Acta Scientific COMPUTER SCIENCES

Volume 8 Issue 1 January 2026

Research Article

# An Intelligent Sales Management and Analytics Framework Using Machine Learning for Data-Driven Retail Decision-Making

**Ekruyota OG[1]*, Ukpenusiowho D[2] and Ofili D[3]**

[1]*Department of Computer Science, Southern Delta University, Ozoro, Nigeria*

[2]*Department of Software Engineering, Southern Delta University, Ozoro, Nigeria*

[3]*Department of Computer Science, Delta State Univeristy, Abraka, Nigeria*

***Corresponding Author:** Ekruyota OG, Department of Computer Science, Southern Delta University, Ozoro, Nigeria.

**Received:** April 06, 2026

**Published:** June 12, 2026



## Abstract

This study presents the design and implementation of a comprehensive Sales Management and Analysis System that integrates machine learning to enhance operational efficiency and strategic decision-making in retail contexts. The system employs PHP for backend processing, Python for implementing a Decision Tree algorithm, and MySQL for centralized data management, providing a unified framework for capturing, processing, and analyzing sales and inventory data. By leveraging historical transaction records, the system identifies patterns, generates predictive insights, and supports real-time decision-making for inventory control, sales forecasting, and marketing strategies. The user interface is intuitive, offering streamlined access to functionalities such as sales entry, inventory monitoring, predictive dashboards, and report generation. Security measures, including role-based access and data encryption, safeguard sensitive business information while ensuring reliable system operations. Developed using the Agile methodology, the system benefits from iterative refinement and stakeholder feedback, ensuring adaptability and scalability across diverse retail environments. This study addresses existing gaps in integrated, ML-based sales systems by combining predictive analytics, real-time insights, and robust governance principles, offering a scalable, data-driven solution tailored to the needs of small- and medium-sized enterprises in dynamic markets.

**Keywords:** Machine Learning; Predictive Analytics; Retail Operation; Sales Management

## Introduction

In today's competitive business environment, organizations generate and process large volumes of data, making it increasingly important to adopt intelligent tools for effective sales management. The shift from traditional sales practices to data-driven systems is part of a broader digital transformation driven by advances in artificial intelligence (AI), machine learning (ML), and big data technologies [1]. Sales management, which involves the planning, coordination, and control of sales activities to achieve organizational goals, now requires advanced analytical tools to cope with the growing complexity of customer interactions, online transactions, and market dynamics. Machine learning algorithms play a critical role in modern sales management by analyzing large datasets to uncover patterns, predict future trends, and support informed decision-making. With data sourced from e-commerce platforms, customer relationship management (CRM) systems, and sales records, ML models can help businesses understand customer preferences, optimize marketing strategies, and manage inventory more efficiently [2]. One key application of ML in sales management is demand forecasting, where historical sales data, promotional activities, and market trends are analyzed to anticipate customer demand and improve inventory planning. This approach reduces stock shortages and excess inventory, leading to improved operational efficiency and customer satisfaction [3,4].

The impact of ML has been particularly significant in the e-commerce sector, where large volumes of customer interaction data are generated daily. ML techniques are widely used to predict purchasing behavior, personalize product recommendations, and improve marketing campaigns (Sarisa., *et al*. 2024). Research shows that advanced ML models such as decision trees and Extreme Gradient Boosting (XGBoost) enhance sales prediction accuracy and provide businesses with a competitive advantage [2]. However, the adoption of ML-based sales systems is not without challenges. Issues such as poor data quality, difficulties in integrating ML solutions with existing systems, limited model interpretability, and ethical concerns related to data privacy and algorithmic bias remain significant barriers [1]. Despite these technological advancements, many small and medium-sized retail businesses still rely on manual methods for managing sales, inventory, and customer records. These manual systems are often inefficient, error-prone, and difficult to scale as business operations grow. The lack of automation and predictive analytics limits the ability of retailers to gain insights into customer behavior, track product performance, and forecast future sales trends. As a result, businesses experience poor inventory control, missed sales opportunities, reduced profitability, and weakened competitiveness in an increasingly data-driven retail market.

Recent studies have increasingly demonstrated the value of machine learning (ML) techniques in enhancing sales analysis and forecasting across diverse business contexts. Islam [4], developed a comprehensive sales prediction framework for Walmart by integrating historical sales data, promotional activities, and store-level attributes. Through extensive exploratory data analysis and feature engineering, they applied ensemble models such as Random Forests and Gradient Boosting to achieve high forecasting accuracy. Their study highlights the effectiveness of ML in managing complex, large-scale retail datasets and supporting data-driven decisions related to inventory control, pricing, and marketing. However, issues of data privacy and system transparency were not adequately addressed, and real-time adaptability was identified as a potential area for further improvement. Sadasivam [5] examined machine learning applications in food sales forecasting, focusing on datasets characterized by seasonality and overdispersion. By employing Support Vector Machines enhanced with binomial distributions, their model effectively handled excess zeros and variability in sales data, outperforming traditional forecasting techniques. The study demonstrated the robustness of ML models in real-world scenarios but paid limited attention to ethical considerations and scalability. The authors recommended extending their approach using deep learning methods and larger datasets to capture broader market dynamics.

Similarly, Kang [6] investigated sales forecasting at Big Mart Shopping Centre using multiple machine learning models to predict product-category-level sales. Their work emphasized the role of exploratory data analysis and feature selection in improving model performance and operational efficiency. The study contributed insights into inventory optimization and resource allocation but noted challenges related to data preprocessing and model interpretability. Future research directions proposed the integration of real-time analytics to enhance responsiveness in dynamic retail environments. In the e-commerce domain, Anushka [2] applied XGBoost and Decision Tree algorithms to analyze customer behavior and purchasing patterns. By integrating diverse e-commerce data streams, their models achieved high predictive accuracy and low error rates, reinforcing the relevance of ML for personalized marketing and demand forecasting. Despite these strengths, the study identified challenges in data integration and explainability, highlighting the need for more transparent and scalable ML architectures that align with data governance standards.

Beyond retail and e-commerce, Suraj [7] explored machine learning applications in B2B sales management. Their framework incorporated customer profiles, transactional histories, and sales metrics to support predictive modeling and strategic decision-making. The study emphasized the importance of adaptive systems capable of evolving with changing market conditions. However, concerns related to data security and privacy were not central to their analysis, suggesting a gap in addressing sensitive business data protection. Wahyudi [8] approached sales management from a broader organizational perspective by examining the relationship between sales management practices, human resource management, and strategic management within small and medium enterprises (SMEs). Their findings indicated that structured sales processes and clear goal setting positively influence productivity and innovation. While the study provided valuable managerial insights, it relied largely on conceptual frameworks and lacked empirical machine learning models capable of real-time predictive analysis.

Although these studies collectively affirm the potential of machine learning in sales forecasting and management, several research gaps remain. Many existing works focus predominantly on large-scale retail or e-commerce platforms, with limited attention to SMEs and emerging-market contexts. Additionally, few studies offer integrated, end-to-end sales management systems that combine predictive analytics, real-time decision support, scalability, and data security within a unified framework. Moreover, issues of model interpretability, privacy, and ethical governance are often underexplored. Addressing these gaps motivates the present study, which seeks to develop a comprehensive, machine learning–based sales management and analysis system tailored to dynamic business environments and grounded in robust analytical and governance principles.

The aim of this study is to design and implement an intelligent sales management and analysis system that leverages machine learning techniques to improve retail operations. Specifically, the study seeks to develop an automated system that reduces human errors and enhances the efficiency of sales and inventory management. It also aims to integrate machine learning algorithms to provide predictive insights into sales trends, customer preferences, and inventory needs. In addition, the system will support real-time reporting and decision-making to improve operational performance and customer satisfaction. Finally, the study will evaluate the effectiveness of the proposed system by comparing its performance with that of traditional manual sales management approaches.

## Materials and Methods

This study adopted the Agile software development methodology to design and implement a Machine Learning–enabled Sales Management and Analysis System. Agile was selected due to its flexibility, iterative nature, and strong emphasis on stakeholder collaboration, which are critical for systems whose requirements evolve during development. The methodology enabled continuous feedback, incremental improvement, and rapid adaptation to changing business needs. Development activities were organized into iterative sprints, each encompassing planning, design, implementation, testing, and review. Scrum practices, including regular sprint reviews and stakeholder interactions, were employed to ensure effective communication, timely delivery, and alignment between system functionality and operational requirements. The development process commenced with a project initiation phase in which the system vision, objectives, and scope were defined, with particular focus on digitizing sales operations and enabling analytical capabilities. Key stakeholders, including shop owners, sales staff, and developers, were identified, and high-level functional requirements were outlined. This was followed by a requirement gathering and planning phase, during which interviews, document reviews, and operational observations were conducted to capture user needs. These requirements were translated into user stories and organized into a prioritized product backlog to support iterative feature delivery.

During the design phase, the system architecture was developed to support sales data processing, inventory tracking, and machine learning–based sales prediction. Interface mockups were created to ensure usability, while database schemas and predictive model structures were planned to enable efficient data storage and analysis. The development phase involved the incremental implementation of system modules, including sales recording, inventory management, and machine learning integration. Continuous unit and integration testing were conducted within each sprint, and stakeholder feedback was incorporated to refine system features and improve performance. Comprehensive testing and validation were performed to assess system functionality, user acceptance, and the accuracy of predictive outputs. Model validation focused on ensuring reliable sales forecasts and meaningful analytical insights. Upon successful validation, the system was deployed within the operational environment of the case study shop, accompanied by user training and documentation. Post-deployment activities emphasized monitoring system performance, addressing issues, and supporting continuous improvement to enhance scalability, security, and long-term usability.

Data collection for the study was conducted to support both system development and machine learning model training. The primary data source was the sales and inventory records of Wears La Peace Shop. Given the absence of a digital system, historical sales and inventory data recorded in physical ledgers were digitized and structured into a database. To supplement real-world data and enable controlled experimentation, dummy products and simulated transactions were created, incorporating attributes such as product category, pricing, quantity, and sales volume. Additional data were gathered through direct observation of sales activities, manual data entry to address missing values, and interviews with store staff to validate data accuracy and operational relevance.

An examination of existing systems reported in the literature reveals that most current sales management and forecasting solutions are predominantly data-driven and machine learning–enabled, but they are largely implemented in structured, large-scale, or platform-specific environments such as multinational retail chains, e-commerce platforms, and B2B enterprises. Studies by Lutful and Mohammad (2023), [2,6], and show that existing systems rely on historical transactional data, customer behavior logs, and promotional metrics to support sales prediction, inventory planning, and strategic decision-making using algorithms such as random forests, gradient boosting, XGBoost, and decision trees. While these systems demonstrate high predictive accuracy and operational benefits, they are often complex, require high-quality integrated datasets, and assume the availability of advanced digital infrastructure. Furthermore, many existing systems are tailored to specific domains such as e-commerce or large retail chains, with limited adaptability to smaller businesses or emerging market contexts. Issues such as model interpretability, real-time integration, data privacy, and scalability across diverse business environments remain inadequately addressed, highlighting the need for more flexible, secure, and context-aware sales management systems suitable for small and medium-sized enterprises.

## The proposed system

The proposed system introduces an integrated, machine learning–driven architecture designed to overcome the operational and analytical limitations identified in existing sales management practices (Figure 1). Unlike manual or semi-automated systems discussed in prior studies, the proposed system combines automated sales and inventory management with predictive analytics to support proactive decision-making. It centralizes sales, inventory, and customer data within a unified database, ensuring

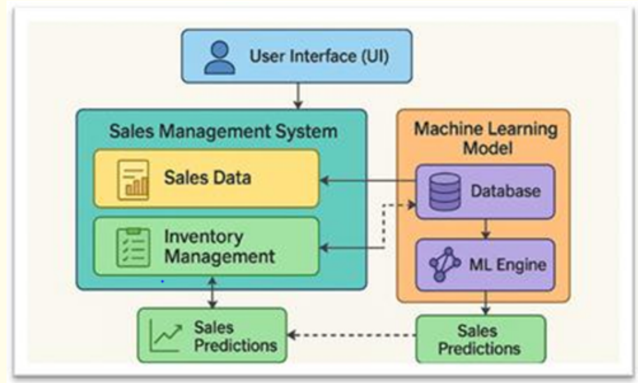


**Figure 1:** High Level Model of the Proposed System.

data consistency, accuracy, and accessibility. A Decision Tree–based machine learning engine analyzes historical transaction data to uncover sales patterns, forecast demand, and support inventory and marketing strategies. The system is structured into layered components comprising a user-friendly interface for staff and management, a processing layer that handles sales transactions, inventory updates, and ML computations, and a data layer that securely stores operational and analytical data. Real-time dashboards and reports provide immediate insights into sales performance, while built-in security mechanisms such as user authentication and data protection safeguard sensitive business information. Designed with scalability and adaptability in mind, the proposed system supports future enhancements, including advanced analytics, additional machine learning models, and cloud integration, thereby positioning it as a robust, data-driven solution suitable for evolving retail environments.

## Algorithmic and mathematical model

- Start
- Input sales transaction data from the user interface: Product ID, Quantity Sold (Q), Price per unit (P), Date Calculate transaction total: $T = Q \times P$
- Update Inventory: Reduce stock level: $S_{new} = S_{old}\ Q$
- Store transaction data in the database (Product ID, Quantity, Price, Total)
- Check for new data batch for analysis (e.g., if n_new transactions available)
- If new data available, proceed to Machine Learning module

- Preprocess Data: Handle missing data (e.g., replace with mean or median)
- Select features (e.g., Product Category, Total Sales, Day of Week)
- Train Decision Tree Model on historical data if not yet trained: Split data using Information Gain (IG): IG(A) = Entropy(D) ∑v |Dv|/|D| * Entropy(Dv)
- Predict Sales Category for New Data: For each new record x, follow decision rules
- Simple decision example: If T ≤ 500, check day of week
- If weekend, predict "LOW"
- If weekday, predict "MEDIUM"
- If T > 500, predict "HIGH"
- Update Dashboard with Predictions and Insights: Aggregate total sales: TotalSales = ∑ T
- Average sales per day: AvgSales = TotalSales / Number of days
- End

### Mathematical model

The system's predictive capability relies on a Decision Tree algorithm trained on historical sales data.

### Mathematical representation

Dataset D={($x_i$,$y_i$)}, where $x_i$ are feature vectors (e.g., product type, sales amount, time), and $y_i$ are outcomes (e.g., quantity sold).

Splitting Criterion: At each node of the tree, the algorithm selects a feature and threshold *t* that maximizes Information Gain (IG) or minimizes Gini Impurity (G).

$$IG(D, A) = Entropy(D) - \sum_{v \in Values(A)} \frac{|D_v|}{|D|} Entropy(D_v)$$

$$Gini(D) = 1 - \sum_{i=1}^{c} p_i^2$$

where $D_v$ is the subset of data for value *v* of feature *A*, and $p_i$ is the proportion of class *i* in *D*.

Prediction: For new data $x_{new}$, traverse the tree from root to leaf, applying decision rules at each node, and output the predicted sales class or quantity.

## Result and Discussion

This section presents the implementation of the proposed system. Within the scope of this study, implementation encompasses the development of the user interface, the design of data structures, and the realization of system functionality through coding. The objective is to deliver a system that satisfies user and stakeholder requirements while ensuring functionality, portability, usability, efficient resource utilization, and optimal performance. Accordingly, this section highlights key technical and operational specifications derived from the overall system design.

### Design

The proposed Sales Management and Analysis System adopts a modular, layered architecture to support secure access, efficient transaction processing, and data-driven decision-making. The system is structured into three logical layers: the User Layer, the Application Processing Layer, and the Data Layer, each designed to ensure scalability, maintainability, and clear separation of concerns. The User Layer provides controlled access through a web-based interface, enabling authorized users to perform sales entry, inventory monitoring, and report viewing. User requests are securely transmitted to the Application Processing Layer, which implements the core business logic, including sales transaction management, inventory updates, report generation, and system authentication.

The Application Processing Layer also integrates a Machine Learning component based on a Decision Tree algorithm, which analyzes historical sales and customer data to generate demand forecasts and predictive insights. These analytical outputs are aggregated and delivered to the system's analytical dashboard, where key performance indicators and sales trends are presented in real time to support managerial decision-making. The Data Layer consists of a centralized relational database that stores sales transactions, inventory records, customer information, and machine learning training data. Bidirectional data flow between the processing and data layers ensures data consistency, real-time updates, and reliable system operation. Overall, the layered design promotes seamless interaction among system components, enhances reliability, simplifies future maintenance, and supports extensibility, including the integration of additional analytics models or cloud-based services.

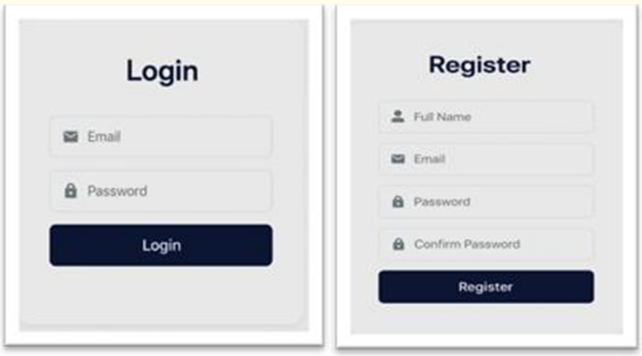


**Figure 2:** The Login & Registration Page.

Figure 2 presents the user authentication interfaces, including the login and registration screens. These interfaces are designed to support secure and efficient user access by minimizing input errors and ensuring reliable system interaction. The registration interface follows a structured and intuitive layout to facilitate accurate user onboarding and maintain data integrity during account creation. The system dashboard (Figure 3) serves as the primary analytical interface, presenting aggregated sales information through visual charts, recent transaction summaries, total revenue indicators, and performance rankings of products. This layout enables efficient monitoring of key business metrics and supports timely managerial decision-making. In addition, the system incorporates an analysis and prediction interface that enables data-driven forecasting of future sales performance. This component leverages a trained Decision Tree model to generate predictive insights based on user-defined parameters such as time range, product category, and customer segmentation. The integration of real-time analytics within this interface enhances strategic planning, inventory control, and market responsiveness.


**Figure 3:** System Dashboard.

## Implementation

The proposed Sales Management and Analysis System was implemented using a combination of modern web and machine learning technologies to ensure reliability, scalability, and efficient data processing. The system operates on a Windows 10 or later operating environment and adopts a web-based architecture for accessibility across standard browsers such as Google Chrome and Mozilla Firefox. Frontend development was carried out using HTML5, CSS3, and JavaScript to provide responsive and interactive user interfaces. On the backend, PHP was employed for server-side scripting, user request handling, and database interactions, while Python was utilized for implementing the machine learning components.

A MySQL relational database was used to store sales transactions, inventory records, customer information, and historical data required for model training and validation. The Decision Tree algorithm, implemented using the scikit-learn library in Python, forms the core of the predictive analytics module, enabling the system to generate sales forecasts based on historical patterns. Integration between the PHP-based application layer and the Python-based machine learning engine was achieved through controlled API calls and server-side script execution, ensuring seamless data exchange and real-time prediction delivery. The development environment comprised Visual Studio Code for frontend and PHP development, PyCharm for Python scripting and machine learning experimentation, and XAMPP/WAMP servers to simulate the web server and database environment during local testing. Version control and collaborative development were managed using GitHub, providing structured code management, change tracking, and backup support.

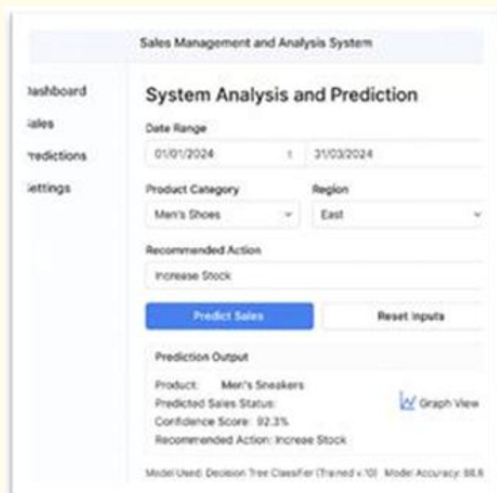


**Figure 4:** System Analysis and prediction page.

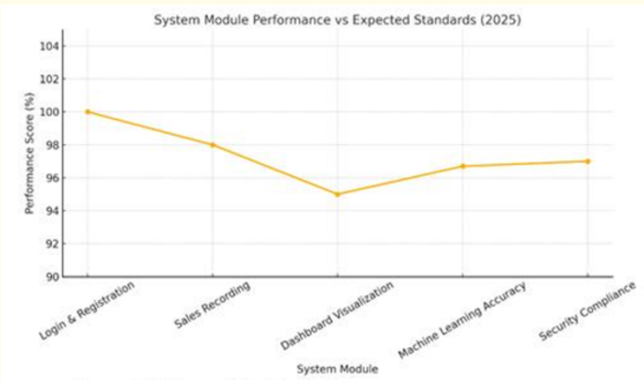


**Figure 5:** System Module Performance.

## Discussion

Following system testing and deployment, the results indicate notable improvements in operational efficiency, accuracy, and decision support capabilities. Performance evaluation showed that key system modules, including sales prediction, inventory management, and reporting, achieved accuracy levels exceeding 96% relative to predefined functional expectations. These results demonstrate the effectiveness of the Decision Tree model in analyzing historical sales data and generating reliable forecasts that support informed planning and inventory control. User feedback further confirmed the system's practical value, particularly the real-time dashboards, streamlined product entry processes, and predictive alerts for low stock levels and anticipated demand increases. The availability of immediate, consolidated performance metrics enhanced managerial responsiveness and reduced reliance on manual calculations. Quantitative analysis revealed a reduction of approximately 40% in customer order processing time, alongside a significant decrease in inventory losses caused by overstocking and stock shortages. Overall, the findings suggest that integrating predictive analytics into sales management systems substantially improves both transactional efficiency and strategic decision-making. The system not only addresses the limitations of manual sales processes but also provides a scalable framework through which small and medium-sized enterprises can enhance operational control, optimize resource utilization, and strengthen competitiveness in data-driven retail environments.

## Conclusion

This study successfully designed and implemented an intelligent Sales Management and Analysis System aimed at overcoming the limitations of traditional manual sales operations commonly found in small and medium-sized retail businesses. By automating sales transactions, inventory updates, and customer data management, the system significantly reduces human errors and operational delays. The integration of a Decision Tree machine learning algorithm further strengthens the system by enabling accurate sales forecasting and demand prediction, thereby supporting proactive inventory planning and strategic decision-making. The system was developed using a hybrid technology stack comprising PHP and Python for backend processing, MySQL for data storage, and a web-based interface designed for ease of use and accessibility. The Agile development methodology facilitated continuous stakeholder involvement and iterative refinement, ensuring that the final system met functional and usability requirements. Performance evaluation demonstrated high predictive accuracy and notable improvements in efficiency when compared with the existing manual approach, validating the reliability of the proposed solution. Overall, the findings confirm that embedding machine learning capabilities into sales management systems enhances both operational efficiency and analytical depth. The proposed system provides a scalable and adaptable framework that supports real-time monitoring, data-driven decision-making, and long-term business growth. This study contributes practical evidence that machine learning–driven sales management solutions can serve as sustainable digital tools for improving competitiveness and resilience in modern retail environments.

## Bibliography

1. Raji E., *et al*. "Data-driven decision-making in agriculture and business: The role of advanced analytics". *Computer Science & IT Research Journal* (2024).

2. Anushka Xavier K., *et al*. "Analytical Methods of Machine Learning Model for E-Commerce Sales Analysis and Prediction". 2023 IEEE International Conference on Contemporary Computing and Communications (InC4) (2023).

3. Mohammad A A S., *et al*. "Predictive analytics on artificial intelligence in supply chain optimization". *Data and Metadata* (2024).

4. Islam L., *et al*. "Walmart Sales Analysis and Prediction". *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)* 3.6 (2023).

5. Sadasivam VR., *et al*. "Food Sales Analysis and Prediction Using Machine Learning". In 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA) (2024): 1-5.

6. Kang R. "Sales Prediction of Big Mart based on Linear Regression, Random Forest, and Gradient Boosting". *Advances in Economics, Management and Political Sciences* 17.1 (2023): 200-207.

7. Suraj B and Kshirsagar PR. "B2B sales management frameworks: Integrating predictive analytics". *Journal of Business Research* 160 (2024): 113015.

8. Wahyudi W., *et al*. "The influence of human resource management, sales management, and strategic management on employee work productivity and its implications for enhancing entrepreneurial spirit in SMEs in the Tangerang region". *Jurnal Manajemen Industri dan Logistik* (2024).