STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
*Stat., Optim. Inf. Comput.*, Vol. **16**, September 2026, pp 2039–2057.
Published online in International Academic Press (www.IAPress.org)

# A Hybrid Machine Learning and Rule-Based Recommendation Framework: A Probabilistic Approach for Reliable Customer Next-Transaction Prediction

Giovanny Theotista [1], Moch. Fandi Ansori [2,*]

[1] *Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Bandung, Indonesia*
[2] *Department of Mathematics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia*

**Abstract** This study presents a hybrid recommendation framework that combines machine learning-based probabilistic prediction with rule-based decision logic for forecasting customers' next transactions in digital banking. It aims to help in customer retention and personalization, as well as cross-sell strategies in settings where the behavior of the user is heterogeneous and transaction data are largely imbalanced. The machine learning component estimates behavioral consistency probabilities, whereas the rule-based component translates these probabilities into personalized product recommendations according to customer segmentation and business constraints. The principal methodological contribution is the proposed inverse transform probability mechanism, which reinterprets predicted probabilities as behavioral consistency scores across active, dormant, and new customer segments. This conversion makes it possible to have a uniform probabilistic understanding of the states of the users while maintaining the economic significance of the recommendation decisions. The model performance was evaluated using the area under the precision-recall curve (PR-AUC) because it is a good performance metric for imbalanced classification problems. The empirical results demonstrate good predictive ability and temporal robustness, achieving an average PR-AUC of 0.82 over the evaluation period. Furthermore, the probabilistic framework is in line with the concept of expected monetary value (EMV) under a binomial state-price interpretation and links the prediction of behavior to the financial decision-making process. The proposed framework provides an interpretable probabilistic methodology for next-transaction prediction under realistic synthetic banking scenarios and establishes a foundation for future validation using real banking transaction data. The results obtained on synthetic transaction data suggest that the inclusion of behavioral consistency in probabilistic models can enhance the stability and practicality of recommending systems for financial institutions.

**Keywords** Rule-based recommendation framework; Next-transaction prediction; Behavioral consistency; Imbalanced classification; Digital banking analytics.

**AMS 2010 subject classifications** 62H30, 68T05, 91G80

**DOI:** 10.19139/soic-2310-5070-3707

## 1. Introduction

The modern banking system operates in an environment characterized by rapid digitalization, intense competition, and high volumes of transactions [1, 2]. Therefore, financial institutions must create data-driven strategies to better understand customer behavior and improve the cross-selling potential of banking products. In countries such as Indonesia, which are in the process of developing digital economies, recent economic policy directions have highlighted the importance of stimulating domestic financial activity and ensuring improved liquidity circulation

*Correspondence to: Moch. Fandi Ansori (Email: mochfandiansori@lecturer.undip.ac.id). Department of Mathematics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

ISSN 2310-5070 (online) ISSN 2311-004X (print)

within the financial sector. Consequently, banks are heavily investing in analytical models that can help them detect latent behavioral patterns and suggest the right financial products to their customers.

Customer transaction behavior is inherently heterogeneous. While some customers have consistent and active engagement levels with financial activity, others have sporadic interactions or are dormant for a long time [3]. Such heterogeneity naturally leads to class imbalance and temporal inconsistency in transaction datasets, which makes the tasks of predictive modeling and recommendations extremely difficult [4, 5]. Traditional predictive approaches often center on active customers and assume that transaction probabilities can be directly deduced from past activity [6, 7, 8]. However, this is a less reliable assumption when consumers' financial behavior is dynamically altered by policy changes, market conditions, or seasonal economic cycles.

Several studies have examined cross-selling and customer behavior in financial services. Kaishev et al. [9] proposed a probabilistic optimization model for selecting customers in the cross-selling model, and Altun and Yucekaya [10] developed models to maximize the yield of financial products. Machine learning techniques have been studied in the past to predict customer behavior. For example, Boustani et al. [11] revealed that deep learning designs could boost the prediction capabilities of cross-selling contexts, whereas Basten and Juelsrud [12] discovered that customer loyalty and the continuity of enduring relations were significant influences in banking systems in terms of price and product acceptance.

Theoretically, the mathematical modeling of the financial decision-making process has been emphasized. The modelling of decision-making of the bank's behavior in channelling loans has been studied in a discrete-time structure [13, 14] and systemic risk framework [15]. An alternative decision-making model related to value-at-risk was used to predict the risk measure in energy risk data [16]. The importance of this combination of behavioral interpretation and probabilistic modelling in financial analytics is worth giving attention to these approaches.

Despite recent advances in machine learning for financial recommendation, relatively little attention has been given to probabilistic models that explicitly capture behavioral consistency while maintaining interpretability under heterogeneous customer behavior and severe class imbalance. To address this gap, this study proposes a hybrid machine learning and rule-based recommendation framework for customer next-transaction prediction. The main contributions are threefold: (i) the introduction of a behavioral consistency-based prediction framework that models stable customer behavior rather than isolated transaction occurrences; (ii) an inverse transform probability formulation that provides a unified probabilistic interpretation across different customer behavioral states; and (iii) a longitudinal evaluation that demonstrates robust predictive performance under realistic synthetic banking scenarios. These contributions provide an interpretable probabilistic foundation for personalized recommendation in digital banking. The proposed framework is also consistent with probabilistic classification under imprecise conditions and contributes to modern data-driven decision support systems in digital banking [6, 8].

The remainder of this paper is organized as follows. Section 2 presents the proposed methodology, including data simulation, feature engineering, model training, inverse transform probability formulation, and evaluation framework. Section 3 reports the experimental results and discusses the performance of the proposed recommendation system. Finally, Section 4 discusses the results, and Section 5 concludes the paper and outlines future research directions.

## 2. Methodology

### *2.1. Synthetic Data Generation*

In this study, simulated transactional data were used to emulate actual customer behavior in digital banking. The dataset spans 32 months (January 2023-August 2025) and is intended to reproduce the trends of the macroeconomy and the heterogeneity of the behavioral data according to published financial data by Bank Indonesia [17], the Indonesia Financial Service Authority [18], McKinsey & Company [3], Google-Temasek-Bain [19], and the Boston Consulting Group (BCG) [20]. The temporal stability and probabilistic consistency of the rule-based recommendation framework were evaluated based on this data.

The initial stage involves assuming the statistical and structural building blocks of the simulation dataset. To ensure initial complete reproducibility, all random number generators were initialized with a known seed

(`random_state = 610`). The fixed random seed ensures that the synthetic dataset can be reproduced exactly under the same implementation. The source code used to generate the synthetic dataset and reproduce the experimental pipeline is available from the corresponding author upon reasonable request. The customer base was seeded with 500 distinct customer accounts and increased every month following a Gaussian distribution, where $\mu = 40$ and standard deviation $\sigma = 5$, to indicate a low customer acquisition trend, which would be typical of early stage digital financial institutions.

A normal distribution with parameters 38 and 5 transactions per month was used to model the transactional frequency per customer. These parameters are in line with the empirical evidence of the Federal Reserve Bank of Atlanta [21] and the Federal Reserve Bank of St. Louis [22], which discovered between 31 and 45 monthly transactions per active consumer. This spectrum in the surging digital payment adoption in Indonesia is also affirmed by additional observations by the Indonesian Financial Service Authority [23]. This probabilistic structure is assured to ensure that the growth of customers and the transacting activity are stabilized in time, realistic variability in the performance, and macroeconomic rationality.

Twenty types of digital transactions were simulated, including day-to-day banking operations such as fund transfers, QR payments, bill payments, and investment-based products such as mutual fund investments and time deposits. Some investment-related products depend on three financial indicators: CASA (current and savings account balance), AUM (assets under management), and OSBAL (operational savings balance). These indicators represent the financial ability of the customer and were only selectively attached to investment-type transactions to preserve causal integrity of the simulation.

To mimic realistic temporal changes in customer transaction behavior, the simulation also includes a seasonal probability modeling mechanism that varies the transaction-type distribution based on recurring macroeconomic and behavioral cycles. The probability vectors for each month were generated using a Dirichlet distribution to obtain an even stochastic distribution among all product categories, which was then weighted by season. Seasonal modulation of cyclic intensity variations for specific types of products. For example, donations and credit card payments have a higher frequency during the year-end period (November to December), corresponding to charity campaigns and holiday expenses. Insurance premium payments are high in January and February because of the annual renewal of the policy. Similarly, mobile top-ups, foreign currency transfers, etc., surge around mid-year holiday periods (June–July), while mutual fund investments and time deposits increase in the period of March–April and September, when corporate bonuses and religious holiday allowances are dispensed. After the seasonal weighting factors, all probability vectors were renormalized to ensure that the total probability mass was equal to one.

To simulate customer base growth and product diversification, new customers and transaction volumes were introduced every month. The number of new customer information files (CIFs) was between approximately 2,500 and 3,000 per month, and the total monthly transactions were between 10,000 and 50,000. Each customer was assigned a stochastic financial profile sampled from a log-normal distribution. Transaction amounts were generated conditionally: investment-linked products were proportional to AUM or OSBAL (between 1% and 20%), while general transactions were sampled from smaller log-normal distributions with right-skewed variability (mean 10.5, standard deviation 1.5). Each transaction was randomly timestamped in its month and later aggregated at the end of the month (`date_pr`) to form a structured time series representation.

To simulate the conditions of banking data in the real world, behavioral sparsities were introduced by using controlled missingness. Two layers of sparsity were employed in this study. First, random month omission was applied using up to 12 random months for each CIF to drop months so that natural dormancy cycles for each CIF were generated. Second, omission was randomly used at the transaction level to remove approximately 5% of transactions to create incomplete records of activities. This two-sided solution left the unbalanced dataset natural since inactive behavior was the most prevalent, as it could be viewed as a realistic banking data setting where active customers are a minority. These imbalanced situations are essential for determining the soundness of the model in imbalanced learning theory [4, 5, 24, 25].

After combining the data from all the months, the final dataset consisted of approximately 290,000 records and eight key variables. Table 1 summarizes the structure of the dataset and the variable descriptions.

Table 1. Description of variables in the transaction dataset

| Variable | Description |
|---|---|
| `cif` | Unique customer identifier |
| `product_type` | Type of transaction |
| `transaction_time` | Actual transaction timestamp |
| `amount` | Transaction value |
| `casa_balance`, `aum_balance`, `osbal_balance` | Financial indicators |
| `date_pr` | Monthly reference period (period-end) |

The dataset was saved as `dummy_data_transaction.csv` to maintain reproducible experiments and further model training. Beyond its numeric characteristics, the dataset was designed to capture three important properties of the behavior of digital banking: temporal variability in customer behavior caused by economic cycles and the structure of the season, wealth stratification in terms of heterogeneous financial balances, and natural class imbalance caused by irregular customer engagement. Together, these characteristics provide a real environment, but a controlled one, to test the probabilistic inverse transform learning and PR-AUC-based consistency in a rule-based recommender system. With the data structure finalized, the next stage was to engineer some of the features to capture temporal and behavioral consistency.

The objective of the synthetic data generation is not to reproduce the confidential transaction records of a particular financial institution, but rather to preserve the principal statistical characteristics observed in digital banking, including heterogeneous customer behavior, seasonal transaction patterns, product-specific eligibility constraints, wealth stratification, and natural class imbalance. Because customer-level banking data are inaccessible owing to privacy restrictions, validation was performed by comparing the aggregate behavioral characteristics of the simulated data with publicly available banking statistics and industry reports. Although rare customer behaviors, fraudulent activities, unexpected economic shocks, and institution-specific operational practices cannot be fully replicated, the resulting synthetic dataset provides a realistic, reproducible, and controlled experimental environment for evaluating the proposed recommendation framework.

### *2.2. Feature Engineering*

The processing of raw transaction histories into statistical and behaviorally interpretable predictors is known as feature engineering. This process follows the concepts of interpretable and transparent learning [26, 8, 27]. The first phase of feature extraction is temporal aggregation on short and long horizons.

Each customer product was examined as a time series to understand the multi-horizon trends of transactions. The transactional dataset was transformed into a monthly panel for each customer-product combination. For each combination $(i, j, t)$ corresponding to customer $i$, product $j$, and month $t$, historical summaries of the transaction amounts and activity flags were calculated to develop the analytical base. Short-term (3-month) and long-term (12-month) rolling windows were used in the study to capture the persistence and volatility of behavior. Multi-horizon design gives the model the ability to identify not only stable engagement patterns but also short-term anomalies caused by market dynamics or a change of policy.

Based on these aggregate measures, several behavioral features were created to describe the variability and stability of customer activities. Core-derived features are the basis for the representation of behavior. The engineered variables for each product target are listed in Table 2.

The selected features were designed to capture complementary aspects of customer financial behavior across different temporal scales. The variables recent and sum3m represent immediate customer responsiveness and short-term transaction momentum, whereas slope, avg12m, and std12m characterize the long-term behavioral trends, habitual engagement, and transaction stability. Together, these features enable the proposed probabilistic model to distinguish persistent behavioral patterns from temporary fluctuations while maintaining financial interpretability.

In addition to transaction-level variables, liquidity-related financial variables, such as OSBAL, CASA, and AUM, were included to record behavior-specific variables in retail investors. These are indicators whereby the application

Table 2. Behavioral features constructed from transaction history

| Feature | Description | Behavioral Purpose |
|---|---|---|
| *recent* | Transaction value in the last month | Captures immediate responsiveness |
| *slope* | 12-month exponential smoothing trend | Detects gradual shifts |
| *std12m* | Standard deviation of 12-month history | Quantifies volatility |
| *avg12m* | Average of 12-month transaction values | Represents habit strength |
| *sum3m* | Total transaction value over 3 recent months | Measures short-term momentum |

was only done on investment-type transactions to reduce spursiosity and preserve the interpretability within the financial environment.

Cross-selling and target integrity rules were introduced to ensure that the recommendation logic was realistic. Specifically, it is not possible to recommend product $j$ to customer $i$ in the actual period $t$ because of practical marketing and preservation of probabilistic validity [9, 10]. This ensures that only eligible products are recommended, which is formally defined as follows:

$$D_{\text{valid}} = \{(i, j, t) \mid y_{i,j,t} = 0,\ j \in P\}.$$

The hybrid feature structure ensures short-term responsiveness and long-term interpretability. Short-term features, such as *recent* and *sum3m*, reflect instant reactions to events or marketing campaigns, whereas long-term indicators, such as *slope*, *avg12m*, and *std12m*, reflect stable behavioral tendencies and financial stability. This two-scale representation increases the strength and understandability of the estimated probabilities in different behavioral states, which is part of the stability objective of the rule-based recommendation framework.

After the analytical feature set was established, the model training and selection processes were performed, which are detailed in the following sections.

### *2.3. Model Development and Selection*

The model training framework focuses on the fairness, balance, and interpretability of the algorithms. After the feature engineering phase is completed, this phase helps build the complete machine learning pipeline that will be used to test the probabilistic consistency under class imbalance.

To ensure adaptive learning and temporal robustness, model training was performed at the end of the month from July 2024 to June 2025, which corresponds to a 12-month rolling evaluation window before the main recommendation period (August 2024–July 2025). Each month, 20 models for each specific product corresponding to different types of transactions were trained and validated independently, resulting in 240 trained models throughout the entire horizon. The probabilistic consistency over time is ensured by this scheme of retraining every month so that the system adapts continuously to the changing customer behavior and the macroeconomic conditions.

Rather than predicting individual transaction occurrences, the proposed framework learns whether customers maintain consistent behavioral states over time. Accordingly, the target variable represents behavioral consistency between two consecutive months.

A target variable, which is the consistency of behavioral repetitions between one month and another, was developed before the application of the inverse transform probability formulation. Let $x_{i,j,t}$ be a binary transaction indicator of the customer $i$ and product $j$ in the month $t$, and thus $x_{i,j,t} = 1$ when the transaction is made and $x_{i,j,t} = 0$ otherwise. The behavioral consistency $Y_{i,j,t}$ is assigned as follows:

$$Y_{i,j,t} = \begin{cases} 1, & \text{if } x_{i,j,t-1} = x_{i,j,t}, \\ 0, & \text{if } x_{i,j,t-1} \neq x_{i,j,t}. \end{cases}$$

For illustration, suppose a customer uses a mobile payment service in January and again in February. In this case,

$$x_{t-1} = 1, \qquad x_t = 1,$$

so that

$$Y_t = 1,$$

This indicates consistent behavior. Similarly, if the customer does not use the service in either month,

$$x_{t-1} = 0, \qquad x_t = 0,$$

then

$$Y_t = 1,$$

This also represents behavioral consistency because the customer's inactive state is maintained. In contrast, if the customer changes from using the service to not using it, or vice versa,

$$(1, 0) \text{ or } (0, 1),$$

then

$$Y_t = 0,$$

indicating a behavioral transition between consecutive months.

A value of 1 implies consistent behavior (Buy $\rightarrow$ Buy or Not Buy $\rightarrow$ Not Buy), and 0 implies inconsistent behavior (Buy $\rightarrow$ Not Buy or Not Buy $\rightarrow$ Buy). This formulation aids a model to acquire stable behavior in contrast with raw transaction frequency, which supports the inverse transform probability framework to ensure that the model is interpretable within the active, dormant, and new customer sectors.

Standardizing and stabilizing the dataset prior to model training was performed by data pre-processing. The preprocessing pipeline guarantees normalization, feature stability, and financial outlier robustness. A summary of the transformations performed on the data is as follows:

**Drop Columns** $\rightarrow$ **Numeric Imputer** $\rightarrow$ **Outlier Clipper** $\rightarrow$ **Robust Scaler** $\rightarrow$ **L1-regularized feature selection** $\rightarrow$ **Mutual Information**.

The modular architecture assists in diminishing the impacts of noise and extreme values of transactions and enhances reproducibility and interpretability [28, 29, 30].

Mutual Information (MI) was employed as a filter-based feature selection method to quantify the statistical dependency between each candidate predictor and the behavioral consistency target. Candidate features were ranked according to their Mutual Information scores, and the highest-ranked features were retained for model training. Since the recommendation framework is trained independently for each transaction product and retrained monthly, the subset of retained features may differ across products and retraining periods according to the observed data characteristics.

Transaction-history features (e.g., *recent*, *sum3m*, *slope*, *avg12m*, and *std12m*) constitute the primary behavioral descriptors because they capture customer engagement over multiple temporal horizons. Financial indicators such as CASA, AUM, and OSBAL are incorporated only for investment-related products, where they provide additional information regarding customers' financial capacity and investment behavior. Consequently, the proposed framework combines behavioral history with financial characteristics whenever such information is economically meaningful.

Following the preprocessing, a benchmarking step was undertaken to compare the performance of the two baseline models with regard to the balanced weighting of the classes. Five predictive algorithms were compared.

1. Logistic Regression [31, 32]
2. Random Forest Classifier [33, 34]
3. Gradient Boosting [35, 36]
4. XGBoost [37, 38]
5. LightGBM [39, 40]

Stratified 3-Fold Cross Validation was applied to each model, and the area under the precision-recall curve (AUCPR) was the primary measure applied because of its sensitivity to skewed class distributions [4, 5, 41]. The two models with the highest mean AUCPR were shortlisted for optimization and testing. This strategy ensures that the process of tuning the most promising algorithms is performed without compromising their computational performance.

To overcome the class imbalance tendency of transactional datasets, three complementary resampling strategies were considered [25, 42]:

- Random Under-Sampling (RUS): removes the extra majority samples;
- Random Over-Sampling (ROS): creates duplicates of minority samples;
- Synthetic Minority Oversampling Technique (SMOTE): Generates new minority examples by using the nearest-neighbour interpolation technique.

Each resampling method was integrated separately into the preprocessing pipeline to assess the effect of each technique on the recall-precision trade-offs and model stability. However, including these sampling methods increases the ability of the model to generalize under severe imbalances, which are typical in banking data [24].

For each product target $j$, the selected models were evaluated based on a hold-out test set. The best model with the highest mean AUCPR across the cross-validation folds was selected according to the following:

$$M^* = \arg\max_{M_k \in \mathcal{M}} \frac{1}{K} \sum_{k=1}^{K} \mathrm{AUCPR}_{M_k},$$

where $\mathcal{M}$ is the candidate model set and $K$ is the number of cross-validation folds.

Hyperparameter optimization using `RandomizedSearchCV` [43] was then performed to maximize the AUCPR on the different folds. The search space consisted of the number of features to retain following mutual information filtering, sampling policies, and part of the key algorithmic parameters, such as $C$, $n_{\text{estimators}}$, $max\_depth$, and $learning\_rate$. Performance improvements were measured as follows:

$$\Delta AUCPR = AUCPR_{\text{tuned}} - AUCPR_{\text{benchmark}}.$$

The final model selection rule is defined as

$$M^* = \begin{cases} M_{\text{tuned}}, & \text{if } \Delta AUCPR > 0, \\ M_{\text{benchmark}} & \text{otherwise.} \end{cases}$$

This adaptive selection mechanism ensures an appropriate balance between predictive accuracy, temporal robustness, and interpretability [25]. The final configuration is selected based on the principle of selective retention. If the tuned configuration improves the precision-recall performance while keeping the AUCPR stable over the 12-month evaluation window, it is adopted as the final model; otherwise, the balanced weight benchmark model is retained. This strategy focuses on the long-term interpretability and temporal stability of the model rather than marginal improvements in numerical quality, which is in line with responsible machine learning practices [41, 25, 42, 44, 45].

It should be noted that the objective of the present study is to evaluate the probabilistic prediction component that underlies the proposed recommendation framework rather than to introduce a new recommendation ranking algorithm. The final recommendation list is constructed by combining the predicted probabilities with the behavioral segmentation and business rule constraints described in the next subsection. Consequently, the experimental comparison focuses on alternative supervised learning models for probability estimation, including Logistic Regression, Random Forest, Gradient Boosting, XGBoost, and LightGBM. Comparisons with recommendation-specific methods, such as collaborative filtering, matrix factorization, and popularity-based recommendation algorithms, constitute an interesting direction for future work once real customer transaction datasets become available.

From a computational perspective, the proposed framework trains one predictive model for each product category during every monthly update cycle. Because these models are independent, the training process is naturally

parallelizable and can be distributed across multiple processing cores or computing nodes. Furthermore, model retraining is performed offline, whereas the online recommendation stage only requires probability prediction and rule-based ranking, resulting in a relatively low inference cost. Monthly retraining was selected as a practical compromise between adapting to evolving customer behavior and limiting the computational overhead. In large-scale banking environments with substantially larger customer populations, distributed computing or incremental learning techniques could further improve scalability while preserving the proposed recommendation framework.

### *2.4. Behavioral Consistency Transformation*

After model selection, the probabilistic framework was reorganized to preserve the interpretability across heterogeneous behavioral states. Let:

$$p_{i,j,t+1} = \text{Predict\_Prob}(y_{i,j,t+1} = 1 \mid x_{i,j,t})$$

The predicted probability $p_{i,j,t+1}$ measures the likelihood that customer $i$ performs transaction $j$ in the next period. However, the recommendation framework proposed in this study aims to quantify behavioral consistency rather than merely the occurrence of transactions. Therefore, the prediction should reflect whether the future behavior is consistent with the customer's current behavioral state.

Let $Y_{i,j,t} \in \{0, 1\}$ denote the customer's current behavioral state, where $Y_{i,j,t} = 1$ indicates that the customer is currently active for product $j$ and $Y_{i,j,t} = 0$ otherwise. Behavioral consistency is achieved when the predicted behavior agrees with the current behavioral state. Consequently, the probability of behavioral consistency can be written as

$$P(\text{Consistency}) = Y_{i,j,t}p_{i,j,t+1} + (1 - Y_{i,j,t})(1 - p_{i,j,t+1}).$$

Since $Y_{i,j,t}$ is binary, this expression simplifies to

$$\hat{p}_{i,j,t+1} = \begin{cases} p_{i,j,t+1}, & Y_{i,j,t} = 1, \\ 1 - p_{i,j,t+1}, & Y_{i,j,t} = 0, \end{cases}$$

which is precisely the inverse transform probability adopted in this study. This transformation defines a behavioral consistency score derived from the predicted probabilities. It is not intended to recalibrate the probability estimate but rather to reinterpret it with respect to the customer's current behavioral state. Active customers receive recommendations with the highest consistency likelihood, whereas dormant or new customers receive suggestions emphasizing globally underutilized products [8, 11, 12]. The transformation should therefore be interpreted as a behavioral consistency score rather than a calibrated probability estimate.

Let $\mathcal{J}$ denote the set of all products and $y_{i,j,t} \in \{0, 1\}$ denote the usage indicator of customer $i$ for product $j$ in month $t$. Let $V_{i,j,t} \geq 0$ denote the transaction volume or frequency for customers $i$ and product $j$ at time $t$. Global product usage is denoted by $u_{j,t} \geq 0$, which represents the number of users, counts or total transaction volume. The normalized global contribution of product $j$ is defined as

$$c_{j,t} = \frac{u_{j,t}}{\sum_{k \in \mathcal{J}} u_{k,t}} \in [0, 1],$$

where smaller values indicate globally underutilized products in a given country, and vice versa.

At the customer level, diversification is measured using the personal portfolio share

$$s_{i,j,t} = \frac{V_{i,j,t}}{\sum_{k \in \mathcal{J}} V_{i,k,t}} \in [0, 1],$$

where smaller values indicate products that are relatively underutilized by the customers. Eligibility constraints are captured through a business mask $\text{eligible}(i, j, t) \in \{0, 1\}$, with the eligible product set defined as

$$\mathcal{J}_{i,t}^{\text{elig}} = \{j \in \mathcal{J} : \text{eligible}(i, j, t) = 1\}.$$

The recommendation list length is denoted by $K$ (here, $K = 5$).

Customers are segmented at time $t$ into three mutually exclusive behavioral groups. A customer is classified as *Active* if at least one product is used during the period, i.e.,

$$\exists j \in \mathcal{J} : y_{i,j,t} = 1.$$

A customer is classified as *Dormant/New* if no product usage is observed,

$$\forall j \in \mathcal{J} : y_{i,j,t} = 0.$$

A customer is classified as *Heavy* if the customer is active and either product coverage or transaction intensity exceeds the predefined thresholds:

$$\frac{|\{j \in \mathcal{J} : y_{i,j,t} = 1\}|}{|\mathcal{J}|} \geq \tau_{\text{cov}} \quad \text{or} \quad \sum_{j \in \mathcal{J}} V_{i,j,t} \geq \tau_{\text{vol}}.$$

For heavy users, the recommendations prioritize customer-level diversification. Products with the smallest personal portfolio shares are selected first.

$$\mathcal{L}_1 = \text{Bottom\_K} \left\{ s_{i,j,t} : j \in \mathcal{J}_{i,t}^{\text{elig}} \right\}.$$

If fewer than $K$ items are obtained, the remaining slots are filled with globally underutilized products as follows:

$$\mathcal{L}_2 = \text{Bottom\_K} \left\{ c_{j,t} : j \in \mathcal{J}_{i,t}^{\text{elig}} \setminus \mathcal{L}_1 \right\}.$$

The final recommendation list is therefore

$$\text{Rec}_i^{\text{heavy}}(t+1) = \mathcal{L}_1 \cup (\mathcal{L}_2 \text{ up to } K - |\mathcal{L}_1|).$$

Optionally, a seasonality-aware adjustment can be applied prior to computing global contribution scores. In this case, global usage $u_{j,t}$ is replaced by the seasonally adjusted quantity $u_{j,t}/s_{j,t}$ before normalization, producing a seasonality-adjusted score $c_{j,t}^{(\text{sea})}$. To maintain consistency awareness when ties occur in ranking $s_{i,j,t}$ or $c_{j,t}$, ties are resolved by prioritizing products with higher predicted probabilities $\hat{p}_{i,j,t+1}$.

Overall, customers are segmented into *Active*, *New/Dormant*, and *Heavy* categories. Active customers receive Top-$K$ recommendations ranked by $\hat{p}_{i,j,t+1}$ (excluding products used at $t$). Dormant or new customers get Bottom-$K$ globally underutilized products derived from $c_{j,t}$. Heavy users are given diversification-based recommendations, prioritizing underutilization on the individual's part and filling the remaining slots with globally underutilized products.

This is parallel to the idea of the binomial Expected Monetary Value (EMV) [46], where the gain/loss states are symmetrically under the expectation of probabilities:

$$EMV = pG + (1-p)L.$$

Here, $\hat{p}$ will be similar to $p$, except that it will consider the stability of behavior and not the probability that the occurrence of the raw outcomes will take place. This formulation preserves the financial interpretability of the probabilistic transformation while maintaining the consistency between behavioral prediction and monetary decision-making. The next stage is to assess whether this consistency is empirically valid using longitudinal performance measures.

From a practical perspective, the transformed probability should be interpreted as a behavioral consistency score rather than a calibrated probability estimate. Unlike conventional calibration techniques, such as Platt Scaling or Isotonic Regression, which modify predicted probabilities to better match observed event frequencies, the proposed transformation preserves the original probability estimates and reinterprets them according to the customer's current behavioral state. Consequently, active and inactive customers can be evaluated within a unified probabilistic framework while maintaining both interpretability and compatibility with any underlying prediction algorithm.

### *2.5. Performance Evaluation*

System evaluation was performed monthly between July 2024 and June 2025, leading to longitudinal PR–AUC curves for the measurement of temporal robustness. The evaluation period spans multiple recurring transaction cycles, allowing the monthly retraining strategy to adapt to evolving customer behavior under changing macroeconomic and seasonal conditions. However, seasonality is captured implicitly through repeated model updating rather than through explicit seasonal predictors.

Because the dataset has a natural distribution of imbalance, in which active transactions occur rarely, PR–AUC was chosen as the main evaluation metric [4, 5, 41, 47]. This choice is supported by He and Ma [24], who showed that the PR–AUC is more accurate in detecting discrimination towards minority classes than the PR–AUC in imbalanced learning scenarios. Before introducing the empirical results, the mathematical formulation of the precision/recall evaluation framework is presented.

To calculate the precision and recall, a classification threshold $\tau$ used, and the values are as follows:

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad \text{Recall} = \frac{TP}{TP + FN}.$$

The continuous precision- recall area under the curve is defined as

$$\text{PR-AUC} = \int_0^1 \text{Precision}(\text{Recall})\, d(\text{Recall}).$$

In discrete predictions, this integral is approximated in a Riemann sum of ranked recall levels:

$$\text{PR-AUC} \approx \sum_{k=1}^{n-1} \text{Precision}_k \left(\text{Recall}_{k+1} - \text{Recall}_k\right),$$

where $(\text{Precision}_k, \text{Recall}_k)$ are the sampled coordinates of precision and recall in increasing recall. In the case of imbalanced datasets, the baseline precision is equal to the prevalence of the positive class as follows:

$$\pi_+ = \frac{N_+}{N_+ + N_-}.$$

A classifier that performs randomly produces PR-AUC $\approx \pi_+$. Therefore, any PR–AUC value that is substantially higher than $\pi_+$ is considered to capture meaningful behavioral patterns above random chance. The interpretive ranges adapted from [28, 29] are presented in Table 3.

Table 3. Interpretation of PR-AUC values.

| **PR-AUC Range** | **Interpretation** | **Description** |
|---|---|---|
| $\approx$ Baseline ($\pi_+$) | Random | No learning signal |
| 0.3–0.5 | Moderate | Weak pattern detection |
| 0.5–0.7 | Strong | Balanced precision–recall trade-off |
| $> 0.7$ | Excellent | High predictive power |

Temporal robustness was evaluated using longitudinal stability analysis. To measure the consistency across time during the evaluation period, a stability index $S(t)$ was defined as follows:

$$S(t) = \frac{1}{M} \sum_{m=1}^{M} (\text{PR-AUC})_m.$$

A stable or upward-trending $S(t)$ indicates stable model performance despite macroeconomic variations. This behavior is in line with the cyclical financial behavior reported in [17, 18, 3, 19, 20] and confirms the interpretability of the inverse transform probability framework over time.

Although both the PR-AUC and PR–AUC represent classification discrimination, the PR-AUC is more sensitive to severe class imbalance. Their approximate analytical relationship [29] is given by

$$\text{PR-AUC} \approx \frac{\text{PR–AUC} - \pi_-/2}{1 - \pi_-}.$$

When $\pi_+$ is small, that is, in a situation of severe imbalance, PR-AUC gives a more severe estimation of the model performance. Hence, it is especially suitable for the current financial context, in which inactive customers comprise the dominant part of the dataset.

Consistently high PR-AUC values throughout the evaluation period indicate that the proposed framework maintains reliable behavioral discrimination under severe class imbalances while preserving temporal stability. The consistency of the PR-AUC over the 12-month analysis period confirms that the inverse transform version and the behavioral segmentation rules are semantically consistent over time. Therefore, the system offers an optimal trade-off between statistical accuracy, interpretability, and economic consciousness under dynamic financial conditions [17, 18, 3, 19].

## 3. Results and Discussion

This section presents the principal empirical findings obtained using the suggested rule-ecommendation framework. It begins with an analysis of the synthetic transaction dataset and proceeds to the analysis of the best performance of the model across product categories and time. Finally, the behavioral implications of the results (output of the recommendations) are examined in this study concerning different groups of customers.

### *3.1. Synthetic Dataset Characteristics*

The synthetic dataset covers the period from January 2023 to August 2025 and includes transaction data of 20 digital banking products in the UAE. The dataset was created to mimic real customer behavior, with heterogeneity added to the transaction frequency, monetary value, and financial indicators. A sample of the raw structure of the transactions is presented in Table 4.

Table 4. Sample of raw transactions.

| CIF | Product Type | Transaction Time | Amount | CASA | AUM | OSBAL | Date_PR |
|---|---|---|---|---|---|---|---|
| 100001 | cash_withdrawal_code | 2023-01-08 | 25,800 | – | – | – | 2023-01-31 |
| 100001 | mobile_top_up | 2023-01-21 | 42,000 | – | – | – | 2023-01-31 |
| 100001 | standing_instruction | 2023-01-01 | 115,600 | – | – | – | 2023-01-31 |
| 100001 | time_deposit | 2023-01-12 | 4,464,000 | 62,211,000 | 58,151,000 | 0.0 | 2023-01-31 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

Several important features can be observed in the raw dataset structure. First, there is a significant variation in transaction behavior among different customers, reflecting varied levels of financial engagement. Second, there is a considerable difference in transaction values and frequencies across product types, reflecting heterogeneity in financial needs and usage patterns. Third, financial indicators such as CASA, AUM, and OSBAL are only available for investment-related products, which also maintain causal consistency between financial wealth and investment activities.

Monthly aggregates were calculated to gain better insight into transaction patterns at the product level. Table 5 summarizes the number of transactions and the amount of money spent on the selected products in January 2023.

Aggregated statistics confirm that certain products, such as cash withdrawal codes and account statement downloads, occur frequently, whereas others seem to occur less often but involve higher monetary values than the former. These differences point to the existence of both high-frequency routine transactions and lower-frequency financial commitments in the dataset.

Table 5. Monthly product aggregates (January 2023).

| Date_PR | Product Type | Transaction Count | Amount Sum |
|---|---|---|---|
| 2023-01-31 | account_statement_download | 2268 | 2,680,209,000 |
| 2023-01-31 | bill_payment | 869 | 994,003,000 |
| 2023-01-31 | cash_withdrawal_code | 2960 | 3,442,701,000 |
| 2023-01-31 | cheque_book_request | 1314 | 1,431,803,000 |
| ⋮ | ⋮ | ⋮ | ⋮ |

At the system level, there were natural fluctuations over time in the monthly totals of all products. The total number of transactions and the total amount for the selected months in 2023 are presented in Table 6.

Table 6. Monthly total transactions (2023).

| Date_PR | Monthly Transaction Total | Monthly Amount Total |
|---|---|---|
| 2023-01-31 | 16,616 | 4.25B |
| 2023-02-28 | 15,468 | 4.25B |
| 2023-03-31 | 17,815 | 6.28B |
| 2023-04-30 | 16,250 | 5.25B |
| 2023-05-31 | 24,830 | 3.13B |

The patterns of variation in transaction counts and volume of dollars spent indicate that there are seasonal patterns that are affected by holidays, billing cycles, and consumer spending behaviors. Such temporal fluctuations are often encountered in digital banking environments and provide a realistic basis for evaluating the stability of predictive models.

### *3.2. Predictive Performance by Product Category*

For our modelling, each product category was modelled separately based on the machine learning pipeline described above. Table 7 summarizes the predictive performances of the selected products in June 2025.

Table 7. Model performance per product (June 2025).

| Target | Model | Test PR-AUC | Transaction Name | Period |
|---|---|---|---|---|
| cash_withdrawal_code_target | Random Forest | 0.954 | cash_withdrawal_code | 2025-06-30 |
| bill_payment_target | Random Forest | 0.940 | bill_payment | 2025-06-30 |
| mutual_fund_investment_target | XGBoost | 0.938 | mutual_fund_investment | 2025-06-30 |
| credit_card_payment_target | LightGBM | 0.937 | credit_card_payment | 2025-06-30 |
| split_bill_p2p_target | Logistic Regression | 0.924 | split_bill_p2p | 2025-06-30 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

The prediction results showed strong classification among most product classes. Several models attained a PR-AUC score of more than $0.93$, which provides evidence of excellent discriminatory ability to distinguish consistent behavioral patterns when there was an imbalance in classes. The predictability of administrative and routine payment products was the highest, owing to their regular and routine use trends.

Conversely, lifestyle-based services and optional transactions produced slightly worse but still reasonable PR-AUC values, which shows that they rely on momentary behavioral urges that create greater variability and make them more complex to predict.

### *3.3. Temporal Robustness Analysis*

To assess long-term stability, the values of the PR-AUC were tracked over time (months) and products. The PR AUC heatmap related to all product types and evaluation periods is shown in Figure 1.

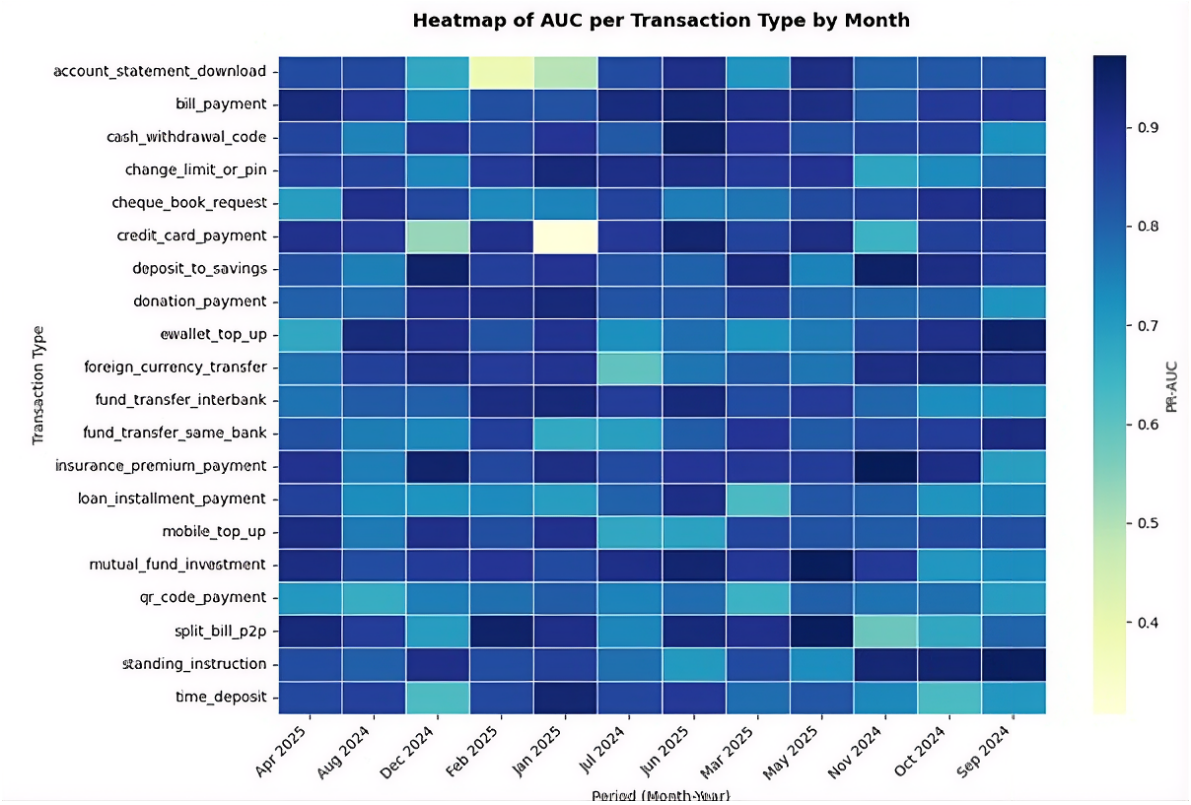


Figure 1. Heatmap of PR-AUC per transaction type by month.

The PR-AUC of most product categories fell within the range of $0.75$ to $0.95$, indicating that they have a high ability to predict most of the time. In the rare instances when we observe a temporary decline in certain products (such as credit card payments at the beginning of the year), the general trend appears to show that the model will be a good predictor across the entire evaluation span.

Correlation analysis was also useful in determining the relationship between model performance and product usage. Figure 2 presents the correlation of the PR-AUC and transaction share as well as the transaction amount share.

The analysis results show that there are several salient relationships. The correlation between the PR-AUC and share of transaction frequency is rather negatively correlated ($r = -0.55$), which means that the high traits of transaction frequency can blur behavioral differentiation in a diverse user base. Compared to the negative relationship between the PR-AUC and share of transaction amount, a rather weak association, $r = -0.22$ implies that high-value products are associated with the introduction of a small amount of behavioral variability. On the other hand, it is positively correlated with transaction frequency and monetary contribution ($r$ or have a positive value of $0.51$), implying that popular services contribute significantly to the total financial activity.

These empirical findings support the statement that PR-AUC represents behavioral consistency, not just statistical accuracy; thus, the issue is fit to evaluate recommendation frameworks that are exposed to disproportionate financial datasets.

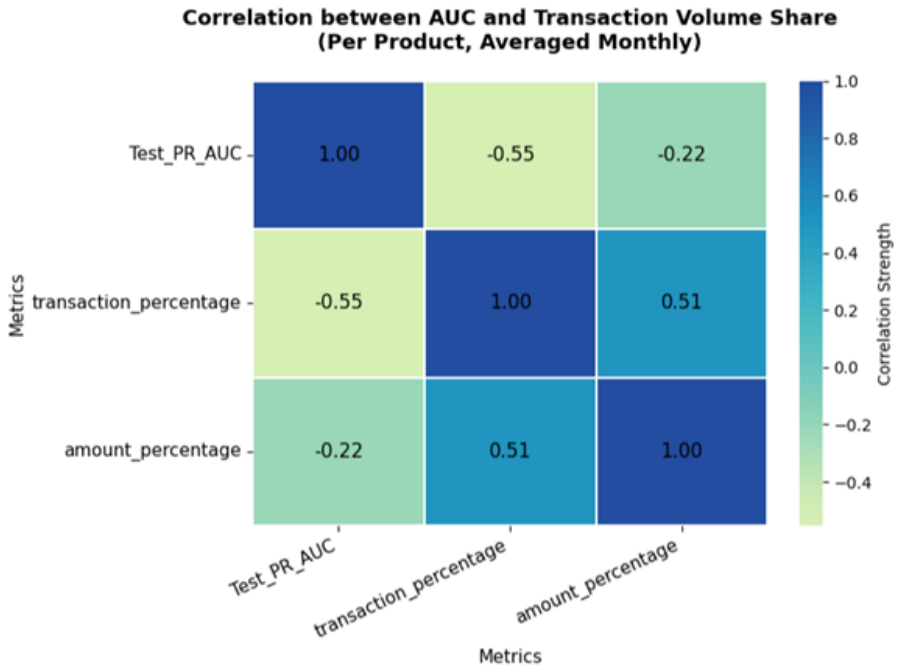


Figure 2. Correlation between PR-AUC and transaction/amount share.

### *3.4. Longitudinal Performance Evaluation*

The overall system stability can also be observed from the monthly average values of PR-AUC presented in Table 8 and Figure. 3.

Table 8. Monthly average PR-AUC.

| Period | Average AUC |
|---|---|
| 2024-07-31 | 0.805 |
| 2024-08-31 | 0.814 |
| 2024-09-30 | 0.811 |
| 2024-10-31 | 0.819 |
| 2024-11-30 | 0.814 |
| 2024-12-31 | 0.801 |
| 2025-01-31 | 0.813 |
| 2025-02-28 | 0.828 |
| 2025-03-31 | 0.823 |
| 2025-04-30 | 0.834 |
| 2025-05-31 | 0.842 |
| 2025-06-30 | 0.850 |

The average PR-AUC was always in the range $0.80$–$0.85$, which belongs to the excellent predictive category for imbalanced classification problems. In addition, the noted progressive increase means that monthly retraining enables the models to adjust to adapting patterns of transactions. The changes in consumer behavior during seasons might be attributed to the slight changes noticed during some months.

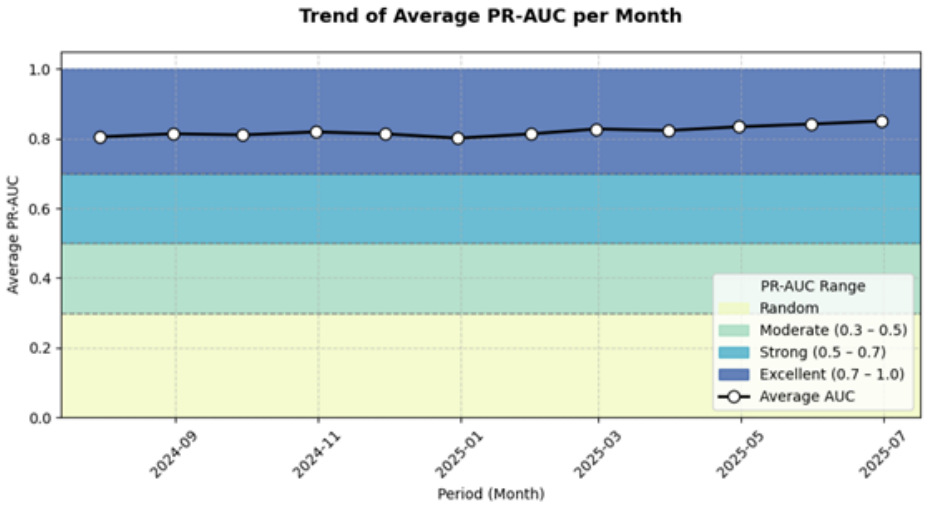


Figure 3. Trend of average PR-AUC per month.

We conclude that the longitudinal evaluation confirms that monthly retraining enables the proposed framework to maintain a stable predictive performance despite temporal variations in customer behavior. These findings support the robustness of the probabilistic consistency framework under realistic, dynamic banking conditions.

### *3.5. Recommendation Behavior Across Customer Segments*

To measure the interpretability of the system and its personalization ability, the results of the recommendation were assessed in reference to three ideal target groups: active, heavy, and dormant/new users. Active users usually have a recurring pattern of transactions in various products; hence, the system supports products that support current trends and initiate complementary financial services, including administrative and investment services.

Heavy users, who have a high rate of transactions and consume a wide range of products, are given non-uniform recommendations based on the fill-to-$K$ strategy, which prioritizes under-represented products without losing the relevancy of the recommendation to the financial habits of the user. Customers that are either dormant or newly registered do not have any past transaction history; consequently, the recommendation engine is more dependent on aggregate signals and population-level trends to generate recommendations. Correspondingly, the lack of friction omnipresent services, such as QR payments, fund transfers, and cash-withdrawal codes, are suggested to be the point of contact and onboarding. A summary of the behavioral explanation of the recommendation plans for each segment is presented in Table 9.

Table 9. Segment comparison summary.

| Segment | Behavioral Characteristics | Recommended Product Types | Model Rationale |
|---|---|---|---|
| Active | Consistent monthly activity; 3–5 product types | Administrative tasks + investments + routine payments | Reinforces habits and expands product adoption |
| Heavy | High frequency and high diversity | Bill payment, insurance, top-up, ewallet, QR | Combines high probability and diversification |
| Dormant/New | Little to no history; recently onboarded | High-frequency simple financial services | Supports onboarding and reactivation |

These findings suggest that the specified recommendation framework successfully combined predictive accuracy with the interpretability of behavior. The system supports statistically robust yet economically salient recommendations that facilitate customer-specific digital banking services owing to the integration of probabilistic consistency and segment-aware logic.

## 4. Discussion

The proposed framework combines supervised machine learning with interpretable rule-based recommendation to predict customers' next transactions in digital banking. Unlike conventional prediction models that estimate only the probability of future transactions, the proposed inverse transform probability reformulates prediction outputs into behavioral consistency scores, providing a unified probabilistic interpretation of active, dormant, and new customer segments. The experimental results demonstrate stable predictive performance across different product categories, suggesting that behavioral consistency provides a useful basis for personalized financial recommendation.

From a practical perspective, the proposed framework is designed to accommodate gradual changes in customer behavior through monthly model retraining. Because product-specific models are trained independently, the training process can be parallelized and performed offline, whereas online recommendation generation only requires probability prediction followed by rule-based ranking, resulting in a relatively low inference cost. However, monthly retraining represents a compromise between computational efficiency and adaptability. Abrupt behavioral changes caused by major life events may not be reflected immediately because sufficient post-event observations are required before the models can adapt to them. Future research may therefore investigate online learning, incremental learning, or event-driven updating strategies to improve responsiveness.

The principal limitation of this study is the exclusive reliance on synthetic transaction data. Although the simulated dataset was carefully designed to reproduce important statistical characteristics reported in banking studies, including heterogeneous customer behavior, wealth stratification, seasonal transaction dynamics, and class imbalance, it cannot fully represent the complexities of real banking environments. Likewise, seasonality is incorporated explicitly only during data generation and is captured implicitly through monthly retraining rather than through explicit seasonal predictors. Consequently, the proposed framework should be regarded as a methodological contribution evaluated under a controlled experimental environment, with validation using real banking transaction datasets remaining an important direction for future research.

The present study also emphasizes interpretability through behaviorally meaningful feature engineering and Mutual Information-based feature selection. However, model-specific explainability techniques, such as SHAP values or gain-based feature importance, were not investigated because the framework consists of multiple, independently retrained, product-specific models. Similarly, fairness analysis could not be conducted because the synthetic dataset intentionally excluded the protected demographic attributes. Future work should therefore investigate feature attribution, fairness-aware recommendation strategies, and explicit seasonal modeling using real banking datasets whenever appropriate privacy regulations permit.

Finally, the empirical evaluation focuses on the predictive component of the recommendation framework. Although the PR-AUC is well-suited for highly imbalanced transaction data, it does not directly evaluate recommendation quality. Recommendation-oriented metrics such as Precision@K, Recall@K, MAP, and NDCG, together with comparisons against collaborative filtering and other recommendation-specific approaches, require customer interaction logs or online deployment data and therefore remain important directions for future investigation.

## 5. Conclusion

This study developed and evaluated a rule-based recommendation model for next-transaction prediction in digital banking. The proposed system combines probabilistic learning with the principles of behavioral consistency to produce easy-to-interpret and economically meaningful recommendations for the restaurant industry. Through a synthetic dataset simulating customer heterogeneity and seasonal and financial transaction dynamics, the framework demonstrates high predictive reliability and temporal robustness.

First, the simulation design can replicate the important features of real-world digital banking environments, including class imbalance, multi-segment behavioral diversity, and seasonal variation in financial activity. The

dataset structure enables the model to reflect long-term patterns in behavior while maintaining meaningful relationships from transactional and financial investment-related indicators.

Second, the machine learning pipeline, which involved feature engineering, behavioral consistency targets, imbalance mitigation, monthly model retraining, and longitudinal performance evaluation, consistently exhibited good predictive performance. Empirical results show the average results of PR-AUC, $0.83 \pm 0.09$ and $0.97$ stability index, revealing the high discrimination capability and robustness over time despite the variations in macro-economy and seasons.

Third, the inverse transform probability formulation is essential for preserving the interpretability of other states for customers. The framework aligns probabilistic predictions with the principles of Expected Monetary Value (EMV) by reformulating the predicted probabilities with respect to behavioral stability. This alignment assists financially in deciding to support cross-selling strategies and customer engagement initiatives, which makes sense.

Finally, the recommendation engine exhibited impressive personalization skills across user lines. Active customers are to be recommended in a manner that strengthens stable structures of engagements, heavy customers are to be recommended a more diversified approach with a fill-to-K strategy, and dormant or newly acquired customers are to be recommended straightforward and high-frequency services that will enable them to reactivate. These findings indicate that the system can not only forecast the consistency of behavior but also implement action-oriented and product segment-aware recommendations.

The suggested rule-recommendation framework is a sound, interpretable, and cost-effective method for predicting the next transaction in online banking. Its timelessness in customer segments and product lines signifies a high chance of implementation in the financial institution world to enhance customer interaction, retention, and cross-selling performance.

Although the simulation was designed to reproduce important statistical characteristics reported in digital banking studies, the proposed framework has not yet been validated using proprietary customer transaction data. Therefore, the present study should be regarded as a methodological contribution that demonstrates the feasibility of the proposed probabilistic recommendation framework under controlled conditions. Future work will focus on validating the approach using real banking datasets subject to appropriate privacy and regulatory requirements.

## Funding

This research is supported by Dana Selain APBN Universitas Diponegoro through the Riset Kolaborasi Undip (RKU) scheme 2026, under grant number 306-196/UN7.D2/PP/V/2026.

## Author Contributions

Giovanny Theotista: Conceptualization, methodology, data simulation, software, formal analysis, visualization, and writing–original draft. Moch. Fandi Ansori: Conceptualization, methodology, validation, writing–original draft, writing–review and editing. All the authors have read and approved the final version of the manuscript.

## Conflict of Interest

The authors declare no conflicts of interest.

## Data Availability

The data used in this study were synthetic data generated for experimental purposes. The implementation code and synthetic dataset used to reproduce the results are available from the corresponding author upon reasonable requests.

## REFERENCES


1. C. C. Aggarwal, *Recommender Systems: The Textbook*. Springer Cham, 2016.
2. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, Massachusetts: Morgan Kaufmann, 2012.
3. J. Sengupta, K. Lam, and D. Desmet, "Digital banking in asia: Winning approaches in a new generation of financial services," McKinsey & Company, Report, 2014, asia Financial Institutions (Accessed: 1 March 2026)). [Online]. Available: https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/pdf/2014%20digital%20banking%20in%20asia%20-%20winning%20approaches%20in%20a%20new%20generation%20of%20financial%20services.pdf
4. H. R. Sofaer, J. A. Hoeting, and C. S. Jarnevich, "The area under the precision-recall curve as a performance metric for rare binary events," *Methods in Ecology and Evolution*, vol. 10, no. 4, pp. 565–577, 2019.
5. J. T. Hancock, T. M. Khoshgoftaar, and J. M. Johnson, "Evaluating classifier performance with highly imbalanced big data," *Journal of Big Data*, vol. 10, no. 42, 2023.
6. F. Provost and T. Fawcett, "Robust classification for imprecise environments," *Machine Learning*, vol. 42, pp. 203–231, 2001.
7. G. Theotista, M. Febe, and M. S. Ryan, "Analysing market dynamics: Revealing obscured patterns in lq45 stocks using ward's hierarchical clustering," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 19, no. 1, pp. 163–172, 2025.
8. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, Jul. 2009.
9. V. K. Kaishev, J. P. Nielsen, and F. Thuring, "Optimal customer selection for cross-selling of financial services products," *Expert Systems with Applications*, vol. 40, no. 5, pp. 1748–1757, 2013.
10. Y. Altun and A. Yucekaya, "A probabilistic approach to maximize cross-selling revenues of financial products," *American Scientific Research Journal for Engineering, Technology, and Sciences*, vol. 79, no. 1, pp. 1–14, 2021, accessed: 1 March 2026. [Online]. Available: https://asrjetsjournal.org/American_Scientific_Journal/article/view/6748
11. N. Boustani, A. Emrouznejad, R. Gholami, O. Despic, and A. Ioannou, "Improving the predictive accuracy of the cross-selling of consumer loans using deep learning networks," *Annals of Operations Research*, vol. 339, pp. 613–630, 2024.
12. C. Basten and R. Juelsrud, "Cross-selling in bank-household relationships: Mechanisms and implications for pricing," *The Review of Financial Studies*, vol. 39, no. 6, pp. 1751–1784, 06 2026.
13. M. F. Ansori and S. Khabibah, "The role of cost of loan in banking loan dynamics: Bifurcation and chaos analysis," *Barekeng*, vol. 16, no. 3, p. 1031 – 1038, 2022.
14. M. F. Ansori and S. Hariyanto, "Analysis of banking deposit cost in the dynamics of loan: Bifurcation and chaos perspectives," *Barekeng*, vol. 16, no. 4, p. 1283 – 1292, 2022.
15. M. F. Ansori, N. Sumarti, K. A. Sidarto, and I. Gunadi, "An algorithm for simulating the banking network system and its application for analyzing macroprudential policy," *Computer Research and Modeling*, vol. 13, no. 6, p. 1275 – 1289, 2021.
16. B. P. Josaphat, M. F. Ansori, and K. Syuhada, "On optimization of copula-based extended tail value-at-risk and its application in energy risk," *IEEE Access*, vol. 9, p. 122474 – 122485, 2021.
17. Bank Indonesia, "Payment system statistics," 2023.
18. Otoritas Jasa Keuangan, "Digital financial reports," 2023.
19. Google, Temasek, and Bain & Company, "e-conomy sea 2023: Reaching new heights: Navigating the path to profitable growth," Google, Temasek, and Bain & Company, Research Report, 2023, accessed: 15 February 2026. [Online]. Available: https://www.temasek.com.sg/content/dam/temasek-corporate/news-and-views/resources/reports/google-temasek-bain-e-conomy-sea-2023-report.pdf
20. Boston Consulting Group, "Southeast asian consumers are driving a digital payment revolution," Boston Consulting Group, BCG Report, May 2020, accessed: 15 February 2026. [Online]. Available: https://web-assets.bcg.com/img-src/BCG-Southeast-Asian-Consumers-Are-Driving-a-Digital-Payment-Revolution-May-2020_tcm9-247804.pdf
21. K. Foster and M. Hitczenko, "2025 survey and diary of consumer payment choice: Summary results," Federal Reserve Bank of Atlanta, Research Data Report 26-1, May 2026, accessed: 15 February 2026. [Online]. Available: https://www.atlantafed.org/research-and-data/surveys/survey-and-diary-of-consumer-payment-choice
22. L. J. Locke and N. Chalise, "The bank on national data hub: Findings from 2023," Federal Reserve Bank of St. Louis, Research Report, Nov. 2024, accessed: 15 February 2026. [Online]. Available: https://www.stlouisfed.org/community-development/bank-on-national-data-hub/bank-on-report-2023
23. Otoritas Jasa Keuangan, "Indonesian banking statistics," 2024.
24. H. He and Y. Ma, Eds., *Imbalanced Learning: Foundations, Algorithms, and Applications*. Hoboken, NJ: John Wiley & Sons, 2013.
25. M. Abdelhamid and A. Desai, "Balancing the scales: A comprehensive study on tackling class imbalance in binary classification," *arXiv*, 2024.
26. K. A. Sidarto, M. Syamsuddin, and N. Sumarti, *Financial Mathematics*. Bandung, Indonesia: Penerbit ITB, 2017.
27. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. VanderPlas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011, accessed: 20 February 2026. [Online]. Available: https://jmlr.org/papers/v12/pedregosa11a.html
28. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
29. A. Fernandez, S. Garcia, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, *Learning from Imbalanced Data Sets*. Cham, Switzerland: Springer, 2018.
30. P. Chapman, "Crisp-dm 1.0: Step-by-step data mining guide," 2000, accessed: 20 February 2026. [Online]. Available: https://api.semanticscholar.org/CorpusID:59777418
31. M. Okabe, J. Tsuchida, and H. Yadohisa, "F-measure maximizing logistic regression," *Communications in Statistics – Simulation and Computation*, vol. 53, no. 5, pp. 2554–2564, 2024.
32. B. T. T. My and B. Q. Ta, "A modification of logistic regression with imbalanced data: F-measure-oriented lasso-logistic regression," *ScienceAsia*, vol. 49S, pp. 68–77, 2023.

33. M. Jiang, Y. Yang, and H. Qiu, "Fuzzy entropy and fuzzy support-based boosting random forests for imbalanced data," *Applied Intelligence*, vol. 52, no. 5, pp. 4126–4143, 2022.
34. C. Chen, A. Liaw, and L. Breiman, "Using random forest to learn imbalanced data," University of California, Berkeley, Technical Report, 2004, accessed: 20 February 2026.
35. J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, "Boosting methods for multi-class imbalanced data classification: an experimental review," *Journal of Big Data*, vol. 7, p. 70, 2020.
36. G. Velarde, M. Weichert, A. Deshmunkh, S. Deshmane, A. Sudhir, K. Sharma, and V. Joshi, "Tree boosting methods for balanced and imbalanced classification and their robustness over time in risk assessment," *Intelligent Systems with Applications*, vol. 22, p. 200354, 2024.
37. U. Rawat and B. Rawat, "A comprehensive study of boosting algorithms for class imbalance dataset," in *Smart Systems and Wireless Communication*, ser. Smart Innovation, Systems and Technologies, R. Chaki, A. Cortesi, S. DasGupta, and S. Saha, Eds., vol. 433. Singapore: Springer, 2025, pp. 259–271, proceedings of SSWC 2024. [Online]. Available: https://doi.org/10.1007/978-981-96-1348-9_21
38. H. Park and H. Kim, "Rrboost: a new ensemble method for classifying imbalanced data," *Knowledge and Information Systems*, vol. 67, pp. 8983–9005, 2025.
39. X. Zhao, Y. Liu, and Q. Zhao, "Improved lightgbm for extremely imbalanced data and application to credit card fraud detection," *IEEE Access*, vol. 12, pp. 159 316–159 335, 2024.
40. A. A. Khan, O. Chaudhari, and R. Chandra, "A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation," *Expert Systems with Applications*, vol. 244, p. 122778, 2024.
41. M. B. McDermott, H. Zhang, L. H. Hansen, G. Angelotti, and J. Gallifant, "A closer look at auroc and auprc under class imbalance," in *Proceedings of the 38th International Conference on Neural Information Processing Systems*, ser. NIPS '24. Red Hook, NY, USA: Curran Associates Inc., 2024.
42. J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. 10, pp. 281–305, 2012.
43. L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
44. T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 785–794. [Online]. Available: https://doi.org/10.1145/2939672.2939785
45. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: a highly efficient gradient boosting decision tree," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 3149–3157.
46. G. Theotista, M. Febe, and Y. Marshelly, "Development of expected monetary value using binomial state price for stock investment decisions," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 17, no. 3, pp. 1703–1712, 2023.
47. Y. Jiang, Q. Pan, Y. Liu, and S. Evans, "A statistical review: Why average weighted accuracy, not accuracy or auc?" *Biostatistics & Epidemiology*, vol. 5, no. 2, pp. 267–286, 2021.