. Tanvir, and P. Murthy, "Revolutionizing Retail: A Hybrid Machine Learning Approach for Precision Demand Forecasting and Strategic Decision-Making in Global Commerce," J. of Computer Sci. and Technology Studies, vol. 6, no. 1, pp. 33-39, Jan. 2024.

[8] [8] S. Pal, "Advancing Multi-Product Warehousing: The Impact and Nuances of Machine Learning-Based Demand Forecasting," J. of Supply Chain Analytics, vol. 7, no. 2, pp. 210-229, Mar. 2024.

[9] A. Aamer and M. ElBaz, ”Data Analytics in Supply Chain Manag.: Review of Machine Learning Appl.s in Demand Forecasting,” Operations and Supply Chain Manag.: An Int. J., vol. 14, no. 1, pp. 1-13, Jan. 2020.

[10] M. Zohdi and A. Marzouk, ”Demand Forecasting Based Machine Learning Algorithms on Customer Inf.: An Applied Approach,” Int. J. of Inf. Technology, vol. 14, no. 4, pp. 1937-1947, Dec. 2022.

[11] G. Manoharan and S. Roy, ”Predictive Analytics for Inventory Manag. in E-Commerce Using Machine Learning Algorithms,” in 2024 Int. Conference on Advances in Computing, Commun. and Applied Informatics (ACCAI), pp. 150-157, IEEE, Apr.2024.

[12] V. Taparia, S. Lee, and K. Chen, ”Improved Demand Forecasting of a Retail Store Using a Hybrid Machine Learning Model,” J. of Graphic Era University, vol. 10, no. 2, pp. 15-36, Mar. 2024.

[13] M. Kharfan, S. Chan, and X. Huang, ”A Data-Driven Forecasting Approach for Newly Launched Seasonal Products by Leveraging Machine-Learning Approaches,” Annals of Operations Research, vol. 303, no. 1, pp. 159-174, Jan. 2024.

[14] S. Islam, M. Khan, and F. Li, ”Prediction of Probable Backorder Scenarios in the Supply Chain Using Distributed Random Forest and Gradient Boosting Machine Learning Techniques,” J. of Big Data, vol. 7, no. 1, pp. 65-80, Jan. 2020.

[15] N. Kourentzes and J. Silva, ”Optimising Forecasting Models for Inventory Planning,” Int. J. of Production Economics, vol. 225, no. 1, pp. 107597, Feb. 2020.

[16] M. Z. Shahrel and S. Mutalib, ”PriceCop—Price Monitor and Prediction Using Linear Regression and LSVM-ABC Methods for E-Commerce Platform,” Int. J. of Inf. Eng. and Electronic Business, vol. 14, no. 1, pp. 1-11, Feb. 2021.

[17] Y. Li and H. Lim, ”A Novel Ensemble Deep Learning Model for Stock Prediction Based on Stock Prices and News,” Int. J. of Data Sci. and Analytics, vol. 13, no. 2, pp. 139-149, Mar. 2022.

[18] S. Kulshrestha, V. Park, and T. Xuan, ”Study for the Prediction of E-Commerce Business Market Growth Using Machine Learning Algorithms,” in 2020 5th IEEE Int. Conference on Recent Advances and Innovations in Eng. (ICRAIE), pp. 1-7, IEEE, Nov. 2020.

[19] H. K. Hameed, ”Deep Learning Algorithm for Predicting Various Retail Store Sales Using LSTM and ARIMA,” Dijlah J., vol. 6, no. 4, pp. 123-139, Dec. 2023.

[20] G. Ban and Y. Nguyen, ”Personalized Dynamic Pricing with Machine Learning: High-Dimensional Features and Heterogeneous Elasticity,” Manag. Sci., vol. 67, no. 9, pp. 5549-5568, Sept. 2021

[21] S. Bajpai, K. Sharma, and B. Kumar Chaurasia, “Detection of Anomalous Bitcoin Transactions in Blockchain Using ML”, *EAI Endorsed Trans IoT*, vol. 10, Aug. 2024.