Segmentation Analysis and Customer Churn Prediction in Delivery Services Using RFM, K-Means, and Random Forest for Retention Strategy Planning

Toni Prabowo, Muhammad Syahputra Novelan

Abstract

Fierce competition in the delivery and logistics service industry demands targeted customer retention strategies to prevent customer churn that threatens revenue margins. However, customer management practices typically treat customer value segmentation and churn prediction as separate, isolated tasks, leading to inefficient allocation of retention resources. This study aims to integrate transaction data-driven descriptive and predictive modeling to identify customer value segments, predict churn risk, and formulate a prescriptive retention strategy matrix at the JNE Binjai Timur 2 Agency. The secondary dataset comprises 19,389 clean transactions from 4,625 unique customers spanning January 2, 2024, to December 31, 2025. We extracted, log-transformed, and standardized Recency, Frequency, and Monetary (RFM) features using StandardScaler, then applied unsupervised K-Means clustering. To address churn imbalance (defined as inactivity >90 days), class weighting was applied during Random Forest Classifier training. The K-Means clustering results (k=2, Silhouette Score = 0.512) validated two value tiers: Cluster Tier 1 (High-Value, 17.47% of customers) accounted for 69.47% of total revenue (IDR 444.08 million), while Cluster Tier 2 (Low-Value, 82.53% of customers) contributed 30.53% of revenue. The Random Forest model showed superior discriminative performance on the test set (N = 822), with an AUC-ROC of 0.830, Accuracy of 91.61%, Recall of 99.59%, and F1-Score of 95.52%, with Active_Days (0.1580) and Recency (0.1312) as the dominant predictors. Integrating value tiers and churn predictions revealed potential financial exposure within the High-Risk group of IDR 243.52 million (38.09% of total revenue). This study yields four managerial retention intervention priorities (P1–P4), including immediate win-back offers for High-Value High-Risk customers and priority VIP/pick-up services for High-Value Low-Risk customers.

Keywords: Delivery Services, Customer Churn, RFM Segmentation, K-Means Clustering, Random Forest, Class Weighting, Retention Strategy.

Toni Prabowo[1]
[1,2] Master of Information Technology, University Pembangunan Panca Budi, Medan, Indonesia
e-mail: toniprabowohsb@gmail.com[1]

Muhammad Syahputra Novelan[2]
e-mail: putranovelan@dosen.pancabudi.ac.id[2]
3rd International Conference Epicentrum of Economic Global Framework (ICEEGLOF)
Theme: Cross-Sector and Cross-Disciplinary Strategies for Achieving the SDGS
https://proceeding.pancabudi.ac.id/index.php/ICEEGLOF

Introduction

Machine learning algorithms and computational modeling have proven effective in mapping user perceptions, preferences, and satisfaction levels within digital ecosystems. This approach has proven reliable, as shown by the use of Naive Bayes and K-Nearest Neighbors (K-NN) to predict customer satisfaction [1] and analyze mobile application reviews [2]. The combination of lexicon-based methods and Naive Bayes has also been shown to automatically classify public opinion across various evaluation domains [3], [4].

The success of such modeling reinforces the vital role of consumer behavior-based data analytics in accurately responding to business dynamics and economic issues. When analyzing public responses to economic challenges such as inflation, Naive Bayes efficiently groups sentiment patterns using concise training datasets [5]. Meanwhile, for complex, lengthy narrative texts, Bidirectional Long Short-Term Memory (BiLSTM) architectures can map consumer perception patterns and demand gaps with precision [6]. These developments show that understanding customer behavior is a crucial predictive tool in strategic business decision-making.

Machine learning-based consumer behavior analytics is becoming increasingly critical in the express delivery and logistics industry, which is characterized by intense competition and narrowing profit margins driven by the rapid growth of the digital economy [7]. In this dynamic business climate, strengthening customer retention strategies serves as the primary foundation for safeguarding long-term revenue stability [7], [8]. Financially, acquiring new customers costs 5 to 25 times more than retaining existing ones [9]. Consequently, customer churn poses a severe threat that directly reduces revenue margins, wastes marketing budgets, and damages corporate reputation [8], [10].

The delivery service sector possesses unique dynamics, including fluctuating transaction frequencies, high sensitivity to pricing and service quality, and a pronounced customer propensity to switch providers [7]. Nevertheless, most firms still use coarse-grained management practices that neglect heterogeneity in individual customer value [7]. This unsegmented approach risks inefficient retention resource allocation and may lead to the attrition of high-value customers due to a lack of focused, data-driven interventions [7].

To address the challenges of undifferentiated customer management, Recency, Frequency, and Monetary (RFM) analysis maps customer transaction recency, frequency, and monetary value [11], [12]. Integrating RFM features with unsupervised machine learning algorithms such as K-Means clustering has proven effective in grouping customers into homogeneous segments without requiring initial class labels [11], [12]. Although K-Means is optimal for establishing value hierarchies, a purely clustering-based approach cannot proactively predict future churn risk [7]. Therefore, supervised ensemble learning methods such as Random Forest, which offer high precision in predicting churn across complex datasets, are required [7], [8].

Despite the respective advantages of both methods, prior literature highlights a critical research gap: customer segmentation and churn prediction are predominantly executed as separate, isolated tasks [7]. Pure segmentation models fail to detect impending attrition signals, whereas standalone churn prediction models lack the economic context of customer value [7]. Without an integrated framework, organizations either waste retention resources on low-value customers or delay intervention for high-value customers at risk of churning [7], [8].

To bridge this gap, this study proposes an integrated framework that simultaneously combines RFM modeling, K-Means clustering, and Random Forest classification within the delivery services industry. This framework extracts RFM features from historical transactions, clusters customer value segments via K-Means, and classifies churn risk levels using Random Forest. The primary objective of this research is to identify critical customer cohorts specifically high-value customers with high churn risk enabling management to develop differentiated retention strategies, prioritize resource allocation, and strengthen sustainable customer loyalty [7], [8].

Literature Review

In the highly competitive digital economy, understanding customer behavior is a fundamental pillar of business sustainability and consumer loyalty [13]. One established analytical framework for evaluating customer transaction patterns is the Recency, Frequency, and Monetary (RFM) model [13]. This model assesses three key dimensions: Recency measures the time since the most recent transaction, Frequency measures transaction regularity within a specified period, and Monetary measures the customer's total financial contribution [13], [15]. Integrating these three dimensions converts raw transaction logs into structured behavioral profiles to identify engagement levels and Customer Lifetime Value [13], [16]. Because RFM variables have disparate scales and units, data preprocessing via normalization techniques such as Min-Max Scaling or RobustScaler is essential to ensure balanced weighting across variables without undue dominance by large numerical features [13], [15].

Manual customer segmentation based on RFM scores is often subjective and computationally inefficient for large-scale datasets; thus, unsupervised machine learning algorithms such as K-Means Clustering are widely used [13]. K-Means Clustering partitions customer data into k homogeneous segments by minimizing within-cluster variance (Within-Cluster Sum of Squares) and maximizing inter-cluster separation [13], [15]. To determine the optimal number of clusters, the Elbow Method is commonly combined with Silhouette Score evaluations and the Davies-Bouldin Index to ensure statistically compact and well-separated cluster structures [13], [15], [17]. Combining RFM and K-Means has successfully categorized customer bases into distinct strategic segments, such as Loyal Customers, Potential Loyalists, At-Risk Customers, and Lost Customers [13], [15], [17].

Customer churn, defined as the cessation of service utilization by customers, poses a critical threat to corporate revenue and profitability [16], [18], [19]. Management literature consistently underscores that retaining existing customers is far more cost-effective than acquiring new ones [19]. To proactively anticipate customer defection, supervised machine learning models are deployed to classify churn risk based on historical transaction logs [16]. Among diverse algorithms, Random Forest, based on ensemble bagging, performs well because it aggregates decisions across many decision trees, reduces overfitting, and handles complex variable interactions [18], [19]. Class imbalance, a common issue in churn modeling, can also be addressed through resampling techniques such as SMOTE or SMOTEENN, combined with feature selection, to optimize F1-score performance [18], [19].

Integrating descriptive segmentation based on RFM and K-Means with predictive prescriptive modeling based on Random Forest establishes a comprehensive Customer Relationship Management (CRM) framework [13], [14]. RFM segmentation provides contextual understanding of customer economic profiles and value, while the Random Forest algorithm precisely pinpoints individuals at high churn risk within those segments [15], [18], [19]. This data-driven synergy empowers management to design targeted retention strategies, such as providing tailored incentives for Loyal Customers and executing reactivation campaigns for At-Risk Customers before they switch to competitors [13], [15], [18], [19].

Research Methodology

3.1 Framework and Research Flow

This study applies an integrated analytical framework combining descriptive modeling (unsupervised machine learning) and predictive modeling (supervised machine learning) on transaction data from the JNE Binjai Timur 2 Agency. This framework is designed to overcome the limitations of isolated analyses by mapping the economic value hierarchy of customers while simultaneously predicting churn risk with high precision. The research workflow was systematically executed through the following stages:

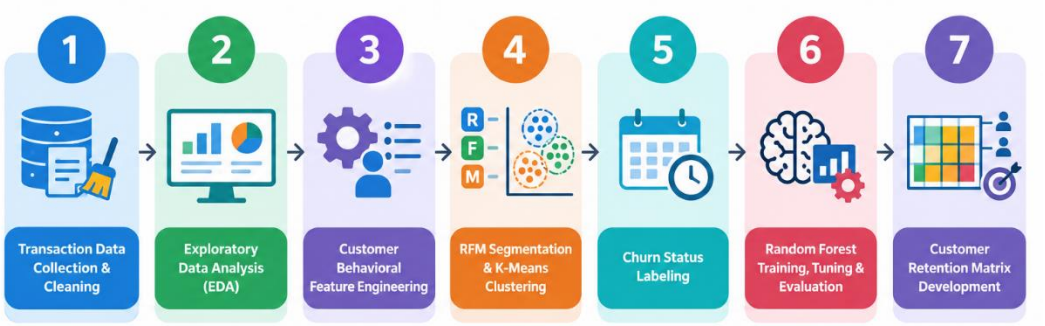


Figure 1. Research Flow

1. Transaction Data Collection & Cleaning: Collect transaction data and handle missing values, duplicates, and inconsistencies.
2. Exploratory Data Analysis (EDA): Analyze transaction patterns, customer behavior, and data distributions.
3. Customer Behavioral Feature Engineering: Create relevant customer features while preventing data leakage.
4. RFM Segmentation & K-Means Clustering: Segment customers based on Recency, Frequency, and Monetary value using rule-based RFM and K-Means clustering.
5. Churn Status Labeling: Define customer churn status using a time-based observation window.
6. Random Forest Training, Tuning & Evaluation: Train and optimize the Random Forest classifier and evaluate its predictive performance.
7. Customer Retention Matrix Development: Combine customer value segments and churn probabilities to develop targeted retention strategies.

3.2 Research Object and Data Sources

The research object uses secondary transaction data from the JNE Binjai Timur 2 Agency spanning January 2, 2024, to December 31, 2025. The data-cleaning procedure removed 1,125 invalid records (missing values, negative monetary amounts, duplicate airway bills, and unidentifiable customer IDs) from an initial pool of 20,514 raw records. The filtering yielded 19,389 clean transactions across 4,625 unique customers. Key variables include tracking/airway bill number, date, service type, weight, quantity, destination, customer ID, payment method, insurance status, and shipping fee.

3.3 Data Preprocessing and RFM Feature Extraction

Data preprocessing began with text standardization, numeric datatype conversion, aggregating rare payment methods into an "OTHER" category, and encoding the insurance variable as a binary variable $Insurance_Used \in \{0, 1\}$. To prevent data leakage, customer behavioral features were dynamically extracted using a fixed reference date (*snapshot date* / T_{ref}). From the transaction logs prior to T_{ref}, the three core Recency, Frequency, and Monetary (RFM) variables were computed:

Recency **(R_i)**: Time elapsed (in days) between customer i most recent transaction (T_{last}, i) and the cutoff date (T_{ref}):

$$R_i = T_{\text{cutoff}} - T_{\text{last},i} \quad (1)$$

Frequency **(F_i)**: Cumulative count of unique transactions (airway bills) completed by customer i:

$$F_i = \sum_{j=1}^{N_i} 1 \quad (2)$$

Monetary (M_i): Cumulative total shipping expenditures made by customer i:

$$M_i = \sum_{j=1}^{N_i} \text{Cost}_j \quad (3)$$

In addition to the three baseline RFM variables, feature engineering yielded 13 auxiliary features to enrich the representation of transaction behavior, including: Avg_Amount (mean cost per transaction), Total_Weight (total parcel weight), Avg_Weight, Total_Qty, Insurance_Rate (proportion of insured shipments), N_Destination (count of distinct destination cities), N_Services (count of distinct service types used), Active_Days (number of unique transaction days), Tenure (customer relationship duration), *Trx_Interval* (*Tenure*/F_i), *Recency_Ratio* (R_i/(Trx_Interval+1)),

as well as activity indicators for the trailing 90-day window (Freq_Recent and Monet_Recent) and payment method proportions (Pay_CASH, Pay_COD, etc.).

3.4 Customer Segmentation Using K-Means Clustering

K-Means segmentation clusters the customer base into homogeneous segments in an unsupervised manner. Given that the empirical distributions of R, F, and M variables in logistics typically exhibit significant right-skewness, the RFM features were first transformed using the natural logarithm $\ln(1 + x)$ to suppress the influence of extreme outliers:

$$X_{\text{trans}} = \ln(1 + X_{\text{RFM}}) \quad (4)$$

Subsequently, the transformed variables were standardized using Z-score scaling (StandardScaler) to equalize variable scales and dispersions prior to Euclidean distance computation:

$$Z = \frac{X_{\text{trans}} - \mu}{\sigma} \quad (5)$$

The determination of the optimal number of clusters (*k*) was evaluated over the range $k \in [2, 8]$ by combining the Elbow Method based on the reduction in Within-Cluster Sum of Squares (WCSS):

$$\text{WCSS} = \sum_{k=1}^{K} \sum_{x \in C_k} |x - \mu_k|^2 \quad (6)$$

and the Silhouette Coefficient (S_i):

$$S_i = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (7)$$

where a(i) denotes the mean intra-cluster distance and b(i) denotes the mean nearest inter-cluster distance. The resulting clusters (k=2, Silhouette Score = 0.512) were then ranked based on a composite value $\text{Score} = \text{Scaled}_F + \text{Scaled}_M - \text{Scaled}_R$ to establish customer value tiers (Cluster Tier 1: High-Value and Cluster Tier 2: Low-Value).

3.5 Time-Based Churn Window Definition

To ensure statistical validity and prevent data leakage, churn status was defined using a non-contractual temporal approach. The cutoff date was set 90 days before the terminal observation date (CHURN_WINDOW_DAYS = 90), on October 3, 2025. All predictive features were strictly extracted from transactions occurring prior to this cutoff date. Customers were labeled as having churned (*Churn* = 1) if they engaged in zero transactions between the cutoff date and the end of the observation window (October 3, 2025 – December 31, 2025); conversely, customers who continued transacting were labeled as active (*Churn = 0*).

3.6 Handling Class Imbalance

Within the training dataset (N = 4,106), transaction dormancy produced a highly skewed class distribution: 89.8% churned customers (N = 3,686) and 10.2% active customers (N = 420). This imbalance was addressed using Cost-Sensitive Learning by enabling the automated class weighting parameter (class_weight="balanced_subsample") within the Random Forest algorithm. Class weights were computed inversely proportional to class frequencies across individual bootstrap samples, imposing heavier penalties on minority-class classification errors.

3.7 Model Training and Hyperparameter Optimization

The dataset was partitioned into training (80%) and testing (20%) subsets using Stratified Random Sampling (N_{test} = 822). The Random Forest Classifier was trained as a bagging ensemble with hyperparameter optimization performed via RandomizedSearchCV (25 iterations) and 5-Fold Stratified Cross-Validation. The optimal parameter combination obtained comprised: n_estimators = 500, max_depth = 5, min_samples_split = 5, min_samples_leaf = 4, and max_features = 'log2'. The final classification decision was determined via majority voting:

$$\hat{Y} = \arg max_c \sum_{b=1}^{B} I(T_b(x) = c) \quad (8)$$

where *B = 500* decision trees, $T_b(x)$ represents the prediction of the *b*, and *I(.)* denotes the indicator function.

3.8 Optimal Decision Threshold Determination

To maximize minority class detection on imbalanced data, the decision threshold was adjusted away from the default 0.5 cutoff. We objectively identified an optimal threshold $\text{best_thr} = 0.114$ based on the Precision-Recall/F1-Score curve derived from Out-of-Fold Cross-Validation on the training set.

3.9 Model Performance Evaluation and Retention Strategy Formulation

We evaluated Random Forest classification performance on the independent test dataset (N=822) using a Confusion Matrix, Accuracy, Precision, Recall, F1-Score, and Area Under the ROC Curve (AUC-ROC), alongside Gini Feature Importance analysis to quantify predictor contributions. Predicted churn probabilities were subsequently categorized into three risk tiers: High Risk ($P \geq 0.70$), Medium Risk ($0.40 \leq P < 0.70$), and Low Risk ($P < 0.40$). Cross-tabulation between customer value segments (K-Means Tier / RFM) and churn risk tiers was mapped into a prescriptive retention matrix. This integrated framework underpins the estimation of financial exposure (revenue at risk) and the operational formulation of four managerial retention priorities (P1–P4).

Results

4.1 Data Description and Exploratory Data Analysis (EDA)

This study utilized secondary transaction records from the JNE Binjai Timur 2 Agency covering January 2, 2024, to December 31, 2025. Data cleaning filtered 19,389 valid transactions from 20,514 raw records by removing 1,125 invalid entries, encompassing 4,625 unique customers. Longitudinal EDA showed monthly active customers growing from about 190 (early 2024) to over 400 (late 2025). Operationally, shipments were dominated by routes to Jakarta and regional destinations across North Sumatra and Aceh, primarily using the REG23 (Regular) service and cash payment methods (CASH).

4.2 Rule-Based RFM Customer Segmentation

Extracting Recency (R), Frequency (F), and Monetary (M) metrics across the 4,625 unique customers produced population averages of 331.51 days for Recency, 3.10 shipments for Frequency, and IDR 118,783.77 in Monetary expenditure per customer, with total cumulative agency revenue reaching IDR 639,248,530.00. Rule-based segmentation utilizing RFM quantiles divided the customer base into seven behavioral cohorts, as detailed in Table 1.

Table 1. Distribution and Characteristics of Rule-Based RFM Customer Segments

RFM Segment	Customer Count	Proportion (%)	Average R (Days)	Average F (Orders)	Total Monetary (IDR)	Revenue Contribution (%)
Champions	940	20.32%	85.22	10.19	IDR 323,091,050	50.54%
At Risk	966	20.89%	534.83	6.55	IDR 155,751,570	24.36%
Loyal Customers	869	18.79%	223.03	1.88	IDR 68,276,530	10.68%
New / Promising	603	13.04%	105.90	1.00	IDR 30,655,980	4.80%
Lost	459	9.92%	627.14	1.00	IDR 23,926,200	3.74%
Hibernating	425	9.19%	453.71	1.00	IDR 21,527,180	3.37%
Potential Loyalist	363	7.85%	290.73	1.00	IDR 16,020,020	2.51%
Total	4,625	**100.00%**	**331.51**	**3.10**	**IDR 639,248,530**	**100.00%**

The RFM findings confirm the empirical validity of the Pareto Principle (80/20): the Champions segment (20.32% of customers) accounts for the absolute majority of agency revenue at 50.54% (IDR 323.09 Million). Conversely, the At Risk cohort is a critical concern: it comprises the largest single population share (20.89% of customers) and contributes a substantial 24.36% of revenue (IDR 155.75 Million), yet shows an extremely high mean Recency (534.83 days), indicating severe, unaddressed attrition risk.

4.3 Customer Segmentation and Validation Using K-Means Clustering

To validate customer value tiers through an objective, data-driven approach, we applied unsupervised K-Means clustering to the RFM features after natural log transformation $ln(1+x)$ and StandardScaler standardization. Evaluating cluster solutions over $k \in [2,8]$ demonstrated that k=2 achieved the peak Silhouette Score of 0.512, indicating optimal separation and structural cohesion. The profile of the two value tiers produced by K-Means clustering is detailed in Table 2.

Table 2. K-Means Cluster Profiles of JNE Binjai Timur 2 Agency Customers (k=2)

Cluster Label	Value Tier	Customers	Proportion (%)	Average R (Days)	Average F (Orders)	Average M (IDR)	Total Monetary (IDR)	Revenue (%)
Cluster Tier 1	High-Value	808	17.47%	129.67	18.18	549,600.77	444,077,420	69.47%
Cluster Tier 2	Low-Value	3,817	82.53%	349.98	1.23	51,132.07	195,171,110	30.53%
Total / Average	—	4,625	**100.00%**	**311.46**	**4.19**	**138,215.90**	**639,248,530**	**100.%**

Cluster Tier 1 (High-Value) represents the core customer segment. Although this cohort comprises only 808 accounts (17.47% of total customers), it generates 69.47% of total agency revenue (IDR 444.08 million), averaging 18.18 transactions per customer. In contrast, Cluster Tier 2 (Low-Value) encompasses the vast majority of accounts (82.53%) but contributes only 30.53% of revenue (IDR 195.17 Million), with an average transaction frequency of 1.23. The cross-tabulation between rule-based RFM segments and K-Means Cluster Tiers in Table 3 shows that 594 of 940 Champions fall into Tier 1, while 892 of 966 At Risk customers fall into Tier 2. Critically, however, 74 At Risk customers were identified within Cluster Tier 1, demonstrating that high-value financial contributors are currently undergoing severe transaction delays.

Table 3. Cross-Tabulation Matrix of RFM Segments against K-Means Clusters

RFM Segment	Cluster Tier 1 (High-Value)	Cluster Tier 2 (Low-Value)	Total Customers
Champions	594	346	940
At Risk	74	892	966
Loyal Customers	109	760	869
New / Promising	30	573	603
Lost	0	459	459
Hibernating	1	424	425
Potential Loyalist	0	363	363
Total	808	3,817	4,625

4.4 Churn Prediction Modeling Using Random Forest and Performance Evaluation

Applying the 90-day time-window churn definition (CHURN_WINDOW_DAYS = 90) prior to the October 3, 2025 cutoff on the training dataset (N = 4,106) produced an inactive class distribution of 89.8% churned customers (3,686 individuals) and 10.2% active customers (420 individuals). This class imbalance was compensated for by enabling Cost-Sensitive Learning (class_weight="balanced_subsample") in the Random Forest model. The model was trained using the optimal hyperparameters (n_estimators: 500, max_depth: 5, min_samples_split: 5, min_samples_leaf: 4, max_features: 'log2') identified via RandomizedSearchCV. The optimal decision threshold was established at *best_thr* = 0,114 based on the maximum F1-Score on Out-of-Fold Cross-Validation. The classification performance on the test dataset (N = 822) is summarized in Table 4.

Table 4. Classification Performance Random Forest Model on Test Data (N = 822)

Evaluation Metric	Decimal Value	Percentage (%)	Operational Description
Accuracy	**0.9161**	**91.61%**	Proportion of total correct classifications (TP+TN = 753 / 822)
Precision (Churn)	**0.9176**	**91.76%**	Accuracy of positive churn predictions (TP / (TP+FP) = 735 / 801)
Recall (Sensitivity)	**0.9959**	**99.59%**	Proportion of actual churners correctly captured (TP / (TP+FN) = 735 / 738)
F1-Score (Churn)	**0.9552**	**95.52%**	Harmonic mean of Precision and Recall

Evaluation Metric	Decimal Value	Percentage (%)	Operational Description
AUC-ROC	**0.8303**	—	Model discriminative power across various decision thresholds

Based on the test set Confusion Matrix (N = 822), out of 738 actual churned customers, the model correctly identified 735 (True Positive / TP = 735) incurring a minimal Type II error rate with only 3 missed churners (False Negative / FN = 3). Among the 84 active customers, 18 were correctly classified (*True Negative* / TN = 18) while 66 were flagged as False Positives (FP = 66). In non-contractual logistics, a Recall of 99.59% is strategically vital, ensuring that nearly all customers prone to switching providers are flagged early. Concurrently, high Precision (91.76%) and Accuracy (91.61%) prevent excessive false alarms, safeguarding retention budgets against wasteful deployment. Gini Feature Importance analysis identified the top five churn predictors, as shown in Table 5.

Table 5. Top Five Predictors (Top Feature Importance) of the Random Forest Model

Rank	Predictor Variable	Impurity Weight (Gini Importance)	Operational Description
1	Active_Days	0.1580	Total unique days on which the customer transacted
2	Recency	0.1312	Elapsed time (days) since the customer's last transaction
3	Frequency	0.1252	Cumulative total of unique transactions (airway bills)
4	Total_Qty	0.1023	Cumulative total number of packages/items dispatched
5	N_Destination	0.0845	Number of unique destination cities/districts

The prominence of temporal metrics (Active_Days and Recency) shows that, in delivery service operations, declining engagement serves as an earlier and more sensitive attrition indicator than a reduction in cumulative monetary expenditure.

4.5 Value Segmentation and Churn Prediction for Retention Strategies

To eliminate inefficiencies stemming from isolated analyses, churn probabilities estimated by the Random Forest classifier were partitioned into three risk tiers (High: $P \geq 0.70$, Medium: $0.40 \leq P < 0.70$, Low: $P < 0.40$) and crossed with historical spending metrics (Revenue at Risk), as shown in Table 6.

Table 6. Distribution of Churn Risk Levels and Financial Exposure (Revenue at Risk)

Churn Risk Level	Probability Range	Customer Count	Customer Proportion (%)	Revenue at Risk / Total Monetary (IDR)	Revenue Contribution (%)
High Risk	$P \geq 0.70$	2,766	59.81%	IDR 243,525,970	38.09%
Medium Risk	$0.40 \leq P < 0.70$	1,219	26.36%	IDR 94,129,510	14.73%
Low Risk	$P < 0.40$	640	13.83%	IDR 301,593,050	47.18%
Total	—	**4,625**	**100.00%**	**IDR 639,248,530**	**100.00%**

This integration reveals that the High-Risk group constitutes an immediate financial exposure of IDR 243.52 Million (38.09% of cumulative agency turnover). By synthesizing customer value tiers (K-Means Tiers) with churn risk tiers, four actionable managerial retention priorities (P1–P4) were established:

1. Priority 1 (P1 - High-Value High-Risk / "Immediate Recovery"): Focuses on Cluster Tier 1 accounts exhibiting elevated churn probability. Interventions involve direct outreach by a Dedicated Account Manager, personalized win-back offers, expedited grievance resolution, and tailored volume-contract pricing schedules.
2. Priority 2 (P2 - Low-Value High-Risk / "Cost-Efficient Reactivation"): Addresses high-risk accounts in Cluster Tier 2. Interventions rely on low-cost mass reactivation campaigns executed via automated messaging channels (WhatsApp/Email) delivering time-delimited

shipping discount vouchers.

3. Priority 3 (P3 - Medium Risk / "Monitor & Frequency Stimulation"): Centers on medium-risk cohorts (Potential Loyalists). Prescriptive actions encompass loyalty point promotions, service bundle discounts, and periodic automated transaction reminders.
4. Priority 4 (P4 - High-Value Low-Risk / "Retain & Expand"): Targets low-risk Cluster Tier 1 accounts (active Champions). Initiatives include dedicated VIP services (priority parcel pick-up), expedited claim processing, and flexible payment terms (Term of Payment / credit lines).

Conclusions

This study validates that integrating descriptive segmentation (RFM Analysis and K-Means Clustering) with predictive modeling (Random Forest Classifier) effectively overcomes the pitfalls of isolated analyses, creating a robust prescriptive customer retention framework for the JNE Binjai Timur 2 Agency. K-Means clustering (k=2, Silhouette Score = 0.512) confirmed the Pareto Principle: Cluster Tier 1 (High-Value, 808 customers) generates 69.47% of revenue (IDR 444.08 Million) with an average of 18.18 transactions, while Cluster Tier 2 (Low-Value, 3,817 customers) comprises 82.53% of the customer base with a 30.53% revenue contribution (IDR 195.17 Million). In predictive modeling, establishing a 90-day churn threshold paired with Cost-Sensitive Learning and an Optimal Decision Threshold (best_thr = 0,114) achieved exceptional diagnostic performance on unseen test data (N = 822): Accuracy 91.61%, Precision 91.76%, Recall 99.59%, F1-Score 95.52%, and AUC-ROC 0.8303. Temporal variables, specifically Active_Days (0.1580) and Recency (0.1312), were the most decisive churn predictors. Cross-referencing segmentation with risk modeling revealed a financial exposure of IDR 243.52 million (38.09% of turnover) within the High-Risk group and enabled the formulation of four targeted managerial retention priorities (P1–P4) designed to safeguard corporate revenue margins.

References

[1] M. S. Novelan and M. Iqbal, "Comparative Analysis of the Performance of K-Nearest Neighbor (K-NN) and Naive Bayes Algorithms on User Satisfaction Levels of the Tokopedia Application," *Proc. Int. Conf. Comput. Sci. Eng. Soc. Sci. Multi-Disciplinary Stud.*, vol. 1, pp. 218–224, Nov. 2025, doi: 10.64803/cessmuds.v1.37.

[2] S. Susilawati and M. Iqbal, "Penerapan Metode Naïve Bayes Untuk Mengidentifikasi Sentimen Pengguna Pada Ulasan Aplikasi ReelShort di Google Play Store," *SIMKOM*, vol. 10, no. 1, pp. 49–59, Jan. 2025, doi: 10.51717/simkom.v10i1.686.

[3] M. Saputra and M. Iqbal, "Analisis Sentimen Masyarakat Terhadap PON XII Aceh-Sumut Menggunakan Algoritma Naïve Bayes," *SIMKOM*, vol. 10, no. 1, pp. 39–48, Jan. 2025, doi: 10.51717/simkom.v10i1.679.

[4] M. R. Tanjung, M. Iqbal, and Z. Sitorus, "Analisis Sentimen Google Review terhadap Mutu Kualitas Pendidikan pada Perguruan Tinggi STIE Al-Washliyah Sibolga dengan Metode Lexicon dan Algoritma Naive Bayes," *J. Multimed. Dan Teknol. Inf.*, vol. 07, no. 02, pp. 400–412, 2025, doi: doi.org/10.54209/jatilima. v7i02.1549.

[5] S. N. Sofyan and M. Iqbal, "Analisis Sentimen Terhadap Dampak Inflasi Menggunakan Naive Bayes," *Bull. Inf. Technol.*, vol. 6, no. 1, pp. 1–8, 2025, doi: 10.47065/bit.v5i2.1796.

[6] Z. Sitorus, M. Iqbal, D. Nasution, and R. Farta Wijaya, "Penerapan Deep Learning dan Analisis Sentimen terhadap Gap Kompetensi Lulusan Lembaga Pendidikan dan Pelatihan Vokasi terhadap Dunia Kerja dengan Metode Long Short-Term Memory (LSTM)," *Bull. Inf. Technol.*, vol. 6, no. 2, pp. 161–172, 2025, doi: 10.47065/bit.v5i2.2029.

[7] Y. Wu, H. Zhang, and H. Wang, "Customer Segmentation and Churn Prediction in Express Logistics," *Acad. J. Manag. Soc. Sci.*, vol. 15, no. 2, pp. 41–46, Apr. 2026, doi: 10.54097/p1zxfn50.

[8] P. Boozary, S. Sheykhan, H. GhorbanTanhaei, and C. Magazzino, "Enhancing customer retention with machine learning: A comparative analysis of ensemble models for accurate churn prediction," *Int. J. Inf. Manag. Data Insights*, vol. 5, no. 1, p. 100331, Jun. 2025, doi: 10.1016/j.jjimei.2025.100331.

[9] M. Imani, M. Joudaki, A. Beikmohammadi, and H. Arabnia, "Customer Churn Prediction: A Systematic Review of Recent Advances, Trends, and Challenges in Machine Learning and Deep Learning," *Mach. Learn. Knowl. Extr.*, vol. 7, no. 3, p. 105, Sep. 2025, doi: 10.3390/make7030105.

[10] G. Kumar S and R. Kannan, "A Study on the Transportation Industry Customer Churn," *J. Contemp. Res. Manag.*, vol. 19, no. 2, pp. 36–52, 2024.

[11] R. M. Fauzan and G. Alfian, "Segmentasi Pelanggan E-Commerce Menggunakan Fitur Recency, Frequency, Monetary (RFM) dan Algoritma Klasterisasi K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 3, pp. 170–177, Sep. 2024, doi: 10.14421/jiska.2024.9.3.170-177.

[12] A. L. M. Tampubolon, T. M. E. Y. Butar Butar, and S. Rochimah, "Segmentasi Pelanggan Majalah pada Situs Web E-Commerce dengan K-Means++ dan Metode RFM," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 6, pp. 1243–1252, Dec. 2024, doi: 10.25126/jtiik.2024118208.

[13] Y. Syahra, A. Fadlil, and H. Yuliansyah, "Customer Segmentation Using RFM and K-Means Clustering to Support CRM in Retail Industry," *sinkron*, vol. 9, no. 3, pp. 1120–1131, Jul. 2025, doi: 10.33395/sinkron.v9i3.14907.

[14] D. Saputra and U. Subagyo, "Analisis Segmentasi Pelanggan E-Commerce India Berbasis Rfm Dan Profitabilitas Menggunakan K-Means," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 10, no. 4, pp. 6946–6953, Aug. 2026, doi: 10.36040/jati.v10i4.18813.

[15] D. S. Nugraha, I. Thoib, N. Sururi, F. Bayu F, and B. P. Candra, "Segmentasi Pelanggan Berbasis Analisis RFM Menggunakan Algoritma K-Means Clustering," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 2, pp. 1361–1369, Sep. 2025, doi: 10.62712/juktisi.v4i2.641.

[16] G. Mena, K. Coussement, K. W. De Bock, A. De Caigny, and S. Lessmann, "Exploiting time-varying RFM measures for customer churn prediction with deep neural networks," *Ann. Oper. Res.*, vol. 339, no. 1–2, pp. 765–787, Aug. 2024, doi: 10.1007/s10479-023-05259-9.

[17] P. M. K. Mado and H. Hendry, "Implementasi Algoritma Clustering K-Means untuk Segmentasi Pelanggan di E-Commerce," *J. Indones. Manaj. Inform. dan Komun.*, vol. 6, no. 3, pp. 1680–1686, Sep. 2025, doi: 10.63447/jimik.v6i3.1563.

[18] R. E. Ako *et al.*, "Effects of Data Resampling on Predicting Customer Churn via a Comparative Tree-based Random Forest and XGBoost," *J. Comput. Theor. Appl.*, vol. 2, no. 1, pp. 86–101, Jun. 2024, doi: 10.62411/jcta.10562.

[19] R. Krishna, D. Jayanthi, D. S. Shylu Sam, K. Kavitha, N. K. Maurya, and T. Benil, "Application of machine learning techniques for churn prediction in the telecom business," *Results Eng.*, vol. 24, no. September, p. 103165, Dec. 2024, doi: 10.1016/j.rineng.2024.103165.