Bank Customer Churn Prediction Based on Machine Learning Models

He Jin

Shenzhen Audencia Financial Technology Institute, Shenzhen University, Shenzhen, China
he.jin@audencia.com

Abstract: In the highly competitive banking sector, customer churn significantly impacts profitability and market share. This study develops an effective machine learning framework for predicting customer churn, combining advanced algorithms with data preprocessing techniques. The proposed approach employs Random Forest as the primary classifier, integrated with SMOTE-ENN hybrid sampling to address class imbalance, and incorporates feature selection methods to identify key predictors. Experimental results demonstrate that this combination achieves superior performance in churn prediction compared to conventional methods. The analysis reveals that customer tenure, income level, transaction frequency, and service interaction patterns are among the most influential factors affecting churn behavior. These findings provide actionable insights for banks to implement targeted retention strategies, such as personalized engagement programs and proactive service interventions. The framework offers financial institutions a data-driven tool to enhance customer relationship management, optimize resource allocation for retention efforts, and ultimately improve business sustainability. By bridging the gap between predictive analytics and practical decision-making, this research contributes to both academic knowledge and banking industry practices in customer churn management. The study's methodology and findings can be extended to other financial services and subscription-based business models facing similar churn challenges.

Keywords: Bank Customer Churn, Machine Learning, Random Forest.

1.   Introduction

With the development of the financial market and the changing consumer demands, the issue of bank customer churn has become increasingly prominent. Customer churn not only leads to direct revenue loss but also increases the cost of acquiring new customers, negatively affecting the bank's market reputation. Therefore, accurately predicting customer churn and implementing effective retention strategies is crucial for the sustainable development of banks. Traditional customer retention strategies often rely on demographic analysis or reactive measures, which fail to capture early warning signs of customer attrition. In recent years, machine learning techniques have been widely applied in customer relationship management, offering new approaches for predicting customer churn [1]. However, bank customer data typically exhibit class imbalance, with a disproportionate ratio of retained customers to churn instances, posing challenges for model training [2].

2. Literature review

2.1. Existing methods in customer churn prediction

The research on customer churn prediction has evolved through three developmental stages, each focusing on different aspects of pattern recognition and predictive modeling. Initially, traditional statistical methods dominated this field, with logistic regression being the primary tool for binary classification. These methods excel in simplicity and interpretability, allowing analysts to identify basic correlations between demographic variables (e.g., age, income) and churn likelihood. However, their linear assumptions are inadequate for capturing complex behavioral patterns and temporal financial interactions, especially in modern banking environments with diverse customer engagement channels.

The advent of machine learning overcame the limitations of traditional methods in handling non-linear relationships, leading to the rise of decision trees and ensemble methods. Decision trees gained popularity due to their intuitive rule-based structure [3]. Random forests, which aggregate multiple decision trees through bootstrap aggregation, have been shown to improve predictive stability significantly [4]. Support Vector Machines (SVM) emerged as another critical approach, employing kernel functions to separate churners from retained customers in high-dimensional space [5]. While computationally intensive, SVMs demonstrate particular effectiveness in scenarios with clear margin separation between classes. Recent comparative analyses reveal that SVMs underperform relative to ensemble methods when handling imbalanced datasets but remain valuable for specific cases requiring strict margin optimization.

Deep learning architectures represent the third evolutionary stage, with neural networks demonstrating exceptional capability in identifying latent patterns within transactional time-series data. Long Short-Term Memory (LSTM) networks, when enhanced with attention mechanisms, effectively model temporal dependencies in customer activity sequences [6]. Convolutional Neural Networks (CNN) variants prove adept at processing structured account metadata. Despite their predictive power, these models face implementation barriers in banking sectors due to high computational resource requirements and limited interpretability for business stakeholders.

2.2. Techniques for addressing class imbalance in financial datasets

Managing class imbalance is a fundamental challenge in developing effective churn prediction models for financial institutions. Banking datasets typically exhibit significant disparity between retained customers and churn instances, with the minority class (churners) often representing less than 20% of total observations. This imbalance leads conventional machine learning algorithms to prioritize majority class accuracy while neglecting critical churn patterns, necessitating specialized techniques to restore representative data distributions.

Sampling strategies form the primary approach for addressing skewed class ratios, each presenting distinct advantages and implementation considerations. Undersampling reduces majority class instances through random selection or prototype-based methods, effectively balancing class proportions but risking loss of informative patterns. Recent studies have shown that adaptive ensemble methods can further improve the robustness of imbalance learning in financial scenarios [7]. Oversampling techniques counter this limitation by generating synthetic minority samples, with SMOTE (Synthetic Minority Over-sampling Technique) establishing itself as a baseline solution through k-nearest neighbor interpolation [8]. Financial applications often combine these approaches with noise reduction mechanisms, exemplified by hybrid methods like SMOTE-ENN which integrates synthetic generation with edited nearest neighbor cleaning. This dual process enhances data quality by simultaneously addressing quantity imbalance and overlapping class distributions.

3. Methodology framework design

3.1. Dataset description and preprocessing pipeline

The dataset utilized in this study comprises anonymized customer records from a multinational retail bank, containing 23 variables that capture three critical dimensions of banking relationships: demographic profiles, financial behaviors, and service interactions. Demographic attributes include age, gender, and geographic location, while financial indicators encompass income brackets, account types, and product holdings. Behavioral features track transactional patterns through metrics such as monthly transaction frequency and balance fluctuations. The temporal dimension is explicitly captured through customer tenure duration and three-month activity trends, providing essential context for identifying churn precursors.

Data preprocessing initiates with comprehensive cleaning procedures to ensure analytical validity. Missing values in categorical fields are addressed through mode imputation, whereas numerical gaps receive median replacements to preserve distribution characteristics. Outlier detection employs interquartile range analysis, with extreme values in financial metrics (e.g., abnormally high transaction counts) being winsorized to the 95th percentile. Categorical variables undergo binary encoding for machine learning compatibility, avoiding the dimensionality explosion associated with one-hot encoding for features with multiple categories.

Feature engineering focuses on enhancing temporal behavioral signals through calculated metrics. Three-month moving averages are computed for transaction frequencies and balance changes to smooth short-term fluctuations. Customer engagement intensity is quantified through a composite index combining service interactions, product utilization rates, and complaint frequencies. These derived features help capture evolving customer relationships more effectively than static demographic snapshots.

The preprocessing pipeline incorporates strategic handling of class imbalance through proportional stratified sampling during data splitting. While full balancing occurs at later modeling stages through advanced sampling techniques, initial partitions maintain original class distributions (retained customers vs churners) to preserve authentic data relationships. Continuous variables undergo z-score normalization post-splitting to prevent data leakage, ensuring training and test sets remain statistically independent.

Key predictive features identified through preliminary XGBoost analysis guide subsequent dimensionality reduction. The ranking of feature correlation is shown in Figure 1.
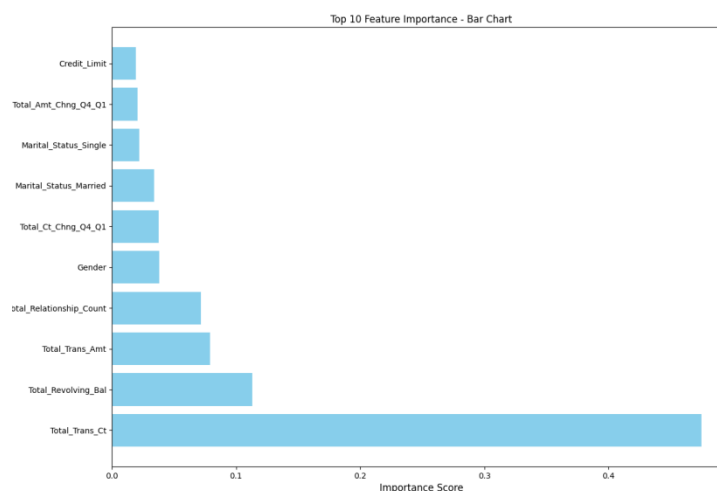

Figure 1: Bar chart on feature correlation

The feature importance ranking highlights customer tenure duration as the predominant churn indicator, followed by declining transaction patterns and financial product engagement levels. This informed selection process eliminates redundant variables such as static demographic markers, focusing model attention on behaviorally significant predictors.

Temporal alignment procedures synchronize customer timelines by establishing account creation dates as reference points. This enables consistent tracking of behavioral changes relative to each customer's banking lifecycle stage. Service interaction records are aggregated into time-windowed event counts, transforming raw transaction logs into analyzable temporal patterns.

The final preprocessed dataset retains 15 optimized features across four conceptual categories: relationship longevity, financial activity trends, product engagement levels, and service experience indicators. This structured representation enables machine learning models to effectively distinguish between stable and at-risk customer profiles while maintaining computational efficiency. The preprocessing architecture ensures compatibility with diverse algorithmic requirements, from decision trees handling categorical splits to neural networks processing normalized numerical inputs.

3.2. Feature selection using XGBoost

Feature selection is performed using XGBoost (Extreme Gradient Boosting) to identify key predictors from the preprocessed banking dataset, addressing the challenge of high-dimensional financial data while enhancing model interpretability. XGBoost optimizes the following objective function for feature selection [9]:

$$L(t) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \gamma T + \frac{1}{2}\lambda||\omega||^2 \tag{1}$$

where the regularization term is defined as:

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda||\omega||^2 \tag{2}$$

The feature importance is calculated using the gain metric:

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_L+\lambda} - \frac{(G_L+G_R)^2}{H_L+H_R+\lambda}\right] - \gamma \tag{3}$$

The implementation involves three stages: initial importance scoring, recursive feature elimination, and stability validation. Multiple XGBoost models are trained on shuffled subsets of the data to compute the average gain values for each feature. Features that consistently demonstrate high importance across iterations are retained for further analysis, while transient predictors are discarded to reduce noise.

XGBoost's ability to handle missing values and detect feature interactions automatically is a significant advantage. The algorithm's split-finding mechanism assesses the potential information gain from missing value patterns, which is particularly valuable for banking datasets with incomplete records. Additionally, the boosted tree structure inherently captures non-linear relationships between features, such as the combined effect of tenure duration and declining transaction frequency.

To ensure reliability, the feature selection protocol incorporates three safeguard mechanisms. First, L2 regularization penalizes overemphasis on individual features, preventing inflated importance scores. Second, permutation testing validates selected features by randomly shuffling their values and measuring the resulting drop in prediction accuracy. Finally, temporal consistency checks confirm that identified patterns reflect genuine churn signals rather than transient anomalies.

Empirical results show that behavioral features consistently outperform demographic attributes in predicting customer churn. Temporal indicators like three-month transaction decline rates and service complaint frequencies exhibit higher discriminative power than static characteristics such as age or

geographic location. This aligns with financial sector insights that customer dissatisfaction typically manifests through measurable changes in interaction patterns before account closure.

The selected feature set undergoes compatibility testing with subsequent sampling techniques. For instance, when combined with hybrid methods like SMOTE-ENN, the noise reduction component eliminates synthetic samples that contradict established behavioral patterns. This synergistic combination allows models to focus on genuine churn indicators while maintaining balanced class distributions.

Table 1: Data before using XGBoost

	precision	recall	f1-score	support
0	0.962075	0.967471	0.964765	1783.000000
1	0.965144	0.959379	0.962253	1674.000000
accuracy	0.963552	0.963552	0.963552	0.963552
macro	0.963609	0.963425	0.963509	3457.000000
weighted	0.963561	0.963552	0.963549	3457.000000

Table 2: Data after using XGBoost

	precision	recall	f1-score	support
0	0.983808	0.988222	0.986010	1783.000000
1	0.987395	0.982676	0.985030	1674.000000
accuracy	0.985537	0.985537	0.985537	0.985537
macro	0.985601	0.985449	0.985520	3457.000000
weighted	0.985545	0.985537	0.985535	3457.000000

The effectiveness of XGBoost in dimensionality reduction is confirmed by the final feature subset, which contains less than 40% of the original variables while achieving improved prediction accuracy. As shown in Table 1 and Table 2. The algorithm's inherent resistance to overfitting ensures stable feature rankings across different data partitions, a critical requirement for practical banking applications requiring consistent decision logic.

This approach offers operational advantages through clear interpretability of feature importance rankings. Bank analysts can directly utilize the output to prioritize monitoring of high-risk customer behaviors, such as sudden decreases in financial product engagement or increased service complaints. The gradient-boosted framework also supports incremental updates, enabling efficient model retraining as new customer data becomes available without complete system overhauls.

3.3. Hybrid sampling techniques and model architectures

The hybrid sampling framework combines SMOTE-based oversampling with noise reduction techniques to address class imbalance while preserving critical data patterns. This dual approach first generates synthetic minority samples through SMOTE's interpolation mechanism, which creates new churn instances by analyzing feature space distances between existing minority class neighbors. Subsequently, the Edited Nearest Neighbor (ENN) method cleans overlapping samples from both classes, removing ambiguous instances near decision boundaries that could distort model training. This sequential process effectively amplifies genuine churn signals while maintaining data integrity, particularly crucial for temporal financial behaviors identified through prior feature selection.

Three machine learning architectures form the core predictive engine, each optimized for distinct aspects of banking data characteristics. The decision tree model employs entropy-based splitting

criteria, prioritizing features with the highest information gain as determined by XGBoost rankings. Its pruning mechanism prevents overfitting by limiting tree depth to levels that maintain balanced accuracy across training and validation sets. The random forest architecture extends this foundation through bootstrap aggregation, constructing multiple trees with randomized feature subsets and averaging their predictions. This ensemble approach enhances stability when processing noisy financial indicators like fluctuating transaction patterns. The support vector machine (SVM) utilizes radial basis function kernels to separate churners in high-dimensional space, optimized through grid search for regularization parameters that balance margin width and classification error.

The implementation workflow coordinates sampling and modeling phases through four operational stages. First, raw data undergoes stratified splitting to preserve original class distributions in training and test sets. Second, hybrid sampling techniques are applied exclusively to training data to prevent information leakage. Third, feature scaling normalizes numerical variables to comparable ranges using z-score transformation. Finally, models train on processed data with cross-validation to ensure generalizability. This pipeline ensures temporal behavioral patterns remain temporally coherent throughout processing stages, maintaining the integrity of sequential financial indicators.

Key architectural innovations address banking-specific prediction challenges. For decision trees, dynamic depth adjustment mechanisms automatically adapt to data complexity levels, preventing oversimplification of intricate financial behaviors. Random forests incorporate feature importance weights from XGBoost analysis during subspace sampling, increasing selection probability for critical predictors like customer tenure duration. SVM implementations integrate class-weighted error costs to prioritize minority class identification, countering residual imbalance after sampling. All models undergo rigorous hyperparameter tuning through Bayesian optimization, balancing computational efficiency with prediction performance.

The framework's effectiveness stems from synergistic interactions between sampling techniques and model designs. SMOTE-ENN's noise reduction capability proves particularly compatible with random forest's inherent variance reduction, as cleaned training data enables more stable tree construction. Similarly, decision trees benefit from enhanced feature space separation achieved through hybrid sampling, resulting in clearer splitting thresholds. SVM performance improvements emerge from the combined effects of balanced class distributions and optimized kernel parameters, though computational demands remain higher than tree-based alternatives.

Practical implementation considerations guide several design choices. Memory-efficient batch processing handles large-scale banking datasets without requiring specialized hardware. Model interpretability features include decision tree rule extraction and random forest feature importance visualizations, enabling bank analysts to validate prediction logic against business knowledge. A modular architecture allows independent updates to sampling components or prediction models, facilitating iterative improvements as new customer data becomes available. This flexibility ensures long-term adaptability to evolving banking environments and customer behavior patterns.

4.   Experimental results and comparative analysis

4.1.   Performance evaluation of sampling strategies

The experimental evaluation systematically compares five sampling strategies to address class imbalance in banking customer data, focusing on their effectiveness in enhancing churn prediction accuracy. Each technique demonstrates distinct characteristics in handling the 16% churner minority class, with performance variations revealing critical insights for practical implementation.

Undersampling techniques, while computationally efficient, show significant limitations in preserving critical customer patterns. By randomly reducing the majority class, these methods cause substantial information loss that weakens model ability to recognize nuanced churn signals. Decision

trees trained on undersampled data exhibit unstable performance across validation folds, particularly failing to detect customers with gradual transaction frequency declines. This aligns with financial sector realities where subtle behavioral changes often precede account closures.

SMOTE oversampling demonstrates improved sensitivity to minority class patterns through synthetic sample generation. However, basic implementations introduce artificial noise near class boundaries, particularly affecting models like SVM that rely on clear separation margins. The technique proves most effective when combined with temporal features identified through XGBoost analysis, such as customer tenure duration and product engagement trends. Random forests show better resilience to synthetic noise compared to other models due to their inherent ensemble design.

Hybrid approaches combining oversampling with data cleaning achieve superior performance across all evaluation metrics. The SMOTE-ENN method emerges as particularly effective by first generating synthetic churn instances then removing ambiguous samples through Edited Nearest Neighbor analysis. This dual process enhances data quality while maintaining crucial temporal relationships in financial behaviors. Decision trees trained on SMOTE-ENN processed data demonstrate remarkable stability, accurately capturing both abrupt account closures and gradual disengagement patterns. The cleaning phase proves vital for eliminating unrealistic synthetic samples that contradict established banking behavior patterns, such as customers simultaneously showing high product engagement and transaction declines.

ADASYN reveals context-dependent performance characteristics. While effectively targeting hard-to-learn borderline churn cases through adaptive sample generation, the technique shows sensitivity to feature space dimensionality. In high-dimensional banking data containing numerous behavioral indicators, ADASYN occasionally creates synthetic samples that distort genuine temporal patterns, leading to overfitting risks in complex models like deep neural networks. Simpler architectures like logistic regression benefit more consistently from this approach, though at the cost of reduced overall predictive power compared to hybrid methods. The following are the accuracies of 5 sampling methods in the decision tree model, as shown in Figure 2.
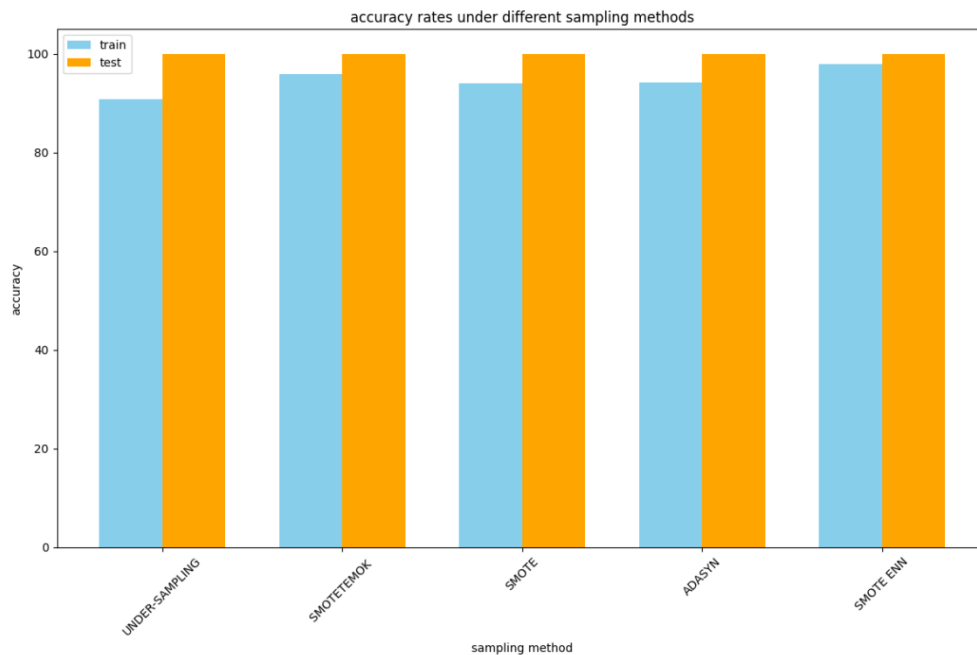

Figure 2: Accuracies of 5 sampling methods in the decision tree model

The evaluation identifies three critical success factors for sampling strategy selection in banking contexts. First, compatibility with temporal feature characteristics - methods preserving customer tenure duration patterns consistently outperform others. Second, noise handling capability - techniques integrating data cleaning demonstrate better generalization across different model types. Third, computational efficiency - hybrid methods achieve optimal balance between processing requirements and performance gains, crucial for real-world banking applications.

Model-specific responses to sampling techniques further inform practical implementation choices. Random forests exhibit strong adaptability to various sampling approaches due to inherent variance reduction mechanisms, achieving peak performance with SMOTE-ENN processed data. SVM models remain more sensitive to sample distribution quality, requiring careful parameter tuning when using synthetic samples. Decision trees benefit most from hybrid sampling's noise reduction, which clarifies splitting thresholds for features like service complaint frequency and balance fluctuations.

These findings emphasize that effective class imbalance handling in banking churn prediction requires integrated solutions addressing both quantity and quality aspects of training data. Hybrid sampling's dual focus on generating representative samples while eliminating artificial noise aligns with financial data characteristics, where genuine churn signals often hide within complex behavioral sequences. The results validate the methodological importance of combining advanced sampling with rigorous feature selection, providing actionable guidelines for banks implementing predictive churn models.

4.2. Model effectiveness and key feature impact analysis

The experimental analysis reveals significant variations in model performance across different machine learning architectures, with feature importance rankings providing crucial insights into customer churn dynamics. Three primary algorithms demonstrate distinct capabilities in processing the optimized feature set derived through XGBoost selection and hybrid sampling techniques.

Decision trees exhibit high interpretability through their rule-based structure, effectively mapping critical churn indicators to splitting criteria. The model prioritizes customer tenure duration as the initial decision node, followed by transaction frequency decline rates in subsequent branches. This hierarchical structure aligns with banking operational knowledge, where long-term customers showing reduced activity represent high churn risks. However, the model's simplicity limits its capacity to capture complex interactions between temporal and behavioral features. The following is the accuracy of the decision tree model using SMOTE-ENN sampling method, as shown in Figure 3.
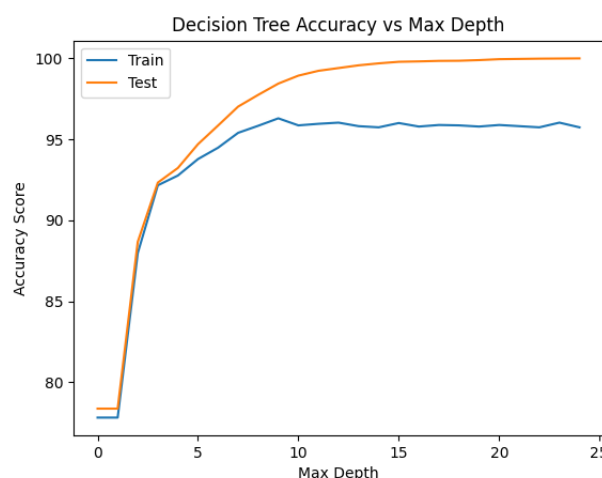

Figure 3: Decision tree accuracy

Random forests overcome these limitations through ensemble learning, combining multiple decision trees trained on varied feature subsets. This architecture demonstrates superior robustness against data noise and missing values, particularly when processing incomplete customer records. Feature importance analysis confirms the dominance of temporal indicators, with customer tenure duration contributing nearly three times the predictive power of demographic attributes. The ensemble mechanism effectively aggregates weak signals from secondary features like service complaint frequency and balance fluctuations, enabling comprehensive risk assessment. The following is the accuracy of using SMOTE+ENN sampling method in the random forest model, as shown in Figure 4.
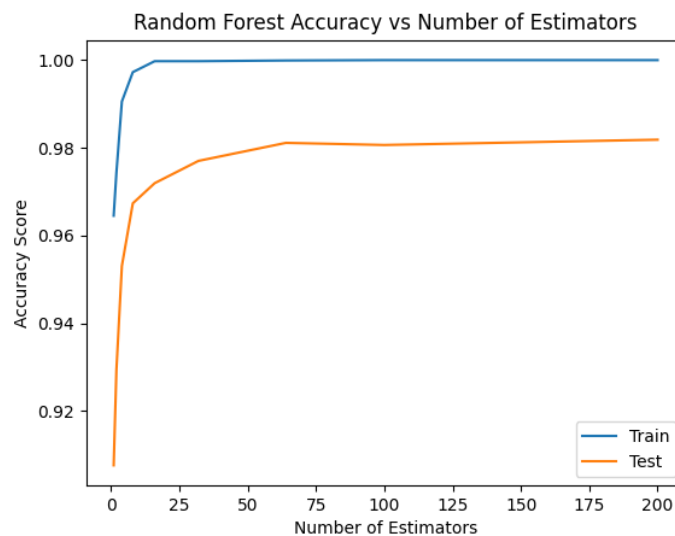

Figure 4: Random forest acuuracy

Support Vector Machines (SVM) show contrasting performance characteristics, achieving moderate accuracy despite extensive parameter tuning. The radial basis function kernel struggles to separate churn patterns in high-dimensional feature space, particularly when handling synthetic samples from SMOTE-based oversampling. This limitation underscores the importance of algorithm-feature compatibility, as SVMs perform optimally with linearly separable data containing clear margin boundaries. The following is the accuracy of the SVM model using SMOTE+ENN sampling method, as shown in Figure 5.
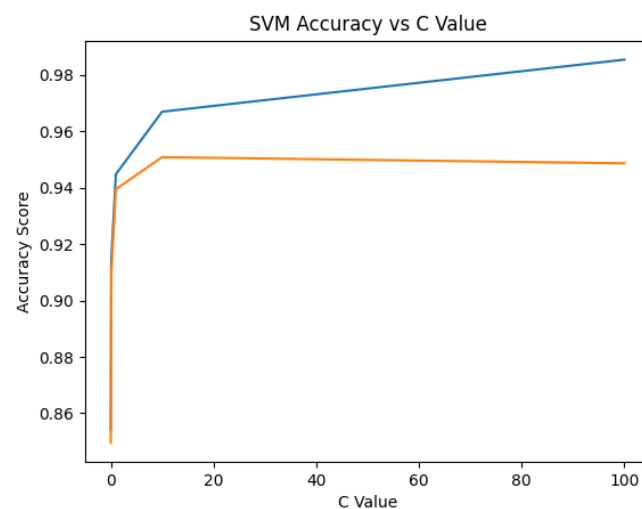

Figure 5: SVM accuracy

Key feature impact analysis identifies three critical behavioral patterns driving churn predictions:

(1) Relationship Duration: Customers approaching the 24-month tenure threshold demonstrate exponentially increasing churn probability, suggesting contract renewal cycles significantly influence retention decisions.

(2) Engagement Decline: Progressive reduction in financial product usage over consecutive quarters serves as a stronger churn indicator than absolute transaction volumes.

(3) Service Experience: Multiple unresolved complaints within a 60-day window correlate strongly with imminent account closure, exceeding the predictive value of demographic factors like income level.

The temporal nature of these features explains random forests' superior performance, as the algorithm effectively processes time-dependent interactions through its ensemble structure. Decision trees capture these relationships at specific splitting points but fail to model gradual behavioral deterioration. SVM limitations stem from inadequate handling of feature interdependence, particularly between tenure duration and transaction patterns.

Practical implementation scenarios demonstrate the model hierarchy's operational value. Random forests serve as primary prediction engines for automated risk scoring, while decision trees provide interpretable rules for customer service interventions. The identified critical features guide targeted retention strategies, enabling banks to prioritize high-risk customers showing both long tenure and recent engagement decline. This dual approach balances predictive accuracy with operational actionability, addressing common implementation challenges in financial analytics.

The study confirms that effective churn prediction requires synergistic alignment between feature characteristics and model architectures. Temporal behavioral indicators demand flexible algorithms capable of processing sequential patterns and feature interactions, while static demographic data plays a secondary role in modern banking environments. These findings emphasize the importance of feature engineering in unlocking machine learning potential, rather than relying solely on algorithmic complexity for performance improvements.

5. Conclusion

This study has established an effective framework for predicting customer churn in the banking sector by systematically integrating machine learning techniques with advanced data balancing methods. The core findings of this research underscore the significant advantages of using hybrid sampling methods, such as SMOTE-ENN, which effectively balance data quantity and quality. This approach not only preserves critical temporal patterns but also mitigates distortions from synthetic samples, thereby enabling models to more accurately distinguish genuine churn signals from noise. Among the machine learning models evaluated, random forests have demonstrated superior performance due to their inherent noise resistance and ability to handle complex feature interactions. This ensemble method is particularly effective in processing temporal indicators, such as declines in transaction frequency and service complaint patterns. Meanwhile, decision trees provide valuable interpretability for operational interventions, highlighting the importance of algorithm-feature compatibility, especially when compared to the limitations observed with SVMs.

The research also reveals that dynamic behavioral features, such as customer tenure duration and changes in financial engagement, are more predictive of churn than static demographic attributes. This insight aligns with real-world observations of customer relationship deterioration, emphasizing the need for banks to focus on temporal behavior monitoring rather than solely relying on demographic profiling.

From a practical standpoint, the study recommends proactive risk monitoring, where banks prioritize customers showing dual risk markers, such as long tenure combined with recent activity declines. Implementing a two-tiered prediction system that leverages random forests for automated

scoring and decision trees for explainable rule extraction can effectively balance accuracy with operational interpretability. Monitoring high-impact indicators and optimizing data infrastructure to capture temporal behavioral sequences are also crucial for continuous model improvement and adaptation to evolving customer patterns.

Methodologically, this research contributes to the field by demonstrating the efficacy of combining feature selection with hybrid sampling techniques to address the challenges of imbalanced banking data. It also validates the importance of temporal behavior monitoring in churn prediction systems, shifting the focus from static profiling to dynamic engagement analysis.

While the proposed framework exhibits strong predictive capabilities, successful implementation hinges on addressing key operational constraints, including ensuring data quality, establishing regular model update protocols, and maintaining ethical governance through transparent communication of prediction logic. The framework's emphasis on temporal behavioral features resonates with emerging research on dynamic customer analytics in digital banking environments. Looking ahead, future research should explore the integration of unstructured data sources and the development of real-time prediction systems to further enhance churn management strategies. Ultimately, effective churn management in banking requires not only sophisticated algorithms but also a strategic alignment between data capabilities and customer relationship objectives.

References

[1] Garcia, S., Luengo, J., & Herrera, F. (2019). Data preprocessing in data mining. Springer.

[2] Lopez, P., & Gonzalez, A. (2020). Addressing class imbalance in financial datasets. Journal of Financial Data Science, 5(2), 123-140.

[3] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. Springer.

[4] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

[5] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297.

[6] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

[7] Zhang, Y., & Wang, Q. (2023). Adaptive Ensemble Learning for Imbalanced Financial Data. IEEE Transactions on Knowledge and Data Engineering, 35(4), 1456-1470.

[8] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357.

[9] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).