Predictive Sales and Inventory System with Customer Segmentation for Enhance Customer Relationship Management

Nathaniel C. Custodio[1], Harold R. Lucero[2], Melvin G. Torres[3], Rhozzel C. Tayubong[4], Angela D. Quillondrino[5]

School of Information Technology, Colegio de Sta Teresa de Avila

DOI: https://dx.doi.org/10.51584/IJRIAS.2026.110100137

Received: 04 February 2025; Accepted: 13 February 2026; Published: 20 February 2026

ABSTRACT

This study presents the design and development of a Predictive Sales and Inventory System with Customer Segmentation for Enhanced Customer Relationship Management. The system addresses operational inefficiencies in small retail businesses by integrating predictive analytics and clustering techniques. Linear regression was applied for sales forecasting, while K-means clustering was used for customer segmentation to support data-driven decision-making and personalized customer engagement. The platform automates sales transactions, inventory monitoring, and membership management, and supports RFID-based cashless payments. Development followed the Agile Scrum methodology, with modules including product management, transaction processing, analytics, reporting, and backup management. User Acceptance Testing based on the Technology Acceptance Model (TAM) yielded an overall weighted mean score of 4.68, indicating strong user acceptance, with one establishment expressing intent to adopt the system. System evaluation based on ISO/IEC 25010 quality standards produced an overall weighted mean score of 4.18, reflecting satisfactory performance in functionality, reliability, usability, and security. Results demonstrate that the system provides an efficient, browser-based, data-driven POS solution that improves transaction speed, operational accuracy, and business insight generation for Spot777 Coffee.

Keywords: Agile Scrum, Predictive Analytics, K-Means Clustering, Customer Relationship Management, Customer Segmentation, Business Intelligence, POS Platform

INTRODUCTION

Predictive analytics via linear regression and descriptive analytics through K-means clustering are essential in contemporary retail, converting extensive sales and customer information into significant insights. Linear regression enables businesses to forecast future sales patterns based on historical data, allowing managers to anticipate demand, optimize inventory levels, and reduce operational uncertainties, which enhances planning and resource allocation in retail operations [1]. Similarly, descriptive methods, such as K-means clustering, enable businesses to identify distinct customer segments based on purchasing behavior, spending patterns, and product preferences. These insights form the basis for targeted marketing, loyalty programs, and personalized promotional strategies, ultimately enhancing customer engagement and profitability [2].

In today's competitive retail landscape, adopting business analytics is critical for organizations striving to make data-driven decisions that improve operational efficiency and strategic planning. With the increasing availability of transactional and customer data, businesses can analyze purchasing behavior, sales trends, and inventory performance in real time. Predictive analytics and data clustering techniques have emerged as essential tools for transforming raw sales data into actionable insights that guide strategic decisions [3] [4].

Point of Sale (POS) platform play a central role in collecting and managing these data streams, serving as the foundation for analytical modeling. Studies show that integrating RFID technology into POS platform enables real-time tracking of customer transactions and inventory, enhancing operational transparency and analytical accuracy [3]. Embedding analytics tools within POS environments allows businesses to monitor sales performance, forecast demand, and reduce human errors associated with manual data entry. Predictive analytics,

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue I January 2026

in particular, enables retailers to align production and purchasing decisions with anticipated customer demand, minimizing overstocking and shortages. Furthermore, customer segmentation using machine learning algorithms such as K-means or RFM segmentation allows organizations to group customers based on purchasing frequency, spending habits, and product preferences [4] [5]. Such insights support targeted promotions, loyalty programs, and personalized marketing strategies that improve retention and profitability.

This study was conducted at Spot777 Coffee, a growing local café that began as a mobile retail establishment in February 2022 and specializes in locally produced coffee drip bags and tea bags. In April 2023, the café expanded to serve repackaged orders for clients, resellers, and affiliated businesses. Spot777 Coffee officially launched its physical location on December 30, 2023, at Level 3, Novamall Plaza Mall, Novaliches, Quezon City. Founded by Mrs. Con Bendicio Lamsen, the café now offers a variety of coffee, tea, pasta, and affordable meals, primarily serving students and mall patrons. The establishment has implemented a membership program that provides exclusive benefits, including a 7% discount, a personalized mug, and a complimentary drink every month for one year. The café serves between thirty and sixty customers daily, with thirty active members enrolled at the time of this study.

According to an interview with Mrs. Lamsen on October 24, 2024, Spot777 Coffee currently manages all transactions manually. This approach presents several challenges that impact efficiency, accuracy, and overall business performance. Handwritten orders and payments slow down transaction processing, create long lines during peak hours, and reduce customer satisfaction. Manual computation of prices, discounts, and change increases the likelihood of errors, potentially resulting in financial losses or customer complaints. The absence of an automated monitoring system prevents real-time tracking of daily sales, forcing employees to manually compile records at the end of each day, which hampers management's ability to make timely, informed decisions. Additionally, the café does not issue receipts, which may lead to disputes, and inventory is tracked manually, increasing the risk of overstocking or shortages. The café also only accepts cash and e-wallet payments, limiting convenience for customers who prefer card-based or cashless options.

To address these challenges, this study proposes the design and development of a web-based POS system with integrated inventory management and RFID-enabled cashless membership payments. By leveraging RFID technology, customers can make secure, cashless payments via their membership cards, reducing wait times and eliminating cash handling. The system will generate real-time, detailed reports and analytics, providing management with valuable insights into sales trends, inventory status, and customer preferences. These data-driven capabilities will support informed decision-making, allowing Spot777 Coffee to adapt quickly to market changes, optimize operations, and enhance overall business performance, ultimately promoting sustained growth and profitability.

Related Studies

The study of Rodriguez, M.Z. et.al., entitled "Clustering Algorithm: A comparative approach". This research systematically compares nine well-known clustering methods in the R language, assuming normally distributed data. The study uses artificial datasets with tunable properties like the number of classes and separation between classes to account for data variations. The sensitivity of the clustering methods to their parameter configuration is also evaluated. The results indicate that the spectral approach performs well with default configurations. The default configurations were not always accurate, and a random selection of parameter values can improve performance. The study guides the choice of clustering algorithms. The study discusses the importance of clustering algorithms in machine learning and their applications in various domains. It also addresses the challenge of choosing a suitable clustering approach and compares different methods using real-world and artificial data. The study uses a diversified set of artificial, normally distributed data. It evaluates performance based on the similarity between known labels and those found by the algorithm. It uses metrics like the Jaccard index, Adjusted Rand index, Fowlkes-Mallows index, and Normalized mutual information [6].

According to the study by Mendoza, A. R. entitled "POINT OF SALE SYSTEM WITH INVENTORY FOR ARM'S FOOD AND DELICACIES,". This study developed a point of sale (POS) system that manages inventory for ARM's Food and Delicacies, a privately owned company. The existing manual system was

inefficient, causing issues with tracking sellable items, checking expiration dates, managing sales records, and generating stock reports. The new system aims to improve efficiency, speed up transactions, and provide better data management. The system uses a computerized approach, improving the accuracy and accessibility of records compared to the manual system. It integrates a barcode scanner and provides daily, monthly, and yearly reports [7]

According to Noveda, J., et al., entitled "Leyte Normal University: Supply Inventory Management System". This study investigates the significance of technology adoption for efficient inventory management in the context of the Philippines' supply office at Leyte Normal University (LNU). It uses action research to design, implement, and evaluate a technology-based system to address current inefficiencies in the manual processes of the Supply Office at Leyte Normal University. The study employs theoretical frameworks such as the two-bin system, just-in-time theory, and vendor-managed inventory. It uses surveys with Likert scales and open-ended questions to gather data. The goal is to create a system that improves efficiency, streamlines procurement processes, minimizes costs, and ensures resource availability. The survey results indicated that the implemented Supply Inventory Management System was considered relevant and beneficial by the employees of the LNU Supply Office, with 80% of participants indicating the system successfully meets their unique demands. The study concludes that the system has become a crucial tool for the office by increasing operating efficiency, streamlining inventory management procedures, and enabling informed decisions through real-time updates and automated inventory tracking.

In the study of Maltezo, M. R. C. et al., and Technological University of the Philippines, entitled "Integrated Port eManagement and Monitoring System using radio frequency identification," Technological University of the Philippines, 2021. This study aimed to automate port services by monitoring and supervising the increasing number of passengers and improving services like passenger and crew tracking and generating real-time voyage reports. The system developed included modules for passenger ticketing, check-in/boarding, baggage billing, and tracking passenger accommodation, using RFID technology and an EPC Class 1 Gen2 standard for identification and authentication. The system was successfully tested, demonstrating the localization capabilities of RFID technology for improved port services. The study concluded that using RFID as an alternative to manual ticketing systems, with added features like accommodation allocation and monitoring, successfully eliminated the need for manual processes, reducing queues and human error, and generating precise reports. The online ticketing system integrated with RFID technology achieved 100% read range authentication, tracking, and mapping of RFID tags [8].

DESIGN AND METHODOLOGY

Methodology

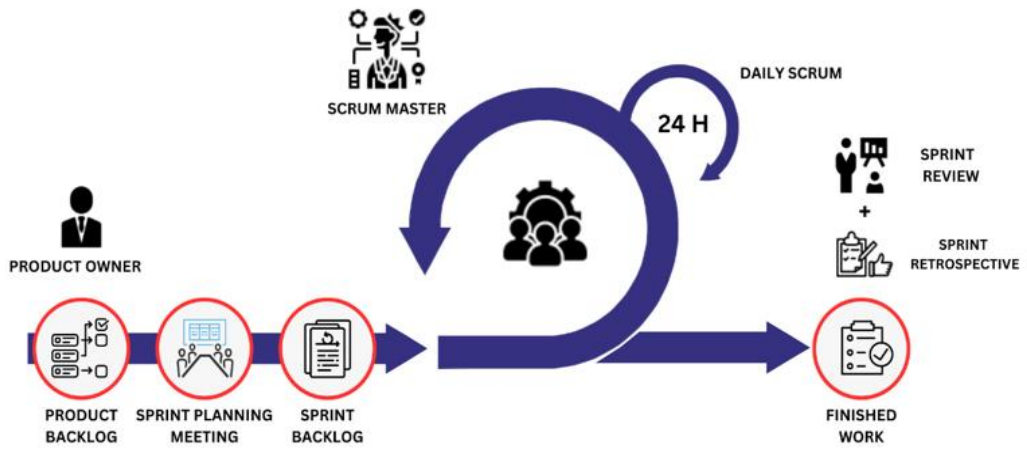

Figure 1. Agile Scrum Methodology

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue I January 2026

The Agile Scrum methodology was adopted as the software development framework due to its iterative, incremental, and feedback-driven nature. The project was divided into time-boxed sprints lasting two to four weeks. Each sprint included sprint planning, task prioritization, development, testing, and review.

Daily stand-up meetings were conducted to monitor progress and identify blockers. Sprint reviews and retrospectives were performed after each cycle to evaluate deliverables and improve subsequent iterations. This approach ensured continuous refinement and alignment with user requirements.

Linear Regression Algorithm

Linear regression was implemented to predict sales by modeling the relationship between time and historical sales data. This approach identifies trends and estimates future sales, providing a quantitative basis for forecasting. The slope m, representing the rate of change, and the intercept b, representing the baseline sales, were calculated using standard formulas (Equations 1 and 2) and applied in the regression model (Equation 3) to generate predictions.

Equation 1. Linear regression slope formula

$$m = \frac{n\sum(XY) - (\sum X)(\sum Y)}{n\sum(X^2) - (\sum X)^2}$$

Equation 2. Intercept formula

$$b = \frac{\sum Y - m\sum X}{n}$$

$$y = mx + b$$

Through this implementation, the system can analyze sales behavior over time and forecast future demand, supporting informed inventory planning and strategic decision-making. Integrating linear regression enhances the system's analytical capability by enabling data-driven insights that improve operational efficiency and promote sustainable business growth.

K-Means Clustering Algorithm

The K-means clustering algorithm was implemented to segment customers based on purchasing behavior. Each transaction was represented as a two-dimensional feature vector consisting of item quantity and transaction value.

Cluster assignments were determined by minimizing Euclidean distance to centroids, which were iteratively recalculated as cluster means until convergence.

The resulting clusters support behavioral profiling, targeted promotions, and customer relationship strategies(Equation 4).

$$d(x, \mu) = \sqrt{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2}$$ Equation 4. Euclidean distance formula

Centroids were iteratively updated as the mean of points in each cluster (Equation 5):

Equation 5. Centroid Update formula

$$\mu = \frac{1}{|C_i|}\sum_{x \in C_i} x$$

The process repeated until centroids stabilized, forming consistent clusters. By embedding these formulas, the system efficiently handles large POS datasets, revealing customer purchasing patterns and supporting profiling, visualization, and targeted marketing.

K-means integration enhances analytics, improves customer relationship management, and guides data-driven decisions.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue I January 2026

System Architecture

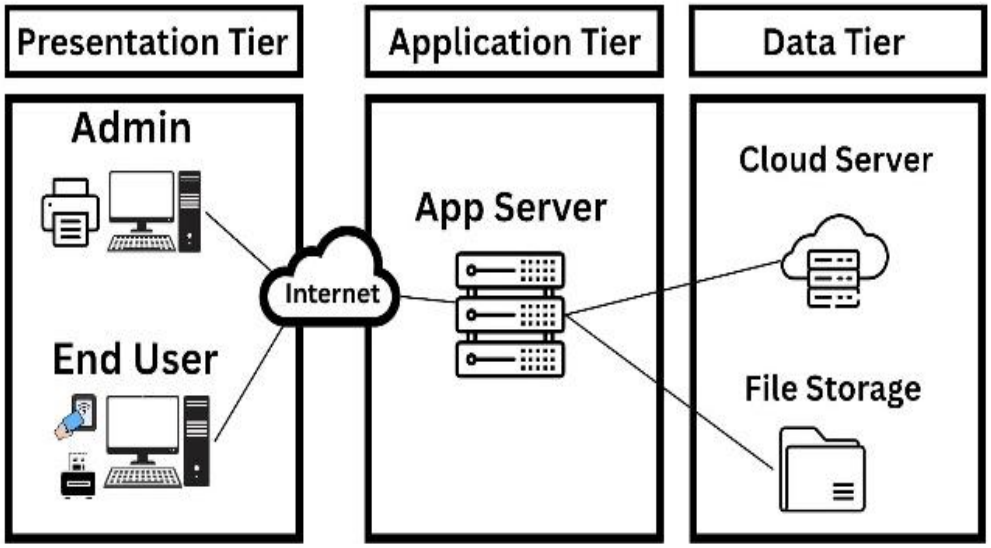

Figure 2. Sytem Architecture of the Developed System

The figure below illustrates the system architecture of the proposed solution. An internet-connected database server stores and manages essential data, including sales, transactions, and user information. The owner accesses a back-office terminal with an inkjet printer for reports, while staff use a POS terminal equipped with a thermal printer and RFID reader for transactions and membership management. Both terminals connect to the server online, with owners having full administrative access and staff limited to transaction processing. The system activates upon user login, streamlining administrative and sales tasks. Development and operation utilize Visual Studio Code 2024, XAMPP v3.3.0, printers, an RFID reader, and a USB cash drawer, ensuring efficient POS functionality.

RESULT AND DISCUSSION

The Developed System

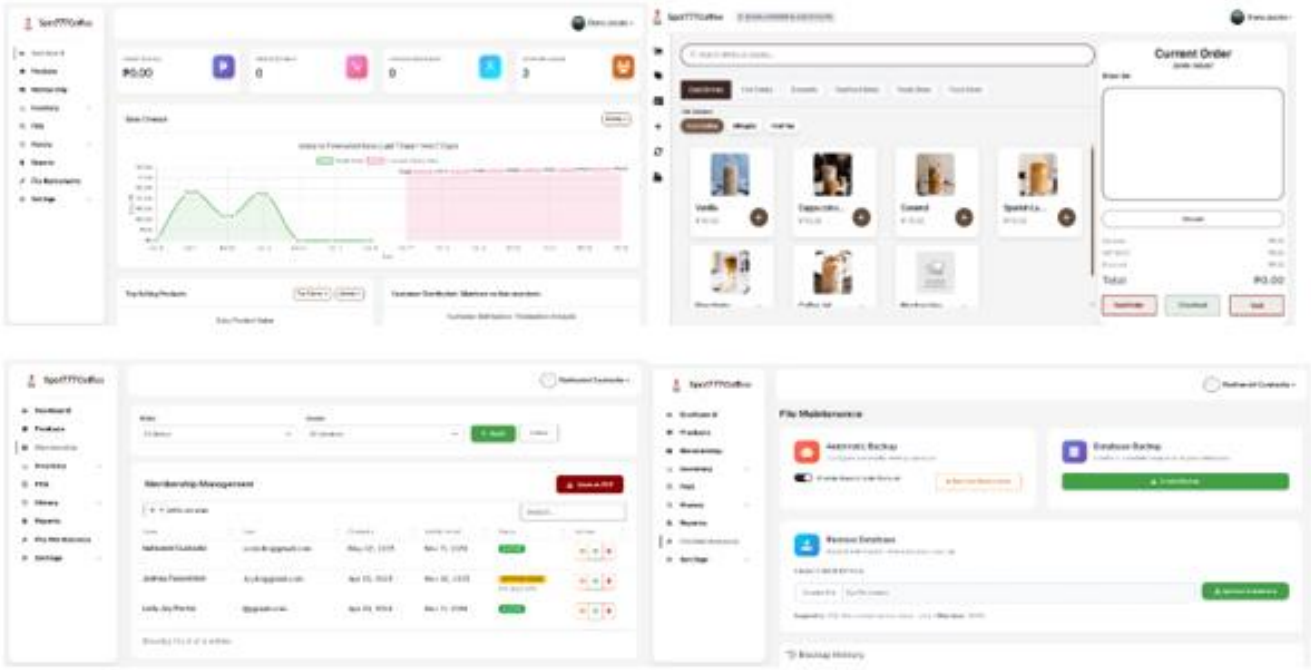

Figure 3. The Developed System

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue I January 2026

The developed Predictive Sales and Inventory System integrates POS automation, real-time inventory monitoring, sales forecasting using linear regression, and customer segmentation through K-means clustering within a unified web-based platform. Implementation results show improved transaction speed, reduced manual errors, and enhanced visibility of sales and stock data through dashboards and automated reports. The system's analytics features support data-driven decision-making by providing demand forecasts and customer behavior insights, while RFID-based membership functionality enables faster and more convenient cashless transactions.

Linear Regression Algorithm Implementation

Figure 4. Sample Sales Prediction

The figure above illustrates the historical sales data alongside the six-month forecast generated by the developed POS system. Applying the linear regression model to the observed data, the system projects a gradual decline in sales for the upcoming period. This trend suggests that without strategic intervention, sales performance may continue to diminish. Consequently, these forecasts serve as a critical decision-support tool, guiding business owners in the implementation of targeted marketing and operational strategies to improve future performance.

Utilization of K-Means Clustering Algorithm

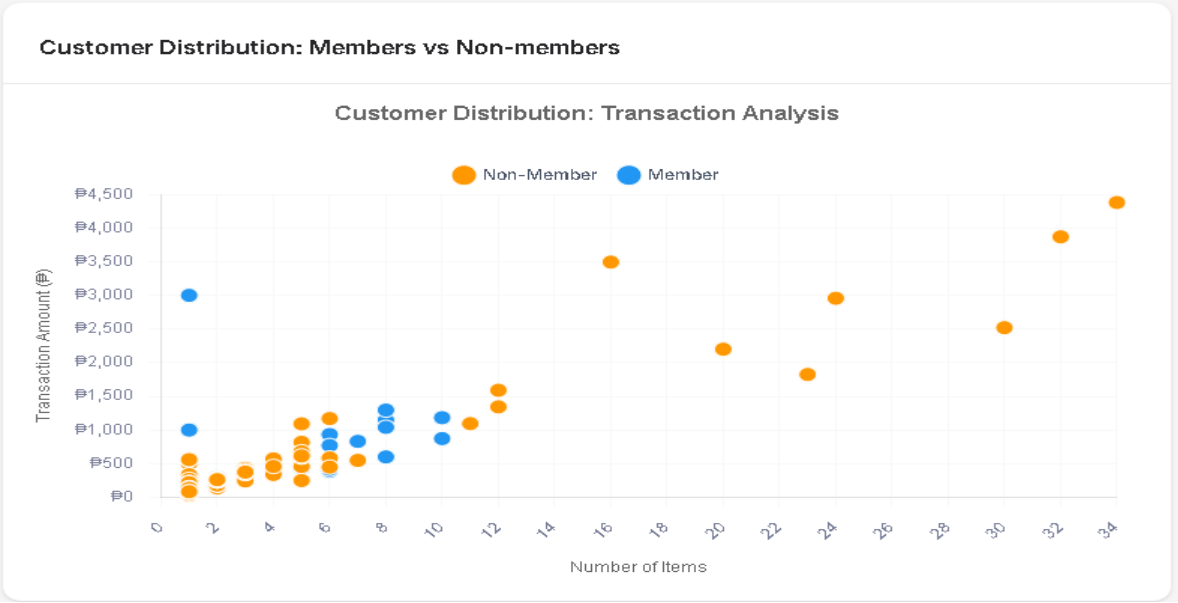

Figure 5. Custumer Segmentation

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue I January 2026

The K-Means clustering algorithm was implemented in the system's transaction analytics module to automate customer segmentation based on purchasing behavior. Each POS transaction was represented as a two-dimensional feature vector (x, y), where x denotes the number of items purchased and y represents the transaction amount. These features were extracted directly from the transaction database and used as inputs for clustering.

The system initialized centroids and assigned each transaction to the nearest one using the Euclidean distance (Equation 4). Centroids were updated iteratively as the mean of their cluster points until convergence. This produced clusters of transactions with similar patterns, supporting customer analysis and decision-making.

User Acceptance Test Result

Table 1:User Evaluation Test Result

#	**Indicator**	**Weighted Mean**	**Interpretation**	**Rank**
1	System's Perceived Usefulness	4.75	Strongly Agree	1
2	System's Perceived Ease of Use	4.76	Strongly Agree	2
3	System's Attitude Towards Use	4.68	Strongly Agree	3
4	System's Behavioral Intention to Use	4.52	Strongly Agree	4
	Overall Weighted Mean	4.68	Strongly Agree	

The User Acceptance Test was evaluated using the Technology Acceptance Model (TAM) with participation from sixteen (16) employees of Spot777 Coffee, including the business owner, manager, cashiers, and crew members, ensuring representation across operational roles. The results indicate a high level of overall user acceptance of the developed system.

Perceived Usefulness obtained a weighted mean of 4.75 (Strongly Agree), showing that users believe the system improves efficiency, productivity, and effectiveness in sales and inventory operations. Respondents highlighted faster transactions, automated reports, and real-time inventory tracking as key benefits. Perceived Ease of Use recorded a weighted mean of 4.76 (Strongly Agree), indicating that the interface is intuitive and functions are easy to learn and operate, even for non-technical users.

Attitude Toward Use scored 4.68 (Strongly Agree), reflecting a positive overall user perception and satisfaction with the system. Behavioral Intention to Use yielded 4.52 (Strongly Agree), demonstrating strong willingness among respondents to continue using and recommending the platform.

Overall, the TAM composite mean of 4.68 confirms strong acceptance and readiness for adoption, indicating that the system is both useful and user-friendly for daily retail operations.

System Evaluation Test Result

Table 2:System Evaluation Test Result

#	**Criteria**	**Weighted Mean**	**Interpretation**	**Rank**
1	Functional Suitability	4.51	Strongly Agree	2
2	Performance Efficiency	4.35	Agree	4
3	Compatibility	4.25	Agree	5

4	Usability	4.58	Strongly Agree	1
5	Reliability	4.15	Agree	6
6	Security	4.13	Agree	7
7	Maintainability	4.13	Agree	7
8	Postability	4.47	Agree	3
	Overall Weighted Mean	4.32	Strongly Agree	

The system evaluation was conducted using the ISO/IEC 25010 software quality model, with assessment provided by seven (7) industry experts. The overall weighted mean of 4.32 (Strongly Agree) indicates that the developed system meets recognized software quality standards and is considered suitable for operational deployment.

Among the evaluated criteria, Usability ranked highest with a weighted mean of 4.58 (Strongly Agree), showing that experts found the interface clear, learnable, and efficient for end users. Functional Suitability followed with 4.51 (Strongly Agree), confirming that the system adequately delivers the required core functions for sales processing, inventory control, forecasting, and customer segmentation.

Portability ranked third with 4.47 (Agree), indicating that the system can be deployed across different environments with minimal difficulty. Performance Efficiency (4.35, Agree) and Compatibility (4.25, Agree) suggest that the platform performs responsively and works well with required hardware and software components.

Meanwhile, Reliability (4.15, Agree), Security (4.13, Agree), and Maintainability (4.13, Agree) received slightly lower but still positive ratings, indicating stable operation and acceptable protection and support for updates, with opportunities for further strengthening in error handling, security controls, and documentation.

Overall, the ISO/IEC 25010 evaluation confirms that the system demonstrates strong quality characteristics, particularly in usability and functional coverage, while remaining robust and maintainable for practical use.

CONCLUSION

The developed Predictive Sales and Inventory System integrates core modules to enhance business operations and customer relationship management. Product Management enables efficient handling of item information, while the Sales Transaction module automates real-time transaction recording, reducing manual errors and improving data accuracy. The Membership Management module supports customer profiling and loyalty programs, strengthening personalized engagement. Through the Business Analytics module, the system evaluates sales trends and customer behavior to generate predictive insights that guide inventory planning and marketing strategies. Inventory levels are automatically updated after each transaction to prevent overstocking or shortages, while the Report Generation module produces clear, printable summaries to support managerial decision-making. System integrity and data security are ensured through role-based access control and user authentication, complemented by a Backup and Restore module that safeguards data continuity. Overall, the system meets its design objectives by providing a comprehensive, data-driven POS solution that improves operational efficiency and customer relations.

ACKNOWLEDGMENTS

First and foremost, we praise and thank God Almighty for His abundant blessings, wisdom, and strength, which guided us throughout the completion of this capstone project. His grace enabled us to overcome the challenges and obstacles along the way.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue I January 2026

We would like to express our deepest gratitude to the Dean of the School of Information Technology of Colegio de Sta. Teresa de Avila, for his unwavering guidance, patience, and invaluable support throughout the course of our study. His insightful feedback, encouragement, and dedication have greatly contributed to the success of our research. We are truly grateful for his time, effort, and commitment in helping us to achieve our goals and for inspiring us to strive for excellence in our academic journey.

We sincerely thank Mrs. Con Bendicio Lamsen, the owner of Spot777, for her generosity, support, and cooperation in providing the information we needed for our study. Her assistance and kindness greatly contributed to the completion of our research. Also, we wholeheartedly thank our friends and classmates for their support and encouragement throughout our study. We are also grateful to our professors for their guidance and valuable insights, and we acknowledge Colegio de Sta Teresa de Avila for providing a supportive environment that helped us complete this study. Most of all, we sincerely thank our loving parents for their unconditional love, sacrifices, and constant support which helped us finish our study.

REFERENCES

1. H. P. S. M. F. I. Ghorban Tanhaei, "Predictive analytics in customer behavior: Anticipating trends and preferences," vol. 17, 2024.
2. O. N. Akande, H. B. Akande, E. O. Asani and B. T. Dautare, "Customer Segmentation through RFM Analysis and K-means Clustering: Leveraging Data-Driven Insights for Effective Marketing Strategy," 2024.
3. S. R. &. L. Vuyyuru, "Unlocking future consumer insights: Using predictive analytics and AI to shape proactive retail strategies," European Journal of Advances in Engineering and Technology, p. 110–115, 2023.
4. J. Tang, "Unlocking Retail insights: predictive modeling and customer segmentation through data analytics," Journal of Theoretical and Applied Electronic Commerce Research, vol. 2, no. 20, p. 59, 2025.
5. R. K. &. N. N. Koppichetti, " Customer Retention with Predictive Analytics in the Retail Industry," International Journal of Multidisciplinary Research and Growth Evaluation, vol. 03, p. 776, 2022.
6. M. Z. Rodriguez et al., "Clustering algorithms: A comparative approach," PLoS One, 2019.
7. A. R. Mendoza, "POINT OF SALE SYSTEM WITH INVENTORY FOR ARM'S FOOD AND DELICACIES," International Journal of Advanced Research in Computer Science, vol. 14, no. 1, p. 1371, 2019.
8. Proceedings of the 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, 2018.