

Article

RFM-Net: A Convolutional Neural Network for Customer Segment Classification

Kadriye Filiz Balbal ¹  and Derya Birant ^{2,*} ¹ Department of Computer Science, Dokuz Eylul University, Izmir 35390, Turkey; kadriyefiliz.balbal@deu.edu.tr² Department of Computer Engineering, Dokuz Eylul University, Izmir 35390, Turkey

* Correspondence: derya@cs.deu.edu.tr

Abstract

Customer Segment Classification is a machine learning task in marketing analytics that involves assigning customers to predefined categories using features derived from historical transactional data. However, conventional approaches, such as statistical and clustering-based algorithms, may face challenges in fully capturing the nonlinear relationships in customer data, which can lead to limited insights and suboptimal segmentation outcomes. This paper introduces RFM-Net, an approach that integrates Deep Learning with Recency, Frequency, and Monetary (RFM) analysis for customer segment classification. By leveraging RFM features as input and labeled customer segments as output, we designed a specialized Convolutional Neural Network (CNN) model tailored for classification tasks. In the proposed method, labels are generated by a rule-based logic from RFM scores and then used as supervised ground truth. Accordingly, learning an expert-defined mapping is employed to model customer segmentation, rather than discovering a new segmentation structure. The proposed method enables businesses to classify customers into strategically meaningful segments such as Champions, Loyal Customers, At Risk, and Hibernating, thereby facilitating effective and targeted marketing strategies. Unlike traditional CNN architectures, RFM-Net offers a more compact, lightweight, and computationally efficient model with fewer layers and parameters, supporting improved interpretability and reduced risk of overfitting. Experimental results conducted on a real-world dataset demonstrated the effectiveness of RFM-Net with an accuracy of 94.33%. The results of this study showed a relative average increase of 13.17% compared to the results reported in previous studies on the same dataset. The core contribution of this research lies in combining the powerful generalization capabilities of deep learning with the effectiveness of RFM analysis, offering a robust solution for data-driven customer relationship management.



Academic Editors: Jae Joon Ahn and Sukjun Lee

Received: 8 January 2026

Revised: 20 February 2026

Accepted: 23 February 2026

Published: 25 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.**Keywords:** machine learning; customer segmentation; classification; convolutional neural network; customer behavior analytics

1. Introduction

Customer Segmentation plays a significant role in enabling personalized marketing strategies, boosting customer satisfaction, and enhancing overall business outcomes. Recency, Frequency, and Monetary (RFM) analysis has proven to be a powerful technique for understanding customer value and behavior in marketing analytics [1]. It evaluates customer engagement across three behavioral metrics: Recency value (how recently they bought something), Frequency value (how often they make a purchase), and Monetary

value (how much they spend). These metrics enable organizations to classify customers based on their similar purchase behaviors.

Despite its widespread use, traditional RFM-based segmentation relies primarily on rule-based logic, which may limit its ability to capture certain nonlinear patterns in real-world datasets. In the contemporary business environment, customer behavior can be multi-faceted and sophisticated and influenced by a variety of dynamic factors. Traditional approaches, such as statistical analysis and clustering, may not always fully capture the intricate relationships within data, sometimes leading to simplified categorizations, limited understanding, and incomplete interpretations. They sometimes offer static customer groupings that lack the adaptability required in dynamic market scenarios. These shortcomings have created a need for a potentially more robust, intelligent, insightful, and adaptable segmentation technique using Deep Learning.

In this study, we introduce RFM-Net, a customer segmentation approach that fuses the effectiveness of RFM analysis with the strong learning capability of deep learning. By designing a Convolutional Neural Network (CNN) trained on RFM features, RFM-Net effectively categorizes customers into actionable groups, including Champions, Loyal Customers, Potential Loyalists, Need Attention, At Risk, About to Sleep, and Hibernating. The proposed approach leverages both domain knowledge and data-driven learning, making it highly adaptable to diverse business contexts.

RFM-Net addresses the challenges by synergizing domain-driven features (RFM) with a data-driven modeling approach (deep learning). This combination allows the model not only to preserve the interpretability of traditional RFM analysis but also to enhance it through automated learning and feature extraction, enabling deeper insight into customer dynamics. As a result, this provides businesses with a segmentation solution that is both intuitive for decision-makers and technically advanced for data scientists and analysts. A primary aim of this research is to enable businesses to gain a more profound and data-driven understanding of their customers, thereby facilitating customer-centric marketing strategies that are not only targeted but also efficient and scalable.

Unlike traditional CNN architectures like AlexNet, GoogleNet, VGG, and ResNet; the architecture of RFM-Net is relatively shallow, consisting of only a few layers and a minimal number of parameters. Its lightweight structure is both computationally efficient and highly interpretable, allowing for real-time applications and ease of deployment. Unlike deeper models, RFM-Net is optimized for low-dimensional and tabular RFM data samples to achieve effective generalization without overfitting. It specially adapts the CNN power to structured customer behavior metrics. Moreover, RFM-Net eliminates the need for extensive feature engineering or manual clustering.

The contributions of this paper are fourfold:

- **RFM-Net:** Generating labels through rule-based logic derived from RFM scores and using them as supervised ground truth. Learning an expert-defined mapping to represent customer segmentation instead of discovering a new segmentation structure from the data.
- **Validation and Comparison:** Validated RFM-Net by a high accuracy of 94.33% using a real-world dataset and demonstrated a significant relative increase of 13.17% over previously reported results in terms of classification accuracy.
- **Strategic Insights:** Providing businesses with a reliable, advanced, and powerful solution for customer segmentation in marketing.
- **Benefits:** Unlike traditional CNN architectures, RFM-Net offers a more compact and lightweight framework with fewer layers and parameters, enhanced interpretability, and optimization for customer segmentation using structured RFM data, with reduced

risk of overfitting, enabling end-to-end learning without the need for a separate clustering algorithm.

The rest of the paper is structured as follows: Section 2 provides an overview of existing customer segmentation methods. Section 3 details the research methodology and introduces the proposed RFM-Net model. Section 4 reports the experimental results and comparative analysis, discussing the findings and their practical implications. Finally, Section 5 wraps up the study and highlights avenues for future work.

2. Related Work

This section provides a systematic review of current studies in customer segmentation. Table 1 provides a summary of previous studies [2–10], including the methods employed, their application purposes, the types of tasks addressed (classification or clustering), the performance metrics used, and the regions in which the studies were conducted.

Table 1. Summary of previous studies. C: Classification, K: Clustering, S: Silhouette, A: Accuracy, P: Precision, R: Recall, F: F-Measure, CM: Confusion Matrix, AUC: Area Under Curve. NA: Not Available, √: Available.

Ref.	Year	Region	Methods	Domain	Description	C	K	Period	Metrics
[2]	2025	U.S.	K-means, Hierarchical C.	Retail	Customer segmentation	√		NA	S
[3]	2025	Portugal	K-means, DBSCAN	Bank	Customer segmentation	√		NA	S, A, F
[4]	2025	Slovakia	Statistical methods	Postal Market	Customer segmentation	√		2024	NA
[6]	2024	Serbia	Hierarchical C., GBM, DT, KNN, NB, RF, SVM	E-Commerce	Correspondence analysis	√	√	2022	A, P, R, F, AUC
[10]	2024	Iran	SOM, CART	Hair Care	IoT-CRM	√	√	2017–2018	A, P, R, F
[9]	2022	Brazil	DBSCAN, XGBoost, LR	E-Commerce	Churn analysis	√	√	2016–2018	AUC
[5]	2022	China	K-means, SVM	E-Commerce	Churn analysis	√	√	2017	A, P, R, CM, AUC
[7]	2022	Malaysia	K-means	E-Commerce	Customer segmentation	√		2019–2021	A, time
[8]	2021	China	K-means, Hierarchical C.	Bank	Customer segmentation	√		2020	Density
Proposed		U.K.	RFM, Deep Learning (CNN)	Retail	Customer segment classification	√		2010–2011	A, P, R, F, CM

Some studies in the literature treat customer segmentation solely as a clustering task [2–4,7,8], whereas others address it within a classification framework as well [5,6,9,10]. In [2], researchers clustered customers into different segments as high spenders and seasonal shoppers. In [3], customers were divided into four different groups using clustering tasks, and the importance of this in determining marketing strategies was emphasized. In [4], where a three-dimensional segmentation model was applied, a system-based clustering approach was adopted. The study was implemented under real-world conditions with customers of a company in the postal market. In [5], where B2C e-commerce customers were segmented based on their shopping behavior, the problem situation was treated as both clustering and classification tasks. In [6], machine learning (ML) methods were used to examine the preferences and behaviors of e-commerce customers. The researchers divided customers into three main groups (young, unemployed, and female e-customers; retirees and the elderly; and employed, highly educated, and middle-aged men). After that, classification was performed using the labeled data obtained in the study. There are also studies in the literature that utilize ML methods to separate customers into different categories

based on their similar behaviors, with the goal of recommending the right product to the right customer for long-term profit [7,8].

Some studies have utilized traditional ML methods, such as Support Vector Machine (SVM) [5,6], Naive Bayes (NB) [6], Decision Tree (DT) [6,10], K-Nearest Neighbors (KNN) [6], and Artificial Neural Networks (ANN) [10]. SVM exhibited better prediction performance with higher accuracy than other models in B2C e-commerce customers' churn prediction model [5]. It was also among the machine learning classifiers applied to identify e-customer profiles (clusters) and was one of the algorithms showing the highest overall classification performance [6]. The Gaussian Naïve Bayes (GNB) algorithm was also used to identify the same e-customer profiles, but it showed lower performance in multi-class accuracy and Area Under the Curve (AUC) metrics than ensemble methods, such as Random Forest [6]. The Classification and Regression Tree (CART) technique was applied to determine the importance of 17 key factors affecting customer satisfaction after segmenting IoT customers [10]. The KNN algorithm was among the tested algorithms in e-customer profile classification, but it exhibited lower performance due to the variability in the categorical dataset [6]. Finally, Self-Organizing Map (SOM), a model based on ANN, was applied to group IoT customers according to their device usage patterns [10].

Clustering approaches used in customer segmentation studies vary. Most of them used K-means [2,3,5,7,8], while others employed hierarchical clustering algorithms [2,6,8] or DBSCAN [3]. K-means-based approaches generally aim to segment customer behavior data [5,7] or retail customer value [8]. For example, in [5], K-means was shown to significantly increase the performance of the churn prediction model by categorizing B2C e-commerce customers into three groups according to their shopping behavior. Ref. [7] presented an improved K-means approach, namely SAPK + K-means, to analyze Malaysian e-commerce customer purchasing behavior data and identify the most profitable customer segments. In [3], K-means was employed in the preprocessing stage of the hybrid KM-DBSCAN algorithm, which enables segmenting bank customers into four distinct groups and can handle noisy data. In [2], customer segments were identified using K-means after dimensionality reduction (FAMD) in mixed datasets containing both numerical and categorical variables. In [8], improved algorithms were proposed to address the shortcomings of the K-means algorithm, including the initial value of k and the tendency to fall into a local optimal solution, for retail customer classification and quantification of customer value systems. In [6], the Hierarchical Clustering on Principal Components (HCPC) method was utilized to model and label e-customer profiles based on demographic factors. In [2], Agglomerative Clustering was applied in comparison with K-means to validate segmentation results in mixed datasets. Similarly, a hierarchical clustering algorithm was also used in [8] to achieve higher efficiency.

Studies in the field of customer segmentation and analysis have focused on various objectives that respond to different industry needs and data types. Some of these studies address Customer Churn Prediction [5,9,11] and Customer Life Value [4,7,8], while others address Customer Behavior [2,6,10]. In studies [5,9], the researchers focused on predicting customer churn in e-commerce retail. In [11], the aim was to predict customer retention behavior and predict loyalty. In [9], the critical impact of behavioral factors such as shipping costs, product categories, and customer initial purchase value on churn is investigated and revealed. The study [8] aimed at segmenting retail customers into value-based categories; customer value was quantified, and then a structure supporting CRM decision-making processes was created by applying improved k-means variants. Studies adopting a sustainable development perspective have focused on identifying strategic segments with high potential, low cost, and high relationship value [4,7]. Studies aimed at understanding customer behavior have addressed demographic influences, online preferences, and interactions

within the IoT ecosystem [2,6,10]. A study examining Serbian e-commerce users modeled the relationships between demographic variables and online behavior, showing that the resulting three basic profiles could later be used as labels for classification models [6]. Studies examining the behavioral patterns of IoT users have analyzed the importance of factors affecting customer satisfaction [10]. There are also studies that perform segmentation on mixed datasets containing both categorical and numerical variables; in this context, distinct groups such as seasonal buyers and high-spending customers have been identified [2].

Performance evaluation metrics across studies differ depending on whether the task is classification or clustering. Since the majority of the reviewed studies were focused on classification, they primarily use metrics such as accuracy [3,5,6,11–13], precision [3,5,6,11–13], and recall [3,5,6,11–13]. These metrics are generally calculated using a confusion matrix [5,12,13]. Furthermore, the F1-score [3,6,11–13] and AUROC [5,6,9,12,13], which provide performance evaluations that are insensitive to the classification threshold, have also been widely applied. In clustering problems [2,3], metrics such as Silhouette coefficient, Calinski–Harabasz index (CHI), and Davies–Bouldin index (DBI) were used for quality assessment.

Customer segmentation research conducted across different geographies has yielded important findings about how these models interact with regional market structures, cultural consumption habits, and industry conditions. Studies in the European context [3,4,6] have focused primarily on regulated service sectors and e-commerce; in Slovakia, system-based segmentation models have been applied to corporate customers [4]; in Serbia, the impact of demographic variables on online behavior has been examined [6]; and in Portugal, banking customers have been classified using advanced clustering techniques [3]. Studies in Asia [7,8,12] have focused on topics such as customer value management and purchasing behavior; in China, retail customer value has been segmented using analytical methods [8], and in Malaysia, “smart segments” have been created by modeling e-commerce behavior with K-Means [7]. Among the studies conducted in South America, a comprehensive segmentation and churn forecasting study for the Brazilian e-commerce market stands out, combining transaction records and socio-demographic indicators [9]. Additionally, retail datasets from the United Kingdom [12] and the United States [2] were used as international references when regional data were limited. This diversity demonstrates that customer segmentation models are sensitive to regional context rather than universal and emphasizes the need to carefully evaluate the adaptability of the methods used to different markets [2,4,6].

Customer segmentation has been studied for different regions such as China [8], Brazil [9], Malaysia [7], Serbia [6], Iran [10], Slovakia [4], Portugal [3], the United States [2], and the United Kingdom [12]. These studies demonstrate how customer behavior varies across different geographic and cultural contexts. In Europe, segmentation models were developed for postal and banking services [3,4], while in Asia, retail customer value and e-commerce behavior were highlighted [7,8]. In South America, churn prediction was addressed using multi-source data [9], while datasets from the US and UK were utilized as common reference points for method comparisons [2,12]. This diversity demonstrates that segmentation approaches are sensitive to regional dynamics and that models need to be adapted to different market conditions [4].

Data used in customer segmentation studies generally fall into basic categories such as sales transactions, behavioral logs, demographic attributes, and geo-socio-demographic indicators. Many studies have modeled customer patterns using transactional/behavioral data such as purchase history, product views, and basket and favorite interactions [2,5,7,12]. In addition, demographic data such as age, gender, education, income, and household characteristics have been widely used in both e-commerce user profiles [6] and IoT customer

behavior analysis [8,10]. Geo-socio-demographic data, including indicators such as regional population structure, income level, and urban/rural location, have played a significant role, especially in churn forecasting studies [9]. In some sectors, structured data such as bank customer records [3] or ESG-based sustainability criteria [4] have been integrated into the model to create an objective-focused segmentation.

In addition to the studies discussed above, there are also studies that combine RFM analysis with machine learning techniques for customer segmentation [14–18]. In [16], researchers explored hybrid analytical processes that integrate RFM features with supervised or quasi-supervised learning components. Studies [14,17] incorporated RFM variables into clustering-based frameworks to support managerial decision-making processes. Similarly, in [15,18], RFM-focused segmentation strategies were primarily used within center-based clustering paradigms to divide customers into predefined behavioral layers (e.g., high-value, medium-value, low-value segments). The methodologies of these studies emphasize segment-based optimization and statistical grouping, generally relying on distance measures and the intuitive selection of cluster numbers. Unlike studies that primarily relied on cluster-focused approaches, we adopted a deep-learning-based solution.

Considering the comparison above, conventional statistical and clustering approaches may offer a reasonable solution for managing the customer segments. However, in scenarios where customer behavior exhibits potentially nonlinear or intricate patterns, deep-learning models might offer additional advantages. To explore this possibility, this study proposes RFM-Net—an approach that combines the predictive capabilities of deep learning with the strategic insights of RFM analysis, aiming to augment the capabilities of existing methodologies.

3. Materials and Methods

3.1. Proposed Method (RFM-Net)

This study introduces RFM-Net, a customer segment classification approach that fuses the valuable insights of RFM analysis with the powerful learning abilities of deep learning. RFM-Net incorporates a specialized convolutional neural network architecture that processes the historical purchase data of the customers, transforming raw RFM inputs into meaningful customer segments. This model identifies intricate behavioral patterns, allowing businesses to categorize customers into strategically relevant groups such as Loyal Customers, About to Sleep, Need Attention, and At Risk.

Figure 1 presents the general framework of the proposed approach. The methodology follows a structured pipeline that begins with raw data acquisition and proceeds through preprocessing and feature engineering to prepare the dataset for analysis. After the labeled data are organized, the model is systematically trained and evaluated. The subsequent stages are then followed by generating predictions and ultimately transforming these outputs into actionable decision-making support.

Algorithm 1 outlines the RFM-Net methodology in a formal, step-by-step manner.

Step 1—Data Acquisition: Historical raw customer data is acquired, focusing on marketing transactions such as purchases and returns. This data can be stored in a cloud-based platform to ensure scalability, availability, and efficient storage. Formally, let $D = \{x_1, x_2, \dots, x_n\}$ denote the dataset consisting of n transactions, where each transaction x_i contains fields like customer ID, invoice number, transaction date, quantity, and unit price. For each customer c , the algorithm gathers all of their transactions, as given in Equation (1):

$$T_c = \{x \in D \mid x.\text{CustmerID} = c\} \quad (1)$$

Step 2—Data Preprocessing: The raw data undergoes several pre-processing steps to ensure its quality, integrity, and usability.

- **Feature Selection:** In this step, only the fields relevant to RFM analysis and customer behavior modeling are retained, such as transaction date and amount. Several non-essential columns, such as country, product number, and name, are excluded from the dataset to reduce dimensionality and computational complexity.
- **Data Cleaning:** Several preprocessing operations are applied to ensure the integrity and consistency of the dataset, including the removal of return transactions, the handling of missing values, and the exclusion of irrelevant entries.
- **Data Transformation:** The raw transactional records are aggregated for each unique customer to calculate Recency, Frequency, and Monetary (RFM) metrics, thereby quantifying their purchasing behavior.

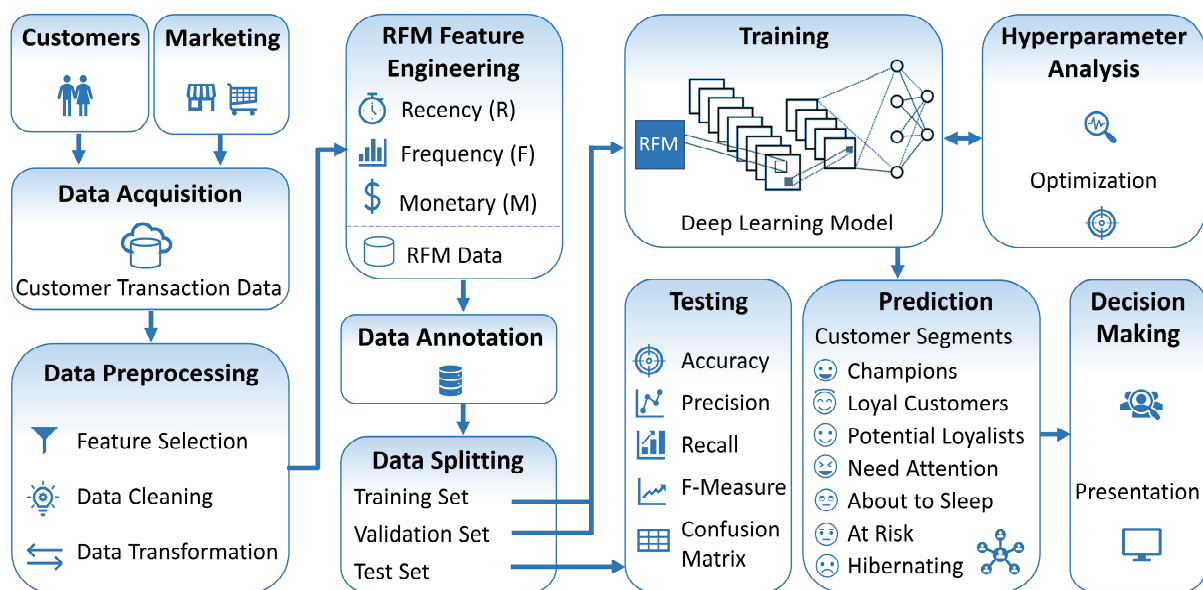


Figure 1. The general overview of the proposed RFM-Net approach.

Recency (R_c) is calculated as the number of days since a customer's most recent purchase (d_{last}^c) relative to the latest date (d_{ref}) in the dataset, as given in Equation (2). This metric helps distinguish between active and dormant customers.

$$R_c = (d_{ref} - d_{last}^c) \cdot \text{days},$$

$$\text{where } d_{ref} = \max(x.\text{Date}) \text{ for all } x \in D$$

$$\text{and } d_{last}^c = \max(x.\text{Date}) \text{ for } x \in T_c$$
(2)

Frequency (F_c) reflects how often a customer has transacted. It is calculated by the total number of distinct purchase events (invoices) associated with the customer c , as given in Equation (3). Higher frequency typically indicates strong engagement with the marketing platform.

$$F_c = | \{ x.\text{InvoiceNo} \mid \forall x \in T_c \} |$$
(3)

Monetary (M_c) represents the total monetary value of all purchases made by a specific customer c over a particular time period. It is derived by summing the product of the quantity and unit price of each transaction associated with the customer, as given in Equation (4). Monetary value quantifies the cumulative financial contribution of a customer to the business. This metric is particularly useful in identifying high-value customers

who generate significant revenue. It helps differentiate between low-spending and high-spending customers, thereby supporting targeted marketing strategies, resource allocation, and personalized service offerings.

$$M_c = \sum_{x \in T_c} (x.Quantity \times x.UnitPrice) \quad (4)$$

At the end of the first step, each customer is represented by a three-dimensional feature vector $x_c = (R_c, F_c, M_c)$. These three features form the core input representation for training the deep learning model.

Step 3—RFM Feature Engineering: Each customer is assigned three numerical scores (R, F, and M), which collectively capture their engagement level and transactional behavior. The continuous RFM metrics are discretized into categorical scores ranging from 1 to 5 according to the user-defined threshold values, allowing the model to normalize the dataset and better capture customer behavior patterns. This process sorts the values, determines thresholds, and divides each metric into five segments. Let $Q_R = \{RT_1, RT_2, RT_3, RT_4\}$ be the thresholds of recency, $Q_F = \{FT_1, FT_2, FT_3, FT_4\}$ for frequency, and $Q_M = \{MT_1, MT_2, MT_3, MT_4\}$ for monetary, each customer's RFM value is mapped to a discrete score between 1 and 5 using these thresholds. For example, a customer who spends a large amount overall would receive a top monetary score (e.g., 5), reflecting strong financial value, whereas one with minimal total spending would be assigned a lower score (e.g., 1), indicating limited contribution. The recency score is assigned inversely—meaning that a lower recency value (i.e., more recent purchases) results in a higher score—while the frequency and monetary scores are assigned directly, with higher values yielding higher scores, as given in Equation (5).

$$R_{score} = \begin{cases} 5 & \text{if } R_c \leq RT_1 \\ 4 & \text{if } RT_1 < R_c \leq RT_2 \\ 3 & \text{if } RT_2 < R_c \leq RT_3 \\ 2 & \text{if } RT_3 < R_c \leq RT_4 \\ 1 & \text{if } R_c > RT_4 \end{cases} \quad F_{score} = \begin{cases} 1 & \text{if } F_c \leq FT_1 \\ 2 & \text{if } FT_1 < F_c \leq FT_2 \\ 3 & \text{if } FT_2 < F_c \leq FT_3 \\ 4 & \text{if } FT_3 < F_c \leq FT_4 \\ 5 & \text{if } F_c > FT_4 \end{cases} \quad M_{score} = \begin{cases} 1 & \text{if } M_c \leq MT_1 \\ 2 & \text{if } MT_1 < M_c \leq MT_2 \\ 3 & \text{if } MT_2 < M_c \leq MT_3 \\ 4 & \text{if } MT_3 < M_c \leq MT_4 \\ 5 & \text{if } M_c > MT_4 \end{cases} \quad (5)$$

At the end of this step, each customer's RFM scores ($R_{score}, F_{score}, M_{score}, c$) are saved to be further used for assigning segment labels. These categorical scores normalize customer activity, making it easier to compare and group customers based on similar behavioral patterns.

Step 4—Data Annotation: Each customer RFM score triplet (R, F, M) is subsequently mapped to a predefined customer segment using rule-based logic. Segment definitions follow established marketing taxonomy, including labels such as “Champions”, “Loyal Customers”, “Potential Loyalists”, “At Risk”, and others, depending on combinations of high or low RFM scores. For example, if the Recency (R), Frequency (F), and Monetary (M) scores are each greater than 4, the customer is classified into the “Champions” segment, which comprises the most active and profitable customers. Similarly, customers with Recency, Frequency, and Monetary scores of $R < 2$, $F < 2$, and $M < 2$ are classified into the “Hibernating” segment, indicating they are low-value customers (likely to be lost). Table 2 presents RFM segment criteria, characteristics, and their corresponding strategy suggestions. We identified seven distinct groups of customers based on their transaction history, frequency, and spending habits, similar to the study [19]. This rule-based labeling offers an interpretable and actionable approach to categorizing customers based on their behavior. It creates a supervised dataset in which continuous RFM values serve as features and customer segments as class labels.

Table 2. RFM segment rules, characteristics, and associated strategy proposals.

Segment	Range of R Values	Range of M and F Average	Characteristics	Strategy Suggestions
Champions	4–5	4–5	Very recent, frequent, and high-spending customers	Offer premium customer services, personalized gifts, and early access to new products
Loyal Customers	2–5	3–5	Frequent and regular buyers—engaged customers	Motivate with loyalty rewards and periodic appreciation messages
Potential Loyalists	3–5	1–3	Recent, not yet frequent but promising	Guide them in discovering relevant products, provide customized communications and product education content
Need Attention	2–3	2–3	Medium recency and frequency—may need re-engagement—unsure if they will make another buy	Implement surveys to understand their needs and offer exclusive experiences
About to Sleep	2–3	1–2	Infrequent buyers, have not visited in a while	Utilize re-engagement emails, tailored discounts, or product bundles
At Risk	1–2	2–5	Long time no purchase, but still have some value (frequency + spend)	Re-engage with recovery campaigns, big discounts, or urgent limited-time offers
Hibernating	1–2	1–2	Least engaged, low-value customers who are likely lost	Provide lower-cost marketing streams, generic bulk offers, or reminders to prevent churn

As given in Table 2, differentiated marketing strategies can be implemented for each customer segment identified by RFM-Net. For the “Champions”, who are the most high-value customers, strategies such as VIP programs, personalized loyalty rewards, and early access to new products can reinforce their satisfaction. “Loyal Customers” represent a stable base and can be motivated with membership programs and periodic appreciation messages to maintain their engagement. The “Potential Loyalists”—recent but not yet frequent buyers—could be cultivated with welcome campaigns, product education content, customized communications, and behavior-based product recommendations. For the customers in the “Need Attention” category, surveys can be implemented to understand their needs. “About to Sleep” customers might be reactivated with re-engagement emails, tailored discounts, or product bundles based on past behavior. The “At Risk” group requires stronger interventions such as targeted win-back strategies, deeper discounts, or urgent limited-time offers. Lastly, “Hibernating” customers, with the lowest engagement and value, may benefit from lower-cost marketing streams, generic bulk offers, or reminders to prevent churn. Designing special communication and promotions according to the purchasing behavior of each group ensures a more efficient allocation of marketing resources, strengthens overall customer relationship management, and maximizes customer lifetime value.

Figure 2 illustrates the customer segmentation grid derived from RFM analysis, where the x-axis represents Recency scores (how recently a customer bought something), and the y-axis combines Frequency and Monetary scores (how often and how much a customer spends). Each dimension is scored on a scale from 1 to 5, with 5 representing the highest level of customer behavior (e.g., most recent, most frequent, or highest spending) and 1 indicating the least desirable behavior. Based on their scores, all customers are organized into seven predefined groups on the grid. Customers positioned in the upper-right quadrant, such as Champions and Loyal Customers, are the most valuable ones (high recency, high frequency/monetary). Customer engagement decreases as one moves leftward and

downward in the grid. Segments located in the lower-left quadrant (e.g., Hibernating) represent customers who have minimal interaction and exhibit low spending behavior, potentially indicating that they are lost customers. The grid provides a clear, quick, and strategic overview of customer behavior, enabling companies to differentiate and interpret customer value and engagement levels at a glance.

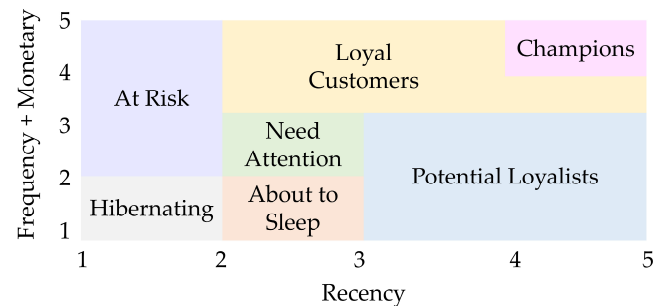


Figure 2. Customer segments in a grid view based on RFM scores.

Step 5—Data Splitting: The annotated dataset is partitioned into three distinct parts: a training set used to build the CNN model, a validation set to monitor training (i.e., early stopping for preventing overfitting), and a test set for evaluating the generalization performance of the model on unseen data.

Step 6—Model Training: A CNN is then trained to learn the complex relationships between RFM patterns and customer segments. The CNN architecture includes layers such as convolutional kernels to capture patterns, pooling layers to decrease data dimensionality, and dense layers to carry out the classification task. The input to the model consists of the numerical values of recency, frequency, and monetary metrics for each customer. The output of the model is a probability distribution over possible customer segments, enabling it to predict the most likely class for each new customer. Model training can be repeated using different hyperparameter configurations in order to improve performance and generalization.

Step 7—Model Evaluation: Once the CNN architecture is trained, the model is evaluated on the test set to predict segment labels for previously unseen customers using standard performance indicators such as accuracy, recall, precision, f-measure, and confusion matrix. These indicators help evaluate the ability of the model to distinguish between customer segments. The algorithm outputs $\hat{Y}$, the predicted customer segment labels for the test instances based on their RFM values, as given in Equation (6).

$$\hat{Y} = \hat{Y} \cup \hat{y}, \quad \text{where } \hat{y} = \text{Model}(\text{sample}) \quad \text{and} \quad \text{sample} = (R, F, M) \quad (6)$$

Step 8—Prediction: The CNN model is used to classify unseen customers into one of the following strategic segments, such as champions, at risk, or hibernating. This predictive capability enables businesses to gain actionable insights for targeted marketing strategies.

Step 9—Decision-Making Support: The final predictions are presented to business decision-makers through an interpretable dashboard or reporting system. Segment-based visualizations and analytics enable marketing teams to make informed decisions regarding campaign design, customer retention, and resource allocation.

Overall, the proposed methodology provides a hybrid solution. By combining the interpretability of RFM analysis with the predictive power of deep learning, RFM-Net offers a scalable, data-driven approach to customer segmentation that is well-suited for real-world applications in marketing, customer relationship management, and personalized recommendation systems. It enables organizations to understand customer behavior

in a structured way, while also leveraging machine learning to automate and scale the segmentation process for real-time applications.

Algorithm 1: RFM-Net: Recency-Frequency-Monetary-based Neural Network

Inputs:

RT , FT , MT : threshold values for recency, frequency, and monetary, respectively

Outputs:

$\hat{Y}$: predicted customer segment labels for the test samples

Begin:

$D \leftarrow \{x_1, x_2, \dots, x_n\}$ where $x_i = (\text{CustomerID}, \text{InvoiceNo}, \text{Date}, \text{Quantity}, \text{UnitPrice})$ // Step 1: Data acquisition

$\text{RFM_Values} \leftarrow \emptyset$ // Step 2: Data preprocessing

$d_{ref} \leftarrow \max(x.\text{Date})$ for all $x \in D$ // Reference date: The most recent date in the dataset

$\text{Customers} \leftarrow \{x.\text{CustomerID} \mid \forall x \in D \text{ and } x.\text{CustomerID} \neq \text{null}\}$ // The set of unique customer numbers

foreach $c \in \text{Customers}$ **do** // Calculate recency, frequency, monetary values

$T_c \leftarrow \{x \in D \mid x.\text{CustomerID} = c\}$ // All transactions belonging to customer c

$d_{last}^c \leftarrow \max(x.\text{Date}) \mid \forall x \in T_c$

$R_c \leftarrow (d_{ref} - d_{last}^c).\text{days}$ // Recency: Days since the customer's last purchase

$F_c \leftarrow |\{x.\text{InvoiceNo} \mid \forall x \in T_c\}|$ // Frequency: Number of unique invoices

$M_c \leftarrow \sum(x.\text{Quantity} \times x.\text{UnitPrice})$ for all $x \in T_c$ // Monetary: Total spending

$\text{RFM_Values} \leftarrow \text{RFM_Values} \cup \{(R_c, F_c, M_c, c)\}$

end foreach

$\text{RFM_Scores} \leftarrow \emptyset$

// Step 3: RFM feature engineering

foreach $(r, f, m, c) \in \text{RFM_Values}$ **do**

// Assign RFM scores (1 to 5) based on thresholds

for i **from** 1 **to** 5 **do**

if $r \leq RT[i]$ **then**

$\text{Rscore} \leftarrow 6 - i$

break

end if

if $f \leq FT[i]$ **then**

$\text{Fscore} \leftarrow i$

break

end if

if $m \leq MT[i]$ **then**

$\text{Mscore} \leftarrow i$

break

end if

end for

$\text{RFM_Scores} \leftarrow \text{RFM_Scores} \cup (\text{Rscore}, \text{Fscore}, \text{Mscore}, c)$

end foreach

$\text{Data} \leftarrow \emptyset$

// Step 4: Data annotation

foreach $(R, F, M, c) \in \text{RFM_Scores}$ **do**

// Rule-based labeling

$\text{AvgFM} = (F + M) / 2$

if $4 \leq R \leq 5$ and $4 \leq \text{AvgFM} \leq 5$ **then** $\text{Segment} \leftarrow \text{"Champions"}$

else if $2 \leq R \leq 5$ and $3 \leq \text{AvgFM} \leq 5$ **then** $\text{Segment} \leftarrow \text{"Loyal Customers"}$

else if $3 \leq R \leq 5$ and $1 \leq \text{AvgFM} \leq 3$ **then** $\text{Segment} \leftarrow \text{"Potential Loyalists"}$

else if $2 \leq R \leq 3$ and $2 \leq \text{AvgFM} \leq 3$ **then** $\text{Segment} \leftarrow \text{"Need Attention"}$

else if $2 \leq R \leq 3$ and $1 \leq \text{AvgFM} \leq 2$ **then** $\text{Segment} \leftarrow \text{"About to Sleep"}$

else if $1 \leq R \leq 2$ and $2 \leq \text{AvgFM} \leq 5$ **then** $\text{Segment} \leftarrow \text{"At Risk"}$

else if $1 \leq R \leq 2$ and $1 \leq \text{AvgFM} \leq 2$ **then** $\text{Segment} \leftarrow \text{"Hibernating"}$

$\text{Data} \leftarrow \text{Data} \cup \{(\text{Recency}, \text{Frequency}, \text{Monetary}, \text{Segment})\}$

Algorithm 1: *Cont.*

```

end foreach
TrainSet, ValidationSet, TestSet  $\leftarrow$  split(Data)           // Step 5: Data splitting
BestParams  $\leftarrow$  HyperparameterTuning(CNN, TrainSet, ValidationSet) // Perform hyperparameter tuning
Model  $\leftarrow$  Train(CNN(BestParams), TrainSet)             // Step 6: Train the CNN model
// inputs  $x \leftarrow$  (R,F,M values), output  $y \leftarrow$  segment label

 $\hat{Y} \leftarrow \emptyset$ 
foreach sample in TestSet                                // Step 7: Test model
     $\hat{y} \leftarrow$  Model(sample)
     $\hat{Y} \leftarrow \hat{Y} \cup \hat{y}$ 
end foreach
End

```

3.2. The Proposed CNN Architecture

In this study, we propose a deep learning architecture called RFM-Net, designed to enhance customer segmentation by integrating the classical Recency, Frequency, and Monetary framework with a custom-built Convolutional Neural Network. The aim of RFM-Net is to help businesses understand customer behavior patterns effectively, thereby enabling them to develop precise, customer-centric marketing strategies. The architecture of RFM-Net is composed of several key layers (input layer, convolutional layer, max-pooling layer, flatten layer, dense layer, and output layer), each of which serves a distinct function to support data-driven customer classification. Each component of the RFM-Net model is described below, highlighting its specific contribution to the customer segmentation process.

Input Layer: This layer receives structured customer data, typically composed of RFM features. The data is reshaped to a format compatible with convolutional operations. Each input sample represents a single customer behavioral profile, forming the foundation for deeper pattern extraction.

Convolutional Layers: These layers apply multiple filters to the features to detect patterns within data. They enable the model to understand how certain RFM feature combinations (e.g., high frequency but low monetary value) might correlate with specific customer segments. The rectified linear units (ReLU) activation function in these layers introduces non-linearity, which helps in modeling complex relationships. This is particularly useful for differentiating subtle variations in customer behaviors, such as identifying “Potential Loyalists” or “Promising” customers. A narrow kernel is used to ensure local feature extraction without overfitting.

Max-Pooling Layer: This layer decreases the spatial size of the feature maps while preserving the most informative ones, thereby enhancing generalizability and reducing noise. It ensures minimizing the impact of small fluctuations in customer data.

Flatten Layer: The multi-dimensional outputs from the convolutional layers are converted into a one-dimensional vector that is suitable for classification. This transformation bridges the convolutional layers and the dense classifier while preserving the learned representations of customer behavior.

Dense Layer (Fully Connected): This layer processes the flattened vector to learn higher-level relationships between RFM patterns. It enables the network to form comprehensive views of customer profiles, such as whether customers belong to the “Champion” group, characterized by consistent spending, or the “Hibernating” group, characterized by minimal activity.

Output Layer: This layer maps the learned features to meaningful customer classes, such as Champions, Loyal Customers, Need Attention, and Hibernating. It uses a softmax activation to assign probabilities across the predefined customer segments. The segment with the top probability value is chosen as the final classification. This output empowers

organizations to design their marketing strategies for each customer group separately with greater precision and personalization.

Through this end-to-end learning process, RFM-Net can accurately categorize each customer into a relevant segment, enabling businesses to develop targeted, customer-centric marketing strategies.

Table 3 presents the architecture of the RFM-Net model, including the types of layers used, their respective output shapes, and the number of trainable parameters at each stage of the network. The architecture begins with an input layer, followed by convolutional layers designed to extract low-level feature patterns from RFM metrics. A pooling layer is then applied to reduce spatial complexity. The process continues with a flattening operation, followed by a dense layer that interprets the extracted features and an output layer that maps them to customer segment predictions.

Table 3. The proposed CNN architecture and parameter summary.

Layer	Output Shape	Number of Parameters
Conv2D	(None, 2, 1, 32)	96
MaxPooling2D	(None, 1, 1, 32)	0
Conv2D	(None, 1, 1, 64)	2112
Flatten	(None, 64)	0
Dense	(None, 64)	4160
Dense (Softmax Output)	(None, 7)	455
Total Parameters		6823 ($\approx$ 26.65 KB)

In the proposed CNN model, each customer is represented by three RFM features (Recency, Frequency, Monetary), which are reshaped into a single-channel 2D tensor of size (3, 1, 1) to comply with the Conv2D input format. The first convolutional layer applies 32 filters with a kernel size of (2, 1) and ReLU activation, followed by MaxPooling2D with pool size (2, 1). Additional convolutional layers use 64 (or higher) filters with kernel size (1, 1), enabling nonlinear feature transformation. After convolution and pooling operations, the feature maps are flattened and passed to a dense layer with 64 neurons and a final softmax output layer corresponding to the customer segments.

Despite its relatively shallow structure, RFM-Net is highly efficient and effective due to its careful task-specific design. The total number of trainable parameters remains minimal, ensuring computational efficiency while preserving model expressiveness. This lightweight structure makes RFM-Net ideal for real-world applications where computational resources may be constrained.

3.3. Comparative Analysis of RFM-Net with Existing CNN Architectures

Although conventional CNN architectures, such as ResNet, AlexNet, VGG, and GoogleNet, have demonstrated superior performance in computer vision tasks, they are not typically designed and optimized for low-dimensional tabular data such as RFM inputs. These models are often overparameterized for tasks like customer segment classification, leading to overfitting, longer training times, and the need for extensive computational resources. The key advantages of RFM-Net over these models can be summarized as follows:

Lightweight and Fast: The architecture of RFM-Net is relatively shallow, consisting of only a few layers and a minimal number of parameters. This lightweight design ensures high computational efficiency, enabling the model to be particularly well-suited for real-time inference and deployment in resource-constrained environments.

Tailored for Tabular Data: Unlike image-centric CNNs, RFM-Net is specifically designed to work with structured data, preserving the semantic relationships between RFM features.

Overfitting Prevention: Deeper models, such as VGG-16 or ResNet-50, may be prone to overfitting when applied to low-dimensional data. RFM-Net addresses this issue through its architectural simplicity and pooling mechanism.

Interpretability: The compact architecture of RFM-Net provides better interpretability than that of deeper and black-box models. This enables easier interpretation and debugging, which is crucial in customer behavior analysis for business applications when understanding model decisions.

End-to-End Learning: By combining feature extraction and classification within a unified framework, RFM-Net eliminates the need for separate clustering algorithms. This end-to-end framework simplifies the segmentation pipeline and improves scalability.

4. Experimental Studies

4.1. Dataset Description

In this study, we utilized the publicly available “Online Retail” dataset [20], which was obtained from the UCI Machine Learning Repository. It is a rich multivariate time-series dataset comprising 541,909 records and 8 variables, collected from a UK-based non-store online retail company. The company primarily sells unique all-occasion gifts, including items such as ceramic homeware, scented candles, novelty mugs, children’s crafts, seasonal decorations, and stationery. These products are typically low-priced, decorative and often purchased in bulk by wholesalers for resale in gift shops or boutique stores. The transactions in the dataset span over a one-year period, from 1 December 2010 to 9 December 2011. The dataset captures customer purchase behavior across 37 different countries, including the United Kingdom, Japan, the United States, Australia, the Netherlands, France, Italy, Spain, Germany, Canada, and several other countries. Due to its sequential nature and temporal granularity, the dataset is well-suited for customer segmentation, market basket analysis, anomaly detection, demand forecasting, and customer lifetime value estimation.

Table 4 presents a structural overview of the dataset, including variable names, data types, brief descriptions, and indications of missing values. Each record corresponds to a line item in an invoice, meaning that a single invoice may have multiple rows corresponding to multiple products purchased in that transaction. Notably, the variables Description, Country, and CustomerID contain missing data, which must be addressed during preprocessing for any segmentation and modeling tasks.

Table 4. The characteristics of the “Online Retail” dataset.

Variable Name	Type	Description	Missing Value
CustomerID	Categorical	A five-digit numerical ID uniquely given to each customer	Yes
InvoiceNo	Categorical	A six-digit integer number uniquely given to each individual transaction	No
InvoiceDate	Date	The timestamp when the transaction occurred	No
StockCode	Categorical	A five-digit numeric identifier uniquely given to each unique product	No
Description	Categorical	The name describing the product	Yes
Quantity	Integer	The count of units for an item within a single transaction	No
UnitPrice	Real	The price charged for one unit of the product in sterling	No
Country	Categorical	The country where each customer resides	Yes

A sample of the dataset is shown in Table 5. Each row in the table represents an individual product item within a customer invoice. As seen in the table, customers often purchase multiple items in a single order. For instance, invoice 536608 includes three

different items purchased by customer 12855 on 2 December 2010. Repeated purchases of the same product (e.g., stock code = 16014) across different invoices and dates indicate recurring demand for specific products in varying quantities. The sample also reflects temporal diversity in transactions, offering insight into customer activity across various points in time throughout the year. Overall, the inclusion of both customer- and product-level details enables a wide range of analyses on customer behavior, sales trends, and product performance.

Table 5. A sample from the “Online Retail” dataset.

Invoice No	Stock Code	Description	Quantity	Invoice Date	Unit Price	Customer ID	Country
536608	22863	Soap dish brocante	6	12/2/2010 09:37	2.95	12855	UK
536608	22962	Jam jar with pink lid	12	12/2/2010 09:37	0.85	12855	UK
536608	22963	Jam jar with green lid	12	12/2/2010 09:37	0.85	12855	UK
537841	16014	Small Chinese style scissor	1000	12/8/2010 15:10	0.32	13848	UK
543991	16014	Small Chinese style scissor	1500	02/15/2011 10:17	0.32	13848	UK
543991	16015	Medium Chinese style scissor	100	02/15/2011 10:17	0.5	13848	UK
543991	16016	Large Chinese style scissor	100	02/15/2011 10:17	0.85	13848	UK
566028	16014	Small Chinese style scissor	1000	09/08/2011 12:58	0.32	13848	UK
549020	84879	Assorted colour bird ornament	160	04/05/2011 15:36	1.45	13525	UK
549020	72801C	4 rose pink dinner candles	48	04/05/2011 15:36	1.06	13525	UK
578785	84879	Assorted colour bird ornament	160	11/25/2011 12:01	1.45	13525	UK
578785	22622	Box of vintage alphabet blocks	2	11/25/2011 12:01	11.95	13525	UK
578785	23395	Belle Jardiniere cushion cover	12	11/25/2011 12:01	3.75	13525	UK
578785	23396	Le Jardin botanique cushion cover	12	11/25/2011 12:01	3.75	13525	UK

Prior to analysis, a series of data preprocessing steps was performed to ensure analytical robustness and relevance. The initial phase involved selecting only the variables essential for RFM analysis. Supplementary fields such as product codes and names and country were omitted, as they were not directly relevant to the RFM framework and would contribute to unnecessary complexity. After that, records lacking customer identification were excluded to address data incompleteness, as they could not be linked to any user behavior. Another critical step involves removing return transactions that are not associated with a corresponding sales record. In other words, transactions whose invoice numbers were prefixed with the letter ‘C’ were eliminated if they did not correspond to an original purchase. Additionally, contextually irrelevant entries such as bank charges, postage, and gift cards were excluded to more accurately reflect actual customer purchase behavior.

Following data cleaning, customer-level metrics were computed: recency was calculated as the number of days since the most recent purchase of the customer; frequency captured the total number of distinct purchases; and monetary value was derived by aggregating total spending across all invoices of the customer. Each individual customer was then assigned discrete R, F, and M scores based on the user-defined threshold values. Specifically, recency was segmented using the bins [10, 30, 50, 150], while monetary values were divided according to the bins [500, 1500, 2500, 5000]. Given the high density of one-time visitors, frequency values were scored using a specialized binning strategy of [1, 2, 4, 6]. Under this scheme, the customers with a frequency score of 1 were assigned to category 1, those with a score of 2 were placed in category 2, those with scores of 3 and 4 were classified into category 3, customers that have the frequency values 5 and 6 were grouped under category 4, and those with a frequency value higher than 6 were included in category 5. These thresholds were determined through an exploratory analysis and iterative empirical experimentation, considering the data distributions, domain knowledge, and the resulting classification accuracy across multiple trials. Finally, a rule-based classification scheme

was applied to the resulting RFM scores to annotate each customer with a corresponding behavioral segment label.

Table 6 illustrates a labeled dataset resulting from an RFM analysis, where each customer has been annotated with a corresponding segment category (e.g., ‘Champions’, ‘Hibernating’) based on their purchasing behavior. Since multiple rows may correspond to a single invoice and each customer may have multiple invoices over time, the raw data are aggregated for customer-level and invoice-level analysis. For instance, as shown in Table 5, customer 13848 made three separate purchases at different times. These transactions included varying quantities and types of products, with a total monetary value of £1255, a frequency of 3 purchases, and a recency value of 92 days from the most recent purchase to the reference date. Based on this aggregated data, RFM scores are assigned according to the user-defined thresholds. For CustomerID 13848, the resulting RFM score is 232, indicating relatively low recency, moderate frequency, and low monetary value. According to predefined segmentation rules, this customer falls into the “Need Attention” segment. In short, the RFM analysis transforms raw sales records into interpretable customer segments, which serve as labels for the classification task.

Table 6. A sample customer segments based on RFM methodology.

CustomerID	Recency	Frequency	Monetary	R-Score	F-Score	M-Score	RFM-Score	Customer Segment
13525	14	2	628.78	4	2	2	422	Potential Loyalists
13848	92	3	1255	2	3	2	232	Need Attention
12855	372	1	38.1	1	1	1	111	Hibernating
13089	2	118	57,385.88	5	5	5	555	Champions
17203	35	6	3563.85	3	4	4	344	Loyal Customers
13194	135	1	60.7	2	1	1	211	About to Sleep
15332	366	4	1661.06	1	3	3	133	At Risk

Figure 3 illustrates the distribution of customer segments, showing the proportion of each group within the overall customer base. The largest portion of the customers (23.70%) falls under the “Potential Loyalists” category, indicating a significant number of recent buyers who have the potential to become long-term loyal users. This is followed by “Loyal Customers” at 18.28%, representing a key group of repeat purchasers. Meanwhile, “Hibernating” customers account for 17.39%, reflecting a substantial portion of inactive and low-engagement users. Champions, the most valuable and engaged customers, account for 14.46%, while “About to Sleep” and “Need Attention” make up 11.30% and 8.51%, respectively. Lastly, the “At Risk” group represents 6.36% of the customer base, highlighting a smaller but still important group with declining activity. This distribution offers critical insights into customer behavior, enabling more effective targeted efforts, from supporting high-potential segments to re-engaging those at risk of churn.

Figure 4 illustrates the outcome of the permutation feature importance analysis, a widely used technique in machine learning for assessing the relative contribution of each input feature to the predictive accuracy of the model. The primary objective of this analysis is to disrupt the relationship between the target variable and a given feature by randomly permuting the feature’s values and then observing how the model’s performance deteriorates as a result. A greater drop in performance signifies a more important feature. According to the results presented, Recency was identified as the most impactful variable, with an importance score of 0.5358, indicating that the time since a customer’s last purchase plays a significant role in the decision-making process. This high value suggests that the prediction is highly sensitive to temporal data, specifically serving as the primary signal for distinguishing between active and churned states. Frequency was ranked as the second most important variable, with a score of 0.4217, suggesting that the number of purchases is

also a strong predictor. This indicates that the model relies on the repetition of purchase behavior to establish patterns of loyalty, implying that ‘how often’ a customer returns is nearly as vital as ‘when’ they last visited. Monetary was identified as the least important among the three variables, with a score of 0.2642, indicating that spending levels are also relevant but play a comparatively smaller role in the prediction process. The hierarchical importance of the RFM variables ($R > F > M$) aligns closely with the Churn Prediction (CP) in the customer analytics domain, where the ‘recency’ of an action is often the strongest indicator of future engagement. Furthermore, this ordering also aligns with the Customer Lifetime Value (CLV) framework, where recency and frequency are typically stronger predictors of future cash flows than monetary value alone, reflecting well-established patterns observed in domain knowledge.

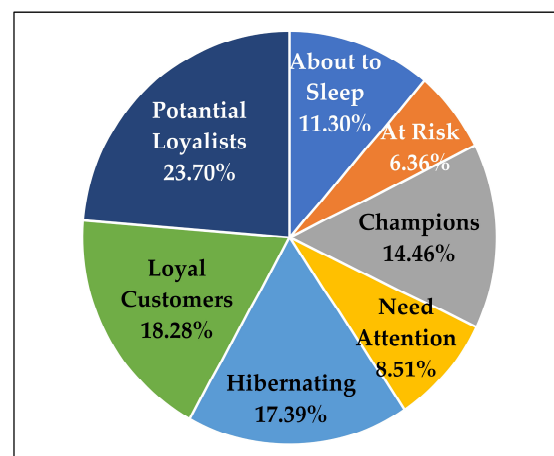


Figure 3. Distribution of customer segments.

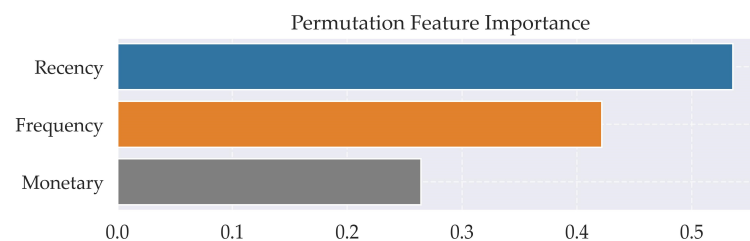


Figure 4. Feature importance analysis.

4.2. Experimental Setup

The proposed method was implemented in Python (version 3.12) using various libraries, including TensorFlow (version 2.20.0), NumPy (version 2.0), Pandas (version 2.0), Scikit-Learn (version 1.8.0), Seaborn (version 0.13.0), and Matplotlib (version 3.9.0). We employed a 10-fold cross-validation procedure to assess the robustness and generalization capability of the proposed model while avoiding data leakage. In this process, the entire dataset was randomly divided into 10 equal-sized and non-overlapping subsets (folds). In each round, one fold (10% of the entire data) was held out as an independent test set, while the remaining nine folds constituted the development set. The development set was further split into training and validation subsets using an 8:1 ratio, corresponding to 80% training and 10% validation data with respect to the full dataset. The training subset was used for model fitting, while the validation subset was used to monitor training (i.e., early stopping). The test fold was used solely for performance evaluation, never involved in training or validation, ensuring that no information leakage occurred. This procedure is performed 10 times, with each fold acting as the test data exactly once. The final performance is then

calculated by averaging the results across all folds, providing a comprehensive and reliable evaluation of the model.

Hyperparameter analysis was conducted as a separate set of experiments. For each hyperparameter configuration (e.g., number of convolutional layers, learning rate, number of filters, and number of folds), the entire 10-fold cross-validation procedure described above was executed independently. Similarly, RFM thresholding was computed by running the complete evaluation process for each different configuration. The preprocessing steps were fitted on the training data and then applied to the corresponding validation and test sets. The approach ensures that preprocessing and thresholding are always based solely on the training data, preventing any distribution leakage. Performance metrics were averaged across folds for each configuration, and comparisons were made between configurations based on these averaged results.

Various assessment metrics were used to evaluate how well the model classifies customers into the correct segments. These metrics include accuracy—Equation (7)—which measures the proportion of correct predictions, as well as precision, recall, and the F-measure—Equations (8)–(10)—which provide more nuanced insights into how well the model performs across different segment classes. The metrics were computed for the multi-class classification task using weighted averaging. Specifically, class-wise accuracy, precision (P_i), recall (R_i), and F-measure values were first calculated, and the final reported metrics were obtained as the weighted average across classes, where weights correspond to the number of samples in each class. This approach accounts for the possible class imbalance in segment counts.

$$\text{Weighted Accuracy (WA)} = \frac{1}{n} \sum_{i=1}^L n_i * \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad (7)$$

$$\text{Weighted Precision (WP)} = \frac{1}{n} \sum_{i=1}^L n_i * \frac{TP_i}{TP_i + FP_i} \quad (8)$$

$$\text{Weighted Recall (WR)} = \frac{1}{n} \sum_{i=1}^L n_i * \frac{TP_i}{TP_i + FN_i} \quad (9)$$

$$\text{Weighted Fmeasure (WF)} = \frac{1}{n} \sum_{i=1}^L n_i * \frac{2 * P_i * R_i}{P_i + R_i} \quad (10)$$

Here, L is the number of class labels, n_i is the number of instances in class i , while n is the total number of instances across all classes. In these formulations, True Positives (TP_i) and True Negatives (TN_i) are correct predictions of positive and negative cases for class i , respectively, while False Negatives (FN_i) and False Positives (FP_i) are incorrect predictions where the model misclassifies negative cases as positive and positive cases as negative. A confusion matrix was also generated to visualize which segments are often confused with others, offering opportunities for model improvement.

4.3. Results

Table 7 presents the performance of the classification model obtained from a 10-fold cross-validation using four key evaluation metrics: Accuracy, Recall, Precision, and F-Measure. The results demonstrated that the model delivered robust outcomes, with accuracy values ranging from 90.78% to 97.23%. On average, the model achieved an accuracy of 94.33%. Precision and recall values closely follow this accuracy trend, reflecting the strong generalizability of the model. These outcomes indicate that the classifier is not only accurate but also maintains a strong balance between sensitivity and specificity.

Table 7. Performance results of the proposed RFM-Net method across all folds.

Fold	Accuracy (%)	Precision	Recall	F-Measure
1	93.55	0.9379	0.9355	0.9354
2	95.85	0.9599	0.9585	0.9585
3	93.32	0.9381	0.9332	0.9334
4	95.16	0.9533	0.9516	0.9517
5	93.78	0.9446	0.9378	0.9383
6	90.78	0.9154	0.9078	0.9065
7	93.55	0.9396	0.9355	0.9341
8	97.00	0.9706	0.9700	0.9701
9	97.23	0.9728	0.9723	0.9723
10	93.07	0.9335	0.9307	0.9288
Avg.	94.33	0.9466	0.9433	0.9429

Figure 5 presents the confusion matrix, which shows the performance of the classification model across different classes. The model demonstrated notably high classification accuracies in customer categories. For instance, the “Champions” group was correctly classified at a rate of 94.3%, with only 5.7% of instances misclassified. Similarly, the “Need Attention” category showed strong performance with 90.5% accuracy. The model performed particularly well in identifying “Potential Loyalists” (96.9%) and “About to Sleep” customers (93.7%), with minor misclassifications distributed across adjacent segments. Overall, the matrix indicates that the classification model is effective in distinguishing between customer segments.

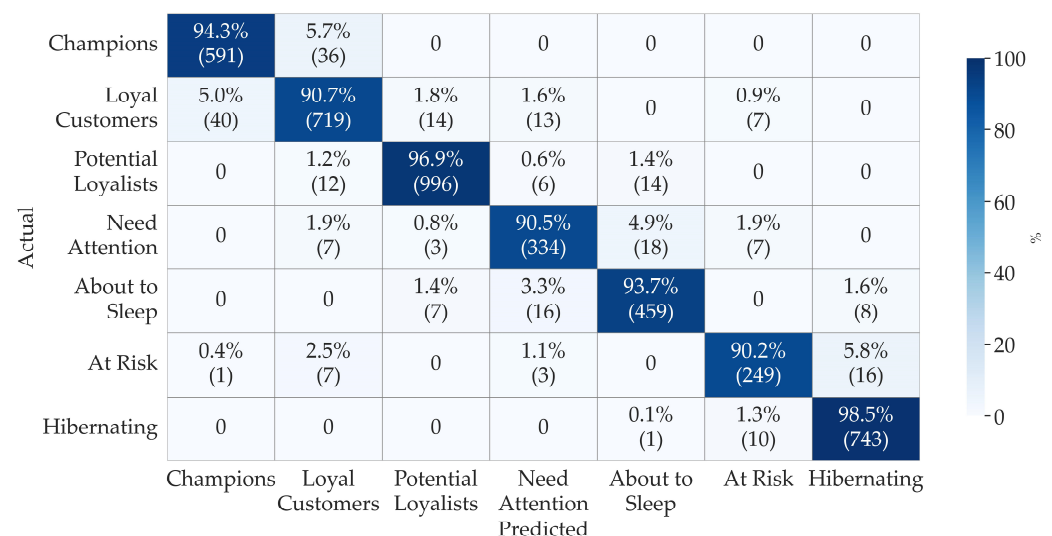
**Figure 5.** Confusion matrix.

Figure 6 illustrates the training and validation loss values over 20 epochs, indicating a consistent improvement in model performance. Both of them decrease substantially, with the training loss dropping from 0.6024 to 0.1549 and the validation loss decreasing from 0.3792 to 0.1502. This trend reflects effective learning and generalization. Notably, from epoch 15 onward, the losses become closely aligned, meaning that the model has reached a stable learning phase. The decreasing gap between training and validation loss toward the final epochs further supports the model’s robustness and its ability to generalize well on unseen data.

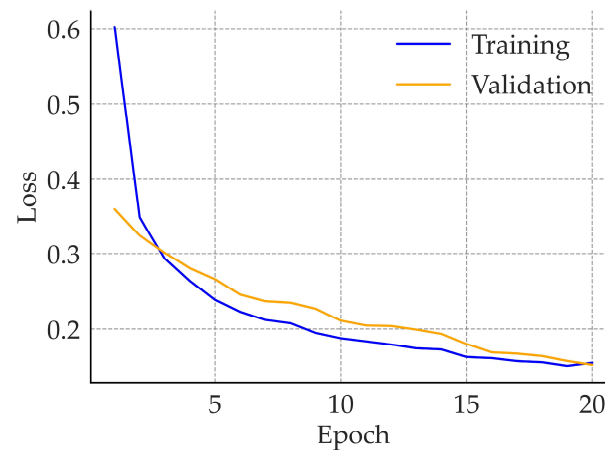


Figure 6. Training—validation loss.

4.4. Sensitivity Analysis

Table 8 presents the results of the sensitivity analysis conducted to evaluate the impact of key hyperparameters on the proposed model performance. This analysis involved systematic testing of multiple values for each parameter. For the number of convolutional layers, the sensitivity analysis explored values ranging from 2 to 7. The highest performance was observed with two layers, yielding an accuracy of 94.33%. As the number of layers increased, performance consistently declined, likely due to overfitting or redundant feature extraction. The learning rate was also tested across a range of values: 0.04, 0.03, 0.02, and 0.01. Among these, a learning rate of 0.01 achieved peak performance in all metrics. Higher learning rates negatively impacted model performance, possibly due to overshooting during the optimization process. For the K-Fold Cross Validation, using 10 folds produced better generalization compared to 5 folds, demonstrating the advantages of a more thorough validation approach. Furthermore, a filter size of 32 was found to be optimal, offering higher accuracy compared to 16 filters. It demonstrates that the model is sensitive to filter size, suggesting that a higher number of filters can enhance feature extraction without overfitting. Finally, an analysis was conducted to examine the impact of different user-defined RFM threshold values on model performance. To provide a comprehensive evaluation, various distinct threshold configurations, ranging from tighter to broader intervals, were tested to determine the optimal discretization strategy. The progressive adjustment of R, F, and M ranges enabled a detailed examination of how threshold granularity influences classification stability. As shown in Table 8, broader threshold intervals resulted in higher accuracy and a more balanced class distribution, better reflecting underlying differences in customer purchasing behavior. Overall, all these results guided the selection of hyperparameters for the final model configuration: two convolutional layers, a learning rate of 0.01, 10-fold cross-validation, and 32 filters.

4.5. Discussion

In this section, the performance of the proposed RFM-Net method is evaluated comparatively with the results reported in prior studies [16,21–34] on the same dataset. As presented in Table 9, the comparison was made using common performance metrics, including accuracy, recall, precision, and F-measure. According to the results, RFM-Net provided a consistent and significant performance advantage over all compared methods in terms of all evaluation criteria. On average, RFM-Net demonstrated a performance increase of approximately 13.17% in accuracy compared to the results reported in state-of-the-art studies. For instance, the proposed method yielded a superior result (94.33%) compared to advanced models, such as PARM (90.00%) [22] and Ret-DNN (90.00%) [25]. In

addition, RFM-Net has also surpassed ensemble learning methods such as Random Forest (87.60%) [16], Gradient Boosting (85.00%) [28], and AdaBoost (73.30%) [21]. Compared to these studies, the RFM-Net method achieved the highest results not only in terms of accuracy but also in all performance metrics, with a precision of 0.9466, a recall of 0.9433, and an F-measure of 0.9429. These results clearly demonstrated RFM-Net's superiority in processing online retail data. As shown in Table 9, standard classification algorithms, such as KNN, DT, and SVM, are more limited in capturing complex and nonlinear relationships in transactional data compared to our deep learning-based approach. In summary, RFM-Net's strong performance compared to other state-of-the-art studies validated the model's ability to distinguish critical classes.

Table 8. Sensitivity analysis for RFM-Net (Online Retail I dataset).

	Accuracy (%)	Precision	Recall	F-Measure
The Number of Convolution Layers				
2	94.33	0.9466	0.9433	0.9429
3	93.47	0.9385	0.9347	0.9339
4	92.95	0.9349	0.9295	0.9283
5	92.51	0.9293	0.9251	0.9233
6	90.96	0.9161	0.9096	0.9075
7	90.85	0.9135	0.9085	0.9083
Learning Rate				
0.04	90.06	0.9080	0.9006	0.8986
0.03	92.05	0.9273	0.9205	0.9208
0.02	92.16	0.9274	0.9216	0.9209
0.01	94.33	0.9466	0.9433	0.9429
K-Fold Cross Validation				
5	92.28	0.9306	0.9228	0.9219
10	94.33	0.9466	0.9433	0.9429
Filter				
16	93.75	0.9412	0.9375	0.9371
32	94.33	0.9466	0.9433	0.9429
Threshold Values				
R [5, 15, 25, 75] F [1, 2, 3, 4] M [250, 750, 1250, 2500]	93.38	0.9382	0.9338	0.9338
R [6, 18, 30, 90] F [1, 2, 3, 6] M [300, 900, 1500, 3000]	93.38	0.9379	0.9338	0.9335
R [7, 21, 35, 105] F [1, 2, 3, 4] M [350, 1050, 1750, 3500]	93.52	0.9396	0.9352	0.9343
R [8, 24, 40, 120] F [1, 2, 4, 5] M [400, 1200, 2000, 4000]	93.75	0.9415	0.9375	0.9368
R [9, 27, 45, 135] F [1, 2, 4, 6] M [450, 1350, 2250, 4500]	94.21	0.9442	0.9421	0.9417
R [10, 30, 50, 150] F [1, 2, 4, 6] M [500, 1500, 2500, 5000]	94.33	0.9466	0.9433	0.9429

Table 9. Comparative results of RFM-Net and previously reported methods on the same dataset.

Ref.	Authors	Year	Method	Accuracy	Precision	Recall	F-Measure
[11]	Devi et al.	2025	XGBoost	92.10	0.9170	0.8990	0.9080
			RF	87.60	0.8620	0.8350	0.8480
[21]	Verma et al.	2025	LR	70.40	0.7000	0.6950	0.7000
			KNN	72.10	0.7300	0.7100	0.7100
			SVM	73.20	0.7450	0.7200	0.7200
			DT	74.20	0.7400	0.7350	0.7400
			RF	72.80	0.7400	0.7150	0.7150
			AdaBoost	73.30	0.7450	0.7200	0.7200
			XGBoost	73.90	0.7450	0.7500	0.7200
[22]	Mohanty et al.	2025	PARM	90.00	0.8480	0.8000	0.8230
[23]	Lv	2025	Optimized BP	87.30	0.8860	-	-
			SVM	78.90	0.8010	-	-
[24]	Hussain Jafri et al.	2025	ESReLU AF	76.00	-	-	-
			Ret-DNN	90.00	-	-	-
[34]	Imani et al.	2025	Tuned_XGB_ADASYN	-	-	-	0.8000
			Tuned_XGB_SMOTE	-	-	-	0.9200
[13]	Mirzaee et al.	2024	LSTM-RNN	92.55	0.9285	0.9002	0.9111
[26]	Pushkarenko and Zaslavskyi	2024	Baseline Model	-	0.8500	0.8500	0.8500
			(SDG) Diffusion Models	-	0.9000	0.9000	0.9000
			SMOTE	-	0.8800	0.8800	0.8800
			ADASYN	-	0.8900	0.8900	0.8900
			Borderline-SMOTE	-	0.8700	0.8700	0.8700
[27]	Lalitha et al.	2024	LR	93.42	-	-	-
			RF (Estimators = 150)	86.57	-	-	-
			KNN Regression ($k = 3$)	78.82	-	-	-
[28]	Asfi et al.	2024	GBM	85.00	0.7800	0.8000	-
[29]	Xu and Hu	2023	DAPT $\rightarrow$ TAPT	74.20	0.7450	0.7280	0.7390
[30]	Mustafa et al.	2023	SVM	-	0.6292	-	-
			LR	-	0.7268	-	-
			KNN	-	0.6778	-	-
			DT	-	0.7327	-	-
			RF	-	0.7593	-	-
			AdaBoost	-	0.7021	-	-
			Gradient Boosting	-	0.7605	-	-
			Collaborative Filtering	-	0.8500	0.7500	0.7900
[31]	Loukili et al.	2023	Content-Based Filtering	-	0.7800	0.6600	0.7100
			Hybrid Filtering	-	0.8200	0.7200	0.7600
[32]	Alrawi and Ajlouni	2022	DT	77.00	0.7700	0.7700	0.7700
			RF	80.00	0.7900	0.8000	0.7900
[33]	Jana	2020	LR	84.00	0.8600	0.7800	0.8200
Average				81.16	0.7947	0.7866	0.8002
Proposed			RFM-Net	94.33	0.9466	0.9433	0.9429

Table 10 presents the performance comparison between the proposed RFM-Net and several baseline models, including logistic regression (LR), naive Bayes, multi-layer perceptron (MLP), k-nearest neighbors, AdaBoost, decision tree (DT), and a tree-based ensemble

method (Bagging (DT)). All models were evaluated under the same experimental protocol (identical preprocessing, RFM thresholds, and segment definitions) to ensure a fair comparison. The results demonstrated that RFM-Net outperformed all baseline models across all evaluated metrics. For instance, while the MLP obtained an accuracy of 85.50%, RFM-Net reached a substantially higher accuracy of 94.33%. Similarly, LR yielded an accuracy of 90.32%, which remains lower than the accuracy obtained by RFM-Net. The tree-based ensemble approach, Bagging (DT), delivered an accuracy of 90.68%, confirming the strength of ensemble-based modeling, yet remaining below the results achieved by RFM-Net. Specifically, RFM-Net improved accuracy by 6.31 percentage points compared to the average baseline performance (88.02%). These results demonstrate the effectiveness of the proposed CNN-based architecture in modeling the RFM feature interactions and improving classification performance.

Table 10. Performance comparison of RFM-Net against baseline machine learning models.

Method	Accuracy (%)	Precision	Recall	F-Measure
Logistic Regression	90.32	0.9030	0.9030	0.9030
Naive Bayes	76.46	0.7850	0.7650	0.7530
Multi-Layer Perceptron	85.50	0.8560	0.8550	0.8520
K-Nearest Neighbors	92.39	0.9240	0.9240	0.9230
AdaBoost	91.72	0.9170	0.9170	0.9170
Decision Tree (DT)	89.07	0.8950	0.8910	0.8900
Bagging (DT)	90.68	0.9120	0.9070	0.9060
Average	88.02	0.8846	0.8803	0.8777
RFM-Net (proposed)	94.33	0.9466	0.9433	0.9429

Although the input consists of only three structured features (R, F, and M), the convolutional architecture can provide advantages in learning localized interaction patterns between these features. Instead of treating R, F, and M as fully independent variables, the convolutional filters act as feature interaction extractors, capturing local dependency structures and nonlinear combinations more effectively than solely global weight updates in a standard MLP. Furthermore, the pooling mechanism enhances robustness by emphasizing dominant interaction patterns while reducing sensitivity to noise. Compared to traditional models such as Regression, which usually assume linearity in their feature space, and tree-based methods that rely on hierarchical splits, the shared-filter mechanism of the CNN acts as an implicit regularizer, which improves generalization performance.

To further strengthen the validity and generalizability of the proposed RFM-Net approach, an additional dataset was incorporated into the experimental evaluation. The publicly available Online Retail II [35] dataset, containing 1,067,371 real-world transactional records from a UK-based non-store online retailer spanning two years (2009–2011), was utilized. The same preprocessing methodology and evaluation metrics were employed to ensure consistency with the primary dataset. In the sensitivity analysis, the same hyperparameter configurations were systematically examined, except for threshold values, which were doubled because the dataset size was also twice that of the main dataset. The results are presented in Table 11. Consistent with the findings from the primary dataset, the best performance was achieved using a shallow convolutional architecture (2 layers) and a lower learning rate (0.01), whereas deeper configurations resulted in performance degradation. Similarly, a higher filter capacity (32) and broader RFM threshold intervals led to improved classification, achieving an accuracy of 95.41%. These findings confirmed that the proposed RFM-Net model maintained strong performance on large-scale, real-

world transactional data. The consistency of results across datasets further supports the robustness and generalizability of the proposed segmentation approach.

Table 11. Sensitivity analysis for RFM-Net (Online Retail II dataset).

	Accuracy (%)	Precision	Recall	F-Measure
The Number of Convolution Layers				
2	95.41	0.9563	0.9541	0.9541
3	95.33	0.9552	0.9533	0.9532
4	93.71	0.9446	0.9371	0.9327
5	92.16	0.9191	0.9216	0.9163
6	91.90	0.9258	0.9190	0.9177
7	85.51	0.8698	0.8551	0.8500
Learning Rate				
0.04	92.34	0.9314	0.9234	0.9235
0.03	93.88	0.9440	0.9388	0.9388
0.02	94.20	0.9464	0.9420	0.9420
0.01	95.41	0.9563	0.9541	0.9541
K-Fold Cross Validation				
5	94.80	0.9508	0.9480	0.9478
10	95.41	0.9563	0.9541	0.9541
Filter				
16	94.95	0.9522	0.9495	0.9493
32	95.41	0.9563	0.9541	0.9541
Threshold Values				
R [10, 30, 50, 150] F [1, 2, 4, 6] M [500, 1500, 2500, 5000]	93.44	0.9378	0.9344	0.9334
R [12, 36, 60, 180] F [1, 2, 4, 6] M [600, 1800, 3000, 6000]	93.83	0.9432	0.9383	0.9377
R [14, 42, 70, 210] F [1, 2, 4, 6] M [700, 2100, 3500, 7000]	94.00	0.9438	0.9400	0.9399
R [16, 48, 80, 240] F [1, 2, 4, 6] M [800, 2400, 4000, 8000]	94.46	0.9479	0.9446	0.9439
R [18, 54, 90, 270] F [1, 2, 4, 6] M [900, 2700, 4500, 9000]	95.23	0.9559	0.9523	0.9525
R [20, 60, 100, 300] F [1, 2, 4, 6] M [1000, 3000, 5000, 10000]	95.41	0.9563	0.9541	0.9541

One limitation of this study is that customer segment labels are generated from predefined RFM rules and subsequently used as ground truth for training the CNN, introducing a degree of circularity. However, although customer segments (labels) are derived from discrete RFM scores (1–5 scale), the CNN model is trained on raw discrete RFM values to preserve the original behavioral measures. The primary objective of employing a deep learning architecture in this context is to create a scalable framework capable of segmenting customers accurately. While segmentation is conceptually grounded in rule-based logic using RFM scores, model training is performed using actual RFM values.

5. Conclusions and Future Work

This study presents a deep-learning-based supervised customer segmentation framework named RFM-Net. In the proposed method, labels are generated through rule-based logic derived from RFM scores. The model, therefore, learns an expert-defined mapping to represent customer segmentation. This approach enables organizations to not only understand customer behavior but also make data-driven decisions for personalized marketing and retention strategies. Unlike conventional clustering or statistical techniques, RFM-Net effectively captures the complex and nonlinear patterns inherent in customer behavior, providing an accurate and actionable classification. The RFM-Net architecture is specifically designed to be lightweight and optimized for structured, low-dimensional inputs, allowing for efficient training, strong generalization, and reduced overfitting. An experimental evaluation on a real-world dataset showed that the proposed method attained a classification accuracy of 94.33%. Furthermore, RFM-Net demonstrated an average relative increase of 13.17% compared to previously reported results on the same dataset. These findings underscore the effectiveness of combining RFM and CNN techniques in driving intelligent customer segment classification solutions.

Future work may include building a web/mobile interface for deploying RFM-Net, thereby enhancing its usability in real-world business scenarios. Such an interface would enable business analysts to upload transactional data, perform an automated segmentation process, and visualize the results dynamically in a dashboard without requiring technical expertise. Implementing this application in a cloud-based platform could enhance its accessibility, scalability, and real-time insights. Furthermore, extending the application with features such as report generation, segment tracking over time, notification-based alerts, and personalized campaign recommendations could significantly enhance its practical value in real-world environments. In addition, future research may focus on improving the interpretability of RFM-Net by integrating post hoc explanation methods such as SHAP [36] and LIME [37], which provide feature-level and locally faithful explanations for individual segment predictions. Since these approaches primarily offer associational insights, incorporating causal explainable AI frameworks, particularly Fuzzy Cognitive Maps (FCMs) with total causal effect analysis [38], could further enable the identification of causal relationships among RFM variables and support more informed strategic decision-making.

Author Contributions: Conceptualization, K.F.B. and D.B.; methodology, K.F.B. and D.B.; software, D.B.; validation, K.F.B.; formal analysis, K.F.B.; investigation, K.F.B.; resources, K.F.B.; data curation, K.F.B.; writing—original draft preparation, K.F.B. and D.B.; writing—review and editing, K.F.B. and D.B.; visualization, D.B.; supervision, D.B.; project administration, D.B.; funding acquisition, K.F.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The “Online Retail” dataset [20] is publicly available in the UCI (University of California Irvine) machine learning repository (<https://archive.ics.uci.edu/dataset/352/online+retail>, accessed on 1 September 2025). The “Online Retail II” dataset [35] is publicly available in the UCI (University of California Irvine) machine learning repository (<https://archive.ics.uci.edu/dataset/502/online+retail+ii>, accessed on 6 February 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADASYN	Adaptive synthetic sampling approach
AdaBoost	Adaptive boosting
AF	Activation function
CART	Classification and regression trees
CNN	Convolutional neural network
DAPT	Domain adaptive pretraining
DBSCAN	Density-based spatial clustering of applications with noise
DT	Decision tree
ELU AF	Exponential linear unit activation function
ESReLU AF	Extended sigmoid ReLU activation function
GBM	Gradient boosting machines
GNN	Graph neural network
GNUS	Gaussian noise upsampling
GRU	Gated recurrent units
KNN	K-nearest neighbors
LeakyReLU AF	Leaky rectified linear unit activation function
LR	Logistic regression
LSTM	Long short-term memory
MLP	Multi-Layer Perceptron
NB	Naive Bayes
Optimized BP	Optimized back propagation
PARM	Para-association rule mining
RELU AF	Rectified linear unit activation function
Ret-DNN	Retail-deep neural network
RF	Random forest
RFM	Recency, frequency, and monetary
RFM-Net	Recency-frequency-monetary-based neural network
RNN	Recurrent neural networks
SMOTE	Synthetic minority over-sampling technique
SDG	Synthetic data generation
SVM	Support vector machine
SOM	Self-organizing map
STPC-PGM	Standing for spatial, temporal, payment, and product category in the probability graphic model
TAPT	Task adaptive pretraining
TARM	Traditional association rule mining
XGB	Extreme gradient boosting

References

1. Christy, A.J.; Umamakeswari, A.; Priyatharsini, L.; Neyaa, A. RFM ranking—An effective approach to customer segmentation. *J. King Saud. Univ. Comput. Inf. Sci.* **2021**, *33*, 1251–1257. [\[CrossRef\]](#)
2. Ufeli, C.P.; Sattar, M.U.; Hasan, R.; Mahmood, S. Enhancing Customer Segmentation Through Factor Analysis of Mixed Data (FAMD)-Based Approach Using K-Means and Hierarchical Clustering Algorithms. *Information* **2025**, *16*, 441. [\[CrossRef\]](#)
3. Yan, X.; Li, Y.; Nie, F.; Li, R. Bank Customer Segmentation and Marketing Strategies Based on Improved DBSCAN Algorithm. *Appl. Sci.* **2025**, *15*, 3138. [\[CrossRef\]](#)
4. Madlenak, R.; Drozdziel, P.; Zysinska, M.; Madlenakova, L. A Systems Perspective on Customer Segmentation as a Strategic Tool for Sustainable Development Within Slovakia's Postal Market. *Systems* **2025**, *13*, 701. [\[CrossRef\]](#)
5. Xiahou, X.; Harada, Y. B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. *J. Theor. Appl. Electron. Commer. Res.* **2022**, *17*, 458–475. [\[CrossRef\]](#)

6. Vrhovac, V.; Orošnjak, M.; Ristić, K.; Sremčev, N.; Jocanović, M.; Spajić, J.; Brkljač, N. Unsupervised Modelling of E-Customers' Profiles: Multiple Correspondence Analysis with Hierarchical Clustering of Principal Components and Machine Learning Classifiers. *Mathematics* **2024**, *12*, 3794. [\[CrossRef\]](#)
7. Tabianan, K.; Velu, S.; Ravi, V. K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data. *Sustainability* **2022**, *14*, 7243. [\[CrossRef\]](#)
8. Fang, C.; Liu, H. Research and Application of Improved Clustering Algorithm in Retail Customer Classification. *Symmetry* **2021**, *13*, 1789. [\[CrossRef\]](#)
9. Matuszelański, K.; Kopczewska, K. Customer Churn in Retail E-Commerce Business: Spatial and Machine Learning Approach. *J. Theor. Appl. Electron. Commer. Res.* **2022**, *17*, 165–198. [\[CrossRef\]](#)
10. Eslami, E.; Razi, N.; Lonbani, M.; Rezazadeh, J. Unveiling IoT Customer Behaviour: Segmentation and Insights for Enhanced IoT-CRM Strategies: A Real Case Study. *Sensors* **2024**, *24*, 1050. [\[CrossRef\]](#)
11. Devi, N.M.; Asha, V.; Dev, P.; Kumar, P. Customer Loyalty and Retention Analysis Using Hybrid Strategy. In Proceedings of the IEEE 3rd International Conference on Inventive Computing and Informatics (ICICI), Bangalore, India, 4–6 June 2025; pp. 472–477. [\[CrossRef\]](#)
12. Khiloun, I.R.; Belmabrouk, K.; Dekhici, L.; Bergmeir, C. Heterogeneous Graph Neural Networks for Product Recommendation on Transactional Retail Data. *Commun. Sci. Technol.* **2025**, *23*, 23–35.
13. Mirzaee, A.; Zeynali, M.; Ghorbanzadeh, A.; Ghorbanzadeh, P. Personal Recommender Model and Predicting Consumer Behavior in Digital Marketing Based on Deep Learning. *Trans. Mach. Intell.* **2024**, *7*, 179–193. [\[CrossRef\]](#)
14. Lewaaelhamd, I. Customer Segmentation Using Machine Learning Model: An Application of RFM Analysis. *J. Data Sci. Intell. Syst.* **2024**, *2*, 29–36. [\[CrossRef\]](#)
15. Liao, J.; Jantan, A.; Ruan, Y.; Zhou, C. Multi-behavior RFM model based on improved SOM neural network algorithm for customer segmentation. *IEEE Access* **2022**, *10*, 122501–122512. [\[CrossRef\]](#)
16. Talaat, F.M.; Aljadani, A.; Alharthi, B.; Farsi, M.A.; Badawy, M.; Elhosseini, M. A Mathematical Model for Customer Segmentation Leveraging Deep Learning, Explainable AI, and RFM Analysis in Targeted Marketing. *Mathematics* **2023**, *11*, 3930. [\[CrossRef\]](#)
17. Cheng, C.-H.; Chen, Y.-S. Classifying the segmentation of customer value via RFM model and RS theory. *Expert. Syst. Appl.* **2009**, *36*, 4176–4184. [\[CrossRef\]](#)
18. Akande, O.N.; Akande, H.B.; Asani, E.O.; Dautare, B.T. Customer Segmentation through RFM Analysis and K-means Clustering: Leveraging Data-Driven Insights for Effective Marketing Strategy. In Proceedings of the IEEE International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2–4 April 2024; pp. 1–8. [\[CrossRef\]](#)
19. Jauhar, S.K.; Chakma, B.R.; Kamble, S.S.; Belhadi, A. Digital transformation technologies to analyze product returns in the e-commerce industry. *J. Enterp. Inf. Manag.* **2024**, *37*, 456–487. [\[CrossRef\]](#)
20. Chen, D.; Sain, S.L.; Guo, K. Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. *J. Database Mark. Cust. Strateg. Manag.* **2012**, *19*, 197–208. [\[CrossRef\]](#)
21. Verma, R.; Rathor, D.; Kumar, S.; Mishra, M.; Baranwal, M. Enhancing Customer Repurchase Prediction: Integrating Classification Algorithms with RFM Analysis for Precision and Actionable Insights. *IIMB Manag. Rev.* **2025**, *37*, 100574. [\[CrossRef\]](#)
22. Mohanty, B.; Champati, S.L.; Barisal, S.K.X. Enhancing Retail Strategies through Anomaly Detection in Association Rule Mining. *IEEE Access* **2025**, *13*, 92376. [\[CrossRef\]](#)
23. Lv, Q. Application and optimization of BP prediction model driven by internet of things in tourism education. *Sci. Rep.* **2025**, *15*, 14698. [\[CrossRef\]](#)
24. Hussain Jafri, S.I.; Al Saedi, A.K.Z.; Elsafi, A.; Abdelguom, G.A.; Ghazali, R.; Javid, I. ESReLU: A Dynamic Activation Function for Enhancing Deep Learning Performance in Recommendations. *Int. J. Intell. Eng. Syst.* **2025**, *18*, 166. [\[CrossRef\]](#)
25. Rudro, R.A.M.; Uddin, M.H.; Aurnob, M.J.A.; Razzaque, R.; Nur, K. Ret-DNN: Predictive Analytics in Retail-An Enhanced Deep Learning Model for Customer Behavior Analysis. *Int. J. Comput.* **2025**, *18*, 1–14. [\[CrossRef\]](#)
26. Pushkarenko, Y.; Zaslavskiy, V. Synthetic Data Generation for Fraud Detection Using Diffusion Models. *Inf. Secur.* **2024**, *55*, 185–198. [\[CrossRef\]](#)
27. Lalitha, S.; Gupta, T.S.; Thelu, A.; Reddy, V.K.; Reddy, V.C. Predictive Modeling for Real-Time Customer Lifetime Value. In Proceedings of the IEEE 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 24–28 June 2024; pp. 1–6. [\[CrossRef\]](#)
28. Asfi, M.; Warsito, B.; Wibowo, A. Enhancing Explainable AI: Leveraging SHAP for Transparent Decision-Making in Machine Learning. In Proceedings of the IEEE Ninth International Conference on Informatics and Computing (ICIC), Medan, Indonesia, 24–25 October 2024; pp. 1–6. [\[CrossRef\]](#)
29. Xu, N.; Hu, C. Enhancing E-Commerce Recommendation using Pre-Trained Language Model and Fine-Tuning. *arXiv* **2023**, arXiv:2302.04443. [\[CrossRef\]](#)

30. Mustafa, S.M.N.; Akhtar, A.; Noronha, J.T.P.; Salman, M.; Baig, M.A. Customer segmentation using machine learning techniques. In Proceedings of the IEEE International Multi-disciplinary Conference in Emerging Research Trends (IMCERT), Karachi, Pakistan, 4–5 January 2023; pp. 1–7. [\[CrossRef\]](#)
31. Loukili, M.; Messaoudi, F.; El Ghazi, M. Personalizing product recommendations using collaborative filtering in online retail: A machine learning approach. In Proceedings of the IEEE International Conference on Information Technology (ICIT), Amman, Jordan, 9–10 August 2023; pp. 19–24. [\[CrossRef\]](#)
32. Alrawi, A.H.; Ajlouni, N. Intelligent Machine Learning Customer Segmentations Algorithm. *Manch. J. Artif. Intell. Appl. Sci.* **2022**, *3*, 1–11.
33. Jana, A.K. A Machine Learning Framework for Predictive Analytics in Personalized Marketing. *J. Artif. Intell. Mach. Learn. Data Sci.* **2020**, *1*, 560–564. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Imani, M.; Beikmohammadi, A.; Arabnia, H.R. Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels. *Technologies* **2025**, *13*, 88. [\[CrossRef\]](#)
35. Chen, D. UCI Machine Learning Repository. Online Retail II Dataset. Available online: <https://archive.ics.uci.edu/dataset/502/online+retail+ii> (accessed on 6 February 2026).
36. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 1–10.
37. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144. [\[CrossRef\]](#)
38. Tyrovolas, M.; Kallimanis, N.D.; Stylios, C. Advancing explainable AI with causal analysis in large-scale fuzzy cognitive maps. *arXiv* **2024**, arXiv:2405.09190. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.