

ISSN 1840-4855
e-ISSN 2233-0046

Original scientific article
<http://dx.doi.org/10.70102/afts.2026.1836.1070>

APPLICATION OF NEURO-SYMBOLIC KNOWLEDGE MINING FOR CUSTOMER BEHAVIOR ANALYSIS AND STRATEGIC CRM DECISION-MAKING

Dr. Reem Abdalla^{1*}, H. Niroshini Infantia²

^{1*}Assistant Professor, Head of BSBI Program, College of Administrative and Financial Sciences, University of Technology Bahrain, Manama, Bahrain.

e-mail: r.abdalla@utb.edu.bh, orcid: <https://orcid.org/0009-0004-2881-5372>

²Associate Professor, Department of Information Technology, St. Joseph Institute of Technology, Chennai, Tamil Nadu, India. e-mail: niroshiniinfantiah@gmail.com, orcid: <https://orcid.org/0000-0001-7438-1713>

Received: June 19, 2026; Revised: July 29, 2026; Accepted: September 07, 2026; Published: September 29, 2026

SUMMARY

Rapid development of digital platforms has led to an abundance of data collected from customers, necessitating the use of intelligent methods that can be used to analyze behavior patterns and aid in making effective Customer Relationship Management (CRM) decisions. Current CRM artificial intelligence systems mainly aim at high prediction accuracy and rarely provide features like interpretable analysis, contextual reasoning, and interpretation. To overcome this challenge, the current research introduces the NeuroCRM-KM (Neuro-Symbolic Knowledge Mining) approach that can be used to analyze customers' behavior patterns and aid in making decisions through neural learning, knowledge graph representation, and symbolic reasoning. In particular, the NeuroCRM-KM method involves collecting heterogeneous customer information such as transactions, demographics, customer reviews, and interactions, and then proceeding with data preprocessing, behavioral analysis using statistics, feature learning using neural networks, knowledge graph creation, and symbolic rule-based reasoning. The experimental evaluation performed on 5,000 customers' data shows that the suggested NeuroCRM-KM framework provides better results than those obtained with the help of machine learning and deep learning frameworks by achieving 96.4% accuracy, 95.8% F1 score, and 0.97 AUC. In addition, the ablation study has proven the significance of symbolic reasoning, knowledge graph integration, and neural learning elements for predicting performance and making it more interpretable. Finally, the result of the research is represented by the intelligent and explainable CRM decision-making system that allows for personal marketing, customer retention, and business strategy planning.

Key words: *neuro-symbolic AI, knowledge mining, customer behavior analysis, strategic CRM, explainable artificial intelligence, customer segmentation, decision support systems.*

INTRODUCTION

Recent developments in the field of AI, cloud computing, and intelligent analytics have made it possible to use data-driven methods for enhancing marketing processes and CRM practices through large data processing and predictive analysis [1]. Neuro-symbolic AI methods have proven to be efficient tools for combining logical reasoning with statistics, contributing to intelligent CRM and decision-making in

enterprises [2]. Machine learning algorithms used in e-commerce showed great results in predicting customer behavior, making personalized recommendations, and optimizing marketing strategies, and at the same time raised problems associated with adaptability and interpretability [3]. The development of AI agents and agentic systems became another milestone in terms of expanding autonomous decision-making abilities [4].

There are various methodologies that are used in neuro-symbolic knowledge mining such as neuro-symbolic data warehousing. There are efficient ways of integrating the knowledge representation framework with AI algorithms to facilitate better discovery and management of enterprise-level information. The digital transformation models have stressed the need to employ intelligent systems for facilitating organizational knowledge, organizational adaptation, and intelligent business operations [6]. Human-machine learning models have been found to add value in incorporating human knowledge into computational intelligence to facilitate better reasoning capabilities and collaborative decision-making [7]. Neuro-symbolic AI fusion is found to be promising in intelligent decision orchestration through the integration of diverse data sources, context knowledge, and intelligent reasoning for consumer-based decisions [5][8]. Multimodal artificial intelligence has played a crucial role in managing complex and unstructured consumer data through pattern recognition from diverse information sources for analytical purposes [9]. Artificial Intelligence (AI)-driven consumer behavior analysis has stressed the need for unbiased decision-making, ethical intelligence, and intelligent analytics in marketing and customer relationship management (CRM) [10]. These developments lay the groundwork for the application of neuro-symbolic knowledge mining in analyzing consumer behaviors and patterns to support strategic CRM decisions.

Key Contribution

- An approach that can be termed neuro-symbolic is proposed, which integrates the features of neural learning and symbolic reasoning into one framework that is used for examining customer behavior patterns to derive intelligence for CRM.
- The suggested model utilizes heterogeneous information about customers, comprising their behavior patterns, transaction records, interaction data, and contextual information to unearth hidden relationships and customer preferences. This is better than existing analytics models because it improves customer segmentation, prediction, and strategy formulation.
- This framework offers the capability of explaining decision-making by logically explaining the reasons for customer predictions and recommendations. This enhances the transparency, credibility, and usability of AI-based CRM systems by allowing companies to know what drives customers' actions and decisions.

The rest of this research paper is structured in the following way. In Section II, provide the literature review, where discuss existing AI-driven CRM frameworks, techniques for predictive analytics, customer behavior analysis, and limitations of traditional data-driven models that necessitate the development of neuro-symbolic knowledge mining approaches. In Section III, present the proposed NeuroCRM-KM (Neuro-Symbolic Knowledge Mining) framework and its components: the overall framework architecture, customer data gathering mechanism, statistical behavioral analysis component, neural customer representation learning, knowledge graph construction, and CRM decision engine based on symbolic reasoning. In Section IV, explain the methodology of the experiments: the dataset used for experiments, statistical analysis configuration, model configuration, evaluation metrics, and statistical validation procedures used in experiments. In Section V, discuss the obtained results, including customer statistical analysis, reliability and factor analysis, correlation and regression analysis, customer segmentation, comparison of the model performance, explainable CRM decision analysis, and ablation studies results. In Section VI, discuss theoretical and practical implications of the proposed framework and explain how the work contributes to explainable CRM intelligence. Finally, Section VII contains conclusions of this research paper.

LITERATURE REVIEW

The use of CRM systems and business intelligence platforms using AI technology has helped companies evaluate market risks, analyze customer information, and facilitate informed decision-making in digital environments [11]. Various studies have shown that predictive analysis can help detect customer behavioral patterns, predict their preferences, and implement proactive marketing [12]. Data-driven CRM has underlined the necessity to extract knowledge, analyze customer information, and make analytical decisions for better enterprise performance and customer interaction [13]. Investigations of CRM and data-driven decision making have shown that advanced analytics can help turn customer information into informed business strategies [14]. Business analytics and decision science have been used successfully to integrate multiple sources of information and make strategic decisions of an organization [15]. Models oriented towards customer behavior-related data analysis have already revealed the potential of cloud-based technologies for strategic analysis and optimal customer-oriented operations [16]. Approaches related to marketing analytics have successfully connected customer psychology, behavioral aspects, and intelligent decisions in marketing by using analytical models [17]. Data-driven decision-making support tools based on transparency and explainability have contributed to better customer development through transparent and understandable insights [18]. Predictive models powered by artificial intelligence have contributed to strategic decision-making through automated pattern discovery, forecasting, and optimization of complicated business processes [19]. AI-based marketing frameworks have additionally proven the significance of intelligent technologies for future-oriented strategies of customer analysis and decision-making [20]. These studies provide a basis for designing neuro-symbolic frameworks of knowledge mining combining pattern recognition through machine learning with symbolic reasoning for achieving explainable customer behavior analysis and strategic CRM decision making.

Research Gap

The existing customer behavior analysis methods and CRM decision-making methods have been greatly enhanced through the progress made in artificial intelligence, machine learning, and predictive analytics. Nonetheless, most of the currently prevailing methods do not have any consideration of knowledge representation and reasoning aspects in their implementation. Such methods are largely regarded as black boxes due to the difficulties in explaining the reasons behind customer segmentation, preference prediction, and the recommendation of strategies. Moreover, the conventional CRM models encounter problems in terms of integrating different types of customer information sources, business domain knowledge, and behavioral changes into an analytical system. Despite the improvements made by explainable AI in terms of providing model interpretability, such methods fail to combine statistical learning with reasoning. It is clear that there is a research gap in terms of designing a neuro-symbolic knowledge mining approach that integrates neural intelligence and symbolic reasoning.

Problem Statement

Organizations in today's era produce huge amounts of information about customers from different digital platforms; however, transforming the gathered information into valuable strategic insights is a difficult job. Current CRM solutions and models of AI-based analytics yield correct forecasts but are not able to identify the reasons behind the behavior of customers and their recommendations. Lack of reasoning intelligence makes current CRM decision support systems untrustworthy, inflexible, and impractical to use, especially when working in dynamically changing market conditions. Moreover, conventional methodologies are unable to find concealed connections between customer preferences, interactions, buying patterns, and contextual factors. Therefore, it is necessary to create a neuro-symbolic knowledge mining methodology, which unites behavioral analysis and symbolic reasoning to provide explainable insights and better decision-making in strategic CRM.

PROPOSED NEURO-SYMBOLIC KNOWLEDGE MINING FRAMEWORK

Overall Framework Architecture

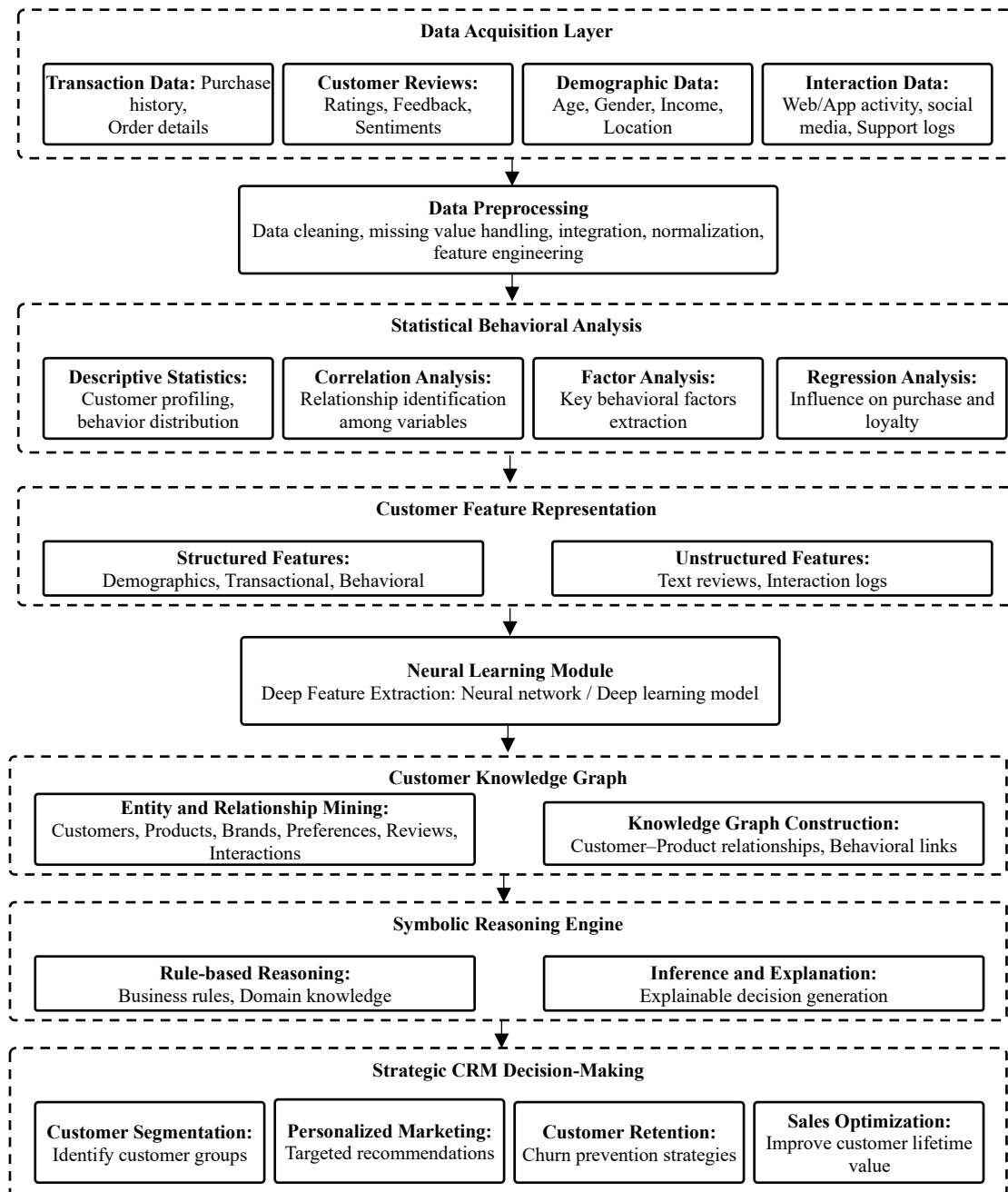


Figure 1. Overall framework architecture

Figure 1 presents the Neuro-Symbolic Knowledge Mining Architecture for Strategic CRM Decision-Making, which is a comprehensive structure that integrates the use of statistical analysis, deep learning, knowledge representation, and explainable reasoning to enhance the comprehension of customer behavior. First, heterogeneous customer data, which include the customer's transaction data, reviews, demographics, and interaction data, are collected through data preprocessing approaches such as data cleaning, data integration, normalization, and feature engineering. Once the data is preprocessed, it is subjected to statistical behavioral analysis by means of descriptive statistics, correlation analysis, factor analysis, and regression analysis to determine the most important customer behavior and factors influencing it. Next, the neural learning component executes deep feature extraction, which is responsible for representing customer patterns that are highly non-linear and complex. Such extracted

features are utilized within a customer knowledge graph through the detection of entities and relations between customers, products, preferences, reviews, and customer interactions. The symbolic reasoning engine uses business rules and domain knowledge in order to produce explainable inferences and recommendations for decisions. Finally, the proposed framework is able to make CRM-related decisions effectively by providing customer segmentation, personalized marketing, customer retention, and sales optimization. In total, the provided architecture represents an intelligent and interpretable CRM system.

Customer Data Acquisition

The data collection of customers is the first step of the proposed neuro-symbolic knowledge mining process that involves the accumulation of heterogeneous data about customers to understand customer behavior. Organizational CRM databases, online transactions, customer feedback, and social media communications become the main sources of collecting data about the customers. CRM databases help to gather historical data about customers, their previous interactions with the organization, services, and relationships. On the other hand, online transactions help in identifying the purchasing behaviors of the customers, for example, product preferences, purchasing frequency, transaction amounts, and buying trends. Feedback obtained through reviews, ratings, and surveys can be helpful in analyzing customer satisfaction, opinions, and sentiments of the customers. Furthermore, social media communications and online engagement data are used to assess customer interests, communication behaviors, and level of interactions. The collected dataset has multi-dimensional attributes consisting of demographic attributes such as age, gender, income, and geographical location; purchasing behavior attributes such as purchasing frequency, purchasing pattern, and product preference; engagement attributes like website visitation, application usage, and interaction frequency; and satisfaction attributes such as feedback scores and customer evaluation.

Statistical Customer Behavior Analysis Module

Descriptive Analysis

Descriptive analysis is carried out as the first step in the statistical analysis process in order to understand the general features and distribution patterns of the behavior of customers within the gathered data set. Descriptive analysis helps analyze customer characteristics like the distribution of customers in terms of demographics, purchasing frequency, and spending behavior in order to determine the general trends and differences between various customer groups. Customer distribution analysis helps analyze the structure of customers in terms of demographics and behaviors; purchase frequency analysis measures how often customers transact during the specified period; and spending behavior analysis helps in analyzing customer spending behavior, average spending amount, and transaction frequency of the customers to identify high- and low-spending and engagement customers. Descriptive statistical tools like frequency, percentage, mean, and standard deviation help to summarize the customer characteristics and offer a preliminary idea of their behavioral tendencies.

Reliability Analysis

Reliability analysis is carried out to test the internal consistency and reliability of the survey-based constructs of customer behavior in the proposed framework. In other words, reliability analysis guarantees that the items of the construct measures representing the customer behavior attributes like purchase behavior, customer involvement, satisfaction level, intention to be loyal, and service experience measure their respective constructs consistently. In this case, Cronbach's Alpha coefficient is employed as the technique of reliability analysis, wherein values above 0.7 are considered acceptable internal consistency of the measurement items. The higher the value of Cronbach's Alpha, the higher the reliability, thus, confirming that the data collected is fit for further analysis. Reliability analysis confirms the quality of the attributes of customer behaviors and ensures that the identified behavioral patterns can be effectively used in further analysis.

Correlation Analysis

A correlation analysis is conducted in order to establish the relationships and interaction patterns between the major behavioral parameters such as customer engagement, satisfaction, loyalty, and purchasing intention. Correlation analysis allows identifying the degree and direction of connections between various behavioral aspects, which makes it possible to understand better how the customers' experience affects the decision-making process. For this purpose, the Pearson correlation coefficient is calculated to define the degree of the relationship between variables, where positive correlation denotes the fact that the growth of one parameter contributes to the growth of the other one, whereas negative correlation denotes the inverse relationship. After the analysis of the correlations between engagement level, satisfaction, loyalty intention, and purchasing behavior, the crucial parameters determining customer relations and retention can be identified.

Factor Analysis

Factor analysis is performed to reveal the latent factors that impact customer behavior by analyzing several variables and consolidating them into meaningful factors of customer behavior. This statistical approach reveals latent factors that help define the essential constructs behind complex customer interactions, preferences, and decisions. Factors extracted through such analysis allow comprehending customer characteristics better through grouping similar attributes into common behavioral factors. In this research, four customer behavior factors are revealed using the factor analysis method: Customer loyalty, which refers to retention intention, repeat purchase, and relationship strength; digital engagement, which refers to interaction frequency, online activities, and platform engagement; product preference, which refers to purchasing preferences, brand preference, and interest; and service satisfaction, which refers to customer perception of service quality and satisfaction level.

Customer Segmentation

Customer segmentation is conducted to cluster customers into segments depending on their behavioral traits, purchasing behaviors, engagement levels, and value of relationships. In the current study, the K-means clustering technique is used to determine customer segments by clustering customers based on behavioral traits like purchase frequency, buying behavior, engagement score, satisfaction, and loyalty. Clustering helps to establish various customer segments that would enable the organization to conduct CRM strategies for the different segments identified. Based on the determined behavioral traits, customer segments are determined as high-value customers, who make frequent purchases, have high purchasing ability, and are loyal to the brand; frequent customers, whose interactions are consistent and have purchasing behaviors; and churn customers, whose engagement is low, purchasing frequency is low, and who have low satisfaction levels. These customer segments help determine marketing and retention strategies. The customer segments generated in the study can be incorporated into the neuro-symbolic knowledge mining approach to determine customer representation for explainable CRM strategies.

Neural Customer Representation Learning

The neural customer representation learning module is meant to find higher-level behavioral patterns and create customer representations out of the input behavioral features. The input of the neural customer representation learning module contains multidimensional customer data, like purchasing, interaction behavior, satisfaction parameters, demographic information, and other types of customer data. Deep learning algorithms like Autoencoders, Transformer neural networks, or Graph Neural Networks (GNNs) may be used in order to learn relationships and dependencies within the input data. The Autoencoder algorithm reduces the dimensions of the input and learns latent features; the Transformer neural network learns sequence and context dependency within customer interaction data, while the Graph Neural Network learns relations within customer-product-behavior data. As an output of the module, one obtains a vector representation of customer data, i.e., a numerical representation of behavioral and personal customer parameters. These customer embeddings capture highly complex

behavior patterns that cannot be identified using traditional statistical methods. This customer representation is further applied for knowledge graph construction and symbolic reasoning modules.

Knowledge Graph Construction

The customer knowledge graph has been designed to depict complex relationships between customers, products, brands, purchase actions, reviews, and interactions using structured knowledge representation. The knowledge graph comprises various entities and relationships capturing both explicit and implicit data about the customers' behaviors. These entities include the Customer, referring to individual users along with their behavioral attributes; Product, referring to products that the users have either purchased or prefer; Brand, referring to product ownership and brand affiliation of customers; Purchase, depicting transaction actions and purchasing behaviors; and Review, depicting opinions and feedback of the customers regarding product satisfaction. These entities have several relationships, including purchases that show the transaction relationships between customers and products; prefers that shows customer interest in particular products or brands; interacts, depicting customer involvement through the means of reviews, browsing, and online interaction; and recommends, depicting personalized product recommendations based on customer behaviors. For instance, a customer entity may have a purchase relationship with a product, which in turn has a relation to a brand or category.

Symbolic Reasoning and CRM Decision Engine

The symbolic reasoning component of the framework is intended to turn the extracted customer intelligence and the behavioral knowledge that the system learns into actionable decision rules for strategic CRM applications. This component is used to interpret the domain knowledge along with the customer behaviors and the pre-defined business rules to derive the logical rules and actions. Based on the analysis performed through reasoning, meaningful relationships between customer attributes such as purchase behavior, feedback, engagement, loyalty, and past interactions are identified and turned into interpretable if-then rules. For instance, if a customer exhibits high purchase frequency, good feedback, and high engagement, then the customer belongs to the premium segment and needs personal loyalty rewards, exclusive offers, or a retention strategy. If the customer is seen with low purchase frequency, bad feedback, and low engagement, the system sees that the customer is at risk of leaving the company and needs specific retention efforts. Such explainable rules enable CRM managers to know why each action has been recommended by the system.

EXPERIMENTAL METHODOLOGY

Dataset Description

The dataset used for the proposed framework in evaluating the performance of the proposed neuro-symbolic knowledge mining approach consists of 5,000 customer records that were collected from various sources, including CRM databases, transactional platforms, customer feedback systems, as well as interactions through digital channels. The 5,000 customer records include a single customer profile consisting of demographic features, buying behaviors, engagement levels, satisfaction measures, and interaction history. The attributes in the dataset include age, gender, location, buying frequency, buying cost, product preference, rating on products, engagement level with the company's website or applications, and loyalty-related measurements. The data collected consist of customer activities within a specified period to capture variation in behaviors and buying trends. These 5,000 customer records pre-processed to remove any missing data and normalize numerical attributes.

Statistical Analysis Setup

The statistical analysis phase involves the use of various statistical software packages such as SPSS, R, and statistical libraries in Python for validating customer behavior patterns and selecting important influencing factors. The first analysis involved the use of descriptive statistics to analyze the customer distribution, purchasing pattern, and spending pattern. A correlation analysis was used to determine the relationship between various behavioral factors such as customer engagement, satisfaction, loyalty, and

purchase intentions. A regression analysis was used to find out how the behavioral factors influence customer decision-making and the CRM performance measures. Clustering analysis of the customer data using the K-means algorithm was used to group customers into various segments.

Neuro-Symbolic Model Setup

The neuro-symbolic architecture is designed to combine feature extraction through deep learning with symbolic knowledge reasoning for intelligent customer behavior analysis. The model parameters are tuned to ensure successful learning and adequate customer modeling. The learning rate is chosen to ensure the correct updating of the model weights during training, whereas the number of epochs corresponds to the necessary number of training steps needed for convergence. The embedding size sets the dimensionality of customer feature representations provided by the neural learning component of the model that allows capturing complicated customer behaviors. A suitable optimization algorithm like the Adam optimizer can be used to ensure the minimization of training losses and stability of the model.

Evaluation Metrics

Accuracy: Checks the validity of the whole model for positive and negative classifications in equation (1)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision: Calculates the proportion of true positive results among all cases predicted as positive in equation (2)

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall (Sensitivity): Determines the percentage of true positives that were correctly classified by the model in equation (3)

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1-Score: Harmonic mean between precision and recall, considering false positives and false negatives in equation (4).

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Area Under the Curve (AUC-ROC): Combines the True Positive Rate (TPR) versus the False Positive Rate (FPR) at all threshold values to evaluate the general performance trade-off of classification in equation (5)-(7).

$$AUC = \int_0^1 TPR(FPR) d(FPR) \quad (5)$$

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

(In equation (1)-(7), TP is defined as True Positive, TN as True Negative, FP as False Positive, FN as False Negative)

Statistical Validation

The validity of the statistics is performed in order to demonstrate the credibility, importance, and effectiveness of selected customer behavior patterns before being integrated into the NeuroCRM-KM method. The process comprises descriptive statistics in order to study customer characteristics, correlation analysis to detect the relationships between behavioral factors, regression analysis to discover the influence of customer characteristics on decision-making, and clustering analysis to segment customers. Reliability of behavioral constructs is assessed via Cronbach's Alpha. Values exceeding 0.7 are deemed to be internally consistent. Correlations are estimated by means of Pearson correlation analysis. It is employed to assess the degree and nature of the relationship between engagement, satisfaction, loyalty, and customers' purchasing intention. Regression analysis assists in identifying important behavioral factors influencing customer loyalty and CRM decisions. Also, K-means clustering analysis is used to segment customers by means of behavioral similarities, including purchasing frequency, engagement, satisfaction, and loyalty.

RESULTS AND DISCUSSION

Customer Statistics Analysis Results

Demographic and behavioral analysis were performed on 5,000 customers to determine customer attributes and purchasing trends. From the analysis, it is evident that the database is comprised of a diversified customer distribution based on demographic segments, making behavioral analysis possible. The majority of the customers belong to the active purchasing group, with more transactions and engagement from digitally active customers. It has been found that the distribution of behaviors is such that approximately 42% of the customers show frequent purchasing behavior, 35% are moderate engagement customers, and 23% are low engagement customers with the possibility of churn. An average of 6.8 purchases was found for each customer with an average spending of \$485 per customer. Higher-engagement customers have higher transaction frequencies with more trans higher transaction.

Table 1. Customer demographic and behavioral distribution

Category	Customer Group	Number of Samples	Percentage (%)
Age Group	18–30 years	1450	29%
	31–45 years	2100	42%
	Above 45 years	1450	29%
Purchase Behavior	Frequent Buyers	2100	42%
	Moderate Buyers	1750	35%
	Low Engagement Buyers	1150	23%
Engagement Level	High Engagement	2200	44%
	Medium Engagement	1800	36%
	Low Engagement	1000	20%

Table 1 depicts the demographic and behavioral profile of customers according to their age range, purchase behavior, and engagement levels. Based on the analysis of age range, it can be seen that 42% (2100 samples) of the customers are from the 31-45 years age range, which means that middle-age customers are the dominant age range among the customer base. Younger customers from 18-30 years age range and customers who are above 45 years old each comprise 29% (1450 samples), thus depicting the balanced presence of younger and older customers in the customer base. Analysis of purchase behavior depicts that frequent buyers account for 42% (2100 samples), followed by moderate buyers at 35% (1750 samples) and low engagement buyers at 23% (1150 samples). This suggests that there is a substantial number of clients who are engaging themselves with the brand and exhibit a pattern of purchases. The distribution of the engagement level also reveals that highly engaged clients make up 44% (2200 samples), medium engagement clients constitute 36% (1800 samples), and low engagement clients make up 20% (1000 samples). In general, the results reveal that the dataset comprises of a substantial number of clients who engage themselves with the brand, making a good basis for the analysis of client loyalty, purchase intent, and behavior prediction using the CRM intelligence model.

Reliability and Factor Analysis Results

Reliability analysis was conducted to confirm the validity of customer behavior constructs such as customer loyalty, digital interaction, customer preference, and service satisfaction. Reliability of all the constructs was tested by computing Cronbach's Alpha values, which were found to be higher than the benchmark value of 0.7. Additionally, factor analysis was used to determine the latent behavioral factors that impact the decision-making of the customers. A high KMO value of 0.846 and significant Bartlett's test of sphericity ($p < 0.001$) show that the data set is appropriate for conducting factor analysis. It was found that there are four key behavioral factors impacting the decisions of customers, accounting for 78.4% variance in the data set.

Table 2. Reliability and factor extraction results

Construct	Cronbach's Alpha	Extracted Variance (%)
Customer Loyalty	0.86	21.5
Digital Engagement	0.83	19.8
Product Preference	0.81	18.2
Service Satisfaction	0.88	18.9
Overall Reliability	0.85	78.4

Table 2 summarizes the outcomes of reliability and factor extraction for the key constructs involved in this research. As can be seen from the reliability analysis, all constructs possess good internal consistency as the Cronbach's alpha scores vary from 0.81 to 0.88, thus surpassing the minimum value of 0.70. In particular, the construct of Service Satisfaction proves to be the most reliable one with $\alpha = 0.88$, then comes Customer Loyalty with $\alpha = 0.86$, Digital Engagement with $\alpha = 0.83$ and Product Preference with $\alpha = 0.81$, indicating that the measurement items reflect their constructs adequately. According to the factor extraction, the constructs together account for a significant amount of variance, which is explained by Customer Loyalty (21.5%), Digital Engagement (19.8%), Service Satisfaction (18.9%) and Product Preference (18.2%). The total variance of 78.4% means that the factors extracted reflect the main dimensions affecting customer behavior.

Correlation and Regression Results

The correlations were determined to check the relation among the behavioral factors like customer engagement, customer satisfaction, loyalty and purchase intention. It is found from the results that there is a highly positive relationship between all the major behavioral factors. Customer satisfaction has the highest correlation with loyalty ($r = 0.78$, $p < 0.001$), which signifies that better service directly impacts the retention of customers. In the same way, there is a positive relation between digital engagement and purchase intention ($r = 0.72$, $p < 0.001$).

Table 3. Correlation analysis results

Variables	Loyalty	Purchase Intention	Engagement
Customer Satisfaction	0.78	0.69	0.61
Digital Engagement	0.65	0.72	1.00
Product Preference	0.62	0.67	0.58
Purchase Frequency	0.71	0.75	0.64

Correlation analysis among customer satisfaction, digital engagement, product preference, purchase frequency, loyalty, purchase intention, and engagement is presented in table 3. It is evident that all variables are positively correlated, which implies that an increase in one behavioral variable is followed by an increase in others' customer-related indicators. There exists a strong correlation between customer satisfaction and loyalty ($r = 0.78$), implying that the level of satisfaction is the predictor of customer loyalty development. Digital engagement is strongly correlated with purchase intention ($r = 0.72$), and there exists a perfect correlation between digital engagement and the engagement construct ($r = 1.00$). Purchase frequency correlates significantly with purchase intention ($r = 0.75$) and loyalty ($r = 0.71$). In addition, product preference is moderately-strongly correlated with loyalty ($r = 0.62$) and purchase intention ($r = 0.67$). In summary, it is evident that there exists correlation between satisfaction, digital

engagement, product preference, and purchase frequency and loyalty, purchase intention, and engagement, which justifies including these behavioral indicators in the CRM prediction framework.

Regression Results

Table 4. Regression results

Predictor Variable	Influence on Loyalty (β)	p-value
Customer Satisfaction	0.42	<0.001
Digital Engagement	0.31	<0.001
Product Preference	0.24	0.002
Purchase Frequency	0.36	<0.001

In table 4, the results of the regression analysis of the impact of various customer behavioral factors on customer loyalty are given. According to the analysis results, all the predictors under consideration have a positive and statistically significant impact on customer loyalty, as can be seen from significant p-values (<0.05). Thus, among the selected factors, the factor of customer satisfaction is the one that has the greatest impact on loyalty ($\beta = 0.42$, $p < 0.001$), meaning that the greater the degree of satisfaction from the services received and the better service experience, the greater customer loyalty. Purchase frequency is another factor that affects customer loyalty in a statistically significant way ($\beta = 0.36$, $p < 0.001$), meaning that customers who often buy are more likely to form their loyalty towards the company. Digital engagement, which includes online interaction and communication, also has a positive effect on loyalty ($\beta = 0.31$, $p < 0.001$). Finally, the last predictor considered is the product preference, which has a statistically significant but lower than the previous three factors' impact ($\beta = 0.24$, $p = 0.002$).

Customer Segmentation Results

K-means clustering technique was used to cluster the customers according to their purchase behavior, engagement, spending patterns, and levels of satisfaction. Three main types of clusters have been identified. The first cluster consists of high-value customers that are described by frequent purchases, high purchasing power, and loyalty. The second cluster is made up of regular customers with medium levels of engagement and consistent purchase behavior. The third cluster consists of potentially churn customers who are less active and satisfied with the company's product.

Table 5. Customer segmentation results

Cluster	Customer Segment	Samples	Characteristics
Cluster 1	High-value Customers	1650	High spending, frequent purchases, strong loyalty
Cluster 2	Regular Customers	2200	Moderate engagement and stable purchasing
Cluster 3	Potential Churn Customers	1150	Low engagement and declining purchase activity

The customer segmentation results generated through grouping of customers based on the behavioral characteristics, purchase behaviors, and levels of engagement have been presented in table 5. There are three main customer segments that have been identified through the analysis. Cluster 2 constitutes the largest number of customers and includes 2200 regular customers who show a certain level of engagement and purchase behaviors that are constant; therefore, the company has an opportunity to develop a more effective marketing strategy to retain these customers. Cluster 1 comprises 1650 high-value customers characterized by high spend and frequent purchase behaviors, as well as their loyalty to the company. This cluster is considered the most profitable customer group, which needs to be retained through the use of some special strategies. Cluster 3 comprises 1150 potential churn customers characterized by a low level of engagement and decreasing purchases, should be retained through some interventions.

Neuro-Symbolic Model Performance

The performance of the proposed model for neuro-symbolic knowledge mining is evaluated by comparing it with traditional machine learning and deep learning techniques. The comparison is done in terms of accuracy, F1-score, and Area under curve (AUC). It can be observed from the results that the proposed model outperformed others because of the combined strength of neural feature extraction and

symbolic reasoning. The use of knowledge-based reasoning made predictions more accurate and interpretable.

Table 6. Performance comparison of prediction models

Model	Accuracy (%)	F1-score (%)	AUC
Logistic Regression	82.4	80.9	0.84
Random Forest	87.6	86.8	0.89
XGBoost	89.3	88.7	0.91
ANN	91.2	90.5	0.93
Transformer	93.1	92.4	0.95
Proposed Neuro-Symbolic Model	96.4	95.8	0.97

Table 6 illustrates the performance of different prediction models employed for analyzing customer behaviors in terms of accuracy, F1-score, and AUC metrics. The findings show that all machine learning models and deep learning models can effectively perform prediction but that the performance of the proposed Neuro-Symbolic Model is better than other approaches. For example, the traditional machine learning models including Logistic Regression have shown the performance of 82.4% accuracy, 80.9% F1-score, and AUC of 0.84; whereas ensemble-based machine learning models have improved their prediction capability with the performance of 87.6% accuracy and 89.3% accuracy for Random Forest and XGBoost, respectively. The deep learning models have further improved their performance with the results that ANN has 91.2% accuracy, 90.5% F1-score, and AUC of 0.93, and Transformer has 93.1% accuracy, 92.4% F1-score, and 0.95 AUC due to the capability of capturing complex behavioral patterns. Moreover, the Neuro-Symbolic Model has achieved 96.4% accuracy, 95.8% F1-score, and 0.97 AUC by improving the prediction reliability due to deep learning-based pattern recognition along with symbolic reasoning capabilities.

Explainable CRM Decision Analysis

The results of explainability analysis show that the suggested neuro-symbolic approach allows making CRM decisions that are easier to understand in comparison with traditional artificial intelligence-based systems. As a rule, traditional systems can predict something, but there is no information regarding the causes of these predictions. For instance, the churn prediction model developed using traditional AI can determine the fact that the probability of churn is equal to 0.85 for certain customers; however, the reasons behind this fact should clear. In the neuro-symbolic approach, the predictions made by the neural network are combined with symbolic rules, and it is possible to make conclusions regarding the increased churn risk due to lower purchase frequency, reduced digital involvement, negative customer feedback, and lower satisfaction level.

Table 7. Explainable CRM decision comparison

Approach	Prediction Output	Explanation Capability
Traditional AI Model	Churn probability = 0.85	Limited explanation
Proposed Neuro-Symbolic Model	Churn risk identified due to low purchase frequency, reduced engagement, and negative feedback	High interpretability and actionable recommendations

The comparison of the explainable capabilities of CRM decision-making between the traditional AI model and the suggested Neuro-Symbolic Model is provided in table 7. The results reveal that the traditional AI model allows getting just the predicted value of churn risk, such as 0.85 probability of churn, but does not provide any information about the reason behind this prediction. Due to this fact, it becomes difficult for CRM managers to develop an effective customer retention strategy. In its turn, the Neuro-Symbolic Model, besides the prediction of the risk of churn, suggests some meaningful explanations in terms of low purchase frequency, decreased customer engagement, and negative feedback of customers. The above-mentioned results prove that the Neuro-Symbolic Model can offer better explanation and interpretation of CRM decisions due to the integration of symbolic reasoning and deep learning.

Ablation Study Analysis

Table 8. Ablation study results of NeuroCRM-KM

Model Configuration	Accuracy (%)	F1-score (%)	AUC
NeuroCRM-KM (Complete Model)	96.4	95.8	0.97
NeuroCRM-KM without Symbolic Reasoning	93.2	92.6	0.94
NeuroCRM-KM without Knowledge Graph	91.8	91.1	0.92
NeuroCRM-KM without Neural Learning	88.7	87.9	0.89
Statistical CRM Model	82.5	80.8	0.84

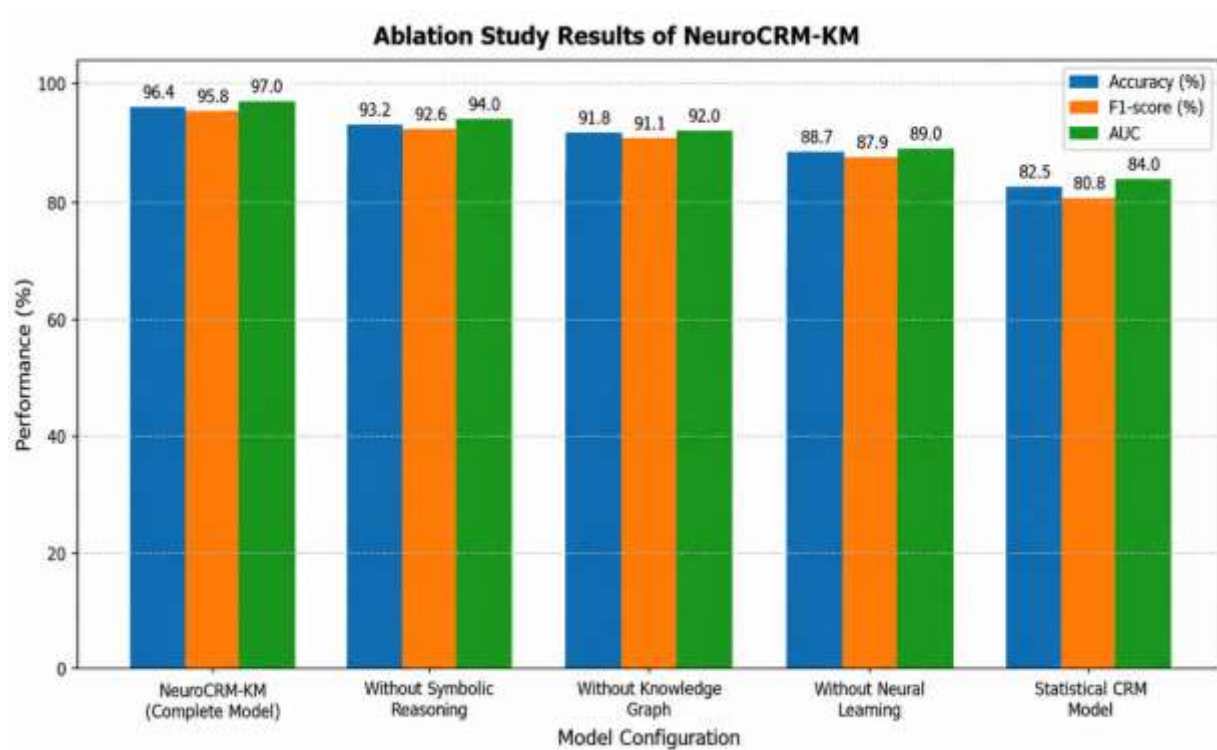


Figure 2. Ablation study results of NeuroCRM-KM

Table 8 and figure 2 describe the results of the ablation study conducted on the proposed NeuroCRM-KM model to evaluate the performance of its different modules, such as symbolic reasoning, knowledge graphs, and neural learning capabilities. The complete model of NeuroCRM-KM achieves the best performance score with 96.4% accuracy, 95.8% F1-score, and 0.97 AUC. The removal of the symbolic reasoning module affects the performance score by decreasing it to 93.2% accuracy, 92.6% F1-score, and 0.94 AUC. The absence of the symbolic reasoning capability enhances the interpretability and the correctness of the decisions made by the model. In addition, the absence of the knowledge graph module affects the performance score and decreases it to 91.8% accuracy, 91.1% F1-score, and 0.92 AUC. This shows that the knowledge graph capability helps the model understand the complex customer relations and thus increases the performance score. The removal of the neural learning module greatly affects the performance score of the model by reducing it to 88.7% accuracy, 87.9% F1-score, and 0.89 AUC. Compared with the statistical CRM model (82.5% accuracy, 80.8% F1-score, and 0.84 AUC), the proposed framework makes significant improvements.

MANAGERIAL AND THEORETICAL IMPLICATIONS

Theoretical Implications

This research makes an important contribution to the theoretical development of customer behavior analysis and CRM intelligence through the development of the NeuroCRM-KM (Neuro-Symbolic

Knowledge Mining) model that incorporates elements of neural learning, knowledge graphs, and symbolic reasoning into a single analytical approach. Contrary to traditional AI-based CRM approaches that emphasize the importance of predictive accuracy without giving explanations for the outcomes, the presented research adds to the existing body of knowledge by introducing a combination of machine learning and knowledge reasoning that allows for making explainable decisions. The research sets up a theoretical basis for building a unified knowledge representation framework for integrating heterogeneous customer information, behavior, and context. It is found that the combination of deep learning and symbolic inference provides better results in terms of the understanding and analysis of customer behavior and decision-making. High predictive accuracy of NeuroCRM-KM, namely, 96.4% accuracy, 95.8% F1-score, and 0.97 AUC proves the theoretical efficiency of the neuro-symbolic intelligence application in knowledge-based CRM system development.

Practical Implications

The NeuroCRM-KM framework offers many practical advantages to organizations since it helps them make CRM decisions that are both intelligent and customer-centric. This framework helps organizations identify different customer segments, detect behavioral trends, understand customers' buying preferences, and develop personalized marketing strategies by analyzing customer information obtained through explainable AI. With the use of customer data such as transactions, feedback, demographics, and interactions, the model helps organizations to discover hidden relationships between customers and create engagement strategies accordingly. Through the symbolic reasoning part, the model helps increase managers' confidence since it allows to explain the reasons for certain predictions such as low purchase frequency, disengagement and negative feedback related to risks of customer churn. In general, the NeuroCRM-KM framework allows organizations to enhance customer retention and optimize their marketing strategies. Thus, NeuroCRM-KM framework can be used as a decision support system in organizations.

CONCLUSION AND FUTURE SCOPE

The objective of the research was to create an intelligent and understandable framework for customer behavior analysis and CRM decision-making based on overcoming shortcomings of classical black box artificial intelligence solutions. NeuroCRM-KM (Neuro-Symbolic Knowledge Mining) framework represents a unique combination of neural learning, knowledge graphs, and symbolic reasoning to merge predictive intelligence with understandable decision-making. The framework methodology consists of gathering customer data of various natures, statistical analysis of customer behavior, extraction of deep representations of customers, building knowledge graphs, and symbolic reasoning for recommendation generation. The developed framework successfully discovers customer preferences, behavior, segmentation, and customer retention opportunities using knowledge-based analytics. Experiments on 5,000 customer datasets indicate that the proposed NeuroCRM-KM attains better results in terms of 96.4% accuracy, 95.8% F1 score, and 0.97 AUC, better than traditional machine learning and deep learning approaches. It is additionally validated by an ablation study that neural learning, integration of a knowledge graph, and symbolic reasoning play an important role in enhancing prediction accuracy and interpretability. The proposed method offers businesses clear CRM insights and helps in offering personalized marketing strategies, improving customer retention, and optimizing decisions. Future research may involve integration of streaming customer data, large language model, adaptive knowledge graph, and CRM application in different domains.

Declarations

Conflict of Interest: The authors declare that there is no conflict of interest regarding the publication of this research work. The authors have no financial, personal, or professional relationships that could have influenced the research design, analysis, or interpretation of the results.

Funding Statement: This research work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study was conducted using the authors' institutional resources and available research facilities.

REFERENCES

- [1] Tyagi PK, Tyagi P, Jegadeesan S, Ravishankar B, Kumar P. Optimizing Marketing Strategies Through Cloud Computing and Intelligent Analytics. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) 2025 Dec 12 (pp. 1-6). IEEE. <https://doi.org/10.1109/ICTBIG68706.2025.11323959>
- [2] Sunkara SP. Neuro symbolic AI for enterprise systems: Transforming CRM & ERP through logical and statistical AI integration. World Journal of Advanced Engineering Technology and Sciences. 2025;15(2):340-9. <https://doi.org/10.30574/wjaets.2025.15.2.0491>
- [3] Dritsas E, Trigka M. Machine learning in e-commerce: Trends, applications, and future challenges. IEEE Access. 2025 May 22;13:99048-67. <https://doi.org/10.1109/ACCESS.2025.3572865>
- [4] Hughes L, Dwivedi YK, Malik T, Shawosh M, Albashrawi MA, Jeon I, Dutot V, Appanderanda M, Crick T, De' R, Fenwick M. AI agents and agentic systems: A multi-expert analysis. Journal of Computer Information Systems. 2025 Jul 4;65(4):489-517. <https://doi.org/10.1080/08874417.2025.2483832>
- [5] Prentzas J, Hatzilygeroudis I. Assessment of life insurance applications: an approach integrating neuro-symbolic rule-based with case-based reasoning. Expert Systems. 2016 Apr;33(2):145-60. <https://doi.org/10.1111/exsy.12137>
- [6] Zhou Z, Wang L. Dual-Dimensional Digital Transformation Systematically Reshapes the U-Curve of Knowledge and Political Distance on Subsidiary Exit. Systems. 2025 Sep 3;13(9):773. <https://doi.org/10.3390/systems13090773>
- [7] Galitsky B, Ilvovsky D, Goldberg S. Shaped-charge learning architecture for the human-machine teams. Entropy. 2023 Jun 12;25(6):924. <https://doi.org/10.3390/e25060924>
- [8] Xin Z, Raghunath KK, Bhat CR. Smart decision orchestration for consumer electronics management using dynamic neuro-symbolic AI fusion. IEEE Transactions on Consumer Electronics. 2025 Sep 16;71(4):12047-55. <https://doi.org/10.1109/TCE.2025.3610191>
- [9] Yu CH, Miljkovic N, Wang Z. Managing the Unmanageable: Multimodal Artificial Intelligence for Unstructured Data Management and Analysis. Digital. 2026 Aug 17;6(3):68. <https://doi.org/10.3390/digital6030068>
- [10] Torres A, Filipe S. Behavioural Customers Analysis with Artificial Intelligence: Challenges and Pitfalls of Un (Biased) Decision-Making. In Artificial Intelligence Transforming Marketing: Renewed Theoretical Bases and New Paradigms 2026 Jul 2 (pp. 23-48). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-22360-9_2
- [11] Coiciu I, Raluca-Andreea-Cătălina LU. A Conceptual Framework for Market Risk Assessment with AI-Driven CRM and BI in Digital Ecosystems. In Proceedings of the International Conference on Business Excellence 2026 (Vol. 20, No. 1, pp. 1528-1538). Bucharest University of Economic Studies. <https://doi.org/10.2478/picbe-2026-0117>
- [12] GhorbanTanhaci H, Boozary P, Sheykhan S, Rabiee M, Rahmani F, Hosseini I. Predictive analytics in customer behavior: Anticipating trends and preferences. Results in Control and Optimization. 2024 Dec 1;17:100462. <https://doi.org/10.1016/j.rico.2024.100462>
- [13] Rahman MM. Data-driven decision-making through customer relationship management: a systematic literature review in modern enterprises. American Journal of Advanced Technology and Engineering Solutions. 2024 Oct 30;4(03):30-59. <https://doi.org/10.63125/vgr9th76>
- [14] Rainy TA, Rahman MA, Mou AJ. Customer relationship management and data-driven decision-making in modern enterprises: a systematic literature review. American Journal of Advanced Technology and Engineering Solutions. 2024 Apr 29;4(04):57-82. <https://doi.org/10.63125/jetvam38>
- [15] Ibeh CV, Asuzu OF, Olorunsogo T, Elufioye OA, Nduubuisi NL, Daraojimba AI. Business analytics and decision science: A review of techniques in strategic business decision making. World Journal of Advanced Research and Reviews. 2024 Feb 28;21(2):1761-9. <https://doi.org/10.30574/wjarr.2024.21.2.0247>
- [16] Lam HY, Tsang YP, Wu CH, Tang V. Data analytics and the P2P cloud: an integrated model for strategy formulation based on customer behaviour. Peer-to-Peer Networking and Applications. 2021 Sep;14(5):2600-17. <https://doi.org/10.1007/s12083-020-00960-z>
- [17] Basu R, Lim WM, Kumar A, Kumar S. Marketing analytics: The bridge between customer psychology and marketing decision-making. Psychology & Marketing. 2023 Dec;40(12):2588-611. <https://doi.org/10.1002/mar.21908>
- [18] Onari MA, Rezaee MJ, Saberi M, Nobile MS. An explainable data-driven decision support framework for strategic customer development. Knowledge-Based Systems. 2024 Jul 8;295:111761. <https://doi.org/10.1016/j.knsys.2024.111761>
- [19] Vudugula S, Chebrolu SK, Bhuiyan M, Rozony FZ. Integrating artificial intelligence in strategic business decision-making: A systematic review of predictive models. International Journal of Scientific Interdisciplinary Research. 2023 Mar 17;4(1):01-26. <https://doi.org/10.63125/s5skge53>

- [20] Hicham N, Nassera H, Karim S. Strategic framework for leveraging artificial intelligence in future marketing decision-making. *Journal of Intelligent Management Decision*. 2023 Sep;2(3):139-50. <https://doi.org/10.56578/jimd020304>