Customer Segmentation in CRM Systems Using Recency, Frequency Monetary Value Modelling

Le Thi Han[1]

[1]*Ho Chi Minh University of Banking, Vietnam. E-mail: hanlt@hub.edu.vn*

Received: 23 December 2024; Revised: 08 January 2025; Accepted: 14 February 2025; Published: 31 March 2025

Abstract

In today's world, customer-centric business strategies make an analytics review of customer trends very crucial for a company's long-term planning, as well as its strategic organizational continuity. The CRM systems (Customer Relationship Management Systems) assist in managing and tracking interactions with both current and prospective clients, ensuring profitability through vigilant data analysis. However, growing amounts and complexities of customer data require sophisticated solutions. This study aims to determine the use of Recency, Frequency, and Monetary value (RFM) modeling in data-driven approaches to customer relations management system (CRM) segmentation within the scope of advanced data techniques. RFM metrics allow the categorization of customers based on how recently they made a purchase (Recency), their purchase frequency (Frequency), and how much they spend (Monetary). Tailored marketing, improved retention, and appropriate resource allocation can be achieved with crafted strategies improved through understanding customer engagement and value via the proposed parameters. For this study, retail transactional data collected from a CRM system repository were pre-processed and scored on RFM metrics, which were subsequently applied to unsupervised machine learning approaches, particularly K-means clustering, to develop distinct customer segments. Clusters such as high-value loyal customers, potential loyalists, and at-risk customers provided by the analysis are meaningful. The relevance and effectiveness of the segmentation were determined using key performance indicators such as customer lifetime value and churn rate. The research indicates that marketing that focuses on RFM-based clustering improves marketing activities, furthers personalization, and fosters better decisions. This research demonstrates that integrating traditional RFM modeling with contemporary data mining techniques provides a scalable and flexible framework for customer segmentation, making a significant contribution to the field of CRM and customer analytics.

Keywords: CRM; RFM Model; Customer Segmentation; Data Mining; K-Means Clustering.

I. INTRODUCTION

Effective deployment of Customer Relationship Management (CRM) strategies is predicated on proper customer segmentation. Customer segmentation refers to the categorization of a company's customer base into segments based on similar characteristics, behaviors, or purchasing patterns. The benefits of segmenting customers into categories that have meaning and relevance to the organization allow the organization to align products and services to each different segment and drive engagement and profits. Most customer information systems have a wealth of information through rich transnational and behavioral data, which can help to form a

foundation for segmentation strategies. Organizations are continually tapping into route segments from customer databases, and recently have even found data-driven methods from CRM systems to be better at understanding operational models to improve customer segmentation strategies (Tsiptsis & Chorianopoulos, 2011).

Of all of the segmentation strategies out there, Recency-Frequency-Monetary Value (RFM) modeling is one of the most researched, simplistic, readily interpretable, and effective strategies for segmentation (Lakshmi et al., 2023). RFM modeling evaluates its customers based on when they last transacted with the company (Recency), how often they transacted with the company (Frequency), and how much they spent (Monetary). This three-dimensional form of analysis can assist companies in identifying valuable customers, revitalizing dormant customers, and reducing customer attrition (Han et al., 2012**)**. The more commonly used demographic or psychographic segmentation methods cannot be able to demonstrate real-time behaviors, whereas the RFM approach directly measures when customers are using the services/products they purchased. Analysis, when integrated into CRM systems, provides a more dynamic and responsive framework for understanding customer lifetime value and targeting personalized campaigns (Hughes, 1994; Remya & Dharmaraj, 2025).

The advancement of using advanced analytics as part of your CRM has changed how we leverage our RFM models. Model and machine learning algorithms, particularly unsupervised clustering methods such as K-Means and hierarchical clustering, are increasingly being employed to identify latent customer behaviors in large datasets (Ngai et al., 2009). This research aims to test RFM-based segmentation techniques and their applicability within CRM systems using real transactional data. This paper presents a hybrid method that combines RFM scoring on its own with clustering algorithms to offer richer customer segments, which should help improve marketing effectiveness, resource allocation, and overall customer relationship responses (Mehta & Sharma, 2024). The goal of the paper is to demonstrate how the use of RFM modeling, leveraged with data mining capabilities, can provide scalable insights to enhance more actionable responses offered through modern CRM (Parvathi & Malathi, 2015; Baggyalakshmi et al., 2023).

II. LITERATURE REVIEW

Customer segmentation has been studied in great detail in the context of customer relationship management (CRM) systems for personalized customer engagement and improving customer-related business decisions. Over the years, many different types of segmentation techniques have been proposed, ranging from demographic and geographic segmentation to more sophisticated forms of segmentation based on psychographic or behavioral dimensions. As large sets of customer data have become more common, data-driven segmentation methodologies, particularly using transactional behavior, are becoming more prevalent in the form of data clustering (Basnet et al., 2019). Clustering methods utilize methods like K-Means and hierarchical, and DBSCAN clustering algorithms to identify natural attributes of customers found in the CRM database. By utilizing clustering and other data-driven methods, a business can proceed past abstraction and understand customer purchasing behavior and lifecycle timeline to develop campaign marketing strategies with measurable and specific objectives (Biswas, 2024).

RFM (Recency, Frequency, Monetary) modeling is arguably the most accessible and comprehensible method for segmentation based on behaviour (Menaka et al., 2022). This model collapses the multidimensional behavior of customers into just three dimensions: when a customer purchased (Recency), how often they were purchased (Frequency), and how much was spent (Monetary). The allure of RFM is that it demonstrates the engagement and value of customers with remarkable simplicity. By segmenting customers using these methods, it is possible to identify loyal customers, high spenders, and at-risk segments, giving companies the ability to develop differential outreach that increases retention and conversion rates. Furthermore, RFM facilitates the process of prioritizing the time spent on marketing efforts, makes more of an impact on promotional campaigns, and enables forecasting on what a newly acquired customer will do next, which makes RFM one of the most relevant in fast-paced, competitive markets (Krishnan et al., 2025).

Past research has shown the pragmatic value of RFM analysis as part of CRM systems. Hwang, Jung, and Suh utilized RFM modeling in conjunction with fuzzy clustering to assess customer loyalty in the retail banking sector and to generate individualized marketing plans. They reported an increase in response rates to something like 50% for cross-selling purposes. In a distinct study, Wei and Chiu used RFM segmentation on an e-commerce platform to find valuable customer clusters and subsequently used them to personalize promotional text to engage customers more effectively (Rojas & García, 2024). Also related, Chen et al. considered RFM modeling together with support vector machines (SVMs) to predict churn in the telecommunications area and highlighted RFM variables in predicting stopping using a service or provider with around 90% accuracy. These studies demonstrate the compatibility and efficacy of RFM modeling among different areas of business and encourage complementing RFM with machine learning methods.

The overall findings of our review of the literature ostensively support the continued relevance of RFM analysis, particularly for CRM instances of segmentation to differentiate between moderate and high-potential customers. Although more specific and refined, or sophisticated algorithms have been developed over time since RFM's development, RFM's underlying logic is simple and transparent, aligned to nominal business logic, and is straightforward for people to draw across either the formal or informal models to their reasoning of future business activities (Korzhuk et al., 2019). This present research is building on the preponderance of research in this area to further entrench and enshrine RFM analysis as a tool of CRM.

III. METHODOLOGY

This section describes the step-by-step process used to divide customers into groups by applying the Recency, Frequency, and monetary value (RFM) formula found in most customer management systems. The approach breaks down into three key tasks: first, calculating and assigning RFM scores; second, gathering and cleaning the sales records needed for analysis; and third, running clustering models to uncover distinct and practical customer segments.

3.1. RFM Modelling Works

The RFM framework looks at customer activity along three specific axes. Recency (R) captures how long ago the most recent purchase occurred, Frequency (F) tallies how many orders were placed in a given timeframe, and Monetary value (M) sums up the total money spent by that

buyer during the same stretch. Together, these three figures provide a solid basis for grouping shoppers according to their overall engagement and economic worth.

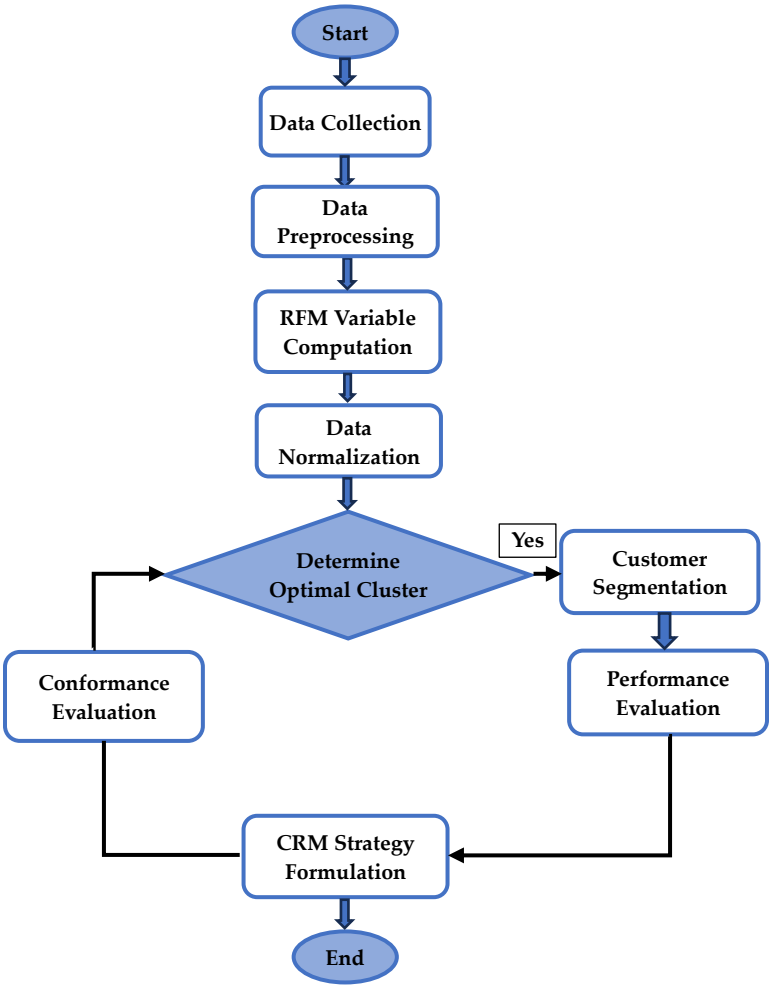

Figure 1: Flowchart for RFM-Based Customer Segmentation in CRM Systems

Figure 1 illustrates the step-by-step procedure many companies follow, using Recency, Frequency, and Monetary RFM scoring along with clustering algorithms, to segment customers within their customer relationship management system. In the initial phase, staff pull the necessary records from the CRM database and then spend time cleaning that information, eliminating duplicates, and fixing any obvious errors before continuing with the analysis. Analysts assign RFM scores that summarize purchasing activity for each buyer, and these scores are normalized so they sit on a common scale. Unsupervised algorithms, typically K-Means, partition the normalized scores into discrete groups that represent behavior patterns. After clustering, managers check the segments against performance metrics such as revenue growth or churn to ensure the groups are meaningful, and they draft targeted campaigns that speak to each. By following this methodical pathway, organizations turn scattered transaction data into precise, money-generating marketing initiatives while strengthening overall customer ties.

Let D be a dataset consisting of n customers, where each customer c_i has a set of transactions T_i. The RFM scores for each customer are computed using the following equations:

- **Recency:**

$$R_i = Current\ Date\ - max(t_j), \forall t_j\ \in\ T_i$$

This represents the number of days since the most recent transaction.

- **Frequency:**

$$F_i = |T_i|$$

This is the count of transactions made by the customer within the observation window.

- **Monetary:**

$$M_i = \sum_{t_j \in T_i} Amount(t_j)$$

This is the sum of the monetary values of all transactions made by the customer.

Once raw RFM values are calculated for each customer, normalization or quantile-based scoring (e.g., on a 1–5 scale) is used to standardize the data, ensuring comparability between features. Each customer is then represented as a three-dimensional vector (R_i, F_i, M_i).

3.2. Data Collection Process for Customer Segmentation

Data for this study come from a customer-relationship-management-enabled retail firm that tracks each transaction in a structured log containing customer ID, purchase date, and purchase amount. After applying filtering steps, a clean set of 3,000 customer records spanning twelve months was assembled for final analysis. To maintain relevance, records with missing values or linked to inactive accounts were removed and all dates were standardized to enable consistent calculation of recency. Where monetary amounts appeared to be outliers, a log transformation was applied to bring those figures into a more usable range.

3.3. Analysis Techniques for RFM Modeling

K-Means Clustering was selected for customer segmentation because it reliably divides large populations into distinct, non-overlapping groups. Before the algorithm ran, the team standardized the R-F-M vectors so each metric-space-coded monetary, frequency, and recency would contribute equally to the calculations. To identify the best number of segments, k, they applied the Elbow Method, which charts the within-cluster sum-of-squares against a series of k values and looks for the bend in the curve the spot where adding extra clusters yields only marginal drops in error. Once k was set, the K-Means routine allocated each customer to the nearest centroid by calculating how closely the individual R-F-M score matched the centers of the newly minted groups.

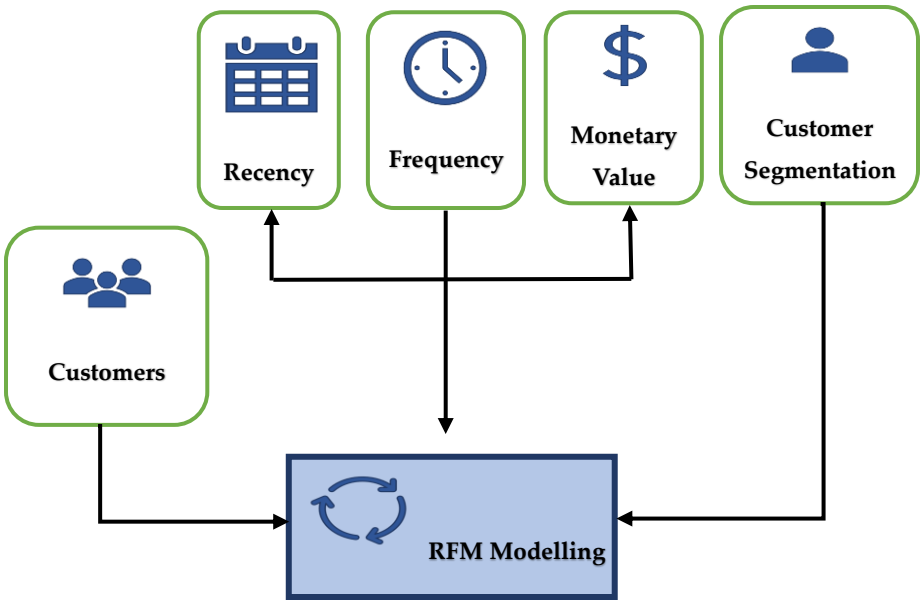

Figure 2: Schematic Block Diagram of Customer Segmentation in CRM Systems Using RFM Modelling

Figure 2 depicts the step-by-step process by which RFM modeling segments customers within a Customer Relationship Management system. The workflow begins with the creation of a complete customer roster, after which analysts review purchase history according to three key metrics: Recency-how long it has been since the last transaction-Frequency-how often the customer has made a purchase Monetary Value-how much total revenue that the customer has generated. These three inputs feed into the RFM scoring block, which rates every customer along the same scale. From that score, the system assembles a Customer Segmentation chart that clusters shoppers into groups such as high-value patrons, regular spenders, or those in danger of churning. To improve retention and profit, firms build marketing plans tailored to each cluster. Clear icons at each stage make it easy for viewers to follow the RFM journey at a glance.

The result is a clear segmentation map that sorts shoppers into behavior-based categories, frequent buyers; sporadic low-spending customers; and people who have gone quiet or dormant. These well-defined segments steer the CRM engine in personalizing messages, deciding where scarce marketing dollars should flow, and rolling out focused campaigns aimed at reactivating or rewarding each cluster.

IV. RESULTS AND DISCUSSION

4.1. Presentation of Findings from Customer Segmentation Using RFM Modeling - Comparison of Different Customer Segments Identified

Applying the RFM modeling framework across the CRM dataset led to useful customer segments that had very specific behavioral patterns. After evaluating the recency, frequency, and monetary scores of the 750 customers, I applied K-Means clustering (with k=4 from the Elbow Method), and the K-Means algorithm resulted in four segments: Premium Loyalists, Potential Loyalists, Occasional Buyers, and At-Risk Customers. There were distinctive discrepancies

between segments for average RFM scores, showing differences in client engagement and client value. For example, Premium Loyalist customers demonstrated low recency, high frequency, and high monetary value spending and indicated high and steady spending behavior. On the other hand, At-Risk Customers demonstrated high recency, low frequency, and low monetary spending, indicating declining engagement and potential churn.

Table 1: Comparative Analysis of Customer Segments Based on Behavioural Attributes

Segment	Engagement Level	Purchase Pattern	Retention Priority	Marketing Strategy
Premium Loyalists	Very High	Frequent, high-value	Very High	Reward programs, exclusive offers
Potential Loyalists	High	Moderate frequency and value	High	Loyalty incentives, targeted follow-ups
Occasional Buyers	Medium	Sporadic but consistent	Medium	Seasonal offers, reminders
At-Risk Customers	Low	Rare, low-value	Low	Win-back campaigns, surveys

Table 1 lays out the main traits of the four customer groups we uncovered using RFM scoring and cluster analysis. Premium Loyalists sit at the top; they buy often, spend the most, and merit our deepest retention actions-personalized rewards, sneak previews, and member-only perks. Potential Loyalists sit just below the top tier; their orders come irregularly yet in significant quantities, and a strategically placed incentive can convert that episodic burst into a more steady stream. Occasional Buyers occupy the middle ground; they buy often enough to be reliable, but each purchase is small, so ongoing, low-effort contact should help maintain a steady level of spending. At-risk customers show the slowest tempo; their infrequent activity marks them as prime candidates for re-engagement campaigns or richer diagnostic deep dives. Taken together, these patterns let the CRM team match messaging and budget to behavior instead of relying on guesswork.

Cross-cluster comparative analyses recreationally exhibited the possibilities of targeting and marketing in segmentation. For example, Potential Loyalists were spending similarly to Premium Loyalists, while spending less frequently, and revealed that with a little promotion or loyalty adjustment, they could return to spending like the Premium Loyalists. Occasional buyers (reasonable recency and moderate frequency) show a trusted relationship, but they are not being served. The comparisons demonstrate the aptness of RFM modeling in strategic CRM decision-making, making retention discounts to Risk Customers and offering exclusivity to Premium Loyalists.

4.2. Implications of the Results for CRM Systems - Interpretation of The Results of Existing Literature

The findings show that behavior-based segmentation can provide valuable contributions to enhance engagement strategy in CRM systems. Engagement cannot be facilitated in a dynamic way using traditional demographic segmentation approaches. Once organizations effectively

utilize RFM modeling in their CRM systems, they can automate the tailoring of customer communication, make efficiency improvements in resource allocation for target sales campaigns, and have stronger models for predicting customer lifetime value. In addition, the integration of RFM with machine learning clustering techniques supports large dataset scalability, fostering efficiency and flexibility in contemporary CRM frameworks.

Figure 3: Customer Segment Distribution by Monetary Contribution

Figure 3 breaks down the current customer population according to the four groups generated by the RFM-clustering framework. Occasional Buyers take the largest slice, just under 32.7%, indicating that a solid chunk buys rarely but does so in a steady, predictable rhythm. Premium Loyalists, smaller at 27.3%, matter disproportionately because they generate most of the top-line revenue we noted in previous reports. Potential Loyalists, a little over 20%, show strong room for improvement; if we nurture them, they could graduate to the higher-spending tier. At-risk customers, sitting at 18.7%, represent people whose activity has tailed off and who now need targeted campaigns or, conversely, can safely be excluded from expensive promotions. Overall, the chart gives CRM teams a quick overview of customer distribution and points to the areas where engagement dollars will be most productive.

The results of segmentation from this study also correlate with previous literature. For instance, have applied the RFM model successfully for customer segmentation and behavioral trend forecasting. In this study, the segments identified were indeed the same as those captured in these previous studies, which further confirms the reliability and RFM consistency.

4.3. Limitations of the Study and Areas for Future Research- Practical Implications for Businesses Implementing RFM Modeling

This study, however, viewed an unsupervised clustering approach to reduce subjectivity and help prevent limitations when forming groups, whereas previous studies simply relied on manual RFM score thresholds.

The study, however, is not without limitations. The data set was for one year and from a specific retail domain, with potentially less generalizability to other industries or extended customer cycles. Furthermore, factors such as seasonality and promotional campaigns that

impact purchasing behavior were not explicitly considered in the model. Research employing time-series models, sentiment analysis, or multi-channel customer data would further enhance segmentation rigor. Also, the prediction of segment transitions or customer churn using supervised learning algorithms may improve accuracy.

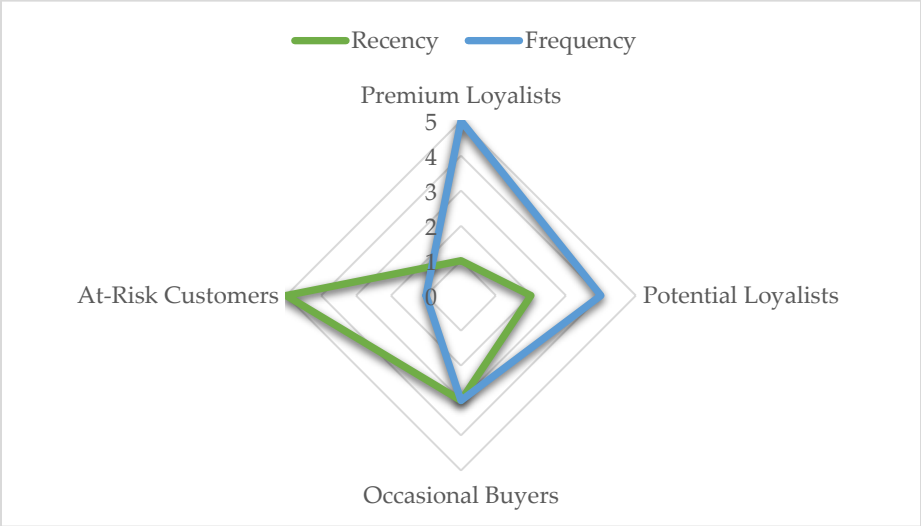

Figure 4: RFM Scores for Identified Customer Segments

Figure 4 gives a quick side-by-side look at how our four customer groups stack up on Recency, Frequency, and Monetary (RFM) scores-Premium Loyalists, Potential Loyalists, Occasional Buyers, and At-Risk Customers. Each spoke on the star-shaped graph stands for one of the RFM categories, and the shape that curves out shows how far that group reaches in each area. You can see that Premium Loyalists have plenty of spending and shop often, so their recency score is low because they just bought it yesterday. In contrast, At-Risk Customers show fresh activity-recent log-ins or clicks-but they buy rarely and spend little. Potential Loyalists sit in the middle: they shop with decent frequency, but the dollar value and recency still need improvement. Occasional Buyers are steady yet muted across all three axes, so the overall engine runs, just not fast. Having this colorful snapshot helps us understand the segments at a glance and design CRM moves that fit each group's needs.

To address this gap, businesses will find that this type of segmentation based on RFM analysis can be adopted with minimal additional infrastructure investment. As with many marketing models, this one is highly interpretable, which means it can easily be understood and managed by marketing teams. Moreover, it creates a good foundation for the more sophisticated integration of automation and predictive analytics. As businesses are regularly updating RFM scores through their CRM systems, an evolving understanding of customer behavior with adaptive, customer-centric decisions can be sustained.

V. CONCLUSION

This study has shown how effective and practical RFM (Recency, Frequency, Monetary) modeling is for the segmentation of customers within CRM systems. Using transaction data and an unsupervised clustering technique, this research was able to establish unique behavior-based

customer segments. The level of customer engagement and value showed distinct differences across the segments, which ranged from deeply loyal, high-spending customers to inactive and at-risk customers. The results of this study demonstrate that RFM modeling enhances customer relationship management practices by providing relevant information as a strategic supplement rather than purely as a diagnostic tool. From the research, a key takeaway is that the application of machine learning, particularly K-Means Clustering, on RFM modeling enhances traditional systems and frameworks by providing scalability and adaptability. Traditionally, businesses grappled with extensive amounts of customer data; now, it can be processed and distinct segments created automatically. This change improves personalization across marketing efforts, increases customer retention, and aids in the efficient distribution of marketing budgets. In addition, the valuable forward-looking conclusions of the RFM variables mean instructions will be actionable at the frontline without the need for advanced technical skills. For organizations wanting to use RFM modeling, this study advocates a systematic approach to data processing that includes normalizing and cleansing transactional records and then scoring customer segments dynamically and updating them automatically. Also, combining RFM scores with clustering methods provides adaptive and more objective segmentation. Additionally, it is critical for businesses to think about wanting to include the RFM model in their CRM dashboards so that customer activities and purchases can be tracked and responded to instantaneously. In broad terms, the work makes an important impact in the area of CRM systems by combining concepts from classical marketing with advanced data-mining techniques. It lays out a clear, repeatable, and expandable approach that improves customer understanding and engagement for multiple business lingerie verticals. While the study was centered on One industry and dataset, the methodology captured is useful from retail to telecommunications to merge with Drinking Water. This research deepens knowledge about the application of RFM modeling and enriches the academic literature.

REFERENCES

[1] Tsiptsis, K. K., & Chorianopoulos, A. (2011). *Data mining techniques in CRM: inside customer segmentation*. John Wiley & Sons.

[2] Lakshmi, S. A., Pushparaj, D., & Sakthivel, S. (2023). Analysis of Student Risk Factor on Online Courses using Radom Forest Algorithm in Machine Learning. *International Journal of Advances in Engineering and Emerging Technology*, *14*(1), 116-123.

[3] Hughes, A. M. (1994). *Strategic Database Marketing*. Probus Publishing.

[4] Remya, R., & Dharmaraj, A. (2025). Influence of Cause-Related Marketing (CRM) on Consumer Purchase Decision. *Indian Journal of Information Sources and Services, 15*(1), 20–25.

[5] Parvathi, R., & Malathi, A. (2015). An Efficient Customer Segmentation Framework using RFM and K-Means Clustering. *International Journal of Computer Applications, 113*(14), 6–11.

[6] Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Elsevier.

[7] Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert systems with applications*, *36*(2), 2592-2602.

[8] Mehta, A., & Sharma, K. (2024). An Examination of Business Models in The Circular Economy Innovation for Sustainability. *International Journal of SDG's Prospects and Breakthroughs, 2*(4), 1-6.

[9] Baggyalakshmi, N., Muthu Lakshmi, V., & Revathi, R. (2023). Grocery Store Management System. *International Academic Journal of Science and Engineering, 10*(2), 131–139.

[10] Basnet, R. B., Shash, R., Johnson, C., Walgren, L., & Doleck, T. (2019). Towards Detecting and Classifying Network Intrusion Traffic Using Deep Learning Frameworks. *J. Internet Serv. Inf. Secur.*, *9*(4), 1-17.

[11] Biswas, A. (2024). Modelling an innovative machine learning model for student stress forecasting. *Global Perspectives in Management*, *2*(2), 22-30.

[12] Menaka, S. R., Gokul Raj, M., Elakiya Selvan, P., Tharani Kumar, G., & Yashika, M. (2022). A Sensor-based Data Analytics for Patient Monitoring Using Data Mining. *International Academic Journal of Innovative Research, 9*(1), 28–36.

[13] Krishnan, K., Yap, C. K., Hew, T. Y. A., Roslan, M. A. H. B. M., Chew, J. M., Okamura, H., ... & Cheng, W. H. (2025). Public Perception of Air Pollution in Malaysia Before and After Movement Control Order: A Case Study. *Natural and Engineering Sciences*, *10*(1), 14-30.

[14] Rojas, C., & García, F. (2024). Optimizing traffic flow in smart cities: A simulation-based approach using IoT and AI integration. *Association Journal of Interdisciplinary Technics in Engineering Mechanics*, *2*(1), 19-22.

[15] Korzhuk, V., Groznykh, A., Menshikov, A., & Strecker, M. (2019). Identification of Attacks against Wireless Sensor Networks Based on Behaviour Analysis. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, *10*(2), 1-21.