International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 OPEN ACCESS
Available Online at : www.ijsrcseit.com
doi : https://doi.org/10.32628/CSEIT25112857

Enhancing Supply Chain Resilience through Machine Learning-Based Predictive Analytics for Demand Forecasting

Vikas Prajapati
Independent Researcher

ARTICLEINFO

Article History:

Accepted : 05 May 2025
Published: 09 May 2025

Publication Issue
Volume 11, Issue 3
May-June-2025

Page Number
345-354

ABSTRACT

In the era of global supply chain complexities, ensuring resilience in supply chain operations is critical for minimizing disruptions and maintaining efficiency. Traditional demand forecasting methods are unable to adapt to changing market conditions, which leads to inefficient inventory control and resource allocation. In the retail industry, precise demand forecasting is crucial for maximizing inventory control, reducing stockouts, and enhancing financial decision-making. This study employs Extreme Gradient Boosting (XGBoost) to enhance sales prediction accuracy using Walmart sales data. The dataset is partitioned into training and testing sets in an 80:20 ratio, and the XGBoost model is fine-tuned to achieve optimal performance. Experimental results indicate superior predictive accuracy, with an R^2 of 95.51%, a minimal MAE of 0.0024, and an MSE of 4.79. Visualization techniques, including density curves and correlation heatmaps, provide deeper insights into feature relationships and data distribution. The findings demonstrate the robustness of XGBoost for demand forecasting, offering a data-driven approach for retailers to enhance operational efficiency and strategic planning.

Keywords—Supply Chain Resilience (SCRes), Demand Forecasting, Risk Mitigation, Predictive Analytics, Machine Learning (ML).

Introduction

Due to rapidly developing world markets and increasing supply chain intricacy businesses now recognize supply chain resilience (SCRes) as their main competitive advantage factor for staying adaptable[1][2]. Organizational supply chain resilience entails its ability to survive unexpected events along with its recovery process and operational reintegration and maintenance of business continuity service quality and financial stability[3]. When supply chains go for more tiers, geographical locations, and industries the company needs reliable and robust supply system for flexibility more of more conditions[4][5].

Agile supply chains depend on demand forecasts to enable an organization to propose future product requirements and make the best arrangements for inventory, resources, and production[6]. Despite the extensive use of Traditional methods of demand forecasting, the traditional techniques tend to provide very low accuracy when it comes to predicting future demand in volatile environments[7]. This area has a lot of potential of being powered by both PA which is infused with ML and AI. In this regard, compared to more conventional approaches, ML algorithms have much more of a shot at detecting various complicated patterns and trends by analyzing a vast amount of historical and real-time big data.

The ML-based predictive analytics improves the accuracy of demand forecasts; organizations can effectively manage inventory and avoid frequent stockouts or overstocking and make timely supply chain operations[8]. Therefore, businesses are in a much better position to manage future unpredictability, possible threats, improve the resiliency of the supply chain[9][10]. This paper aims to discuss the capacity of machine learning to enhance demand forecasting for supply chains and the ways in which businesses can use them to enhance supply chain resilience and improve their efficiency in volatile market conditions.

Motivation and Contribution of Study

The development of this research is as a result of the shifting nature of supply chain management and the need to have better demand forecasting. Traditional approaches to forecasting are usually incapable of reacting to changes in the markets for inputs or outputs, supply chain disruptions or any shifts in consumption patterns among consumers. This paper aims to identify the advantages of using ML-based Predictive Analytics for enhancing forecasting tools and supply chain performance. The key contribution of study is as:

- Leverage Walmart sales data spanning three years, covering 45 stores and 99 departments, to identify key factors influencing demand.
- Implement pre-processing for handling missing values, normalization, and categorical encoding to improve model performance.
- Apply feature engineering techniques such as time-based features, lag variables, and moving averages to capture patterns in sales data.
- Utilize machine learning techniques, specifically XGBoost, to predict Walmart sales with high precision.
- Evaluate the model using performance metrics like R^2, MAE, and MSE to ensure reliable demand predictions.

Justification and Novelty

This study addresses the shortcomings of conventional approaches in identifying intricate sales trends, hence defending the use of ML for demand forecasting. By integrating advanced data preprocessing, feature engineering techniques such as time-based features, lag variables, and moving averages, and leveraging XGBoost's high scalability and efficiency, the proposed model enhances prediction accuracy. The novelty lies in its comprehensive data processing pipeline, advanced visualization techniques, and hyperparameter tuning, resulting in superior performance. The study's results may help stores optimize their inventory management, cut down on stockouts, and make better financial decisions using data-driven forecasts.

Structure of paper

This paper is organized in the following way: Section II reviews the literature on demand forecasting and supply chain resilience. Methodology, which includes data collection, preprocessing, and model selection, is discussed in Section III. Section IV presents the experimental results and performance evaluation of various ML models. Section V presents the findings and suggests areas for further research.

Literature Review

Studies have previously been performed to forecast demand for retail industry corporations based on the availability of relevant historical data. various Previous research are:

Zhang et al. (2024) proposed HA-LSTM network combines LSTM layers with multi-head self-attention modules to understand time series data for both local and long-term trends. The results show that compared to state-of-the-art baselines like ARIMA, Prophet, and vanilla LSTM models, the HA-LSTM outperforms them by 15% in MAPE and 12% in RMSE[11].

Iwakin and Moazeni (2024) created and tested a hybrid demand forecast model based on conformal predictions; it uses a CNN and Bi-LSTM to anticipate water demand probabilistically on an hourly basis. The suggested model's efficacy in probabilistic water demand forecasting in practical contexts is confirmed. The results show that deterministic predictions had a significant improvement of 10% while probabilistic predictions had an improvement of 26.7%[12].

Chung, Lee and Yang (2023) combination of K-means, ElasticNet, and GPR to form a hybrid model. This research makes use of GPR, K-means, and ElasticNet to put these methods into practice. The performance of the hybrid model proved the usefulness of using this model as a base for constructing the model for demand forecast under uncertainty with the best MAE of 5.57%[13].

Panda and Mohanty (2023) employed a variety of regression models, including RF, GBR, LGBM, XGBoost, CatBoost Regressor, LSTM, and Bidirectional LSTM. The results show how deep learning models may be used for predicting and emphasize how much better LSTM is than other methods. The corresponding values for the MAE, MAPE, RMSLE, and RMSE are 0.28, 18.83, 6.56%, and 14.18[14].

Miguéis et al. (2022) forecasted the demand for perishable seafood in a representative store of a prominent European retailer. The LSTM networks model had the best overall outcomes in terms of RMSE (27.82), MAE (20.63), and mean positive error (17.86) among machine learning models, which surpassed the performance of baseline and statistical models in terms of prediction accuracy. In conclusion, these models have the potential to increase the lifetime of fresh fish species and enhance customer satisfaction[15].

Shokouhifar and Ranjbarimesan (2022) the proposed model for blood donation/demand forecasting was successfully reproduced using data collected from the Tehran Blood Centre in Tehran, Iran, between February 24, 2020, and October 14, 2021. The findings demonstrate the effectiveness of the suggested model, with a precision of 6.1% for donations and a precision of 6.5% for requests, when comparing the actual and predicted values, respectively. Applying the suggested approach to blood platelet inventory management yielded findings that showed their methodology could withstand comparisons to current uncertainty handling methods, reducing shortage rates by 32.1% and waste rates by 26.6%[16].

Thomas and Vedi (2021) examines the effectiveness of ML models—Multi-Regression, LSTM, and ANN—in enhancing demand forecasting and inventory management using data from a well-known Chinese home appliance shipping business. Reduced supply chain disruptions and improved operational efficiency were two outcomes of LSTM's superior prediction accuracy compared to the other models studied. RMSE values averaged 231.98 across all Local Transshipment Centers (LTCs)[3].

Table I presents a structured comparison of the studies while summarizing key findings and possible future improvements.

TABLE I. SUMMARY OF SUPPLY CHAIN RESILIENCE FOR DEMAND FORECASTING IN DIFFERENT DOMAIN

Reference	Model/Technique	Dataset	Results	Limitation/Future Direction
Zhang et al. (2024)	HA-LSTM (Hybrid Attention-based LSTM)	Predict Future Sales dataset	HA-LSTM outperforms ARIMA, Prophet, and vanilla LSTM, achieving 15% improvement in MAPE and 12% reduction in RMSE	Future work could focus on incorporating external factors like promotions and economic indicators for improved forecasting accuracy
Iwakin and Moazeni (2024)	CNN + BiLSTM with conformal prediction	Hourly water demand dataset	Hybrid model improves deterministic predictions by 10% and probabilistic predictions by 26.7%	Future studies can explore real-time adaptation to demand fluctuations and integration with IoT sensor data
Chung, Lee, and Yang (2023)	K-means + ElasticNet + GPR	U.S. manufacturing company dataset	Achieved best prediction accuracy with MAE: 5.57 by clustering similar data before GPR training	Future research can explore adaptive clustering techniques to further refine accuracy
Panda and Mohanty (2023)	RF, GBR, LightGBM, XGBoost, CatBoost, LSTM, BiLSTM	Food Demand Forecasting (Genpact) dataset	LSTM performed best, with RMSLE: 0.28, RMSE: 18.83, MAPE: 6.56%, MAE: 14.18	Future work could include hybrid models combining multiple regressors for enhanced forecasting performance
Miguéis et al. (2022)	LSTM, FFNN, SVR, RF, Holt-Winters	Fresh fish demand dataset (European retail store)	LSTM provided best results with RMSE: 27.82, MAE: 20.63, Mean Positive Error: 17.86	Future studies may incorporate external environmental factors like weather conditions and holiday effects
Shokouhifar and	Blood	Tehran Blood	Model reduced	Future directions

Reference	Model/Technique	Dataset	Results	Limitation/Future Direction
Ranjbarimesan (2022)	donation/demand forecasting model	Center dataset	shortage and wastage rates by 32.1% and 26.6%, respectively	include extending the model for multi-regional blood supply chains
Thomas and Vedi (2021)	Multi-Regression, LSTM, ANN	Chinese home appliance shipping firm dataset	LSTM had the best predictive accuracy with RMSE averaging 231.98 across LTCs	Future studies could incorporate real-time supply chain disruptions and adaptive learning techniques

Methodology

The methodology for this study involves several key stages shown in Figure 1. The research started by involving data collection from Walmart sales data containing records of 45 stores and 99 departments for three years. Data preprocessing was performed, including handling missing values, label encoding categorical variables, and applying Min-Max normalization to ensure feature uniformity. Feature engineering, including time-based features, lag variables, and moving averages, was applied to enhance model performance. The dataset was split 80:20 for training and testing, and XGBoost was used for demand forecasting due to its scalability and loss minimization. The model's complexity was controlled using regularization parameters, including shrinkage and tree depth limitations. Performance evaluation was conducted using R^2, MAE, and MSE metrics.

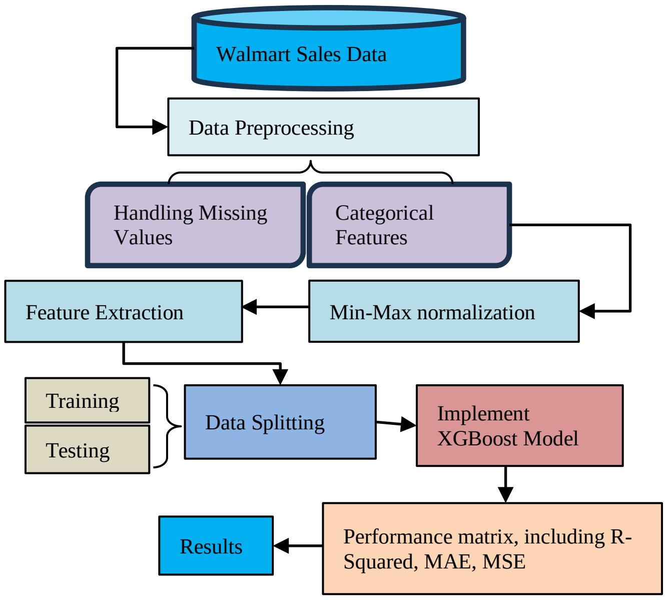

Figure 1. Flowchart for Demand Forecasting in Supply Chain Management

The overall steps of the flowchart for Demand Forecasting in Supply Chain Management are provided below:

Data Collection

The study's dataset consists of Walmart sales data from a US-based retailer. Over a three-year period, it contains sales data for 45 locations and 99 categories. Store number, store size, department, date, regional temperature, fuel costs, consumer price index (CPI), unemployment rate, and holiday indicators are just a few of the variables that have an impact on sales in this dataset. Data analysis and visualization are essential for understanding the insight of data. The visual aspects of data are provided in below:

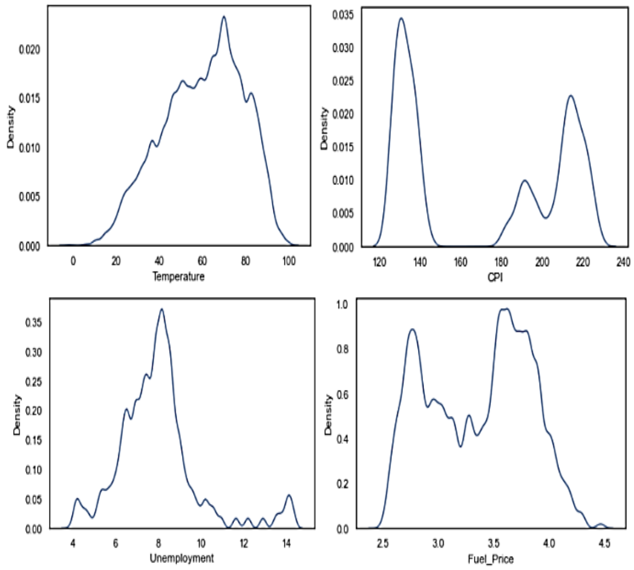

Figure 2. Density Curves for Variables

Figure 2 displays density curves for Temperature, CPI, Unemployment, and Fuel Price. Temperature and Unemployment show unimodal distributions with peaks around 70 and 9, respectively. CPI exhibits multiple peaks, indicating variability, while Fuel Price follows a bimodal pattern. These distributions highlight the variability and clustering in the data.

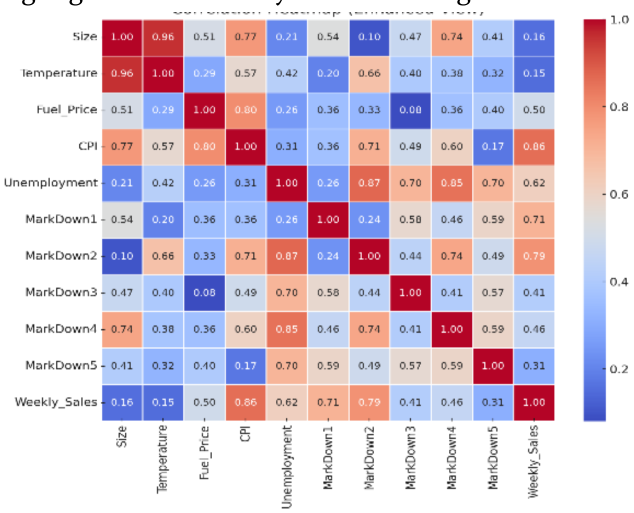

Figure 3. Heat Map of Correlation between Numerical Features

Figure 3 visually represents the relationships between various features in the dataset using a "coolwarm" color scheme, where red indicates strong positive correlations and blue represents strong negative correlations. The diagonal values are all set to 1, signifying perfect correlation with themselves. Annotated numerical values provide precise correlation coefficients for easier interpretation. The heatmap helps identify key relationships, such as how features like CPI and Weekly Sales exhibit a strong positive correlation, while others show weaker or negative associations, aiding in feature selection and predictive analysis.

Data Preprocessing

Data pre-processing is the most important and influential for the generalization performance of a Machine Learning Algorithm. Data Preprocessing proceeds with Handling Missing Values, Min-Max normalization, and Categorical features. In this step, the data were processed in the following ways.

- **Handling Missing Values:** Data has been checked for inaccuracies, missing or out-of-range values. Columns with missing values have been dropped.
- **Categorical features:** Special characters and categorical data are converted to numerical values in order to improve model performance. Using label encoding techniques, the categorical data types are transformed into numerical data types. Label encoding gives each category a distinct number label, transforming categorical data into numerical data.

Min-Max Normalization

Normalization is a "scaling down" transformation of features to ensure they contribute equally to a model[17]. Min-Max Normalization rescales values within a fixed range, typically [0,1] or [-1,1], using the Equation 1:

$$x' = \frac{(x - x_{min})}{(x_{max} - x_{min})} \qquad (1)$$

where X' is the normalized value, and x_{min}, x_{max} are the feature's minimum and maximum values.

Feature Extraction

creating time-based features (e.g., extracting month, week, and holiday indicators from date data), and normalizing numerical attributes like temperature, fuel price, CPI, and unemployment to ensure uniformity in scale. Additionally, lag features were introduced to capture historical trends, and moving averages were computed to smooth out fluctuations in weekly sales. These engineered features helped improve model generalization and capture important patterns in the data.

Data Splitting

The dataset was split into training and test sets using an 80:20 ratio. The training set was used to develop the model, while the test set was reserved for evaluation.

Forecasting with a XGBoost Model

XGBoost is a DT ensemble that uses gradient boosting for great scalability. XGBoost is similar to gradient boosting in that it minimizes a loss function to increase the objective function additively[18]. Equations (2) and (3) demonstrate how XGBoost controls the tree complexity using a variation of the loss function, as DTs are the only base classifiers used by the program.

$$L_{xgb} = \sum_{i=1}^{N} L\left(y_j, F(x_i)\right) + \sum_{m=1}^{M} \Omega(h_m) \quad (2)$$

$$\Omega(h) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2 \quad (3)$$

where T is the total number of tree leaves and ω is the score for each leaf's production. To implement a pruning strategy, this loss function may be added to the split criteria of decision trees. Trees whose γ values are greater are less complex. The amount of loss reduction gain needed to separate an internal node is determined by γ[19]. An extra regularization parameter in XGBoost called shrinkage is used to decrease the size of the additive expansion step. Finally, there are various strategies that may be used, such tree depth, to control the level of tree complexity. By simplifying the trees, it can train the models more quickly and use less storage space.

Performance Matrix

Metrics for performance are numerical assessments of how well a model predicts or categorizes data. Using R-Squared, MAE, and MSE, trials were assessed within the framework of machine learning-based demand forecasting. Below you can find a list of regularly used performance metrics:

1) R-Square

R2 is a measure representing the proportion of variation in the outcome that can be attributed to the predictor variables. It shown as Equation (4).

$$R^2 = 1 - \frac{SS_{res}}{SS_{total}} = \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \mu)} \quad (4)$$

The residuals' sum of squares is denoted by SS_{res}, a total of all squares is represented by SS_{total}, y_i is the truevalue, $\hat{y}_i$is the forecasted value, and μ is the mean.

2) Mean Absolute Error (MAE)

MAE quantifies prediction accuracy, similar to RMSE[20]. The average absolute discrepancy between the anticipated and actual values is calculated. It is present as Equation (5).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

y_i represents the actual value, $\hat{y}_i$ stands for the anticipated value, and n signifies the quantity of observations.

3) Mean Squared Error (MSE)

Additionally, it is an average assessment of the prediction errors in a collection. Equation (6) illustrates how the squares of each mistake are put together and then averaged.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

The true value is denoted by y_i n is the number of observations, and $\hat{y}_i$ is the forecast value.

Results And Discussion

A Python software, version 3.9, is executed on a Windows 10 operating system in order to carry out the experiment in accordance with the criteria for demand prediction. The machine in question has 8 GB of RAM. The proposed model experiment results are presented in Table II, and the visual forecasting graph.

TABLE II. EVALUATION OF PROPOSED MODEL FOR DEMAND FORECASTING AMONG PARAMETERS

Performance Metric	Extreme Gradient Boosting (XGBoost)
R2	95.51
MAE	0.0024
MSE	4.79

Table II presents the XGBoost model performance. The model demonstrates exceptional predictive performance, achieving a high R^2 value of 95.51, indicating a strong correlation among forecasted and actual values. Additionally, the model exhibits minimal error, with an MAE of 0.0024, signifying highly precise predictions. Furthermore, the MSE of 4.79 reflects the model's efficiency in minimizing overall prediction errors, making XGBoost a highly effective approach for demand forecasting.

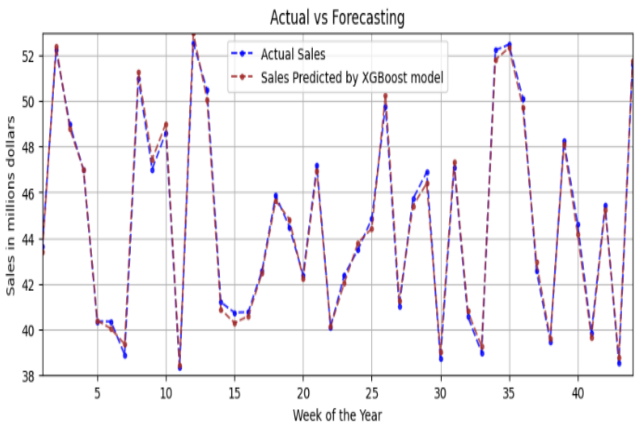

Figure 4. Forecasting with XGBoost Model

Actual vs. Forecasting compares actual sales with XGBoost-predicted sales over 52 weeks, presented in Figure 4. The blue (actual) and red (predicted) dotted lines would be fairly close and the model would be relatively good at predicting the right curve for sales. The sales vary between 38 and 52 million dollars; the fluctuations are identified by the model with low margins of error.

Comparative Analysis

This section discusses various ML models that have been used in the analysis of data for the purpose of demand forecasting in supply chain management. XGBoost provides better results for forecast tasks when compared to both ANN and KNN methods as demonstrated in Table III. XGBoost achieves a high R^2 value of 95.51, indicating that it explains most of the variance in the target variable, while also recording exceptionally low error values with an MAE of 0.0024 and an MSE of 4.79. The performance indicators of ANN demonstrate an R^2 value of 39.58 while showing elevated error metrics of 0.0129 MAE and 6.45 MSE. KNN demonstrates a lower R^2 value of 59.4 but reveals much higher error metrics with 81.99 MAE and 21.32 MSE. All these results collectively show that XGBoost performs better than both ANN and KNN in terms of prediction accuracy and error estimation.

TABLE III. COMPARATIVE ANALYSIS FOR DEMAND FORECASTING ON SALES DATA

Metric	XGBoost	ANN[21]	KNN[21]
R2	95.51	39.58	59.4
MAE	0.0024	0.0129	81.99
MSE	4.79	6.45	21.32

The proposed XGBoost-based demand forecasting model has certain strengths such as high accuracy, flexibility, and ease of managing complicated patterns. Some of the feature engineering strategies used include time-based attributes, lag variables, and moving averages that aid the model in the identification of improved trends. Thus, loss reduction through gradient boosting enhances its capacity to make accurate forecasts. The performance of XGBoost is higher compared to other models which can explain why it is ideal for retail demand prediction that is dynamic and based on the available data.

Conclusion And Future Scope

Demand forecast is challenging in supply chain management because markets are volatile, inventory is limited, and there are many external economic forces. These variations often go unnoticed in most traditional forecasting models hence inhibiting efficiency in the decision-making process. This paper provides a methodology for effective demand forecasting through the application of the XGBoost model, which has a high predictability with R^2 at 95.51%, an MAE, 0f 0.0024, MSE of 4.79. The results yielded showed that the model is efficient in achieving high accuracy in giving out diversified features involved in sale calculation via feature engineering and data pre-processing. The research would achieve better predictive results by including crucial outside elements like consumer sentiment together with macroeconomic trends and promotional effects in addition to its current use of historical sales data. However, the model does not take into consideration the occurrence of demand fluctuations caused by unexpected events. Future work can work on the development of other types of machine learning, deep learning methods together with real-time data feeding approaches for adaptability, other new external variables for model stability and performance.

References

[1]. F. Alfarsi, F. Lemke, and Y. Yang, “The importance of supply chain resilience: An empirical investigation,” Procedia Manuf., vol. 39, pp. 1525–1529, 2019, doi: 10.1016/j.promfg.2020.01.295.

[2]. S. Chatterjee, “Mitigating Supply Chain Malware Risks in Operational Technology : Challenges and Solutions for the Oil and Gas Industry,” J. Adv. Dev. Res., vol. 12, no. 2, pp. 1–12, 2021.

[3]. J. Thomas and V. Vedi, “Enhancing Supply Chain Resilience Through Cloud-Based SCM and Advanced Machine Learning: A Case Study of Logistics,” J. Emerg. Technol. Innov. Res., vol. 8, no. 9, 2021.

[4]. U. Shankar, “Predictive Analytics in Supply Chain Management : The Role of AI and Machine Learning in Demand Forecasting,” vol. 4, no. 3, pp. 2976–2985, 2024.

[5]. K. Murugandi and R. Seetharaman, “Analysing the Role of Inventory and Warehouse Management in Supply Chain Agility : Insights from Retail and Manufacturing Industries,” Int. J. Curr. Eng. Technol., vol. 12, no. 6, pp. 583–590, 2022.

[6]. V. singh Chandraul and S. kumar Barode, “A Review on Demand and Forecasting in Supply Chain Management,” IJOSTHE, vol. 5, no. 5, Oct. 2018, doi: 10.24113/ojssports.v5i5.76.

[7]. T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, “Predictive Analytics for Demand Forecasting – A Comparison of SARIMA and LSTM in Retail SCM,” Procedia Comput. Sci., vol. 200, no. 2019, pp. 993–1003, 2022, doi: 10.1016/j.procs.2022.01.298.

[8]. S. Punia and S. Shankar, “Predictive analytics for demand forecasting: A deep learning-based decision support system,” Knowledge-Based Syst., vol. 258, p. 109956, 2022, doi: https://doi.org/10.1016/j.knosys.2022.109956.

[9]. M. A. Khan et al., “Effective Demand Forecasting Model Using Business Intelligence Empowered with Machine Learning,” IEEE Access, 2020, doi: 10.1109/ACCESS.2020.3003790.

[10]. S. Pahune and N. Rewatkar, “Cognitive Automation in the Supply Chain: Unleashing the Power of RPA vs. GEN AI,” no. April, 2024, doi: 10.13140/RG.2.2.22528.85761.

[11]. X. Zhang, P. Li, X. Han, Y. Yang, and Y. Cui, “Enhancing Time Series Product Demand Forecasting With Hybrid Attention-Based Deep Learning Models,” IEEE Access, vol. 12, pp. 190079–190091, 2024, doi: 10.1109/ACCESS.2024.3516697.

[12]. O. Iwakin and F. Moazeni, “Improving urban water demand forecast using conformal prediction-based hybrid machine learning

models," J. Water Process Eng., 2024, doi: 10.1016/j.jwpe.2023.104721.

[13]. D. Chung, C. G. Lee, and S. Yang, "A Hybrid Machine Learning Model for Demand Forecasting: Combination of K-Means, Elastic-Net, and Gaussian Process Regression," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 6s, pp. 325–336, 2023.

[14]. S. K. Panda and S. N. Mohanty, "Time Series Forecasting and Modeling of Food Demand Supply Chain Based on Regressors Analysis," IEEE Access, vol. 11, pp. 42679–42700, 2023, doi: 10.1109/ACCESS.2023.3266275.

[15]. V. L. Miguéis, A. Pereira, J. Pereira, and G. Figueira, "Reducing fresh fish waste while ensuring availability: Demand forecast using censored data and machine learning," J. Clean. Prod., 2022, doi: 10.1016/j.jclepro.2022.131852.

[16]. M. Shokouhifar and M. Ranjbarimesan, "Multivariate time-series blood donation/demand forecasting for resilient supply chain management during COVID-19 pandemic," Clean. Logist. Supply Chain, 2022, doi: 10.1016/j.clscn.2022.100078.

[17]. B. Boddu, "Scaling Data Processing with Amazon Redshift Dba Best Practices for Heavy Loads," Int. J. Core Eng. Manag., vol. 7, no. 7, 2023.

[18]. S. Nokhwal, S. Nokhwal, S. Pahune, and A. Chaudhary, "Quantum Generative Adversarial Networks: Bridging Classical and Quantum Realms," in 2024 8th International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence (ISMSI), New York, NY, USA, NY, USA: ACM, Apr. 2024, pp. 105–109. doi: 10.1145/3665065.3665082.

[19]. T. R. Mahesh, V. Vinoth Kumar, V. Muthukumaran, H. K. Shashikala, B. Swapna, and S. Guluwadi, "Performance Analysis of XGBoost Ensemble Methods for Survivability with the Classification of Breast Cancer," J. Sensors, vol. 2022, pp. 1–8, Sep. 2022, doi: 10.1155/2022/4649510.

[20]. S. A. Perez-Rodriguez et al., "Metaheuristic Algorithms for Solar Radiation Prediction: A Systematic Analysis," IEEE Access, vol. 12, no. June, pp. 100134–100151, 2024, doi: 10.1109/ACCESS.2024.3429073.

[21]. 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.