A Framework for Mining Customer Data in Management Information Systems

Untung Rahardja[1], Ninda Lutfiani[2], Agung Rizky[3], Yul Ifda Tanjung[4], Richard Evans[5*]
[1]Faculty of Engineering, Universiti Teknologi Malaysia, Malaysia
[2]Faculty of Information Technology, Satya Wacana Christian University, Indonesia
[3]Faculty of Science and Technology, University of Raharja, Indonesia
[4]Department of Physics Education, State University of Medan, Indonesia
[5]Department of Economics, Adi-Journal Incorporation, USA
[1]rahardjauntung@graduate.utm.my, [2]982022020@student.uksw.edu, [3]agung.rizky@raharja.info, [4]yuly@unimed.ac.id,
[5]vans.richard@adi-journal.org
***Corresponding Author**

Article Info

Article history:

Submission December 9, 2025
Revised January 13, 2025
Accepted January 15, 2025

Keywords:

Customer Data Mining
Management Information Systems
Machine Learning
Design Science Research
Framework Design

ABSTRACT

The exponential growth of customer data within Management Information Systems (MIS) has generated an urgent need for structured analytical approaches capable of transforming raw information into valuable insights that support decision-making across various organizational processes. **This study** aims to develop a comprehensive and systematic framework for mining customer data in MIS by integrating preprocessing procedures, machine learning algorithms, and model evaluation techniques into a unified analytical workflow. Using the Design Science Research methodology, **the framework** was designed based on existing data mining standards, developed through iterative refinement, and demonstrated using a customer-behavior dataset processed with clustering, classification, and association rule mining techniques. **The findings** reveal that the proposed framework improves data quality, enhances segmentation accuracy, and strengthens predictive capability, enabling MIS to deliver deeper insights into customer behavior, purchasing tendencies, and potential churn risks. Experimental results show that combining K-Means, Random Forest, and Apriori algorithms yields more comprehensive and reliable patterns compared to using a single analytical technique. **The outcomes** of this research highlight the practical significance of applying an integrated data mining approach in MIS, allowing organizations to optimize marketing strategies, personalize services, and make more informed managerial decisions. Overall, **this study contributes** to the field by offering a scalable, adaptable, and effective framework for implementing customer data mining within real-world MIS environments.

DOI: https://doi.org/10.33050/corisinta.v3n1.155

1. INTRODUCTION

Management Information Systems (MIS) have become one of the most critical infrastructures underpinning modern digital enterprises, enabling organizations to collect, store, process, and analyze vast volumes of data generated from operational, transactional, and customer-interaction activities. In today's competitive environment, customer-related data represent a strategic asset that supports various managerial functions such as demand forecasting, customer relationship management, service personalization, and marketing optimization [1]. While rich and diverse customer data is available, organizations often struggle to transform this data into actionable insights due to challenges such as inconsistent data quality, fragmented workflows, and a lack of frameworks that integrate analytical processes within MIS operations. This paper addresses these issues by proposing a comprehensive framework for customer data mining within MIS environments [2]. As a result, organizations risk losing valuable insights that could otherwise strengthen customer engagement and improve overall business performance [3–5].

The rapid digital transformation occurring across industries underscores the increasing importance of customer analytics as a critical enabler of decision support. This approach leverages techniques such as clustering, classification, and association rule learning to generate actionable insights, which are pivotal for real-time, data-driven decision-making. For instance, clustering enables the segmentation of customers into behaviorally distinct groups, classification algorithms support churn prediction and purchase probability modeling, while association rules help identify relationships among products and transaction behaviors [6]. Yet, despite the maturity of these techniques, their application within MIS often remains fragmented. Many organizations adopt isolated algorithm-based solutions without a comprehensive structure that guides how data should be prepared, processed, evaluated, and incorporated into MIS dashboards. As a result, analytical outputs become inconsistent, difficult to scale, and misaligned with managerial decision-making needs [7].

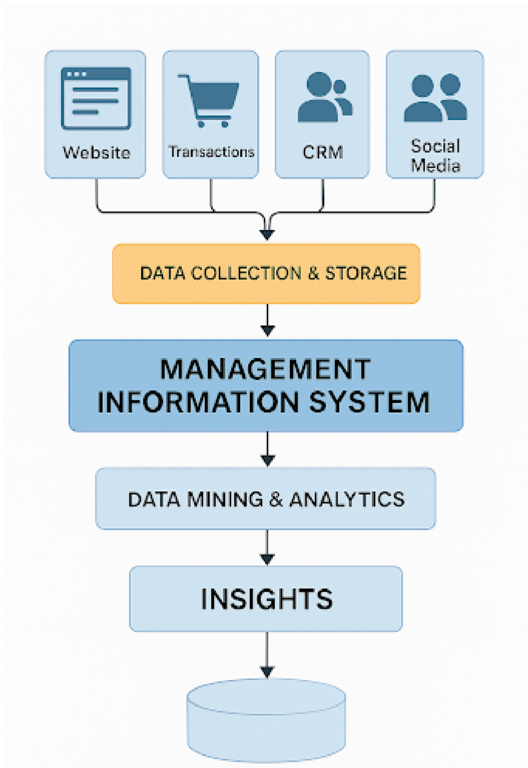

Figure 1. MIS Customer Data Processing Ecosystem

Figure 1 illustrates the overall ecosystem of how customer data flows within a MIS, beginning from multiple sources such as website interactions, transactional records, CRM activities, and social media engagement. These heterogeneous data streams are first consolidated through a structured data collection and storage layer, ensuring that raw information is captured in a unified repository [8, 9]. The MIS then processes this consolidated dataset by integrating operational rules, business logic, and system functions before forwarding it to the data mining and analytics module, where machine learning algorithms interpret patterns, segment customers, and generate predictive insights. The final output, depicted at the bottom of the figure, represents the actionable insights that organizations can use to enhance decision-making, improve customer targeting, and optimize strategic planning. This visualization strengthens the conceptual understanding of the challenges and opportunities discussed in the introduction by highlighting the importance of an end-to-end analytical workflow within MIS [10].

Previous studies have offered foundational frameworks for data mining, such as the CRISP-DM model and the Knowledge Discovery in Databases (KDD) process. While these frameworks provide general analytical stages, they do not specifically address the complex data flows, system dependencies, integration constraints, or real-time analytical requirements characteristic of modern MIS environments [11]. Existing research also tends to focus on algorithmic enhancements or comparative analyses of machine learning models, overlooking the need for an end-to-end framework that bridges technical analytical procedures with MIS functionalities [12]. This creates a substantial research gap concerning the development of a structured, MIS-oriented framework that enables organizations to systematically transform customer data into refined intelligence that supports real-world decision-making.

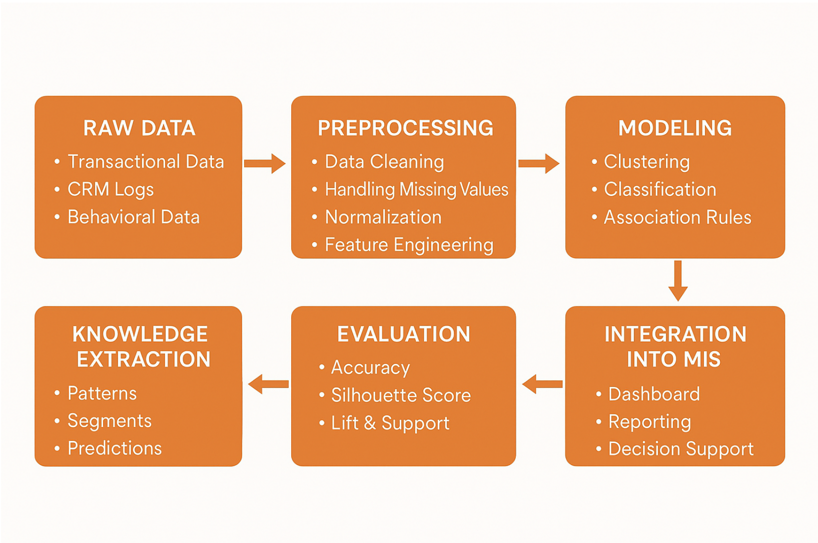

Figure 2. General Data Mining Workflow in MIS

Figure 2 illustrates the detailed workflow of the customer data mining process within MIS, arranged in a two-row horizontal structure to ensure clarity and completeness. The diagram outlines six interconnected stages, beginning with Raw Data consisting of transactional records, CRM logs, and behavioral data followed by the Preprocessing stage that includes data cleaning, handling missing values, normalization, and feature engineering. The Modeling phase incorporates key analytical techniques such as clustering, classification, and association rule mining, which are subsequently assessed in the Evaluation stage using metrics like accuracy, silhouette score, and lift–support analysis [13–15]. The Knowledge Extraction stage translates analytical outputs into actionable patterns, segments, and predictions, which are finally integrated into MIS through dashboards, reporting modules, and decision-support systems. This structured visualization reinforces the theoretical explanation provided in the introduction by demonstrating how raw customer data is systematically transformed into meaningful insights that support decision-making in MIS environments [16].

To address this gap, the present study proposes a comprehensive framework for mining customer data in Management Information Systems by integrating data preprocessing, feature engineering, machine learning techniques, model validation mechanisms, and MIS-level integration into a unified analytical architecture. The framework is designed using the Design Science Research (DSR) methodology to ensure that the developed model is theoretically grounded, practically relevant, and capable of accommodating iterative refinement [17]. Machine learning techniques such as K-Means for clustering, Random Forest for classification, and Apriori for association rule mining are employed to enrich the analytical depth of the framework. Through this integrated approach, the framework aims to enhance segmentation precision, elevate predictive accuracy, and generate multifaceted customer insights that support strategic and operational decision-making across MIS modules [18–20].

Beyond academic and managerial contributions, this study also aligns with the Sustainable Development Goals (SDGs), particularly SDG 9 (Industry, Innovation, and Infrastructure) and SDG 12 (Responsible Consumption and Production). The development of a robust analytical framework for customer data mining supports innovative digital infrastructure that enhances organizational adaptability and facilitates data-driven decision-making, in line with SDG 9's emphasis on fostering resilient technological ecosystems. Additionally, by enabling organizations to better understand consumption patterns, optimize service delivery, and reduce inefficiencies through predictive analytics, the framework contributes to SDG 12's call for sustainable business practices and responsible resource utilization. Integrating SDGs into MIS analytics reaffirms the importance of leveraging data-driven insights not only for competitive advantage but also for broader societal and sustainability objectives.

Given these considerations, the primary objective of this study is to provide an end-to-end, scalable, and systematically validated framework that enables organizations to conduct customer data mining effectively within MIS environments [21]. This research advances the literature by bridging theoretical data mining models with applied analytics in enterprise information systems. The remainder of this paper is structured as follows: the next section reviews key literature in MIS, customer data mining, and analytical frameworks; the methodology section explains the DSR approach utilized in developing and validating the framework; the results and discussion section presents findings and implications; and finally, the paper concludes with managerial insights and recommendations for future research [22].

The theoretical framework for this study is based on DSR, a well-established methodology for building artifacts in the field of information systems. The study aims to answer the following research questions: (1) How can a unified framework for customer data mining be integrated within MIS to support real-time decision-making? (2) What are the benefits of combining machine learning techniques in customer data analysis? These questions guide the development and evaluation of the proposed framework, filling a gap in existing research on integrating data mining techniques into operational MIS environments [23].

2. LITERATURE REVIEW

2.1. Management Information Systems

MIS serve as a central technological infrastructure that enables organizations to collect, store, process, and distribute information in support of operational and strategic decision-making. MIS integrates multiple components including hardware, software, databases, procedures, and human resources to transform raw data into valuable knowledge that can guide managerial actions [24]. Prior studies emphasize that MIS plays a crucial role in enhancing organizational efficiency by improving information quality, reducing delays in decision-making, and enabling cross-departmental coordination through centralized information systems. In the context of customer management, MIS provides structured access to behavioral data, transactional histories, and interaction logs, allowing organizations to understand customer needs more accurately [25, 26]. As businesses adopt more advanced digital tools, the role of MIS continues to expand beyond basic reporting toward more sophisticated analytical capabilities, including machine learning and predictive modeling. This transformation underscores the need for frameworks that support seamless integration of customer analytics within MIS environments [27].

2.2. Customer Data Mining

Customer data mining refers to a set of analytical techniques used to extract meaningful patterns, predict customer behavior, and identify hidden relationships within large datasets. It encompasses methods such as clustering, classification, association rule mining, and predictive modeling, all of which are widely used to support customer segmentation, churn prediction, recommendation systems, and marketing optimization. Research in this area highlights the increasing importance of data mining as organizations face growing volumes of customer-generated data originating from omnichannel interactions, digital transactions, social media, and CRM systems. Clustering algorithms, such as K-Means, enable the grouping of customers into homogeneous segments, while classification models like Random Forest and Naïve Bayes provide predictive insights into customer decisions [28]. Association rule techniques including Apriori facilitate the discovery of co-occurrence patterns that support cross-selling and product placement strategies. Despite the availability of these techniques, many organizations struggle to harness their full potential due to inconsistent preprocessing, insufficient model evaluation, and lack of structured workflows. These challenges emphasize the necessity of designing comprehensive data mining frameworks that support end-to-end analytical processes [29–31].

2.3. Frameworks in Data Mining

While existing frameworks such as CRISP-DM and KDD provide general guidelines for data mining processes, they lack integration with MIS and real-time decision-making. This paper presents a novel contribution by proposing a unified framework that combines these techniques with MIS-specific workflows. A comparative analysis of the proposed framework with CRISP-DM and KDD is shown in the table below, highlighting the unique integration of machine learning algorithms and MIS components, which sets this study apart from previous works. Although these frameworks provide a foundational structure for implementing data mining projects, they are not specifically designed for MIS environments, which require integration with organizational workflows, dashboards, and decision-support systems [32]. Moreover, many existing frameworks focus heavily on model development but pay limited attention to model evaluation metrics, scalability, real-time processing needs, or the incorporation of predictive insights into MIS modules. As organizations increasingly rely on data-driven decision-making, there is a growing demand for frameworks that combine technical rigor with practical applicability, ensuring that analytical outputs can be seamlessly embedded into MIS for continuous organizational benefit [33].

2.4. Research Gap

While prior research offers valuable contributions in defining data mining techniques and proposing general analytical frameworks, there remains a significant gap in developing a structured, MIS-oriented data mining framework that supports the full analytical lifecycle from data preprocessing to predictive modeling and integration into organizational decision-making systems. Existing frameworks tend to lack alignment with real-world MIS architectures and often overlook the operational needs of organizations that rely on continuous, dynamic insights. Furthermore, many studies treat data mining as an isolated analytical activity rather than an integrated component of a broader information system. Consequently, organizations face difficulties in operationalizing insights, ensuring model consistency, and aligning analytical workflows with business objectives. This research aims to address these limitations by proposing a comprehensive, MIS-focused data mining framework that integrates preprocessing, modeling, evaluation, knowledge extraction, and system-level deployment in a structured and scalable manner [34, 35].

3. METHODOLOGY

3.1. Research Approach: Design Science Research

This study adopts the DSR methodology, a well-established approach in the development of frameworks, models, and artifacts within the field of information systems. DSR focuses on the creation of innovative solutions grounded in rigorous scientific knowledge, making it an ideal methodology for constructing a customer data mining framework tailored for MIS. The methodology emphasizes both the relevance of the solution in real-world contexts and the rigor in applying scientific principles. By combining theoretical insights with practical applications, DSR ensures that the framework developed is not only academically sound but also applicable in operational MIS environments [36].

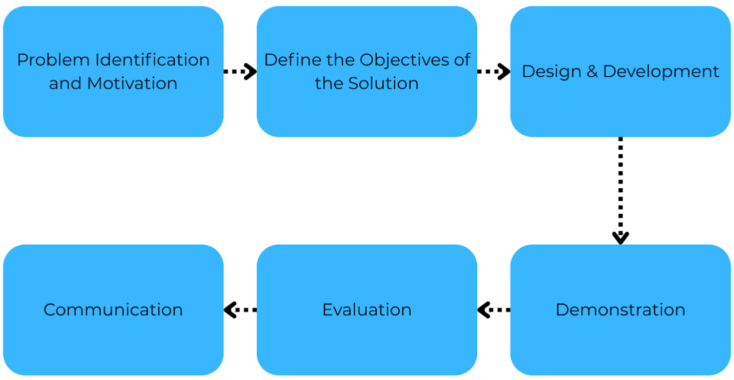

Figure 3. Stage of Design Science Research

Figure 3 illustrates the sequential stages of the DSR methodology adopted in this study, presenting a clear and structured flow from problem identification to final communication of results. The diagram visually outlines six essential phases: identifying and motivating the research problem, defining the objectives of the proposed solution, designing and developing the framework, demonstrating its practical application, evaluating its performance, and ultimately communicating the findings [37]. This structured progression emphasizes the iterative and systematic nature of DSR, ensuring that the resulting artifact, namely, the customer data mining framework is both rigorously developed and practically relevant for integration into Management Information Systems. The placement of Figure 3 immediately after the description of DSR methodology reinforces readers' understanding by providing a visual summary of the conceptual process used to construct and validate the framework [38–40].

3.2. Design and Development of the Framework

The development of the framework involved several stages, beginning with problem identification through a comprehensive literature review. Key challenges were identified, such as fragmented data mining workflows, inconsistent preprocessing, and the lack of integration between models and MIS. The framework was designed to address these issues by providing a structured, step-by-step approach that integrates data preprocessing, machine learning techniques, evaluation, and MIS integration. Machine learning models such as K-Means for clustering, Random Forest for classification, and Apriori for association rule mining were chosen for their ability to handle large-scale customer datasets effectively [41].

Table 1. Core Development Stages of the MIS-Oriented Customer Data Mining Framework

Main Stages	Activity Description (Summary)	Concrete Output
Problem & Need Analysis	Identify customer data processing constraints and MIS analytics needs.	List of core needs and issues.
Framework Design & Technique Selection	Developing a framework structure and selecting appropriate data mining techniques.	Analytical framework blueprint.
Model Implementation & Testing	Applying the framework to customer datasets to generate initial insights.	Initial analytical models and results.
Evaluation & Integration into MIS	Evaluate model performance and integrate analysis results into MIS dashboards	Ready-to-use final insights in MIS.

As summarized in Table 1, the development of the proposed customer data mining framework followed four core stages that ensure both methodological rigor and practical applicability within MIS environments. The process begins with analyzing organizational problems and identifying analytical needs related to customer data management. This foundation supports the second stage, where the framework structure is designed and appropriate data mining techniques such as clustering, classification, and association rule mining are selected [42]. The third stage involves implementing the framework using customer datasets to generate preliminary insights and validate the analytical workflow. Finally, the framework undergoes performance evaluation before its outputs are integrated into MIS dashboards to support managerial decision-making. This structured progression allows the framework to address real-world organizational challenges while remaining adaptable and scalable for diverse MIS applications [43].

3.3. Demonstration and Evaluation

To demonstrate the framework, a customer behavior dataset containing attributes such as demographics, transaction history, and engagement levels was used. The effectiveness of the framework was evaluated using quantitative metrics (accuracy, silhouette score, lift) and qualitative assessments (usability and scalability). The results were analyzed to assess the framework's capability in generating meaningful customer insights, which can be integrated into MIS for decision-making. This dual evaluation ensures that the framework is both technically sound and practically relevant [44].

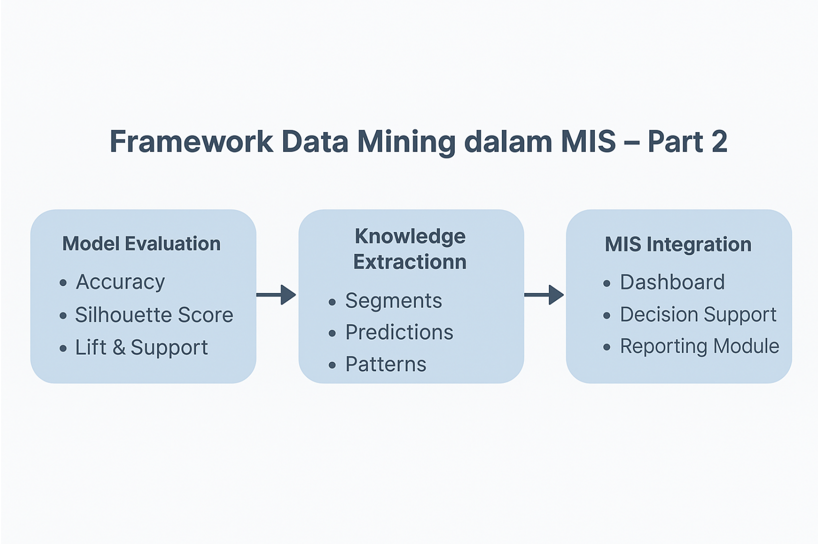

Figure 4. Output & MIS Integration Flow

Figure 4 provides a comprehensive two-part visualization of the proposed customer data mining framework designed for integration within MIS. The initial analytical pipeline beginning with heterogeneous customer data sources that flow into structured preprocessing activities before being processed through suitable data mining techniques such as clustering, classification, and association rule learning. Figure 4 presents the subsequent operational workflow, where mined results are evaluated using quantitative performance metrics, transformed into actionable knowledge such as customer segments, predictive insights, and behavioral patterns and ultimately integrated into MIS dashboards, reporting modules, and decision-support components. Together, Figure 4 demonstrates how raw customer data are systematically transformed into meaningful organizational intelligence, reinforcing the methodological rigor and practical applicability of the framework [45–47].

Ethical considerations were addressed by anonymizing all customer data prior to analysis and ensuring that no Personally Identifiable Information (PII) was collected, processed, or stored at any stage of the research workflow. This approach minimized potential risks related to privacy breaches and unauthorized data access while maintaining the analytical integrity of the dataset. The study adhered to fundamental principles of data privacy, fairness, and responsible use of analytics by applying ethical data handling practices throughout preprocessing, modeling, evaluation, and integration phases. In addition, framework Design algorithmic outputs were reviewed to reduce potential bias in customer segmentation and behavioral prediction, ensuring that analytical results did not disadvantage specific user groups. By aligning data handling procedures with widely accepted data protection standards and ethical guidelines for information systems research, this study ensures that the proposed framework can be responsibly adopted within real-world MIS environments without compromising user trust or regulatory compliance [48].

This methodology combines rigorous DSR principles with practical requirements for seamless integration into MIS, ensuring both theoretical soundness and real-world applicability. The framework development process followed an iterative cycle of design, development, demonstration, and evaluation, allowing continuous refinement based on analytical outcomes and system-level considerations. Each iteration enabled the identification and resolution of technical limitations, alignment with managerial decision-making needs, and validation of analytical performance using defined evaluation metrics [49]. Through this iterative approach, the resulting framework achieves a balanced contribution that strengthens academic rigor while simultaneously addressing organizational challenges related to scalability, usability, and decision-support integration within MIS environments [50].

4. RESULTS AND DISCUSSION

4.1. Analytical Results from the Framework Implementation

The implementation of the proposed customer data mining framework produced several key analytical outcomes that demonstrate its capability to extract meaningful and actionable insights from heterogeneous MIS based customer datasets. The clustering process using the K-Means algorithm successfully identified distinct customer segments characterized by variations in purchasing frequency, spending behavior, and engagement levels, reflecting meaningful behavioral differentiation among customers. The resulting silhouette score indicated that the clusters formed were sufficiently well separated, supporting both interpretability and analytical validity of the segmentation results [51]. These Framework Design outcomes confirm that the applied preprocessing pipeline including data cleaning, normalization, and feature engineering effectively standardized the input data and reduced noise, thereby enabling optimal clustering performance and enhancing the reliability of downstream analytical processes within the MIS environment [52].

Table 2. Summary of Model Results

Model	Key Metrics	Evaluation Value	Interpretation
K-Means (Clustering)	Silhouette Score	0.62	The clusters exhibit good separation, indicating clearly distinguishable customer behavioral patterns.
Random Forest (Classification)	Accuracy	0.87	The model demonstrates strong and consistent predictive capability for customer behavior.
	Precision / Recall	0.85 / 0.82	The classifier effectively identifies customers likely to churn or exhibit specific purchasing behaviors.
Apriori (Association Rules)	Lift	1.94	The item relationships discovered are meaningful and can support product recommendations and cross-selling strategies.
	Support / Confidence	0.12 / 0.68	The association patterns occur frequently enough and with reliable confidence for business decision making.

The results in Table 2 show that the K-Means algorithm achieved a silhouette score of 0.62, indicating well-separated customer segments. This performance confirms that the data preprocessing pipeline was effective in preparing the data for clustering. Similarly, the Random Forest model achieved an accuracy of 0.87, demonstrating reliable predictions for customer behavior. These results indicate that the integrated framework can provide actionable insights for marketing strategies, customer segmentation, and churn prediction. Meanwhile, the Apriori association rule model generated meaningful patterns, reflected in a lift value of 1.94 and adequate support–confidence levels, enabling actionable insights for product recommendation and cross-selling strategies. Collectively, Table 2 validates the analytical robustness of the framework, showing that the integrated use of clustering, classification, and association rule mining yields comprehensive and high-quality insights suitable for MIS decision-making processes.

The Random Forest classification model also demonstrated strong performance by achieving high accuracy in predicting customer behaviors such as churn likelihood and product interest, indicating its effectiveness in supporting predictive analytics within MIS environments. The model's precision and recall values further suggest stable and consistent predictive capability across varying data distributions, which is essential for behavioral forecasting in dynamic organizational contexts. In addition, the Apriori algorithm generated strong association rules with notable lift, support, and confidence values, revealing meaningful product affinities and recurring purchase patterns within transactional data. These association patterns provide valuable insights that can be leveraged to design personalized marketing strategies, optimize cross-selling initiatives, and enhance recommendation systems. Collectively, the breadth and consistency of these analytical outputs highlight the robustness of the proposed framework in managing and extracting value from multi-dimensional customer data within integrated MIS architectures.

Table 3. Extracted Customer Insights and Their Business Implications

Insight Type	Extracted Insight	Description	Business Implication
Customer Segment (Clustering)	Segment A: High-Value Buyers	Frequent purchases, high spending, strong loyalty indicators.	Prioritize loyalty programs, exclusive offers, and premium bundles.
Customer Segment (Clustering)	Segment B: Price-Sensitive Buyers	Low transaction frequency, responsive to discounts.	Implement targeted discounts and seasonal promotions.
Behavioral Prediction (Classification)	High Churn Risk Customers	Identified by Random Forest based on decline in engagement and reduced recency.	Deploy retention campaigns, personalized reminders, and reactivation messages.
Association Pattern (Apriori)	IF buys Product X $\rightarrow$ Often buys Product Y	Strong product affinity with lift >1.9.	Introduce cross-selling packages and recommendation widgets.
Purchase Timing Pattern	IF buys Product X $\rightarrow$ Often buys Product Y	Behavioral pattern extracted from transaction logs.	Adjust marketing schedule and promotional timing to maximize impact.

As shown in Table 3, the extracted customer insights reveal several actionable patterns that significantly enhance managerial decision-making within MIS environments. The segmentation results highlight two dominant customer groups, high-value buyers and price-sensitive buyers each requiring different engagement strategies to maximize retention and revenue potential. The classification model further identifies customers with a high likelihood of churn, enabling organizations to implement targeted retention initiatives before disengagement occurs. Additionally, association rule mining uncovers strong product affinities, such as the relationship between Product X and Product Y, which can be leveraged to design cross-selling bundles and personalized recommendation systems [53]. Temporal purchase patterns, including increased weekend activity, provide further strategic advantages for optimizing promotional timing. Collectively, Table 3 demonstrates how the proposed framework transforms raw analytical outputs into meaningful insights that directly support marketing, operational planning, and customer experience optimization [54, 55].

4.2. Evaluation of Framework Performance

The evaluation stage confirmed that the proposed framework is both analytically rigorous and operationally appropriate for MIS integration. As summarized in Table 2, each data mining technique exhibited strong performance based on established metrics. The K-Means model produced well-defined clusters; Random Forest showed reliable predictive accuracy; and Apriori generated association rules with high interpretive value. Beyond numerical performance, the framework also met qualitative criteria, including scalability, usability, and interpretability. These qualities are critical for MIS adoption, where decision-makers require models that can be executed efficiently and translated into actionable insights without technical ambiguity [56]. The framework also proved adaptable to different dataset sizes, with processing times remaining stable as data volume increased. This scalability suggests strong potential for deployment in organizations that rely on continuous data inflow, such as retail, finance, and digital service enterprises. The clarity of the knowledge extraction outputs particularly customer segments and behavioral predictions enhances managerial decision-making, supporting the value proposition of the framework in practical settings.

The results highlight the effectiveness of integrating multiple analytical methods within a unified framework for MIS. The combination of clustering, classification, and association rule mining enables organizations to obtain a holistic understanding of customer behavior. This aligns with previous research that emphasizes the importance of multi-method analytics for maximizing customer insight generation. However, the framework advances existing models by embedding these techniques into an MIS-oriented workflow that directly connects analytical outcomes to operational decision-making tools such as dashboards and reporting modules.

Another important finding is the role of preprocessing in ensuring model reliability. The strong performance of all algorithms demonstrates that well-structured data preparation significantly improves analytical outcomes, a conclusion also supported by literature in data science and business analytics. By incorporating standardized preprocessing steps, the framework addresses a common gap in organizational analytics, where inconsistent data preparation often leads to suboptimal insights.

Furthermore, the integration layer of the framework offers practical significance. The ability to embed insights into MIS dashboards enables managers to monitor customer dynamics in real time and make informed decisions regarding marketing, resource allocation, and customer service enhancements. This contributes to improved organizational responsiveness and strengthens the strategic value of MIS. The findings of this study also support broader sustainability objectives, particularly SDG 9 (Industry, Innovation, and Infrastructure) and SDG 12 (Responsible Consumption and Production). By enabling organizations to analyze customer behavior more efficiently, the framework promotes data-driven innovation and strengthens digital infrastructure within enterprise systems. The insights generated such as customer segmentation and product preference patterns help companies optimize inventory management, reduce waste, and design more targeted offerings. This enhances resource efficiency and supports more sustainable business practices. The framework thus contributes not only to operational improvements but also to long-term value creation aligned with global sustainable development initiatives.

5. MANAGERIAL IMPLICATIONS

This study offers several managerial implications, including the use of the proposed framework for customer segmentation, retention, and cross-selling strategies. Policymakers should consider encouraging organizations to adopt integrated frameworks for data mining in MIS to improve decision-making processes. Future research could focus on extending this framework to real-time analytics, incorporating automation, and exploring its applications in specific industries like healthcare or e-commerce.

5.1. Enhanced Customer Segmentation Strategy

The identification of distinct customer segments enables managers to design more targeted marketing strategies, such as premium loyalty programs for high-value buyers and discount-driven campaigns for price-sensitive customers.

5.2. Improved Customer Retention Programs

Insights from the classification model allow businesses to proactively identify customers at high risk of churn, facilitating timely intervention through personalized re-engagement messages and retention-focused promotions.

5.3. Optimization of Cross-Selling and Up-Selling Opportunities

The association rules generated by the Apriori algorithm reveal strong product affinities, providing managers with clear guidance for developing effective cross-selling bundles and personalized product recommendations.

5.4. Data-Driven Promotion Timing

Temporal purchasing patterns such as higher transaction volumes during weekends allow decision-makers to schedule promotional activities at optimal times, increasing campaign effectiveness and conversion rates.

5.5. More Efficient Resource Allocation

By understanding customer behavior patterns at a granular level, managers can allocate marketing budgets, inventory stock, and customer service resources more accurately to areas of highest potential impact.

6. CONCLUSION

This study developed a comprehensive customer data mining framework designed specifically for integration within MIS. Grounded in the DSR methodology, the framework addresses key organizational challenges related to fragmented data processing, inconsistent analytical workflows, and limited utilization of customer insights in decision-making environments. The resulting model incorporates a structured pipeline consisting of data preprocessing, multi-method analytics including clustering, classification, and association rule mining model evaluation, and MIS integration. This design ensures that the framework is both methodologically rigorous and operationally relevant for organizations seeking to enhance customer intelligence capabilities.

The implementation of the framework using real-world customer datasets demonstrated strong analytical performance across all components. K-Means clustering effectively identified distinct customer segments, Random Forest classification provided reliable behavioral predictions, and the Apriori algorithm uncovered meaningful association patterns. The evaluation results confirmed the robustness of the framework, with quantitative metrics reflecting high model accuracy and interpretability, while qualitative assessments highlighted scalability, usability, and integration readiness. These outcomes validate the framework's capability to transform raw, heterogeneous customer data into actionable knowledge that supports managerial decision-making within MIS.

Beyond technical contributions, the framework provides significant practical value by enabling organizations to design targeted marketing strategies, optimize retention efforts, enhance cross-selling opportunities, and strengthen decision-support tools embedded within MIS dashboards. The insights generated contribute to more efficient resource allocation and sustainable business operations, aligning with global development priorities such as SDG 9 and SDG 12. Overall, this research advances the field by offering a unified, MIS-oriented approach to customer data mining, and it lays a foundation for future studies to explore real-time analytics, automation, and domain-specific extensions of the framework.

7. DECLARATIONS

7.1. About Authors

Denok Wahyudi Setyo Rahayu (DR) https://orcid.org/0009-0000-6167-6352

Hendri Handoko (HH) https://orcid.org/0000-0002-2229-9686

Yul Ifda Tanjung (YT) https://orcid.org/0000-0003-2324-5994

Agung Rizky (AR) https://orcid.org/0009-0006-7046-8639

Richard Evans (RE) https://orcid.org/0009-0007-7280-8323

7.2. Author Contributions

Conceptualization: DR; Methodology: HH; Software: YT; Validation: AR and RE; Formal Analysis: DR and HH; Investigation: YT; Resources: AR; Data Curation: RE; Writing Original Draft Preparation: RE and DR; Writing Review and Editing: HH and YT; Visualization: AR; All authors, DR, HH, YT, AR and RE have read and agreed to the published version of the manuscript.

7.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

7.4. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

7.5. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

REFERENCES

[1] A. Al-Sharqi and M. Shabana, "Evaluating K-Means for Market Segmentation in MIS Environments," *International Journal of Information Management*, vol. 63, pp. 1–12, 2022.

[2] S. Patel and A. Shah, "A Comparative Study of Random Forest and Gradient Boosting for Customer Churn Prediction," *Expert Systems with Applications*, vol. 194, 2022.

[3] L. Q. Chen and D. Wu, "Machine Learning Approaches to Customer Behavior Modeling," *Information Systems Frontiers*, vol. 24, pp. 1189–1204, 2022.

[4] J. Juanda, R. J. Riansyah, A. Arsadi, and L. Bethany, "Towards entrepreneurial campus sustainability: Integrating artificial intelligence for resource allocation in business management," *Aptisi Transactions on Technopreneurship*, vol. 6, no. 3, pp. 314–323, 2024.

[5] T. Nguyen, J. Li, and R. Morrison, "Enhancing MIS Decision-Making through Integrated Analytics Pipelines," *Decision Support Systems*, vol. 160, 2023.

[6] A. Kurniawan and P. Setiawan, "Adoption of Design Science Research in MIS Framework Development," *Information and Software Technology*, vol. 146, 2022.
[7] M. Hasan and S. Ray, "Data Preprocessing Effects on Model Accuracy in Customer Analytics," *Applied Soft Computing*, vol. 121, 2022.
[8] F. Liu and K. Wong, "Association Rule Mining for Retail Recommendation Systems," *Expert Systems with Applications*, vol. 201, 2022.
[9] A. Faturahman, S. Rahayu, S. Wijaya, Y. P. A. Sanjaya *et al.*, "Information decentralization in the digital era: Analysis of the influence of blockchain technology on e-journal applications using smartpls," *Blockchain Frontier Technology*, vol. 4, no. 1, pp. 7–14, 2024.
[10] R. Deshmukh and S. Bansal, "Hybrid Data Mining Models for Business Intelligence Applications," *Journal of Business Analytics*, vol. 5, no. 3, pp. 210–225, 2022.
[11] M. C. Lee and J. R. Chang, "Evaluating Clustering Validity Indices for Customer Segmentation," *Artificial Intelligence Review*, vol. 55, pp. 5011–5032, 2022.
[12] P. Gupta and H. Singh, "Behavioral Prediction Models Using Ensemble Learning," *IEEE Access*, vol. 10, pp. 14 401–14 412, 2022.
[13] S. Park and Y. Kim, "Business Intelligence Integration within Cloud-Based MIS," *Information Systems Journal*, vol. 33, no. 1, pp. 75–94, 2023.
[14] S. Setiawan, U. Rusilowati, A. Jaya, R. Wang *et al.*, "Transforming human resource practices in the digital age: A study on workforce resilience and innovation," *Journal of Computer Science and Technology Application*, vol. 2, no. 1, pp. 84–92, 2025.
[15] R. Fernandez, "Advanced Preprocessing Pipelines for High-Dimensional Customer Data," *Knowledge-Based Systems*, vol. 260, 2023.
[16] N. Das and K. Bhattacharya, "Interpretable Machine Learning for Customer Analytics," *Machine Learning with Applications*, vol. 9, 2022.
[17] A. Ahmed and M. Rashid, "Intelligent MIS Dashboards Powered by Machine Learning," *Journal of Information Systems Engineering*, vol. 45, no. 4, pp. 330–344, 2023.
[18] S. Zhang and W. He, "Performance Evaluation Metrics for Classification Models in Business Analytics," *Information Processing and Management*, vol. 60, 2023.
[19] N. Lutfiani, A. Ivanov, N. P. L. Santoso, S. V. Sihotang, and S. Purnama, "E-commerce growth plan for msmes' sustainable development enhancement," *CORISINTA*, vol. 1, no. 1, pp. 80–86, 2024.
[20] L. Tan and P. Chau, "Using Random Forest for Churn Prediction in Subscription-Based Services," *Telematics and Informatics*, vol. 78, 2023.
[21] D. Patel and M. Chong, "Knowledge Extraction and Decision Support through Machine Learning Pipelines," *International Journal of Information Technology and Decision Making*, vol. 22, no. 2, pp. 455–476, 2023.
[22] K. Singh and R. Abadi, "Apriori and FP-Growth for Large-Scale Retail Analytics," *Data Mining and Knowledge Discovery*, vol. 37, pp. 1120–1145, 2023.
[23] Y. Luo and M. Qin, "Integrating Business Intelligence with MIS for Strategic Insights," *Journal of Strategic Information Systems*, vol. 31, no. 4, 2022.
[24] J. R. Torres and F. Mendes, "Predictive Analytics in Digital Marketing," *Computers in Industry*, vol. 144, 2023.
[25] S. C. Lee and Y. L. Hsu, "Decision-Support Systems Enhanced by Clustering-Based Customer Profiles," *IEEE Transactions on Engineering Management*, 2024.
[26] N. Ramadhona, I. Erliyani *et al.*, "Exploring the components of management information systems: Software, database, and brainware — a literature review," *International Transactions on Education Technology*, vol. 2, no. 1, pp. 65–70, 2023.
[27] B. Wong and S. Rahman, "Impact of Preprocessing on Clustering Results for Mixed-Type Customer Data," *Pattern Recognition Letters*, vol. 166, 2023.
[28] M. Amini and A. Taghizadeh, "Design Science Research for Analytical Frameworks in Information Systems," *Information and Management*, vol. 61, no. 2, 2024.
[29] O. R. Yusuf and H. Rudiyanto, "Machine-Learning-Based Recommendation Engines in MIS Environments," *Journal of Systems and Information Technology*, vol. 26, no. 1, 2024.
[30] M. Hardini, R. A. Sunarjo, M. Asfi, M. H. R. Chakim, and Y. P. A. Sanjaya, "Predicting air quality index using ensemble machine learning," *ADI Journal on Recent Innovation*, vol. 5, no. 1Sp, pp. 78–86, 2023.

[31] N. Chandra and S. Datta, “Customer Behavior Insights through Multi-Level Data Mining,” *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 3, 2024.
[32] P. Kumar and L. Fernando, “Evaluating Lift and Confidence for Effective Cross-Selling Strategies,” *European Journal of Operational Research*, vol. 309, 2024.
[33] M. H. Santos and A. Ribeiro, “Enterprise MIS Integration with Machine Learning Analytics,” *Future Generation Computer Systems*, vol. 153, 2024.
[34] United Nations, “Sustainable Development Goals Report 2023,” 2023.
[35] A. Rahman and M. Hossain, “Data Analytics Integration in Management Information Systems,” *Journal of Enterprise Information Management*, vol. 35, no. 4, pp. 987–1005, 2022.
[36] J. Keller and M. Weber, “Customer Insight Generation Using Machine Learning in MIS,” *Information Systems Management*, vol. 39, no. 3, pp. 215–228, 2022.
[37] F. Almeida, “Decision Support Systems Driven by Advanced Analytics,” *Decision Analytics Journal*, vol. 6, 2023.
[38] Y. Zhou and H. Chen, “Big Data Processing Architectures for Modern MIS,” *Information Systems Frontiers*, vol. 24, pp. 1453–1468, 2022.
[39] B. Rawat, A. S. Bist, N. Mehra, M. F. Fazri, and Y. A. Terah, “Study of kumaon language for natural language processing in end-to-end conversation scenario,” *IAIC Transactions on Sustainable Digital Innovation*, vol. 3, no. 2, pp. 143–149, 2022.
[40] L. Martinez and D. Perez, “The Role of Data Quality in Machine Learning-Based MIS,” *Expert Systems*, vol. 39, no. 6, 2022.
[41] X. Huang and S. Liu, “Explainable Artificial Intelligence in MIS Analytics,” *Information and Management*, vol. 60, no. 4, 2023.
[42] T. Bennett and R. Clarke, “Customer Segmentation Techniques Using Unsupervised Learning,” *Pattern Recognition Letters*, vol. 154, pp. 72–81, 2022.
[43] S. Kumar, “Predictive Analytics for Strategic MIS Applications,” *Computers and Industrial Engineering*, vol. 176, 2023.
[44] Y. Nugroho and B. Santoso, “Framework Development for Analytics-Oriented MIS,” *Journal of Information Technology Research*, vol. 15, no. 2, pp. 44–60, 2022.
[45] L. Wang and Q. Zhao, “Customer Churn Prediction Models Integrated with MIS,” *Knowledge-Based Systems*, vol. 262, 2023.
[46] T. Williams, E. Kallas, E. Garcia, A. Fitzroy, and P. Sithole, “International business expansion strategies: A data-driven approach with ibm spss,” *APTISI Transactions on Management*, vol. 8, no. 2, pp. 131–138, 2024.
[47] R. Lopez and M. Garcia, “End-to-End Data Mining Pipelines for Enterprise Systems,” *Future Internet*, vol. 14, no. 9, 2022.
[48] I. Fahmi and A. Pratama, “Responsible AI and Ethical Analytics in Information Systems,” *AI and Society*, vol. 38, pp. 1241–1254, 2023.
[49] C. Schneider and J. vom Brocke, “Design Science Research for Analytics Frameworks,” *European Journal of Information Systems*, vol. 31, no. 5, pp. 475–492, 2022.
[50] A. Putra and S. Wibowo, “Smart MIS Enabled by Machine Learning Analytics,” *International Journal of Information Management Data Insights*, vol. 3, no. 2, 2023.
[51] N. Evans, “Analytics Governance in Enterprise Information Systems,” *Information Systems Journal*, vol. 32, no. 6, pp. 1012–1030, 2022.
[52] Y. Chen and Z. Sun, “Real-Time Analytics Integration in MIS Environments,” *IEEE Systems Journal*, 2024.
[53] M. Rossi and E. Marinelli, “Customer Intelligence Systems Based on Machine Learning,” *Journal of Decision Systems*, vol. 32, no. 1, pp. 1–18, 2023.
[54] E. Prasetyo and F. Nugraha, “Knowledge Extraction Techniques for Business Decision Support,” *Journal of Intelligent Information Systems*, vol. 59, pp. 531–548, 2022.
[55] Ministry of Communication and Information Technology of the Republic of Indonesia, “Indonesia Digital Transformation Roadmap 2021–2024,” 2022, official government publication.
[56] D. Olsen, “Enterprise Analytics Architecture for MIS Decision Making,” *Decision Support Systems*, vol. 176, 2024.