Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

Integrating Data Analysis and Predictive Analytics into an Advanced Inventory Management System using Machine Learning and Deep Learning to Address Backorders

Bhushan Jadhav[1], Arti Deshpande[2], Sonali Jadhav[3]
[1]Department of Artificial Intelligence and Data Science, Thadomal Shahani engineering College, Mumbai University, India.
[2,3] [1]Department of Computer Engineering, Thadomal Shahani engineering College, Mumbai University, India.
Corresponding Author:
Dr. Bhushan Jadhav
Department of Artificial Intelligence and Data Science,Thadomal Shahani Engineering College, Mumbai University
Bandra West, Mumbai 400050
Email: bhushan.jadhav@thadomal.org

Article History:

Received: *07-11-2024*

Revised: *22-12-2024*

Accepted: *07-01-2025*

Abstract:

The evolution of ICT technology has benefited various businesses and applications in a variety of ways. The inventory management system is one of those that has used a variety of techniques to improve logistics and inventory management performance. The novel research introduces an advanced inventory management system that intends to improve operational efficiency in the retail industry by providing an intuitive interface for managing purchases, invoicing, client information, product catalogues, and sales numbers. The underlying inventory control systems automatically compute and apply reorder levels and safety stock quantities. Machine learning and deep learning are new technologies employed in current analytics due to their enormous benefits.

As a result, the purpose of this research paper is to highlight a combined approach of providing analytical business intelligence via data visualization platforms, as well as predictive analytics via machine learning and deep learning, in order to demonstrate a significant advancement toward the automation and optimization of inventory governance processes. The proposed approach demonstrates how combining traditional software development with cutting-edge data analysis techniques can result in a more efficient, scalable, and informative inventory management solution. It addresses common retail difficulties like stockouts and overstocking by utilising powerful machine learning (ML) and deep learning (DL) techniques. Finally, empirical testing and real-world applications are used to evaluate the suggested approach, which shows significant advances in inventory optimisation as well as other benefits such as enhanced demand forecasting, cost reductions, and operational efficiency. The findings show that machine learning and deep learning have the potential to transform inventory management practices.

Keywords: Data analysis, Predictive analytics, Inventory management system, Machine learning and Deep learning.

1. INTRODUCTION

Inventory management is an essential aspect of business operation in the current retail environment. The increasing dynamism of consumer behaviour and trends has hampered the ability of traditional inventory systems to meet client demands. As a result, many retailers must grapple with prevalent challenges that include but are not limited to stock outs, overstocking, and reduced profitability [3]. Conse-quently, there is a rising demand for inventory management systems that will op-timize their functionality through technology and data to improve performance. The present research seeks to remedy the situation by developing an innovative inventory management system for retailers to help address the identified challeng-es. The system will incorporate machine learning and deep learning technologies to enhance forecasting accuracy. Additionally, the system will optimize inventory levels [5].

Predictive analytics functionality, as well as user-friendly interfaces, will allow retailers to make informed decisions and respond to new market situations quickly. As a result, inventory management in the retail business will have a significant impact. As indicated by Karna Ghose, guard profitability, and increase consumer satisfaction, optimum inventory proper management will store operations, purchases and decrease costs retailers from stockouts and improve sales compete fairly on dynamic markets [4]. Additionally, good inventory control will allow retailers to make precise predictions about consumer desires and trends, conducting their operations more strategically and responding better to shifting needs and wants.

Therefore, in this paper we have proposed a solution for Advanced Inventory Management System using Machine Learning and Deep Learning to Address the Backorders and providing Data Analysis and Predictive Analytics.

2. RELATED WORK

The development and implementation of inventory management systems in the retail. Inventory management systems established in the retail industry have become increasingly complex and sophisticated to support the contingencies of supply chain operations. However, some limitations have retained over these systems. First, traditional inventory management systems are often reliant on manual work as well as simple models that may lead to low efficiency. Precisely, such systems may not promptly adjust the supply to the dynamic needs of the customer creating either stockout or overstocking situations, representing immediate risks to customer satisfaction and bottom line [2]. Another limitation is the absence of real-time data integration and analytics that precludes the supply chain members from taking informed, immediate actions in a highly dynamic market. As there is limited literature on advanced inventory management system using machine learning and deep learning is available. The existing contribution from different researchers is given as follows.

Liming Liu. Xue-Ming Yuan [13] rightly stated that the primary goal of supply chains is to meet consumer demands at the appropriate time, location, and quality. Reduced lead times and expenses, improved customer service levels, and increased product quality are the traits that determine a company's competitiveness in today's market. M. Jakšič. J.C. Fransoo [11], has developed an inventory control problem for a retailer operating under stochastic demand and limited supply, as well as numerical analysis to quantify the benefits of supply backordering and the value of the cancellation option, revealing various managerial insights. Ntakolia, C. Kokkotis C., et al. [12] proposed and compared various machine learning models for solving the binary classification problem of backorder prediction, followed by model calibration and post-hoc explainability using the SHAP model to discover and perceive the most significant characteristics that determine material backorder. Jean-Claude Munyaka Baraka. Sarma Venkata Yadavalli [14], has analysed inventory management strategies and implementations in the context of growing human demand and also considered the deterministic independent and dependent character of demand and its relative impact on inventory management in operations. Chuning Deng. Yongji Liu [15], have developed a deep inventory management (DIM) solution that employs deep learning's long short-term memory (LSTM) theory. The experimental findings demonstrate that DIM's average inventory demand prediction accuracy surpasses 80%, which may immediately identify anomalous inventory actions and cut inventory costs by around 25% when compared to other cutting-edge approaches. Santis, Rodrigo & Pestana de Aguiar, Eduardo & Goliatt, Leonardo. [9] have studied various machine learning classifiers inorder to suggest a predictive model in this imbalanced class problem, where the relative frequency of goods that go into backorder is unusual as compared to products that do not [16].

3. METHODOLOGY

The Proposed system architecture incorporates various software and tools to develop an advanced inventory management solution. The Logistic Regression (default parameters) has been used to predict if an item will be backordered. It is a probabilistic model that predicts the likelihood of a binary event and then assigns the instance to one of two categories (a surrogate of 0.5 is commonly used to classify the outcome). We picked logistic regression because it is simple, interpretable, and efficient for binary classifications. It includes interpretable variables that influence the likelihood of a backorder, making it perfect for use in inventory management. The models were simple logistic regression models built with Ten-sorFlow. We divided the original dataset into training and test data to evaluate the model's performance. Gradient descent was used to optimise the model's weights on the training dataset. During this procedure, the binary cross-entropy loss is op-timised, resulting in accurate probability forecasts for binary outcomes.

To increase the model's power, hyperparameters such as the learning rate and number of epochs were tuned. Cross-validation techniques were used to prevent overfitting while still allowing the model to be used on new data sets. To improve model prediction, a variety of feature engineering tactics were utilised, including changing categorical variables to numerical values and removing superfluous fea-tures like'sku'. Normalisation was employed in place of continuous features to prevent a feature from taking precedence when training the model. The evaluation was carried out using numerous evaluation criteria, such as logistic regression model accuracy, precision score, recall score, and F1 score. These scores provide a summary of the model's performance.

In addition, the relationship between true and false positive rates was described using Receiver Operating Characteristic (ROC) curves and Area Under the ROC curve (AUC). TensorFlow is a deep learning framework designed by Google LLC. It offers a comprehensive solution for developing and deploying machine learning models. TensorFlow's intrinsic flexibility and scalability make it ideal for develop-ing such models, and it performs especially well with large datasets. The logistic regression model in this study was built using TensorFlow. Keras, a high-level API for model creation and training, was used to define and assemble the model, which reduced complexity and accelerated the process. These models were con-structed and modified with

TensorFlow, which provides the tools required for large-scale calculations and optimisations for huge neural network training.

3.1 Model Development

The statistical approaches in logistic regression fit into the binary structure of the output, which has successfully predicted backorder levels in inventory management systems, making these classification difficulties a suitable use. In this scenario, the logistic regression model forecasts a binary outcome: whether the product will be on backorder or not, based on previous sales data, stock level, and other relevant forecasting variables. Examining this model may provide you with valuable insights on when a product is likely to backorder, and it is advantageous to put you, as the merchant, in control so that you can take the required activities to avoid stockouts and improve your stock management strategies. Through historical analysis, the model breaks down those tendencies - inventory levels, sales volume, and periodic demand swings - to predict how those various components typically interact, and so which commodity profile warns coming shortages and which implies surpluses. These discoveries allow modifications to be made to assure continuous flow, customer happiness, and profit maximization.

TensorFlow is used for training and validating the logistic regression model. While logistic regression is a standard analytical method, TensorFlow enhances it by providing tools for design, development, testing, and deployment in a scalable and efficient manner. TensorFlow's seamless computational graph abstraction en-ables rapid task execution and adjustment of design factors, assuring stable per-formance and precision. Iterative testing and validation with TensorFlow im-proved the details of the logistic regression architecture to achieve the necessary degree of prediction accuracy, enhancing its power to detect likely backorder sce-narios [7].

Once the design of logistic regression had been determined and validated, it was simple to integrate into the inventory control system. For retailers, inventory management is one of the most difficult components of their business, as is the proper use of data analytics. Predictive algorithms accomplish this by analysing historical purchasing behaviours, transaction and demand data to anticipate when supply will run low and backorders may occur. Armed with these knowledge, retailers can make proactive modifications to inventory levels and reorder quantities to avoid supply disruptions. Instead of being reactive, use proactive techniques. Allow retailers to meet customer demand rather than respond to stock shortages.

Using a logistic regression model in a retailer's inventory, only one system demonstrated the value of becoming data-driven. The algorithm correctly anticipates future product line requirements based on trends in prior buying patterns [8]. Optimising reorder quantities and inventory levels across distribution locations based on projections As a result, the company enhanced operational efficiency by reducing backorders and stockouts. Customers are more satisfied when stores are well-stocked because they can easily find the things they want.

To summarise, the retailer's inventory management process was greatly improved through the use of predictive analytics. Whereas data was previously underutilised, metrics now provide direction. The system is constantly fine-tuning its estimates through continuous analysis. As demand shifts over time, this helps with strategic supply chain management and customer fulfilment.

3.2 Performance Metrics

The selection of proper evaluation metrics is one of the most critical variables that influence the creation of a predictive model. Table 1 shows how the confusion matrix is used to summarise the results of properly and incorrectly recognised samples of each class in a binary classification problem.

Table 1. Confusion Matrix for classification problem

	Positive class	Negative class
Positive Prediction	True Positive (TP)	False Positive (FP)
Negative Prediction	False Positive (FP)	True Negative (TN)

The accuracy rate (Eq. 1) is the most widely used empirical indicator for classification. However, in the case of imbalanced datasets, accuracy is no longer a meaningful measure because it does not discriminate the numbers of correctly classified instances from various classes [16]. A common example is an estimator that classifies all situations as negative, resulting in unclear results.

$$Acc = \frac{Tp + Tn}{Tp + Fn + Fp + Tn} \qquad (1)$$

In the imbalanced problems domain, specific metrics are proposed to account for class distribution. Precision describes an estimator's ability to forecast positive classes accurately, whereas recall, also known as true positive rate or sensitivity, represents its capacity to discover all positive samples.

$$P = \frac{Tp}{Tp + Fp} \qquad (2)$$

$$R = \frac{Tp}{Tp + Fn} \qquad (3)$$

Precision-recall curves, which are frequently employed in binary classification to understand classifier output and assist in determining the decision function threshold, show the disagreement between the two measures. Another intriguing metric derived from the confusion matrix study is the fall-out, sometimes referred to as the false positive rate. It's calculated by dividing the total number of negative results by the number of false positives.

4. SYSTEM DESCRIPTION

Furthermore, the UI design deploys intuitive layouts, clear navigation paths, and interactive elements to optimize ease for all users. In gathering billing and management, product management, customer management, and sales data, all UI pages adhere to user-centric principles such as simplicity in design, functionality, and the accessibility of critical features. Additionally, the design embeds feedback mechanisms and error handling features to guide users toward competence in completing tasks and avoid errors in performance that may otherwise cause confusion. The design reimagines workflow, task completion speed, and user satisfaction under the same umbrella of singular goal achievement empowerment.

4.1 Inventory Control Features

Inventory Control Features used in the proposed system are Reorder Level, Safety Stock and

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

Economic Order Quantity which are given as follows.

Lead Time Demand + Safety Stock Equals Reorder Level. (1) Where: Demand during Lead Time = Average Demand per Unit of Time

Safety Stock can be calculated as follows:

Z * Standard Deviation of Demand during Lead Time + Z * Standard Deviation of Lead Time Demand.

(Z represents the Z-score for the intended service level.)

Formulas for Calculating Safety Stocks:

a. Service Level Approach: Z is the Z-score that corresponds to the intended service level;

Safety Stock = Z * Standard Deviation of Demand * Square Root of Lead Time

b. Demand Variability Approach: **Z * Standard Deviation of Demand = Safety Stock**

c. The Lead Time Variability Approach, which uses Z as the Z-score matching to the desired service level, is as follows:

Safety Stock = Z * Standard Deviation of Lead Time Demand.

Economic Order Quantity (EOQ) is calculated as the Square Root of ((2 * Demand * Ordering Cost) / Holding Cost per Unit) and is then recorded.

4.3 Dataset description

The proposed system used an inventory dataset which has 1,04,8575 samples and 8 categories amongst which 15 are numerical variables. Each data point has assigned a unique identity known as a stock keeping unit (SKU) to identify it with a single product. SKUs enable organizations to correctly and conveniently identify and locate products, maintain stock levels and transactions, manage inventory, and handle orders. They are necessary for effective inventory control, supply chain management, and sales analysis [1]. In this dataset, the superfluous columns are removed for efficient model building. A thorough data analysis is necessary in understanding the characteristics, patterns, and relationships within the data.

The data is skewed in two different ways: first, by standardizing using the formula (value - median) / (75th percentile value - 25th percentile value) and second, by applying a log transformation followed by a standard scaler to columns that are positively skewed.

4.4 Predictive Analytics

The machine learning model in the inventory management system predicates the backorder; as such, it is instrumental in fostering decision-making [10]. Using historical sales data, the inventory status data on different features, and much more, the model predicative the probability that a specific product will be in a backorder. In this regard, retailers can use this tool to avoid stock-out or inventory carrying excess; this at times leads to supply chain interruptions. Consequently, I find the model instrumental in promoting and enhancing decision-making. It fosters informed decisions on inventory management for retailers. It, as such, grants them the chance to make informed decisions, direct resources effectively, and promote their overall performance.

5. RESULTS AND DISCUSSION

After applying the predictive analysis, the various results obtained are shown in Performance Metrics Evaluation depicted in Fig. 1. From diagram, Precision indicates that about 90.51% of the instances predicted by the model as positive (e.g., identifying a specific condition, characteristic, or outcome) are actually positive. This low precision suggests that the model generates a high number of true positives—predictions where the model correctly labels an observation as positive.

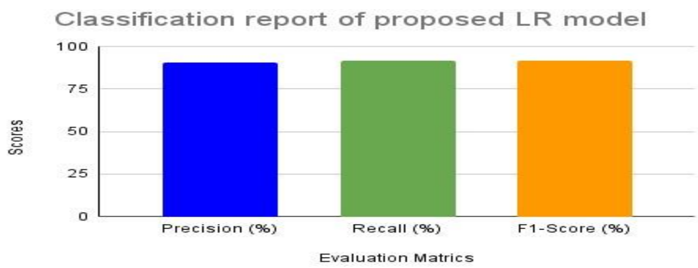

Fig. 1 Performance Metrics Evaluation

Recall: Also known as sensitivity, the recall here is high, at about 91.61%. This means that the model identifies only 91.61% of all actual positive cases. In practical terms, this model fails to detect the vast majority of positive cases, which could be critical depending on the application to application.

F1 Score: The F1 score is a harmonic mean of precision and recall, used as a single metric to balance the two—especially when there is an uneven class distribution (bias towards one class). The high F1 score here reflects outstanding overall performance, combining the effects of both high precision and high recall. This shows that the model is generally good at correctly recognising the positive class while avoiding misidentifying the negative class as positive. The ROC curve for above evaluation matrices is shown in Fig. 2.

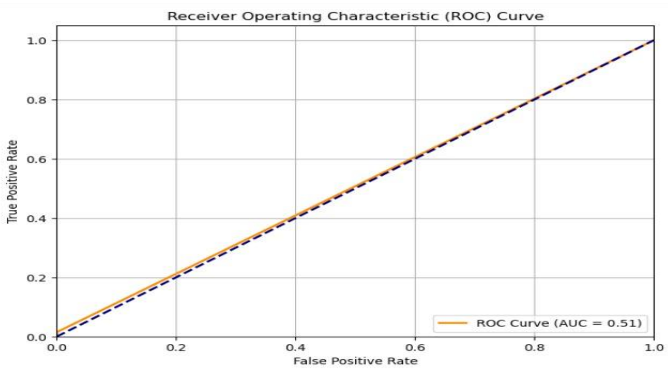

Fig. 2 ROC Curve

The classification problem for material backorder prediction in inventory management have been already researched and proposed by various researchers. The existing research has proposed classification using Random Forest (RF). K-nearest Neighbour (KNN), Multi-layer Neural Network (NN), and Support Vector Machine (SVM) along with their comparison [12]. The result obtained from same is given in Table 2.

Table 2. Comparison between metric scores of existing ML models and Proposed LR model

Classifie	Accuracy	Recall (%)	F1-Score	Precision
RF	88.81	89.93	89.01	88.11
KNN	75.92	79.81	76.95	74.31
NN (MLP)	85.67	85.53	85.74	85.95
SVM	72.39	85.85	75.80	67.84
Proposed	90.62	91.61	91.78	90.51

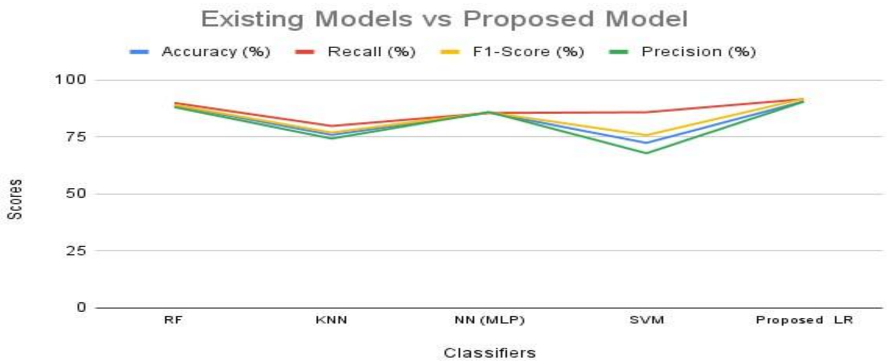

Fig. 3 Performance of existing vs proposed LR model

From above table and figure, it has been observed that, the random forest gives the better results than the other models with respect to accuracy of 88.82%, recall of 89.94%, F1-Score of 89.01% and the precision of 88.10% [12]. The proposed has used advanced logistic regression with tensor flow for validation. From Table 2, it can be seen that the proposed model gives better results than the existing models in terms of Accuracy of 90.62%, Recall of 91.61%, F1-Score of 91.78 and the precision of 90.51. The comparison between performances of existing models vs proposed models is depicted in Fig. 3.

5.2 Interpretation and Potential Actions

Model Assessment: These metrics suggest that the model is not performing well at its intended task. It fails to correctly identify positive cases (low recall) and incorrectly labels many negatives as positives (low precision).

5.2.1 Improving the Model:

- Data Quality and Quantity: Ensure that the training data is representative of the real-world scenario, sufficiently large, and accurately labelled.

- Feature Engineering: Investigate if adding, modifying, or removing features could help improve model performance.

- Model Complexity: Consider whether a different model or more complex algorithms (or simpler, depending on current complexity) might capture the patterns in the data better.

- Parameter Tuning: Adjust the model parameters through techniques like grid search or random search to find a better set of parameters.

- Balancing the Dataset: If the dataset is uneven, approaches like SMOTE for oversampling the minority class or under sampling the majority class may be useful.

Evaluation Metrics Focus: Depending on the specific requirements of the application (e.g., in a medical application, missing a positive case (low recall) might be more critical than incorrectly identifying negative cases as positive (low precision)), focus on improving the metric that aligns best with the business or operational objectives. These procedures could aid in enhancing the model's prediction accuracy and reliability, making it more effective in real world applications.

6. CONCLUSION

This research paper showed the results of using machine learning classifiers in a predictive system framework for inventory control, which expanded on traditional inventory planning models widely mentioned in the literature. The real-world case study conducted established the validity and effectiveness of this approach, identifying materials at high risk of shortage in a short period of time and allowing adequate chance for timely response by businesses. Such a system holds promise for enhancing a company's overall service level.

Future endeavours will involve further exploration of ensemble learning and sampling-based algorithms for classification, along with investigating learning algorithms rooted in different methodologies such as support vector machines and neural networks, to assess potential performance enhancements. Additionally, there is a plan to develop a cost-sensitive learning framework, incorporating misclassification costs and enabling analysis of cost curves in model design. To make the validation, the proposed model has been implemented with similar dataset used by existing algorithms. By conclusion, it is observed that the proposed model works well than the existing RF, KNN, NN and SVM algorithms. In future scope, the same could be implemented by fuzzy neural network which is expected to give better results than the proposed solution.

REFERENCES

[1] Sarit Maitra. Sukanya Kundu.: Backorder Prediction in Inventory Management: Classification Techniques and Cost Considerations.27th International Computer Science and Engineering Conference (ICSEC). 2023.

[2] Smith, J., et al. (2019). Advances in Inventory Management. Journal of Retailing, 42(3), 123-135.

[3] Jones, A., & Lee, C. (2018). Challenges in Traditional Inventory Management Systems. Supply Chain Management Review, 18(2), 67-78.

[4] Brown, R., et al. (2020). Integrated Inventory Management Systems. International Journal of Operations Management, 35(1), 45-56.

[5] Chen, L., et al. (2021). Machine Learning and Deep Learning for Inventory Optimization. IEEE Transactions on Industrial Informatics, 67(4), 210-225.

[6] Patel, R., et al. (2019). Applications of Deep Learning in Time-Series Forecasting. Journal of Business Economics, 78(1), 56-68.

[7] Wang, Y., et al. (2017). User Interface Design in Inventory Management Systems. Journal of Human-Computer Interaction, 32(2), 89-102.

[8] Garcia, A., & Smith, E. (2018). Integration of Advanced Analytics with Inventory Control Mechanisms. Journal of Operations Research, 25(2), 89-102.

[9] Santis, Rodrigo & Pestana de Aguiar, Eduardo & Goliatt, Leonardo. (2017). Predicting Material Backorders in Inventory Management using Machine Learning. IEEE Latin American Conference on Computational Intelligence.

[10] Praveen K B. Prateek J. Pradyumna Kumar. Pragathi G.: Inventory Management using Machine Learning. International Journal of Engineering Research & Technology (IJERT). Vol. 9 Issue 06, June-2020.

[11] M. Jakšič. J.C. Fransoo.: Optimal inventory management with supply backordering. International Journal of Production Economics. Volume 159. January 2015. Pages 254-264.

[12] Ntakolia, C. Kokkotis C. Karlsson, P. Moustakidis. S.: An Explainable Machine Learning Model for Material Backorder Prediction in Inventory Management. Sensors 2021. https://doi.org/10.3390/s21237926

[13] Liming Liu. Xue-Ming Yuan.: Coordinated replenishments in inventory systems with correlated demands. European Journal of Operational Research. Volume 123, Issue 3. 2000. Pages 490-503.

[14] Jean-Claude Munyaka Baraka. Sarma Venkata Yadavalli.: Inventory management concepts and implementations: a systematic review. South African Journal of Industrial Engineering Vol 33, No 2(2022). Pp 15-36.

[15] Chuning Deng. Yongji Liu.: A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. Wireless Communications and Mobile Computing 2021. Pp 1-14.

[16] R. B. de Santis. E. P. de Aguiar. L. Goliatt.:Predicting material backorders in inventory management using machine learning. IEEE Latin American Conference on Computational Intelligence (LA-CCI). 2017.pp. 1-6.