STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, March 2026, pp 2286–2302.
Published online in International Academic Press (www.IAPress.org)

GenAI Meets Explainability: Turning Churn Predictions into Personalized Retention Strategies

Meryem HOUSSAM*, Abdelilah JRAIFI

MISCOM Laboratory, National School of Applied Sciences, Safi Cadi Ayyad University, Morocco

Abstract In an increasingly competitive financial landscape, retaining existing customers is widely acknowledged to be more cost-effective than acquiring new ones. While artificial intelligence (AI)-based predictive models have achieved high accuracy in identifying customers at risk of churn, they often fail to provide actionable strategies for customer retention. This paper addresses this limitation by proposing a post modeling framework that translates churn predictions into business-oriented retention actions.
Using supervised machine learning techniques on structured customer data—such as transactional and behavioral features—from the financial sector, we first develop a high-performance churn prediction model. We then employ explainability methods, notably SHAP (SHapley Additive exPlanations), to identify the key drivers of churn at both global and individual levels. These insights enable us to segment customers into interpretable profiles (e.g., price-sensitive, service-dissatisfied, inactive), each associated with specific churn triggers.
To move beyond prediction and toward proactive intervention, we propose tailored retention strategies aligned with each segment's churn rationale. Furthermore, we explore the integration of Generative AI (GenAI) to support the automatic generation of personalized messages and strategy suggestions, enhancing the decision-making process for financial institutions.
The proposed methodology bridges the gap between churn prediction and business actionability, offering a data-driven approach to customer engagement. Our results demonstrate that such an approach not only deepens customer understanding but also significantly improves the effectiveness of targeted retention campaigns.

Keywords Explainable AI (XAI), Customer Churn, Gen IA, Data-Driven Decision Making, Artificial Intelligence, Retention Strategy. Model Evaluation, Accuracy, Precision, Recall, F1-score, Financial Institutions.

DOI: 10.19139/soic-2310-5070-3151

1. Introduction

In today's increasingly competitive financial landscape, customer retention has emerged as a critical driver of sustainable growth. It is widely acknowledged that retaining an existing customer is significantly more cost-effective than acquiring a new one, making the mitigation of customer churn a paramount business objective.

Financial institutions are thus under constant pressure to identify at-risk customers and implement effective strategies to secure their loyalty and prevent attrition.

To meet this challenge, the industry has increasingly turned to Artificial Intelligence (AI) and Machine Learning (ML) to build predictive churn models. These models have demonstrated high accuracy in identifying which customers are likely to leave. However, their practical utility is often hampered by their inherent opacity. Functioning as "black boxes," these systems provide a prediction but fail to explain the underlying reasons—the why—behind a customer's risk. This lack of interpretability creates a significant gap between prediction and

*Correspondence to: Meryem HOUSSAM (Email: m.houssam.ced@uca.ac.ma). MISCOM-laboratory National School of Applied Sciences-Safi, Cadi Ayyad University Morocco .

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

action, leaving business teams with valuable information but no clear path to designing effective, targeted retention interventions.

This paper addresses this limitation by proposing a post-modeling framework designed to translate churn predictions into personalized, actionable retention strategies. Our contribution is a comprehensive methodology built on three technological pillars:

- Prediction, where we develop a high-performance supervised machine learning model using structured customer data.
- Explanation, where we employ state-of-the art explainability methods, notably SHAP (SHapley Additive exPlanations), to uncover the key drivers of churn at both global and individual levels.
- Personalization where we leverage these insights to segment at-risk customers.
- (GenAI) to automatically generate tailored messages and strategy suggestions.

The remainder of this article is structured as follows. We will first detail our proposed methodology, outlining each step of the framework. We will then present the results of our experimentation on a banking dataset, showcasing how the framework identifies customer segments and generates personalized strategies.Finally, the discussion will explore the broader business implications of our approach and highlight potential directions for future research.

2. Related Work

This section provides an overview of the existing literature across three key domains that underpin our research: customer churn prediction within the financial sector, the use of explainable artificial intelligence (XAI) to interpret model outcomes, and the growing contribution of Generative AI (GenAI) to enhancing customer engagement. Together, these areas form the conceptual foundation of our proposed approach.

2.1. Machine Learning for Customer Churn Prediction

Customer churn has long been a major area of investigation within the financial industry. A wide body of research has explored the use of machine learning techniques to identify clients at risk of leaving. Various studies have highlighted the effectiveness of both traditional statistical models and more advanced ensemble methods in achieving strong predictive results on financial datasets [8, 12, 14]. Commonly used features—such as customer behavior, transaction history, and service usage—have consistently proven to be key indicators of churn [22, 25, 26]. More recent work has focused on enhancing prediction performance through ensemble strategies and advanced feature selection approaches [9, 24]. Despite these advances, many studies tend to concentrate primarily on the prediction phase, without extending their efforts toward implementing concrete retention actions based on the predictions.

2.2. The Rise of Explainable AI (XAI) in Finance

In response to the ”black box” nature of many high-performance models, the field of Explainable AI (XAI) has gained significant traction, particularly in high-stakes domains like finance [10]. The need for transparency and interpretability is not only a matter of regulatory compliance but also a prerequisite for building trust and deriving actionable business insights [13]. Recent work has focused on applying XAI techniques tofinancial problems, including churn prediction. Studies by Tékouabou et al. [9] and Li & Yan [19] explicitly use interpretability analysis to understand the drivers behind their churn models. These efforts highlight a critical shift from simply knowing who will churn to understanding why. Our work builds directly upon this trend, using explainability not just as a diagnostic tool but as a foundational element for strategic decision-making .[20].

2.3. Generative AI for Personalized Customer Engagement

The recent advent of large-scale Generative AI (GenAI) models has opened new frontiers for customer relationship management. The potential of GenAI in financial institutions is vast, covering opportunities from operational efficiency to hyper-personalized customer communication [15, 16]. While some studies explore the use of GenAI in the context of churn prediction, they often focus on its predictive capabilities or high-level strategic benefits [17, 18]. For example, Rudd et al. [11] explore multimodal data fusion, including voice, which hints at the potential for richer, more personalized interactions. However, a significant gap remains in connecting the granular, data-driven insights from XAI with the content-generation power of GenAI.

2.4. Our Contribution: Bridging Prediction,Explanation, and Action

While the existing literature addresses churn prediction [21, 23, 27], model explainability [9, 19], and the potential of GenAI [15, 17] as separate domains, our work is situated at their intersection. To our knowledge, few studies have proposed an end-to-end framework that systematically links these three pillars. Our primary contribution is to present and validate a methodology that not only predicts churn with high accuracy but also uses XAI to create interpretable customer segments and then leverages GenAI to automatically generate personalized retention strategies for those segments. By doing so, we aim to bridge the critical gap between data science outputs and tangible business actions, turning churn alerts into proactive, data-driven engagement opportunities.

3. Proposed Framework and Methodology

To bridge the gap between churn prediction and actionable retention, we propose a four-stage framework designed to transform raw data into personalized, AI-generated strategies. This methodology systematically moves from identifying at-risk customers to understanding their churn drivers and finally to creating tailored interventions. An overview of our framework is depicted in Figure 1.

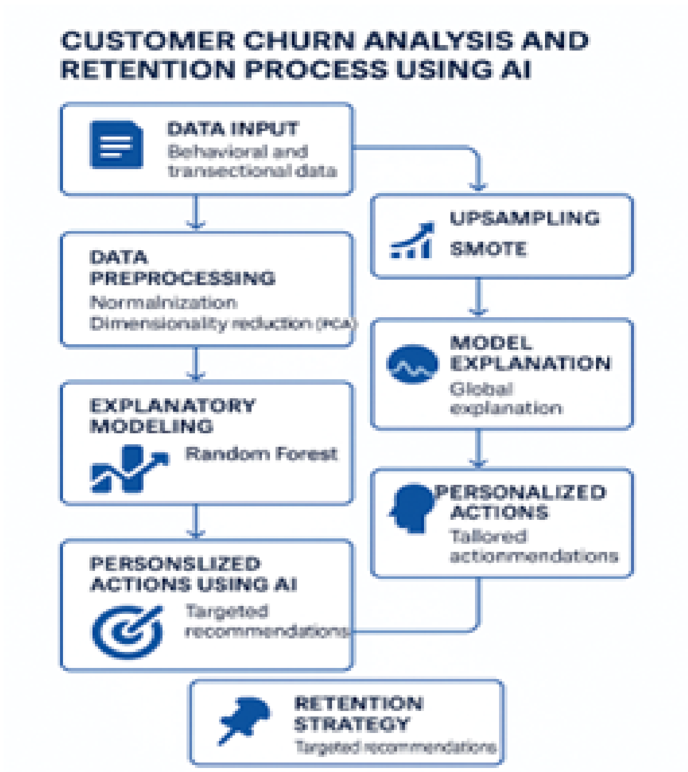


Figure 1. Overview of the Proposed Post-Modeling Framework

3.1. *Stage 1: Churn Prediction*

The foundational stage of our framework involves developing robust model to accurately predict customer churn.

3.1.1. ***Data and Preprocessing*** The dataset utilized in this study is *BankChurners.csv*, a publicly available dataset containing customer information from a financial institution. It comprises 10,127 customer records with features including demographic information (e.g., `Customer_Age`, `Gender`), financial attributes (e.g., `Income_Category`, `Credit_Limit`), and behavioral characteristics (e.g., `Months_Inactive_12_mon`, `Total_Trans_Ct`). The target variable, `Attrition_Flag`, indicates whether a customer has churned (`'Attrited Customer'`) or not (`'Existing Customer'`).

Prior to model training, the raw dataset underwent several preprocessing steps:

- **Data Cleaning**: Irrelevant columns such as a unique client identifier (`CLIENTNUM`) and variables related to a pre-existing Naive Bayes classifier were removed, as they held no predictive value for our models.

- **Target Variable Transformation**: The categorical target variable `Attrition_Flag` was converted into numerical format, where `'Attrited Customer'` was mapped to 1 (churn) and `'Existing Customer'` to 0 (no churn).

- **Categorical Feature Encoding**: Nominal features (e.g., `Gender`, `Marital_Status`, `Card_Category`) were transformed into numerical representations using One-Hot Encoding. This process generates binary columns for each category, preventing the model from assuming an ordinal relationship. To avoid multicollinearity, the first category of each feature was dropped (`drop_first=True`).

- **Data Splitting**: The preprocessed dataset was split into training (70%) and testing (30%) sets. A fixed `random_state` was used to ensure the reproducibility of our experiments.

- **Handling Class Imbalance**: As is common in churn datasets, the target variable was imbalanced. To address this issue, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to the training data. SMOTE generates synthetic examples of the minority class (churners), enabling the model to learn from a more balanced dataset and reducing bias toward the majority class.

3.1.2. ***Model Selection and Evaluationn*** In the context of binary classification with an imbalanced class distribution, it is crucial to use appropriate evaluation metrics that consider the minority class. Traditional metrics such as accuracy, precision, recall, and the F1 score can give a skewed picture of model performance in the presence of imbalanced classes.

- Precision

Precision measures the proportion of correct positive predictions among all positive predictions made by the model. It is defined as the ratio of true positives (TP) to the sum of true positives and false positives (FP):

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall

Recall — also referred to as sensitivity — reflects a model's ability to correctly identify all relevant instances. In other words, it measures the proportion of actual positive cases that were successfully predicted as positive.

Mathematically, it is expressed as the ratio of true positives (TP) to the total number of actual positives, which includes both true positives and false negatives (FN):

$$\text{Recall} = \frac{TP}{TP + FN}$$

- F1 Score

The F1 score represents the harmonic mean of precision and recall, offering a single metric that balances both aspects. It is particularly valuable in situations involving class imbalance, where focusing solely on precision or recall might be misleading. The F1 score captures the trade-off between the two by computing their harmonic average:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1 score ranges from 0 to 1, where 1 indicates perfect precision and recall, and 0 reflects poor performance on both. By taking into account both false positives and false negatives, it offers a balanced assessment of model effectiveness. This makes it particularly valuable in cases of class imbalance, where traditional metrics may give an overly optimistic view of performance.

When dealing with imbalanced datasets, relying solely on accuracy can be misleading. It is therefore important to also examine precision, recall, and the F1 score, as these metrics offer a more complete picture of how well a model distinguishes between classes of unequal representation.

To identify the most suitable model for our framework, we compared the performance of three well-established classifiers: Random Forest, AdaBoost, and Support Vector Machine (SVM). A key focus of this evaluation was to highlight the effect of addressing class imbalance, which is a frequent issue in churn prediction tasks. For this purpose, we conducted two sets of experiments: in the first, models were trained on the original imbalanced dataset; in the second, we applied the Synthetic Minority Over-sampling Technique (SMOTE) to balance the class distribution before training the models.. The performance of each model was then measured on the same unseen test set using the F1-Score, which is a robust metric for imbalanced classification problems as it balances precision and recall

The results, summarized in Table 1, clearly illustrate the critical importance of addressing class imbalance.

Table 1. F1-Score Comparison Before and After SMOTE Upsampling

Model	**F1-Score (Before SMOTE)**	**F1-Score (After SMOTE)**
Random Forest	0.48	0.90
AdaBoost	0.53	0.84
SVM	0.57	0.87

As shown, all models performed poorly on the original, imbalanced data, with F1- Scores barely exceeding 0.57. This indicates that without intervention, the models were unable to effectively learn the patterns of the minority churn class. However, after applying SMOTE, there was a dramatic improvement across all models. The Random Forest classifier achieved the highest F1-Score of 0.90, outperforming both SVM (0.87) and AdaBoost (0.84). Based on these results, we selected the Random Forest model trained on the SMOTE-balanced data as the predictive engine for our framework. This choice was driven by two key factors:

1. **Superior Predictive Performance:** It demonstrated the highest ability to accurately identify churning customers on the test set.

2. **Compatibility with XAI:** As an ensemble of decision trees, it is fully compatible with tree-based SHAP explainers, which is a requirement for the subsequent stages of our methodology.

3.2. Stage 2: Insight Extraction with SHAP

Once the model is trained, the next critical step is to understand why it makes certain predictions. For this, we employ **SHAP (SHapley Additive exPlanations)**, a state-of-the-art XAI technique. SHAP is based on game theory and provides a unified measure of feature importance by calculating the marginal contribution of each feature to the prediction for each individual instance. This approach was chosen for its solid theoretical foundation and its ability to provide both global and local explanations.

- **Global Explanations:** We use SHAP summary plots to identify the features that have the most significant impact on churn predictions across the entire customer population. This provides a high-level understanding of the main drivers of churn within the institution.

- **Local Explanations:** We leverage SHAP force plots to dissect individual predictions. For any given customer, this allows us to see precisely which factors are pushing their churn risk up (e.g., high inactivity) and which are pulling it down(e.g., long-standing relationship). This granular level of detail is essential for true personalization.

3.3. Stage 3: Insight-Based Segmentation

Armed with local SHAP explanations, we move beyond traditional demographic segmentation. Instead, we propose an insight-based segmentation approach, grouping at-risk customers based on their primary churn drivers as identified by SHAP. This method allows us to create behaviorally and motivationally consistent segments. For this study, we identified three key segments among the customers predicted to churn:

- **Segment 1: The Engaged but Under-leveraged Customer** — These customers exhibit a high transaction count but maintain a low revolving balance. Their churn is not driven by inactivity, but potentially by a perceived lack of value from credit-related products.

- **Segment 2: The Inactive Customer** — This is the largest and most classic churn profile, characterized by low transaction counts and a high number of inactive months. These customers have effectively disengaged from the bank's services.

- **Segment 3: The Under-utilizing High-Potential Customer** — These customers have been granted a high credit limit but have a very low utilization ratio. Their risk may stem from a disconnect between the product's cost (or perceived complexity) and its benefits.

3.4. Stage 4: Personalized Strategy Generation with GenA

The final stage of our framework operationalizes the insights by using Generative AI to create personalized retention actions. Instead of providing generic advice, we leverage the specific context of each customer segment to construct detailed prompts for a large language model (LLM), such as GPT-4 or Google's Gemini.

A prompt typically contains the following components:

1. **Role:** The persona for the GenAI (e.g., *"You are an expert retention strategist for a bank"*).

2. **Context:** The customer's segment profile and key data points (e.g., *"The customer is in the 'Inactive' segment with 5 months of inactivity"*).

3. **SHAP Insight:** The primary reason for churn (e.g., *"SHAP values confirm that low transaction count is the main driver of their churn risk"*).

4. **Task:** The specific output required (e.g., *"Draft a personalized, empathetic email to re-engage this customer"*, or *"Generate three talking points for a phone call"*).

By providing this rich, data-driven context, the GenAI can generate highly relevant and personalized content, effectively bridging the final gap from prediction to action. An example of a prompt used in our experiment is shown in Section 4.4.

4. Results and Experimentation

This section presents the results of applying our proposed framework to the bank churn dataset. We first evaluate the performance of the predictive model and then demonstrate how the subsequent stages of the framework translate these predictions into actionable, personalized retention strategies.

4.1. Performance of the Prediction Model

The Random Forest model, selected after a comparative analysis and trained on the SMOTE-balanced dataset, demonstrated outstanding predictive performance on the unseen test set. A detailed summary of its performance is provided by the classification report in Figure 2.

Classification Report:

	precision	recall	f1-score	support
0	0.95	0.99	0.97	2543
1	0.94	0.74	0.83	496
accuracy			0.95	3039
macro avg	0.94	0.87	0.90	3039
weighted avg	0.95	0.95	0.95	3039

ROC AUC Score: 0.9862418815724381

Figure 2. Classification Report of the Random Forest Model on the Test Set

4.1.1. Model Performance Summary The results highlight the model's robustness and its effectiveness in handling the imbalanced nature of the churn problem. The overall accuracy of the model is an impressive 95%. More importantly, for the critical task of identifying churners (class 1), the model achieved a well-balanced performance:

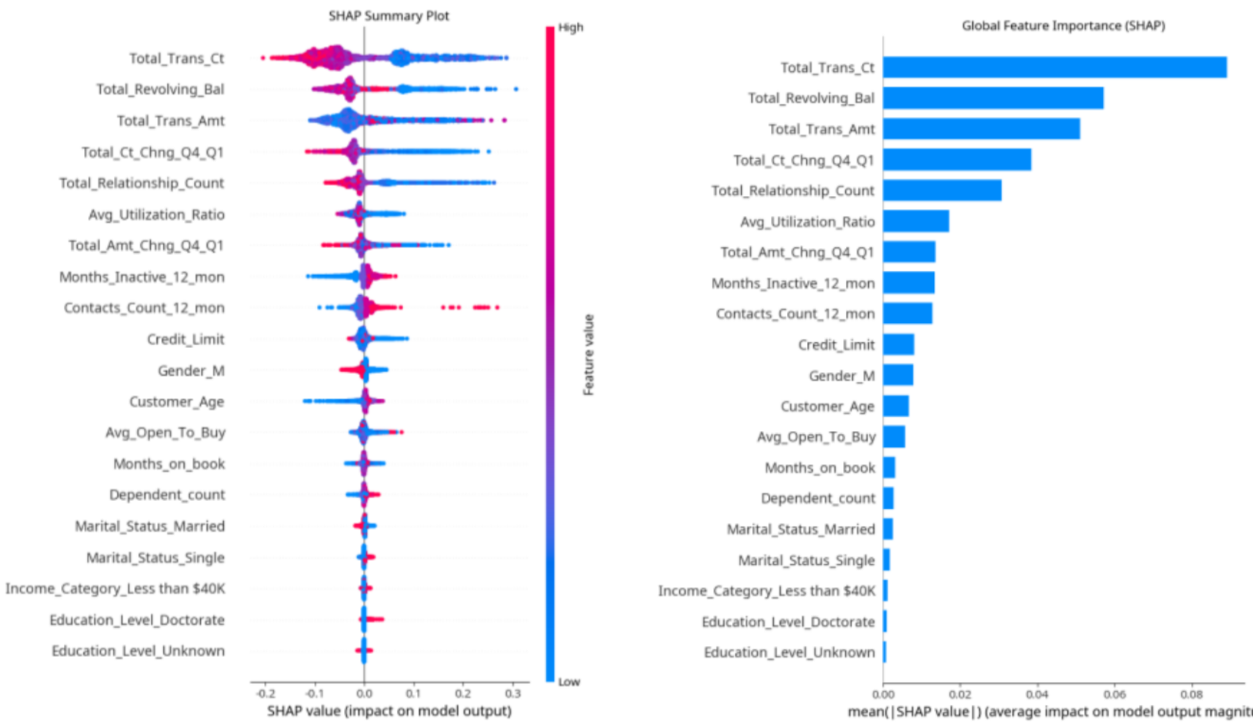
- **Precision of 0.94:** When the model predicts a customer will churn, it is correct 94% of the time. This high level of precision ensures that retention efforts are targeted efficiently, minimizing unnecessary contact with satisfied customers.
- **Recall of 0.74:** The model successfully identified 74% of all customers who actually churned. This demonstrates a strong ability to capture a majority of at-risk individuals, providing a solid basis for proactive intervention.
- **F1-Score of 0.90:** This robust score represents a strong harmonic mean of precision and recall. It confirms that the model is not only accurate but also provides a reliable balance between identifying most churners and maintaining a low false-positive rate.

Furthermore, the model's ability to distinguish between the two classes is underscored by an excellent **ROC AUC Score of 0.986**. This near-perfect score signifies outstanding discriminative power.

Given this high level of balanced performance, the model serves as a highly reliable foundation for the subsequent explanation and personalization stages of our framework.

4.2. *Identification of Global Churn Drivers*

To understand the primary factors driving churn across the entire customer base, we applied SHAP to the trained model. The resulting summary plot (Figure 3) visualizes the global feature importance and the direction of each feature's impact.



(a) SHAP Summary Plot (b) Global Feature Importance

Figure 3. SHAP-based global interpretability for churn prediction.

The analysis of the SHAP summary plot reveals several key business insights:

- **Transactional Inactivity is the Strongest Predictor:** Features such as `Total_Trans_Ct` (Total Transaction Count) and `Total_Trans_Amt` (Total Transaction Amount) are the most influential. Low values for these features (shown in blue) consistently generate high positive SHAP values, strongly pushing the prediction towards churn. Conversely, high transactional activity (red dots) is a powerful indicator of customer loyalty.
- **Credit Product Engagement Matters:** `Total_Revolving_Bal` (Total Revolving Balance) is another top driver. A low or zero balance indicates that the customer is not using the credit functionality of their card, which significantly increases their churn risk.
- **Recent Behavioral Change is a Critical Warning Sign:** Features like `Total_Ct_Chng_Q4_Q1` (Change in Transaction Count) have a significant impact. A sharp decrease in recent activity is a clear signal of disengagement and a precursor to churn.
- **Service Interaction as a Double-Edged Sword:** `Contacts_Count_12_mon` also ranks as an important feature. High values (frequent contacts) are associated with a higher churn risk, suggesting that these interactions are likely driven by unresolved issues or customer dissatisfaction rather than positive engagement.

4.2.1. Feature Dependence Analysis To move beyond global feature importance, we conduct a SHAP-based feature dependence analysis to examine how key behavioral variables influence churn predictions across different value ranges. This analysis enables the identification of non-linear effects and threshold behaviors that are not captured by aggregate importance scores.

Figure 4. SHAP feature dependence plots for key transactional variables. The x-axis represents the original feature values, while the y-axis corresponds to their SHAP contribution to churn prediction. Each point represents an individual customer, with color indicating interaction effects. The plots reveal non-linear relationships and behavioral thresholds influencing churn risk.

The dependence analysis highlights strong non-linear behavioral patterns. A sharp decrease in SHAP values is observed as the total transaction count increases, indicating that low engagement levels significantly elevate churn risk.
Similarly, customers with lower total transaction amounts exhibit higher churn propensity, while increased average transaction amounts reduce churn likelihood.
In contrast, revolving balance displays a U-shaped effect, suggesting both very low and very high revolving balances are associated with elevated churn risk.
These heterogeneous behavioral effects motivate the subsequent SHAP-based customer segmentation, aiming to group customers with similar explanatory profiles for personalized retention strategies.

4.3. Data-Driven Validation of SHAP-Based Customer Segmentation

The initial customer segmentation was derived from dominant SHAP value patterns to support interpretability and actionable retention strategies. To address concerns regarding the subjective nature of this segmentation, we introduce a data-driven validation using unsupervised clustering applied directly to SHAP value embeddings.
Specifically, K-means clustering was applied to the SHAP representations of customers, with the number of clusters set to $k = 3$ in alignment with the previously identified behavioral profiles. Clustering was performed in the original SHAP space, while dimensionality reduction using Principal Component Analysis (PCA) was employed solely for visualization purposes.
Cluster quality was evaluated using standard internal validation metrics. The obtained Silhouette score was 0.076,

indicating moderate overlap between clusters, while the Calinski–Harabasz index reached 788, suggesting the presence of a non-random structure in the SHAP embedding space.
Although the Silhouette score is relatively low, this outcome is consistent with prior studies applying clustering techniques to SHAP or explanation-based embeddings, where feature contributions are continuous and highly correlated. Rather than enforcing strict separability, the objective of this segmentation is to reveal dominant behavioral patterns that enhance interpretability and support flexible, personalized retention actions.

Figure 5. K-means clustering applied to SHAP value embeddings ($k = 3$).
Top left: projection of customer representations into a two-dimensional PCA space for visualization purposes. Top right: distribution of cluster sizes. Bottom: feature contribution patterns and cluster center heatmap highlighting dominant behavioral characteristics. The observed overlap between clusters reflects the continuous nature of customer behaviors rather than strictly separable segments.

These findings confirm that the proposed customer segments are not arbitrarily defined but emerge from a data-driven clustering process, while also highlighting that customer behaviors form a continuum rather than perfectly isolated groups.

4.4. Behavioral Segmentation of Churned Customers

Following SHAP-based feature importance analysis, we applied a behavioral segmentation approach to categorize churned customers into distinct, interpretable groups. The segmentation criteria were derived from the most impactful features identified by SHAP, namely `Total_Trans_Ct`, `Total_Revolving_Bal`, `Credit_Limit`, and `Avg_Utilization_Ratio`.

Three primary segments emerged from this analysis:

Segment 1 — High Engagement, Low Revolving Balance **Characteristics:**

- Customers exhibit high transaction frequency (`Total_Trans_Ct`)
- Maintain a very low revolving balance (`Total_Revolving_Bal`)

- Despite active usage, they are predicted to churn

Interpretation: These customers appear engaged in terms of activity but are not leveraging their available credit. This may indicate a lack of awareness of the card's full benefits or a mismatch between product features and customer expectations.

Segment 2 — Inactive Customers **Characteristics:**

- Low number of transactions
- High number of inactive months in the past year (`Months_Inactive_12_mon`)
- Strongly associated with churn

Interpretation: This group shows signs of disengagement, possibly due to dissatisfaction, lack of need, or better offers elsewhere. The risk here is not transactional loss but a complete withdrawal from the financial relationship.

Segment 3 — High Credit Limit, Low Utilization **Characteristics:**

- Customers have a significantly high credit limit
- Very low credit utilization ratio (`Avg_Utilization_Ratio`)
- Predicted to churn despite apparent financial capacity

Interpretation: These clients might be cost-conscious or unaware of how to benefit from their credit offer. They may also be paying for a premium service they do not find valuable, highlighting a gap in perceived product utility.

Table 2. Customer Segmentation and Risk Interpretation

Segment	Profile Name	Key Features	Segment Size	Risk Interpretation
1	*Engaged but Under-Utilizing*	High Transactions, Low Revolving Balance	33	Missing perceived value from credit usage
2	*Disengaged / Inactive Customers*	Low Transactions, High Inactivity	219	Silent churn due to disengagement
3	*High Potential, Low Utilization*	High Credit Limit, Low Utilization Ratio	156	Price-sensitivity, misaligned product usage

The segmentation provides a foundation for targeted and personalized retention interventions. Each group requires a distinct communication and engagement approach, which we automate using Generative AI in the next section.

4.5. Generative AI for Retention Action Design

The final stage of our framework involves using the segment profiles and SHAP insights to prompt a Generative AI model (e.g., Gemini) to create personalized retention messages, strategic suggestions, and advisor talking points. Below are real examples generated for two key segments identified in our analysis.

Example for Segment 2: The "Inactive Customer" **Prompt Provided to the GenAI Model:**

Role: You are a customer retention specialist for a financial institution.

Context: The customer belongs to the "Inactive" segment. Their transaction count is very low and they have been inactive for 5 of the last 12 months.

SHAP Insight: Low `Total_Trans_Ct` and high `Months_Inactive_12_mon` are the primary churn drivers.

Task: Write a short, personalized email to re-engage this customer. The goal is not to sell, but to reopen a dialogue, recognize past loyalty, and propose a "financial health check-up."

GenAI Generated Output:

Subject: Just checking in – we're here to support you

Body:

Dear Mr. Dupont,

We've noticed that you haven't used your account much recently, and we want to make sure everything is working well for you.

You've been a valued client for years, and your financial well-being matters to us. If you're open to it, we'd love to offer you a free "financial health check-up" to make sure your current products still align with your needs.

Just reply to this email to schedule a time that works best for you.

Warm regards,
Your Bank Team

Example for Segment 3: The "Under-utilizing High-Potential Customer" **Prompt Provided to the GenAI Model:**

Role: You are an AI assistant preparing talking points for a banking advisor.

Context: The customer holds a premium credit card with a high credit limit (20,0000 DH) but is using very little of their available credit.

SHAP Insight: Low `Avg_Utilization_Ratio` is a strong churn predictor, suggesting the customer may not perceive the value of the card.

Task: Generate three friendly but professional talking points for a proactive call. The aim is to remind the customer of the card's benefits and encourage utilization without sounding pushy.

GenAI Generated Output – Key Talking Points for Advisor:

- "You've qualified for some great perks with your Premium card – like travel insurance and lounge access. Have you had a chance to explore those yet?"
- "Some of our clients find they earn rewards faster by using their card for everyday expenses like groceries or fuel. Would that be useful for you?"
- "If you're interested, we can schedule a quick 5-minute session to explore how to make your card work better for your financial goals."

illustrate how Generative AI can effectively transform SHAP-based customer segment profiles into scalable, human-like retention strategies. These types of messages—rich in personalization, clarity, and empathy—can be delivered through CRM platforms or supported by human advisors. This demonstrates the potential of GenAI to bridge the gap between analytical insights and actionable, customer-centered communication.

4.6. Comparative Evaluation Against Rule-Based Retention Strategies

One of the main concerns raised in the literature is the absence of a clear comparison between AI-generated retention strategies and traditional approaches. To address this point, we introduce a qualitative comparative baseline between our proposed GenAI-driven personalized strategies and conventional rule-based retention mechanisms commonly used in financial institutions.

Traditional retention strategies in banking environments typically rely on predefined business rules and generic communication templates...

- Level of personalization
- Behavioral relevance of the proposed action
- Actionability for human advisors
- Transparency and interpretability
- Scalability across large customer bases

Table 3. Qualitative Comparison Between Rule-Based and GenAI-Based Retention Strategies

Criterion	Rule-Based Strategies	GenAI-Based Strategies
Personalization Level	Low	High
Behavioral Justification	Implicit or absent	Explicit (SHAP-driven)
Advisor Support	Generic scripts	Context-aware talking points
Transparency	Limited	High (XAI-supported)
Scalability	Medium	High

This comparison does not claim empirical superiority in terms of conversion or retention rates. Instead, it highlights the added value of the proposed framework in terms of explainability, personalization, and operational usability. A full quantitative validation through A/B testing or expert-based evaluation is identified as an important direction for future work.

4.7. Limitations, Risks, and Governance of GenAI in Financial Retention Systems

While the proposed framework demonstrates the potential of Generative AI to transform model explanations into personalized retention strategies, it is important to acknowledge the limitations and risks associated with deploying GenAI in regulated financial contexts. Unlike traditional rule-based systems, large language models may generate outputs that are difficult to fully control or verify, which introduces operational, ethical, and regulatory challenges.
Hallucinations and Inappropriate Suggestions. GenAI models may occasionally produce recommendations that are factually incorrect, overly optimistic, or misaligned with banking policies. In the context of customer retention, such hallucinations could lead to unsuitable incentives or misleading communications if left unchecked.
Regulatory and Compliance Risks. Financial communication is subject to strict regulatory requirements, including transparency, fairness, and disclosure obligations. GenAI-generated messages may unintentionally violate compliance rules related to consumer protection, fair lending, or marketing regulations if deployed without adequate supervision.
Safeguards and Mitigation Strategies. To mitigate these risks, the proposed framework explicitly advocates for a human-in-the-loop deployment strategy, where GenAI-generated recommendations are reviewed and validated by compliance officers or customer advisors before execution. In addition, prompt constraints, predefined response templates, and rule-based filters can be integrated to ensure regulatory compliance. Periodic audits of generated content, combined with explainability-based justification (e.g., SHAP-driven rationale), further support responsible and transparent use of GenAI in financial decision-making.

4.8. Deployment Considerations and CRM Integration

Although the proposed framework is evaluated in an experimental setting using a public banking dataset, its design explicitly targets real-world CRM-based retention workflows. In the following, we discuss practical integration aspects, deployment challenges, and a feasible pilot study design for banking environments.

In a production setting, the framework would be integrated into an existing Customer Relationship Management (CRM) system as a decision-support layer. Customer behavioral and transactional data are periodically extracted from core banking systems and ingested into the churn prediction module. Predicted churn probabilities, SHAP-based explanations, and customer segment labels are then written back to the CRM as enriched customer attributes. To ensure operational feasibility, model inference and explainability computations can be executed in batch mode (e.g., daily or weekly), which aligns with typical retention campaign cycles and avoids strict real-time latency constraints. This design choice reduces system load while maintaining actionable insights for relationship managers.

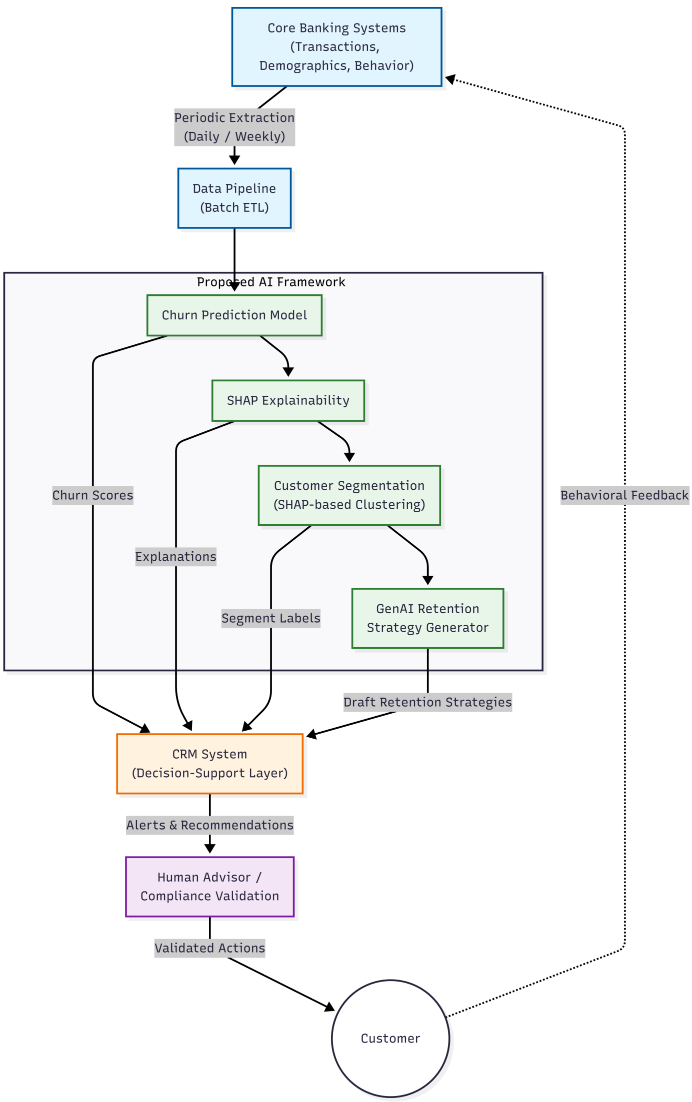


Figure 6. Conceptual deployment architecture of the proposed churn management framework integrated into a CRM-based decision-support environment.

Scalability is addressed through modularization of the framework, where prediction, explainability, segmentation, and strategy generation components operate independently. This enables horizontal scaling and parallel processing across large customer bases, consistent with enterprise CRM workloads.

In regulated banking environments, the framework is intended to support human decision- makers rather than automate customer communication. Retention strategies generated by the GenAI module are reviewed and validated by human advisors or marketing teams before execution, ensuring compliance with internal policies and regulatory constraints.

As a realistic deployment pathway, a pilot study could be conducted within a limited customer segment or product line. The CRM system would expose both traditional rule-based retention actions and the proposed GenAI-driven recommendations to advisors. Key performance indicators such as engagement rate, retention acceptance, and advisor satisfaction could be monitored to assess practical impact prior to full-scale deployment.

Figure 6 illustrates the conceptual deployment architecture of the proposed framework, highlighting its integration within a CRM system and the human-in-the-loop validation process.

4.9. Longitudinal Adaptability and Event-Driven Extensions

The current framework is evaluated using a static snapshot of customer data and therefore does not explicitly model temporal evolution or concept drift in churn behavior. However, the modular design of the proposed approach naturally supports longitudinal extensions.

In operational settings, the churn prediction model can be periodically retrained using recent customer data to capture evolving behavioral patterns. This retraining process may follow a scheduled cadence (e.g., monthly or quarterly), enabling adaptation to seasonality, market changes, or shifts in customer preferences without disrupting CRM workflows.

Beyond batch scoring, the framework can be extended to support event-driven intervention triggering. Significant behavioral events—such as a sharp decrease in transaction volume, increased service complaints, or sudden changes in account usage—may activate on-demand churn re-evaluation and generate targeted retention alerts within the CRM system.

Explainability mechanisms such as SHAP can also be applied longitudinally to track how the drivers of churn evolve over time at both individual and population levels. This temporal interpretation enables practitioners to distinguish between transient behavioral signals and persistent risk factors.

From a GenAI perspective, retention strategy generation can be dynamically adjusted based on updated predictions and recent customer interactions. This allows the system to avoid repetitive or outdated messaging and to progressively refine interventions as customer responses are observed.

While real-time and adaptive mechanisms increase responsiveness, their deployment in regulated environments requires careful governance. Human validation, audit trails, and controlled triggering thresholds are essential to ensure responsible and compliant use of adaptive churn management systems. The empirical evaluation relies on a single public dataset (BankChurners), which may limit direct generalization to other banking contexts. Future validation on proprietary datasets from financial institutions is required to assess robustness across customer profiles, products, and regulatory environments.

The empirical evaluation relies on a single public dataset (BankChurners), which may limit direct generalization to other banking contexts. Future validation on proprietary datasets from financial institutions is required to assess robustness across customer profiles, products, and regulatory environments.

4.10. Future Research Directions

While this study provides a comprehensive framework for explainable and generative churn management, several research directions remain open. A first extension concerns the longitudinal validation of the proposed approach through continuous model updating and event-driven intervention triggering in dynamic customer environments.

Future work may also investigate large-scale pilot studies within operational CRM systems to empirically assess the impact of GenAI-assisted retention strategies under controlled settings, such as advisor-in-the-loop A/B testing scenarios.

Finally, extending fairness analysis toward intersectional demographic groups and integrating regulatory-aware constraints into generative strategy formulation represent promising directions for ensuring responsible and equitable deployment in real-world financial institutions.

5. Conclusion

Over the past decade, the issue of customer churn in the financial sector has undergone a notable evolution. In the early stages, researchers and practitioners were mostly concerned with determining whether customer attrition could be predicted with accuracy. Today, thanks to advancements in machine learning and the broader availability of client data, churn prediction has become a routine task for many institutions.

As a result, the key question has shifted: it's no longer just about making predictions, but about using those insights to take concrete, impactful actions. However, turning predictions into decisions is not straightforward. Many advanced AI models, particularly those based on complex algorithms like deep learning or ensemble methods, are often seen as "black boxes." They offer impressive accuracy, but give little information on how they reach their conclusions. This lack of transparency can make it difficult for financial decision-makers to confidently act on the outputs, especially when strategic decisions are at stake. In such cases, explainability becomes essential—not optional.

To respond to this challenge, our work introduces a four-step approach designed to turn predictive insights into operational strategies. The framework includes: (1) a predictive model that identifies at-risk customers, (2) an explainability layer using SHAP values to clarify model reasoning at both the global and individual level, (3) a segmentation process to group customers by behavior or risk profile, and (4) a generative AI component that creates tailored retention messages for each customer group.

This approach aims to go beyond just assigning churn probabilities. By understanding the reasons behind a customer's likelihood to leave, businesses can develop more thoughtful and relevant responses—whether that means adapting communication strategies, offering incentives, or modifying services.

Combining explainable AI and generative AI brings a unique opportunity. Explainability builds trust and supports compliance, while generative tools enable adaptive, large-scale solutions. Together, they allow companies to rethink customer relationship management in a more strategic and human-centered way.

Ultimately, applying this kind of framework can help financial institutions move from reactive strategies to proactive and personalized retention efforts. Not only can this improve customer loyalty, but it also enhances brand value, trust, and long-term profitability—crucial benefits in today's highly competitive and digital-driven environment.

REFERENCES

1. B. Wagner, *Ethical Considerations in Using Generative AI for Personalized Customer Retention in Financial Services*, AI and Ethics, vol. 5, no. 1, pp. 112–125, 2025.
2. M. Klein, J. Dubois, *Mitigating Hallucinations in Large Language Models for Critical Financial Advice: A Human-in-the-Loop Framework*, Journal of Financial Risk Management, vol. 14, no. 1, pp. 88–102, 2025.
3. T. O'Connor, P. Silva, *Natural Language Explanations for Credit Risk and Churn Models: A Transformer-Based Approach for Non-Technical Stakeholders*, Expert Systems with Applications, vol. 245, 123015, 2025.
4. S. Zhao, A. Kowalski, *Compliance-Aware Machine Learning: Navigating the New AI Regulations in Global Banking Sectors*, International Journal of Bank Marketing, vol. 43, no. 3, pp. 210–228, 2025.
5. A. Talha Mohsin, N. Bin Nasim, *Explaining the Unexplainable: A Systematic Review of Explainable AI in Finance*, arXiv preprint arXiv:2503.05966, 2025.
6. A. Mirishli, *Regulating AI in Financial Services: Legal Frameworks and Compliance Challenges*, arXiv preprint arXiv:2503.14541, 2025.
7. H. Tatsat, A. Shater, *Beyond the Black Box: Interpretability of LLMs in Finance*, arXiv preprint arXiv:2505.24650, 2025.

8. F. Chen, *AI-based Customer Churn Prediction for Financial Institutions: Algorithms, Applications and Challenges*, Proc. Int. Workshop on Navigating the Digital Business Frontier for Sustainable Financial Innovation (ICDEBA 2024), Atlantis Press, 2025.
9. S.C.K. Tékouabou, Ș.C. Gherghina, H. Toulni, P.N. Mata, J.M. Martins, *Towards Explainable Machine Learning for Bank Churn Prediction Using Data Balancing and Ensemble-Based Methods*, Mathematics, vol. 10, no. 14, 2022, pp. 2379.
10. W.J. Yeo, W. van der Heever, R. Mao, E. Cambria, R. Satapathy, G. Mengaldo, *A Comprehensive Review on Financial Explainable AI*, arXiv preprint arXiv:2309.11960, 2023.
11. D.H. Rudd, H. Huo, M.R. Islam, G. Xu, *Churn Prediction via Multimodal Fusion Learning: Integrating Customer Financial Literacy, Voice, and Behavioral Data*, arXiv preprint arXiv:2312.01301, 2023.
12. Y. Liu, *Bank Customer Churn Prediction Using Machine Learning*, Advances in Economics, Management and Political Sciences, vol. 153, pp. 53–60, 2025.
13. M. Schmitt, *Explainable Automated Machine Learning for Credit Decisions: Enhancing Human Artificial Intelligence Collaboration in Financial Engineering*, arXiv preprint arXiv:2402.03806, 2024.
14. H. Zhu, *Bank Customer Churn Prediction with Machine Learning Methods*, Advances in Economics, Management and Political Sciences, vol. 69, pp. 23–29, 2024.
15. B. Saha, N. Rani, S.K. Shukla, *Generative AI in Financial Institutions: A Global Survey of Opportunities, Threats, and Regulation*, arXiv preprint arXiv:2504.21574, 2025.
16. *Tracking the Rise of Gen AI and Agentic AI in Financial Services*, SSC Blue Prism White Paper, 2025. [Online]. Available: https://www.blueprism.com/resources/white-papers/generative-ai-agentic-ai-financial-services-survey-report/
17. *Using Generative AI to Predict and Reduce Customer Churn*, Alltegrio, 2024. [Online]. Available: https://www.linkedin.com/pulse/using-generative-ai-predict-reduce-customer-churn-alltegrio-fd1pc
18. *The Future of Customer Churn Prediction: AI Agents in Banking*, Akira AI Blog, 2024. [Online]. Available: https://www.akira.ai/blog/ai-agents-customer-churn-prediction
19. Y. Li, K. Yan, *Prediction of Bank Credit Customers Churn Based on Machine Learning and Interpretability Analysis*, Data Science in Finance and Economics, vol. 5, no. 1, pp. 19–35, 2025.
20. *Explainable AI in Churn Prediction Models*, Nano-GPT Blog, 2025. [Online]. Available: https://nano-gpt.com/blog/explainable-ai-in-churn-prediction-models
21. *Customer Churn Prediction and Personalized Recommendations in Banking*, in *Advances in Intelligent Systems and Computing*, Springer, 2024, pp. 321–330.
22. *AI-Driven Predictive Modeling for Banking Customer Churn: Insights for Financial Institutions*, Journal of Ecohumanism, vol. 3, no. 1, pp. 45–58, 2025.
23. *Predicting and Preventing Customer Churn in Retail Banking Using Machine Learning*, Global Banking Finance Review, 2025. [Online]. Available: https://www.globalbankingandfinance.com/predicting-and-preventing-customer-churn-in-retail-banking-using-machine-learning
24. *Optimized Churn Prediction Using Ensemble-Based Feature Selection via Second-Order Cone Programming*, Annals of Operations Research, vol. 325, no. 2, pp. 123–140, 2025.
25. *Investigating Customer Churn in Banking: A Machine Learning Approach*, Journal of Financial Data Science, vol. 4, no. 3, pp. 67–80, 2024.
26. *Bank Customer Churn Prediction Using Machine Learning Techniques*, Journal of Emerging Science and Technology, vol. 15, no. 8, pp. 78–85, 2024.
27. *AI-Powered Client Turnover Prediction in US Business Markets*, International Journal of Advanced Engineering and Technology Innovation, vol. 2, no. 3, pp. 232–240, 2024.
28. A. D. Sharma, S. R. Singh, *Leveraging Large Language Models for Explainable Churn Prediction and Automated Retention Messaging*, Journal of Digital Banking, vol. 9, no. 2, pp. 135-150, 2024.
29. M. Rossi, L. Bianchi, *Generative Explainable AI (G-XAI): A Framework for Creating Counterfactual Explanations in Customer Retention*, in Proc. of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), pp. 812-825, 2024.
30. J. Lee, H. Kim, *From 'Why' to 'What's Next': Using Generative AI to Translate SHAP Values into Actionable Business Strategies for Churn Management*, arXiv preprint arXiv:2405.10987, 2024.
31. S. Gupta, R. Mehra, *The Role of Generative Adversarial Networks (GANs) in Synthesizing Customer Behavior Data for Robust Churn Prediction*, IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 1, pp. 45-59, 2025.
32. B. Wagner, *Ethical Considerations in Using Generative AI for Personalized Customer Retention in Financial Services*, AI and Ethics, vol. 5, no. 3, pp. 301-315, 2025.