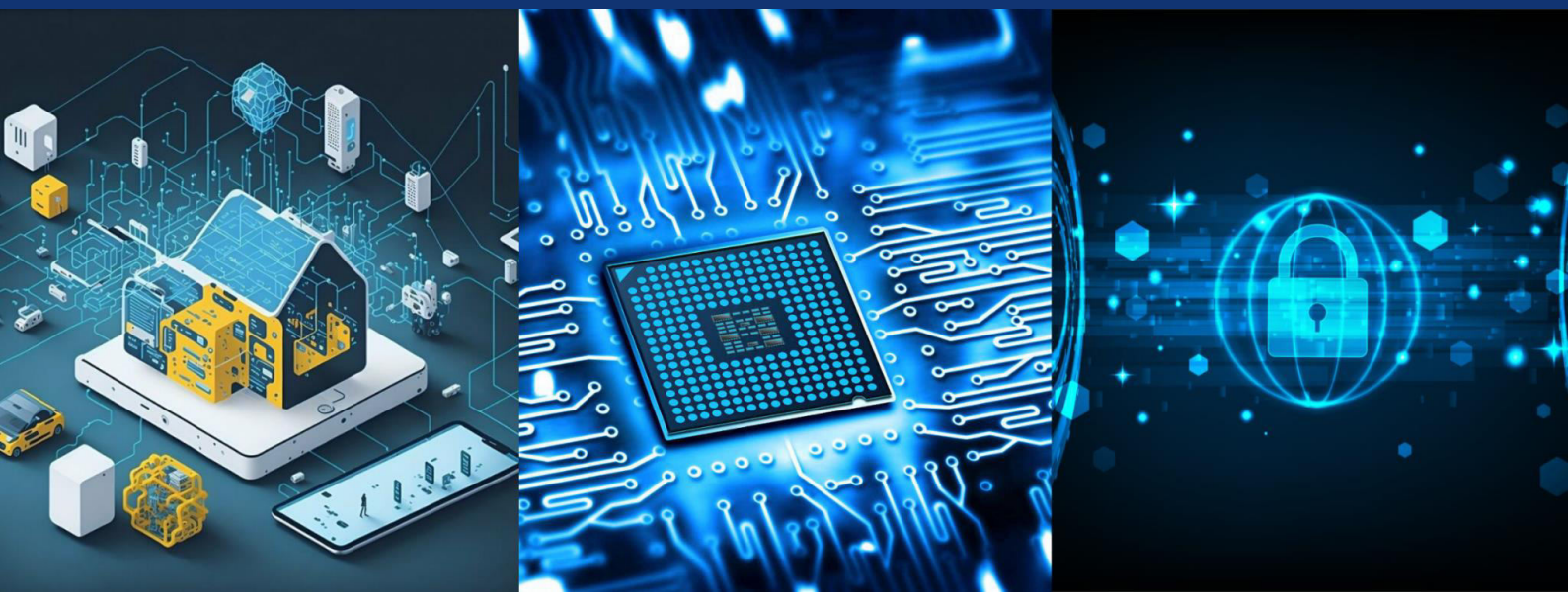
IJIRCCE

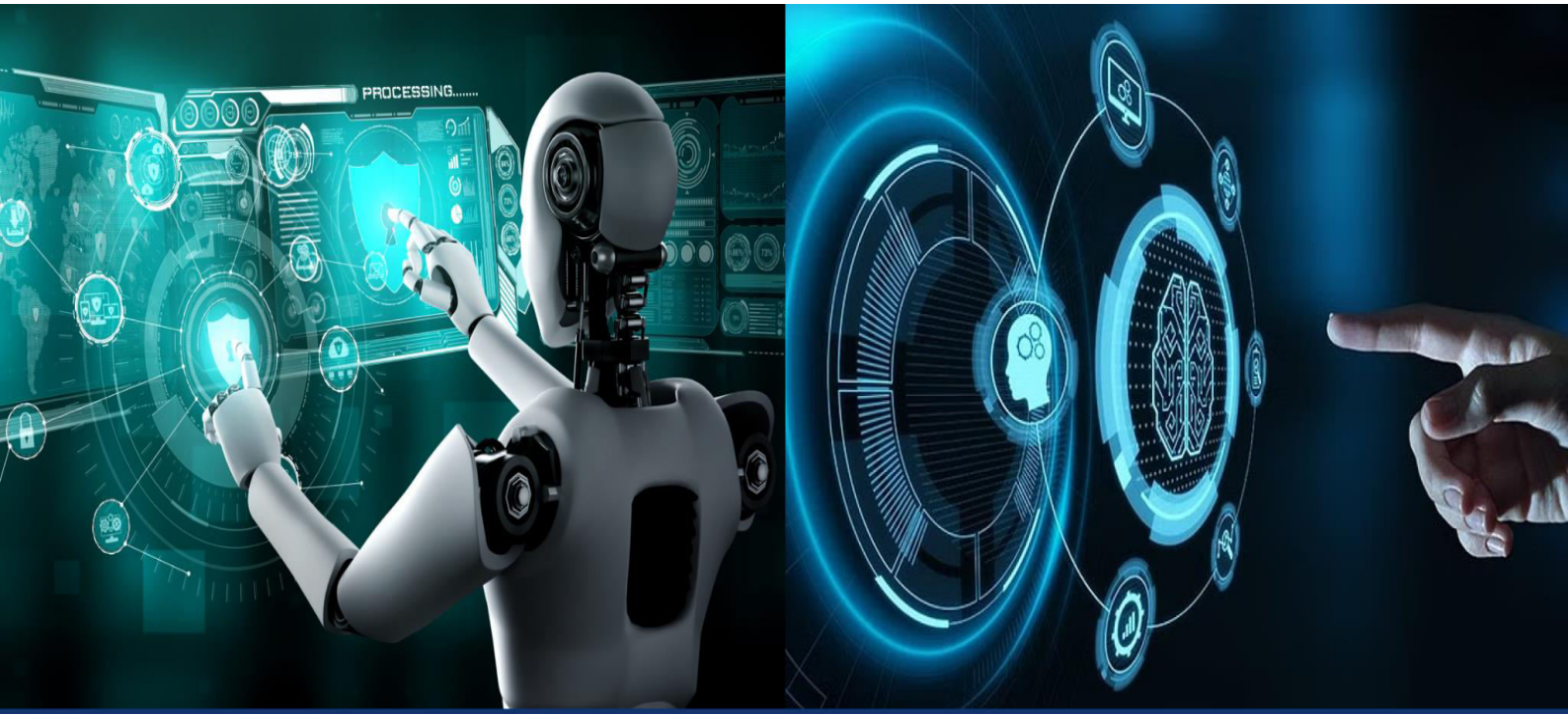
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

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



Impact Factor: 8.771

Volume 14, Issue 3, March 2026

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026 DOI: 10.15680/IJIRCCE.2026.1403109

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Grocery Demand Insight and Stock Planning System

Sauparnesh G[1], Vignesh D[2], Anitha M[3]

UG Student, Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India[1]

UG Student, Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India[2]

Assistant Professor, Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College Coimbatore, India[3]

ABSTRACT: This paper presents an analytical study of grocery demand patterns and inventory optimization in a retail transaction environment. A Random Forest regression model is employed to forecast product demand and support stock planning decisions. Feature engineering techniques are used to derive indicators such as seasonal trends, discount impact, sales velocity, and stock risk levels. Experimental results demonstrate that machine learning–based demand prediction and stock classification effectively identify understock and overstock situations, providing actionable insights for data-driven inventory management and operational efficiency in grocery retail systems.

KEYWORDS: Grocery Demand Forecasting, Stock Planning, Random Forest, Machine Learning, Python.

I. INTRODUCTION

In the competitive grocery retail sector, understanding the relationship between product demand patterns and inventory management is critical for sustainable operational efficiency and revenue growth. Retailers often experience fluctuations in demand influenced by seasonal trends, purchasing behaviour, and promotional strategies, which directly impact stock availability and profitability. This study aims to analyse and forecast grocery demand by distinguishing high-demand products from low-demand items using data analytics and machine learning techniques to support effective stock planning and inventory optimization.

II. DATASET DESCRIPTION

The dataset used in this study consists of structured grocery retail transaction records containing product details, quantity sold, pricing information, discount values, stock availability, payment methods, and transaction dates. The data was pre-processed to handle missing values, correct inconsistencies, and standardize formats to ensure accurate feature engineering and reliable model training for demand forecasting and stock classification.

III. METHODOLOGY

Random Forest Regression was employed to model and forecast grocery demand based on historical transaction data. The model was selected due to its robustness in handling nonlinear relationships among input variables and its ability to provide reliable predictive performance compared to traditional regression techniques.

3.1 Demand Variability

Demand Variability represents the fluctuation in product sales quantity across different time periods. Higher variability indicates inconsistent purchasing patterns, often associated with seasonal trends or promotional impacts. This metric helps identify products with unstable demand that require dynamic stock planning strategies.

3.2 Stock Utilization Ratio

Stock Utilization Ratio measures the proportion of available inventory sold within a given period. A higher ratio reflects efficient stock movement and strong product demand, while a lower ratio indicates slow-moving inventory that may lead to overstock conditions.

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026 DOI: 10.15680/IJIRCCE.2026.1403109

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

3.3 Sales Fluctuation Score

Sales Fluctuation Score serves as a proxy for variability in product demand over time. It captures responsiveness to seasonal trends and promotional pricing changes, helping distinguish stable-demand products from volatile-demand items.

IV. RESULTS AND DISCUSSION

The Random Forest Regression model effectively predicted grocery demand patterns based on historical transaction data. Demand variability emerged as a significant factor influencing stock imbalance, while stock utilization ratio showed strong correlation with product performance. The model demonstrated high predictive capability, supporting data-driven inventory optimization and risk reduction in retail operations.

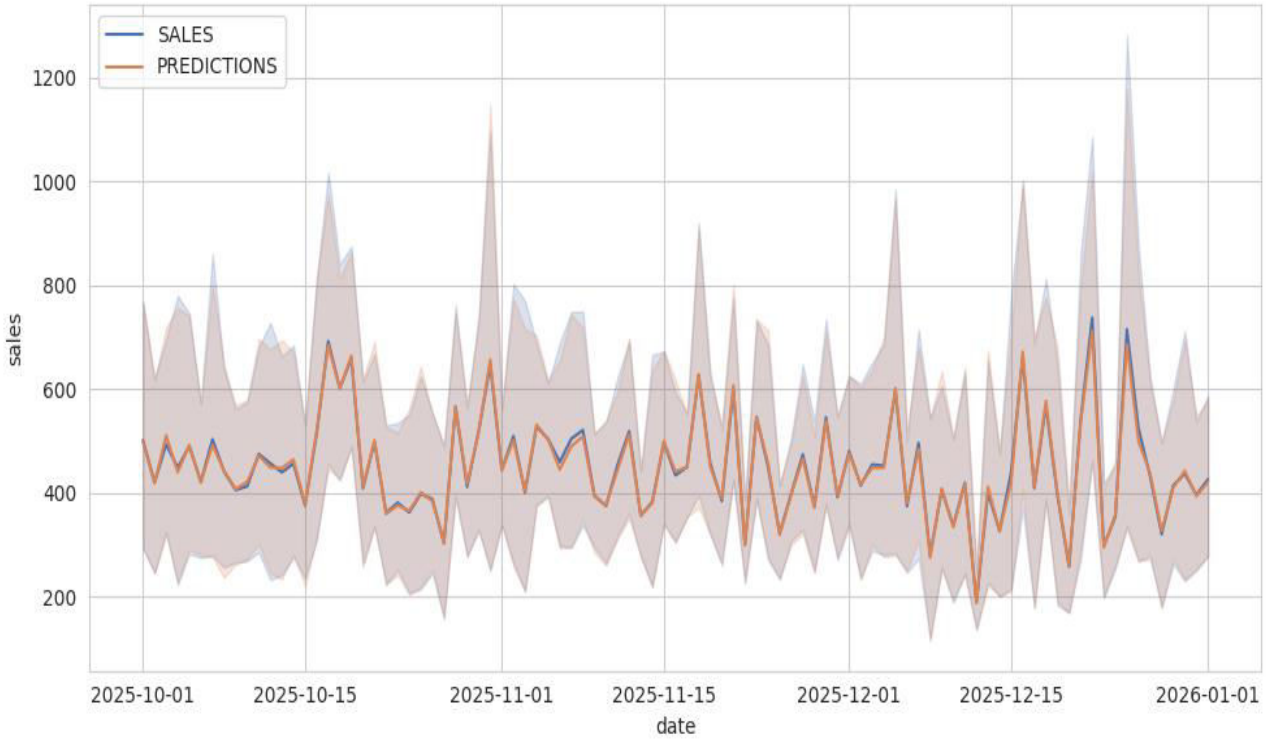


Fig. 1 Sales vs Predictions Over Time

Fig. 1 explanation

The figure illustrates the comparison between actual sales values and predicted sales generated by the Random Forest Regression model over the observed time period. The predicted curve closely follows the actual sales trend, indicating strong model accuracy in capturing demand patterns. Sales peaks and fluctuations are effectively modeled, with only minor deviations due to natural transactional variability. The high R^2 score of 0.9652 confirms excellent predictive performance and validates the suitability of the proposed framework for grocery demand forecasting and stock planning.

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026 DOI: 10.15680/IJIRCCE.2026.1403109

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

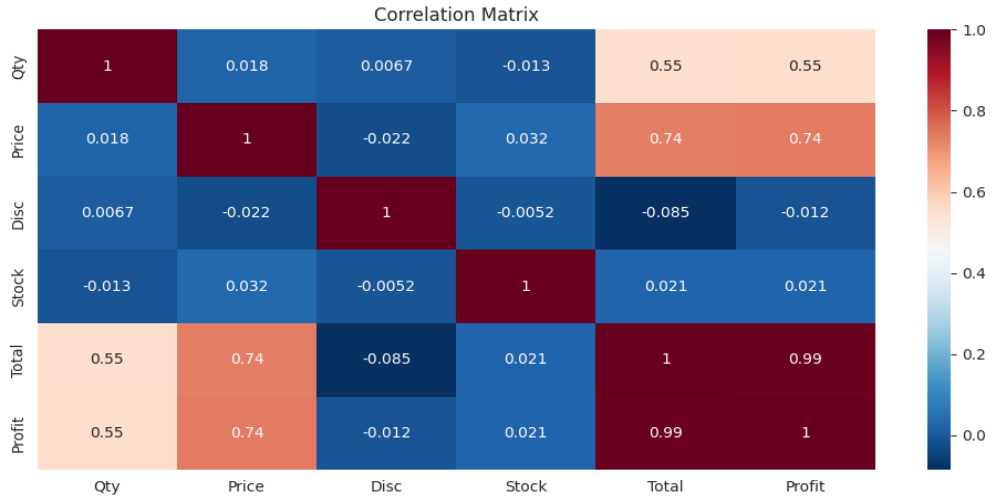


Fig. 2 Correlation Matrix

Fig. 2 explanation

The correlation matrix presents relationships among quantity, price, discount, stock, total sales, and profit. Total sales show strong positive correlation with price (0.74) and quantity (0.55), indicating that revenue is primarily influenced by these variables. Profit demonstrates near-perfect correlation with total sales (0.99), confirming revenue-driven profitability. Discount impact appears weakly negative, suggesting limited overall influence on total revenue generation.

STOCK RISK REPORT:

	Qty	Date	Stock	Avg_Daily_Demand	Weekly_Demand	Status	Weeks_of_Supply
Cashews	186	19	399	9.79	68.53	OVERSTOCK	5.82
Batter Cake	280	29	312	9.66	67.59	OVERSTOCK	4.62
Turmeric Powder	371	39	422	9.51	66.59	OVERSTOCK	6.34
Milk Powder	207	23	349	9.00	63.00	OVERSTOCK	5.54
Onion Powder	186	21	38	8.86	62.00	**UNDERSTOCK**	0.61
Soy Sauce	292	33	645	8.85	61.94	OVERSTOCK	10.41
Frozen Samosa	244	29	280	8.41	58.90	OVERSTOCK	4.75
Ash Gourd	208	25	52	8.32	58.24	**UNDERSTOCK**	0.89
Walnuts	190	23	480	8.26	57.83	OVERSTOCK	8.30
Tulsi Tea	139	17	390	8.18	57.24	OVERSTOCK	6.81
Puffed Rice	327	40	376	8.18	57.23	OVERSTOCK	6.57
Custard Apple	204	25	29	8.16	57.12	**UNDERSTOCK**	0.51
Cornflakes	212	26	239	8.15	57.08	OVERSTOCK	4.19
Green Gram	267	33	279	8.09	56.64	OVERSTOCK	4.93
Spinach	209	26	9	8.04	56.27	**UNDERSTOCK**	0.16

Fig. 3 Stock Risk Classification Report

Fig. 3 explanation

The stock risk report identifies products classified as overstock and understock based on weekly demand and available inventory. Products such as Cashews, Milk Powder, and Soy Sauce are categorized as overstocked due to excess inventory relative to demand. Conversely, Onion Powder, Spinach, and Custard Apple are identified as understocked items requiring immediate replenishment. This automated risk assessment supports proactive inventory control and minimizes potential revenue loss.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Performance Metric	Random Forest
R^2 Score	0.9652
RMSE	32.2537
MAE	11.0209
Accuracy	94.1307

Table. 1 Model Performance

Table 1 summary:
The performance metrics demonstrate strong generalization ability of the Random Forest model. The R^2 score of 0.9652 indicates that 96.52% of the variance in sales data is explained by the model. Low RMSE and MAE values confirm minimal prediction error. The overall accuracy of 94.131% reflects stable and reliable demand forecasting performance suitable for real-world retail applications.

V. CONCLUSION

This study confirms that machine learning-based demand forecasting significantly enhances grocery inventory management and operational decision-making. The Random Forest Regression model achieved high predictive accuracy and successfully identified stock imbalance conditions. The integration of correlation analysis, trend evaluation, and automated stock risk classification provides a comprehensive decision-support framework for modern grocery retail systems. The proposed framework can be extended to real-time monitoring and intelligent replenishment strategies to further optimize supply chain performance.

REFERENCES

[1] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[2] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[3] T. McKinney, "Data structures for statistical computing in Python," Proceedings of the 9th Python in Science Conference, pp. 51–56, 2010.
[4] J. D. Hunter, "Matplotlib: A 2D graphics environment," Computing in Science & Engineering, vol. 9, no. 3, pp. 90–95, 2007.
[5] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed., Morgan Kaufmann, USA, 2012.
[6] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, New York, 2009.
[7] S. Chopra and P. Meindl, Supply Chain Management: Strategy, Planning, and Operation, 6th ed., Pearson, 2016.
[8] C. Chatfield, The Analysis of Time Series: An Introduction, 6th ed., CRC Press, 2004.
[9] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," Proceedings of the International Joint Conference on Artificial Intelligence, pp. 1137–1143, 1995.
[10] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Media, 2019.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details