INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 13, Issue 4, (2024) E-ISSN: 2226-6348

A Multiclass Ensemble Learning Approach for Predicting Customer Churn in Commercial Banks

Chen Shuofeng[1], Asif Mahbub Karim[2], LinLi[3]

[1]PhD Researcher, Binary University of Management & Entrepreneurship, Malaysia； Senior Economist, Guangxi University of Finance and Economics, Nanning, Guangxi, China, [2]Professor, Dean, Binary Graduate School, Binary University of Management & Entrepreneurship, Malaysia, [3]PhD Researcher, Binary University of Management & Entrepreneurship, Malaysia, Senior Economist, Guangxi University of Finance and Economics, Nanning, Guangxi, China

Corresponding Author Email: LinLi, clairelinli@gxufe.edu.cn

To Link this Article: http://dx.doi.org/10.6007/IJARPED/v13-i4/23543 DOI:10.6007/IJARPED/v13-i4/23543

Published Online: 06 November 2024

Abstract

In the era of burgeoning big data and the expansive reach of the Internet, commercial banks are confronted with the challenge of managing an extensive customer base while striving to meet their evolving needs. A nuanced and reliable understanding of consumer preferences is imperative for banks to ensure customer retention and to preemptively address potential churn. This research introduces a sophisticated approach to predict customer churn through the lens of multiclass categorization, leveraging the prowess of ensemble machine learning algorithms. By integrating the strengths of XGBoost, LightBoost, and CatBoost with a bagging ensemble method, our model offers a refined prediction of customer churn, distinguishing between various levels of churn risk. This multiclass ensemble learning framework not only enhances the predictive accuracy but also provides a more granular insight into customer behavior patterns. The efficacy of our model is assessed using the kappa statistic, a robust measure for evaluating the consistency of predictions across multiple categories. Our experimental results reveal that the kappa value of our multiclass ensemble model significantly surpasses that of single-algorithm approaches, indicating a superior predictive performance and reliability. The insights gleaned from our model can inform targeted marketing strategies and customer retention efforts, thereby mitigating the risk of customer churn. Through the application of this multiclass ensemble learning model, banks can achieve a more strategic and informed approach to maintaining customer loyalty and optimizing their service offerings.

Keywords: Customer Churn, Fusion Model, Big Data, Machine Learning

Introduction

Customers are the soul of any industry. In the fiercely competitive world of business, retaining existing customers presents a significant challenge for companies, as attracting new customers incurs higher costs than maintaining existing ones (Khaled et al., 2019; Caigny et al., 2018; Ganesh et al., 2000),Companies that establish long-term relationships with their customers, focusing on their needs, find this approach more profitable than seeking new clients (Reinartz & Kumar, 2003), This is especially true for banks. The rapid development of the internet and the advent of the big data era have led to a swift expansion in the scale of commercial banking operations. Banks have developed a variety of customer touch points, both online and offline, to meet daily service requirements and transactional needs (Bashir et al., 2020). As a result, modern commercial bank customers exhibit characteristics of online consumers, in contrast to traditional bank clients. At the same time, various commercial banks have accumulated massive amounts of customer data, including vital information such as customer attributes, transaction behaviors, feedback, and preferences. Faced with a large customer base, banks need to gain a more comprehensive and accurate understanding of customer needs (Hossain et al., 2018). In practical business operations, it's essential to identify potential customer churn, predict changes in customer funds, and engage in timely and proactive marketing to minimize bank fund outflows.

Predicting customer churn plays a crucial role in retaining customers by identifying those at risk of leaving the company. In the era of big data, scholars have employed various machine learning methods for this purpose, including Random Forest (Tékouabou et al., 2022), Support Vector Machine (Maldonado, 2015), Decision Trees (Caigny et al., 2018), Neural Networks (Zobair et al., 2021) and Ensemble Algorithms (Wang & Yu, 2021).These methods have achieved notable success in industries such as telecommunications (Jain et al., 2020), insurance (Jones & Sah, 2023), healthcare (Shehab et al., 2022) and services (Parvez, 2020).Despite the significant achievements in the field of customer churn prediction, there remain some pressing issues to be addressed as the economy evolves and technology advances. Firstly, the majority of existing research focuses on customer churn within traditional telecommunications and financial sectors, with less attention given to the modern commercial banking sector. Secondly, most commercial bank customer churn prediction models rely predominantly on single-model approaches, with limited use of ensemble model algorithms. Furthermore, the majority of customer churn predictions are based on binary classification, with fewer addressing multiclass prediction challenges. Therefore, the motivation behind our paper is to successfully predict early signs of commercial bank customer churn, enabling the implementation of preventative measures. This study aims to analyze different types of customer transaction behaviors and financial changes in commercial banks, seeking customer characteristics from multiple dimensions. By employing advanced hybrid ensemble learning algorithms and data mining techniques, we intend to develop a customer churn prediction model that balances accuracy and efficiency. Our research findings aim to contribute valuable insights to the theoretical study of customer churn prediction in commercial banks and inform decision-making processes.

Problem Statement

The banking industry is currently facing significant challenges in managing customer relationships and reducing customer churn in an increasingly competitive and dynamic business environment. As customer expectations continue to evolve, banks are under pressure to develop innovative strategies for customer retention (Polas et al., 2020; Bhattacharjee et al., 2019). Recent studies utilizing transactional data from large commercial banks have highlighted the complex nature of customer churn, revealing patterns that traditional binary classification models, which simply categorize customers as churned or retained, fail to capture effectively (Shirazi & Mohammadi, 2019; Zhu et al., 2018). Has research found that the majority of studies relied on single algorithmic approaches for churn prediction, with limited application of advanced ensemble methods (Ullah et al., 2019). Furthermore, while the existing literature extensively covers the telecommunications and insurance sectors, the unique customer dynamics within the modern banking industry remain relatively underexplored (Dalvi et al., 2016).

To address this research gap, the current study proposes a novel multi-class ensemble learning framework designed to unravel the complex layers of churn risk inherent in the banking domain. By integrating advanced predictive algorithms with a robust bagging ensemble technique, this research aims to uncover the intricate behavioral patterns of bank customers. The primary objective is to provide financial institutions with detailed, actionable insights that can inform the development of more precise and effective customer retention strategies. This study not only contributes to filling a significant gap in the existing academic literature but also aligns with the urgent need for banks to adapt to the rapidly evolving landscape of customer relationship management.

Limitations

While this research aims to provide valuable insights into the prediction of customer churn in commercial banks, several limitations should be acknowledged. Firstly, the generalizability of the findings may be influenced by the specific characteristics of the dataset obtained from Xiamen Commercial Bank. The applicability of the developed multiclass ensemble learning model to other commercial banks may require further validation and adaptation.

Secondly, the study primarily focuses on the application of ensemble machine learning algorithms for customer churn prediction. However, the effectiveness of the model in real-world banking operations may be influenced by additional factors such as regulatory constraints, economic conditions, and evolving customer behaviors, which are not fully accounted for in the research.

Furthermore, the research is based on historical customer data, and the dynamic nature of customer preferences and behaviors may introduce uncertainties in the predictive accuracy of the model. The ability to adapt the model to real-time data and evolving customer trends remains a potential challenge.

Acknowledging these limitations is crucial for interpreting the research findings and understanding the scope of applicability of the developed multiclass ensemble learning model for customer churn prediction in commercial banks.

Literature Review

Lin (2012) addressed the customer churn issue in banks by leveraging the stability and strong learning capabilities of BP artificial neural networks. Through the analysis of the correlation between input and output variables, a bank customer churn analysis model was established to identify customers likely to churn (Lin, 2012). The predictive effectiveness of logistic regression, decision trees, and artificial neural networks on bank customer churn, finding logistic regression to be the most effective (Cai-xian & Zhi-rong, 2013). Dong proposed a method for extracting features of customer transaction behavior in fund trading scenarios using decision trees. This approach effectively avoids subjective feature selection and significantly improves the recall rate (Jiyang, 2020). The effectiveness of bank customer churn prediction using six methods: random forests, support vector machines, SBADA (Stochastic Boosting with Adaptive Data Augmentation), logistic regression, regression trees, and multivariate adaptive regression (Gervasi et al., 2020).Given the typical imbalance in commercial bank customer churn data, direct application of statistical prediction methods and traditional classification methods results in poor accuracy. However, random sampling methods can reduce data imbalance by changing the distribution of the dataset. Consequently, He improved the support vector machine model using random sampling methods and compared its predictive effectiveness with the Logistic regression model, finding that this approach significantly enhances model prediction accuracy (Ben-lan, 2014). Researcher utilized data reduction techniques to enhance feature extraction from time series data, further aiding in the improved prediction of customer churn in the banking sector. This approach helps in reducing the time consumed in training imbalanced datasets, as the imbalance in commercial bank customer data leads to unpredictability in minority groups. They validated their method through extensive historical data of bank customer transactions (Britto & Gobinath, 2020). Li & Xie (2020) proposed an imbalanced data approach based on Generative Adversarial Networks (GANs) to address the poor predictive performance of traditional classifiers on minority classes. Their method involves generating minority class samples using GANs to improve the imbalance, thereby enhancing the predictive performance for minority classes (Li & Xie, 2020).Considering the characteristics of bank customer churn. The random forest model with random sampling methods to enhance the performance of customer churns prediction (Verma, 2020). Tran et al. (2023) examined the impact of customer segmentation on the accuracy of customer churn prediction in the banking sector through machine learning models. They employed various machine learning models, such as k-means clustering for customer segmentation, k-nearest neighbors, logistic regression, decision trees, random forests, and support vector machines applied to the dataset for predicting customer churn. The experimental results indicated that the dataset performed well in the random forest model (Tran et al., 2023).

The prediction of customer churn in the banking sector is addressed not only through machine learning methods but also via survival analysis models. Mavri and Ioannou (2008) utilized data from the Greek banking sector, employing proportional hazards models to determine the risks associated with customer churn behavior. Their goal was not to predict whether a specific customer would churn but to understand which characteristics influence the transition behavior of bank customers (Mavri & Ioannou, 2008). Wang et al. (2014) based on survival analysis methods and using a commercial bank's customer churn data within a year as the research sample, analyzed 12 predictive variables that might significantly impact customer churn. The results indicated that seven factors have a significant effect on bank customer churn. Among these, the average monthly deposit, average monthly expenditure, highest deposit balance, account type, average monthly transactions, and response to promotions are negatively correlated with customer churn, while the maximum interval between transactions is positively correlated with customer churn. This suggests that the longer the interval between two transactions initiated by the customer, the greater the likelihood of churn. Banks can apply the Cox model to predict customer churn at a future point in time and take effective measures in advance (Weiqing et al., 2014).

From the aforementioned works, we can identify several areas of concern. First, the limited comparison of machine learning models: Although several studies have compared different machine learning models for predicting bank customer churn, there is a need for a more comprehensive comparison across a broader range of models to identify the most effective approach. Second, the integration of different techniques: While some research integrates techniques such as random sampling with machine learning models to enhance predictive performance, there is a lack of effective studies on integrating multiple techniques to improve the accuracy of predicting customer churn in the banking sector. Third, the application of survival analysis models: Although some studies have explored the use of survival analysis models to understand customer churn behavior in the banking sector, there is a gap in focusing on the practical application and comparison of survival analysis models with traditional machine learning methods in predicting customer churn. Fourth, addressing imbalanced data: The issue of data imbalance in predicting bank customer churn is acknowledged in the literature, but there is a need for further research on advanced techniques such as data augmentation, minority class sample generation, and integration methods to effectively tackle this challenge and improve prediction accuracy.

Research Questions

1.How can an ensemble model be developed to predict different levels of customer churn in commercial banks?
2.What is the predictive performance of a bagging ensemble model integrating XGBoost, LightGBM and CatBoost for multi-class customer churn prediction?

Research Objectives

1.To develop an ensemble learning framework combining XGBoost, LightBoost and CatBoost using a bagging method for multi-class customer churn prediction in commercial banks.
2.To evaluate the predictive performance of the ensemble model in distinguishing different levels of customer churn risk compared to individual algorithms.

Research Methodology

In order to tackle the issue of multi-classification, this study will explore the predictive performance of ensemble algorithm models such as Lightboost, Xgboost, and Catboost. Additionally, the method of Begging will be utilized to integrate multiple algorithms, thereby enhancing the predictive capability and effectiveness of the model.

Begging Integration Method

Bagging (Bootstrap Aggregating) is an ensemble learning method. Its core concept revolves around enhancing the model's generalizability and stability by generating multiple distinct training sets through bootstrap sampling (sampling with replacement) from the original dataset. An independent base learner is then trained on each of these training sets. Ultimately, the results of these base learners are integrated through voting or averaging. This process is depicted in the following diagram:

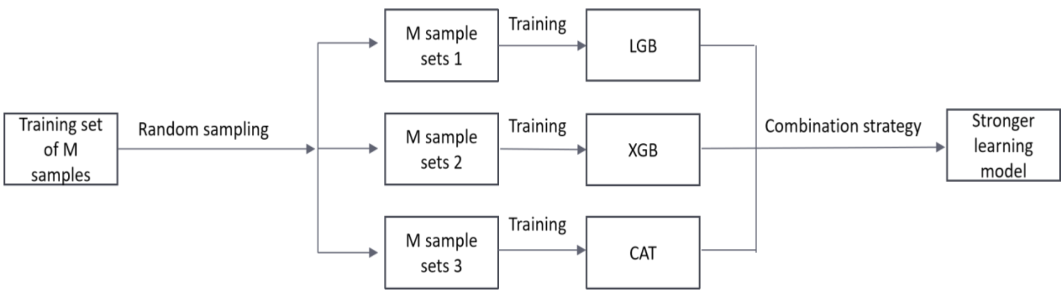

Fig 1. Begging Integrated Model Structure

By employing multiple distinct models, bagging can reduce the risk of overfitting associated with a single model. As bagging utilizes a variety of models, it can lower the model's variance and enhance its stability. Even if a particular base model is affected by noise or outlier samples, the overall performance of the ensemble model can still be well maintained. Moreover, each base model is trained independently, allowing the bagging method to be highly amenable to parallel computation. This significantly saves on training time and improves the computational efficiency of the model. Therefore, in this study, the advantages of this method have been fully leveraged to improve the model's performance.

LightBoost Algorithm Model

LightBoost is an efficient algorithm based on gradient-boosted decision trees. Its core concept revolves around enhancing training speed and accuracy through a Leaf-wise growth strategy and histogram optimization.

The objective function of LightBoost comprises a loss function and a regularization term, with the optimization goal being to minimize the objective function in the following form:

$$obj(\Theta)=\sum_{i=1}^{n}L(y_i, f(x_i))+\sum_{k=1}^{K}\Omega(f_k) \tag{1}$$

In this context, $L(y_i, f(x_i))$ denotes the loss function, which measures the discrepancy between the model's predicted values $f(x_i)$ and the actual labels y_i ; $\Omega(f_k)$ represents the regularization term, aimed at controlling the complexity of the model to prevent overfitting; Θ symbolizes the model parameters. It is evident that LightBoost's objective function optimizes the model parameters by minimizing the sum of the loss function and the regularization term. This approach ensures that the model can better fit the real labels on the training data while maintaining an appropriate level of complexity to enhance its generalization ability.

In the context of multi-class problems, LightBoost can be adeptly handled through either the One-vs-All (OvA) or One-vs-One (OvO) strategies. This enables the rapid and efficient construction of models with superior accuracy, thereby enhancing the precision and efficiency of multi-class tasks.

XGBoost Algorithm Model

XGBoost (Extreme Gradient Boosting) is also an ensemble learning algorithm based on gradient boosting trees, employing a Level-wise growth strategy. It integrates gradient boosting algorithms with regularization techniques, offering both efficiency and accuracy.

The objective function of XGBoost consists of two components: the loss function and the regularization term. Its goal is to minimize the objective function in the following form:

$$obj(\Theta)=\sum_{i=1}^{n}L(y_i, \hat{y}_i)+\sum_{k=1}^{K}\Omega(f_k) \tag{2}$$

In this context, $L(y_i, \hat{y}_i)$ represents the loss function, which is used to quantify the discrepancy between the model's predicted values $\hat{y}_i$ and the actual labels y_i ; $\Omega(f_k)$ denotes regularization, aimed at controlling the model's complexity to prevent overfitting; Θ indicates the model parameters. It is evident that XGBoost optimizes the model parameters by minimizing the sum of the loss function and the regularization term. This approach ensures that the model can more accurately fit the real labels on the training data while maintaining an appropriate level of complexity to enhance its generalization ability. Although similar to

the objective function of LightBoost, XGBoost requires the conversion of categorical features into numerical features, typically using one-hot encoding or label encoding. This conversion can lead to dimensionality explosion and reduced computational efficiency.
In multi-class problems, XGBoost can also be adeptly managed through either the One-vs-All (OvA) or One-vs-One (OvO) strategies. XGBoost is capable of handling high-dimensional sparse data in multi-class issues, offering high accuracy and robustness against outliers. By integrating multiple weak classifiers to construct a strong classifier, it effectively enhances the accuracy and generalization capability of multi-class tasks.

Catboost Algorithm Model

CatBoost is specifically designed for handling classification problems, exhibiting superior parallel processing capabilities during training. When tackling classification issues, CatBoost can automatically process categorical features, achieving both efficiency and accuracy. The design of CatBoost's objective function retains similarities with LightBoost, ensuring the model can be effectively trained and predicted. Therefore, the reason CatBoost and LightBoost share the same objective function is that they both adhere to the fundamental principles and design philosophies of the gradient boosting tree algorithm.
CatBoost optimizes model parameters by minimizing the sum of the loss function and the regularization term. This strategy allows the model to better fit the real labels on the training data while maintaining an appropriate level of complexity to enhance its generalization ability. CatBoost's ability to automatically handle categorical features eliminates the need for one-hot encoding or label encoding, thereby simplifying the steps in feature engineering. Hence, through gradient-based learning algorithms and feature combination optimization techniques, CatBoost can efficiently construct deep decision tree models, improving training speed. In addressing classification problems, it achieves high accuracy and generalization capability, enhancing model performance through the optimization of the objective function and feature combinations. Additionally, it exhibits strong robustness against missing values and outliers, effectively dealing with noise and incompleteness in the data.
In summary, CatBoost plays a role in multi-class problems by automatically processing categorical features and efficiently building models with excellent accuracy, thereby enhancing the accuracy and robustness of multi-class tasks.

Data Analysis Plan

The primary objective of our data analysis plan is to construct a predictive model that accurately identifies potential customer churn in a commercial banking context. This plan is structured around a comprehensive roadmap that encompasses data preprocessing, feature selection, model development, and validation phases. Guided by the insights from "Fig 2. Data Analysis Roadmap," our approach integrates both traditional statistical methods and advanced machine learning techniques to ensure robustness and accuracy in our predictions.

Data Preprocessing

Initially, our dataset will undergo a thorough preprocessing phase to ensure data quality and relevance. This will include handling missing values through imputation techniques, encoding categorical variables, and normalizing numerical features to a common scale. Special attention will be paid to identifying and removing outliers that could potentially skew our analysis.

Feature Selection and Engineering

Using statistical methods for feature selection, we aim to identify the most predictive variables influencing customer churn. This process will help in reducing dimensionality and focusing on the features with significant predictive power. Additionally, feature engineering techniques will be applied to construct new variables that capture complex patterns and interactions among the original dataset features, potentially enhancing the model's predictive capabilities.

Model Development

Our model development phase will employ a novel ensemble learning approach, integrating the strengths of XGBoost, LightBoost, and CatBoost algorithms through a bagging ensemble method. This multiclass ensemble model is designed to categorize customers into different churn risk levels, thus providing a nuanced understanding of churn dynamics. Model hyperparameters will be optimized using cross-validation techniques to ensure the model's generalizability.

Model Validation and Performance Evaluation

The final model's performance will be rigorously evaluated using a hold-out validation set. The kappa statistic will be employed to assess the model's predictive power. It represents a more comprehensive performance metric than mere accuracy (i.e., the proportion of correct classifications), particularly suitable for scenarios involving imbalanced datasets.

Interpretation and Implications

The culmination of our data analysis plan will involve interpreting the model's findings to uncover actionable insights into the factors driving customer churn in commercial banks. This will not only contribute to the theoretical understanding of customer behavior patterns but also provide commercial banks with practical strategies for enhancing customer retention and loyalty.

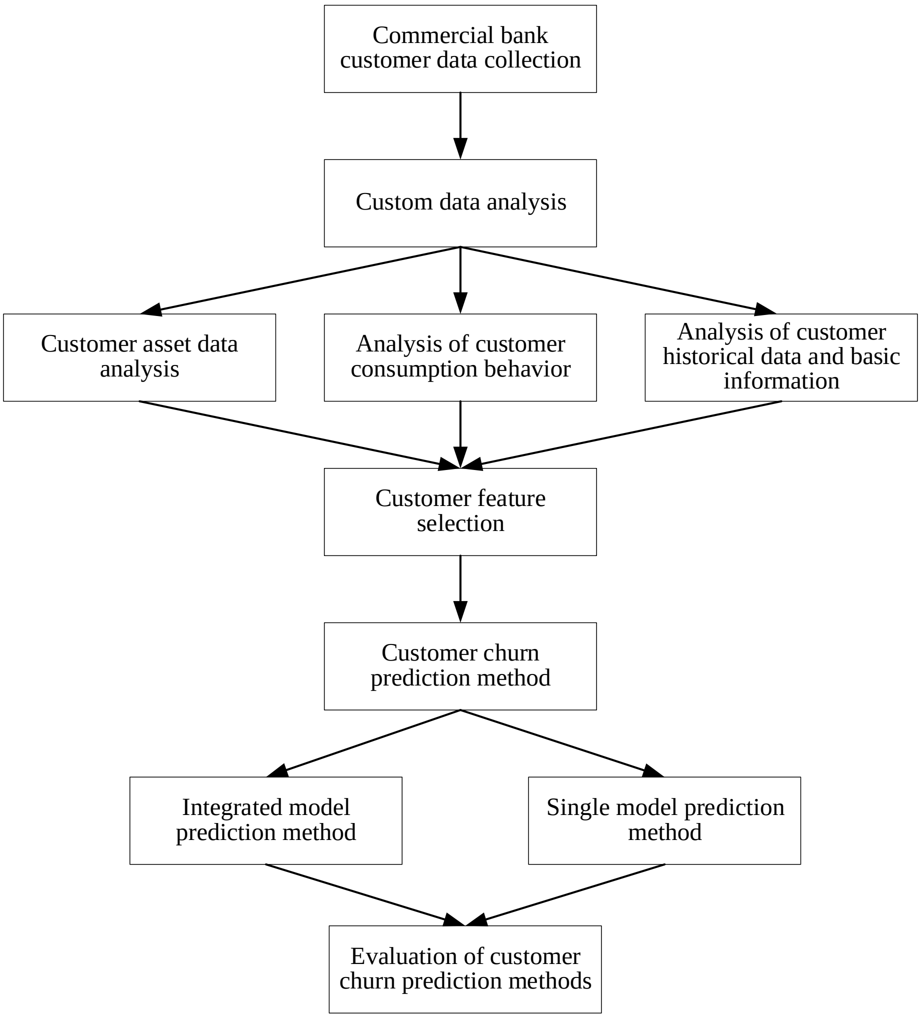

Fig 2. Data Analysis Roadmap

Data Analysis

To assess the effectiveness of ensemble models in predicting multi-class customer churn, we analyzed customer data from Xiamen Commercial Bank in China. We utilized the Kappa score as the evaluation metric, conducting comparative analyses against single models such as LightBoost, XGBoost, and CatBoost.

Data Preparation

In this study, customer data for two quarters within a year were obtained from a commercial bank, with over 490,000 records for the third quarter and over 540,000 for the fourth quarter. The dataset primarily encompasses five dimensions of data, including the month-end point asset, user behavior data, significant historical data of customers, deposit data, and basic customer information, totaling 56 features, as shown in Table 1:

Table 1
Basic Content of Customer Data Dimensions

Characteristic Dimension	Field Name	Meaning
	cust-no	User unique ID

a)The month-end point asset （aum_m(Y)）	X1	Balance of structured deposits at the end of month Y
	X2	Balance of time deposit at the end of month Y
	X3	Balance of current deposits at the end of month Y
	X4	Y month end point financial balance
	X5	Y month end point fund balance
	X6	Y month end time point asset management balance
	X7	Y Loan balance at the end of month
	X8	Large deposit balance at the end of Y month
b)User behavior data(behavior_m(Y))	B1	Y Monthly times of mobile banking and online banking
	B2	Transfer in times in Y month
	B3	Transfer in amount in month Y
	B4	Transfer out times in Y month
	B5	Transfer out amount of month y
	B6	Last trading time
	B7	Number of transactions in the quarter
c)Significant historical data of customers(big_event_Q(Z))	E1	Account opening date
	E2	Online banking opening date
	E3	Mobile app account opening date
	E4	First online banking login date
	E5	First mobile app login date
	E6	Date of first current deposit
	E7	First time deposit business date
	E8	First loan business date
	E9	First overdue date
	E10	Transaction date of the first capital business
	E11	Transaction date of the first bank securities transfer
	E12	First counter transfer date
	E13	First online bank transfer date
	E14	First mobile app transfer date
	E15	Maximum amount transferred out

	E16	Date of maximum amount transferred out
	E17	Maximum amount transferred in
	E18	Date of maximum amount transferred in
d)Deposit data(cunkuan_m(Y))	C1	Deposit product amount of the current month
	C2	Number of deposit products in the current month
f)Customer basic information	I1	Gender
	I2	Age
	I3	Customer level
	I4	Staff of the bank
	I5	Job description
	I6	Loan customer mark
	I7	Number of products held
	I8	Constellation description
	I9	Customer contribution
	I10	Education
	I11	Annual family income
	I12	Industry
	I13	Marital status
	I14	Duties
	I15	QR code acquiring customer
	I16	VIP customers
	I17	Online banking customers mark
	I18	Mobile banking customer mark
	I19	SMS customers
	I20	Wechat payment customers

Defining Online Customer Churn

In the study of multi-class bank customer churn prediction, customers were assigned one of three labels based on certain business rules and the scale of their asset management: -1, 0, and 1. These labels typically represent the customer's churn tendency or status. The specific meanings of the labels are as follows:

-1 : Indicates a decrease in the customer's churn tendency, meaning the customer is more likely to remain within the bank's clientele, with a lower probability of churn.

0 ：Indicates a stable churn tendency, meaning the customer's status is relatively steady, with no clear signs of either churn or improvement.
1 ：Indicates an increase in the customer's churn tendency, meaning the customer has a higher likelihood of leaving the bank in the foreseeable future.

Data Pre-Processing
Based on the assessment of banking operations, this study should particularly focus on five categories of features derived from the aforementioned five dimensions of customer information.

(1)Date information:The date features include the latest transaction time of each quarter and the occurrence time of the user's major time. Generally speaking, the closer the transaction time is to the end of the current transaction, the more likely the user is to be promoted, which means that the user is active. On the contrary, the more likely the user is to be a drain type user.

(2)behavior features ：The numerical behavior characteristics in the table mainly include the number of login times, transfer in times, transfer out times, transfer in amount, transfer out amount, etc.

(3)Customer balance information and deposit information:Customer deposit information and balance information are the key characteristics to judge customer type.

(4)Quarterly change of account funds:This is the main basis for banks to judge whether the loss of customers.

(5)Customer basic information:This is the label of the customer, which has the least impact on the model in practical application.

In this study, feature extraction and selection are comprehensively considered in conjunction with banking business needs, data characteristics, and model selection. Through statistical methods (such as correlation coefficients, analysis of variance, etc.), features are evaluated to select those with a significant impact on the target variable. One-hot encoding is a commonly used data preprocessing technique, primarily aimed at converting categorical variables into numerical variables, thereby facilitating a better understanding and processing of these features by machine learning algorithms. The application of one-hot encoding can enhance the learning of combinational features among the variables.
After the process of feature construction, feature extraction and feature selection, I extracted more than 70000 effective customers and 212 features as the prediction model data.

Model Evaluation
The Kappa coefficient is employed as an evaluation metric for multi-class models in this research. It accounts for the superiority of the model's predictive accuracy over random guessing, offering a more objective assessment of model performance. By comparing the accuracy of the model against that of random predictions, the Kappa coefficient can more accurately reflect the predictive capability of the model. In multi-class problems, there might be an imbalance in the number of samples across different categories, which could affect the evaluation of metrics such as accuracy. The Kappa coefficient accommodates the predictive accuracy across various categories, thereby providing a more comprehensive evaluation of the model's performance across different classes. Hence, it is utilized to gauge the accuracy of multi-classification models. In practical applications, its value generally ranges from [0,1], with a higher coefficient indicating greater classification accuracy of the model.

The calculation method of kappa coefficient can be expressed as follows:

$$k=\frac{p_0-p_e}{1-p_e} \qquad (3)$$

Among them, p_0 is the total classification accuracy;

$$p_e=\frac{a_1\times b_1+a_2\times b_2\cdots a_c\times b_c}{n\times n} \qquad (4)$$

a_i represents the number of class i real samples.
b_i represents the number of samples predicted by class i.
9.5 Experimental Design and Analysis of Experimental Results
1.Divide the dataset into a training set and a test set.
2.Train and predict using three single models: LightGBM, XGBoost, and CatBoost, respectively, and calculate their Kappa score. The main parameters of these models are as follows in the table below:

Table 2
The Main Parameters of the Algorithm

Algorithm	Main parameters
XGBoost	learning_rate: 0.3 n_estimators: 100 max_depth: 6
LightBoost	learning_rate: 0.1 max_depth: -1 n_estimators: 100
Catboost	iterations: 1000 learning_rate: 0.03 depth: 6

3.Utilize the Bagging method to integrate LightGBM, XGBoost, and CatBoost into an ensemble model. Within the bagging framework, the averaging method is employed to determine the weights of each base model. The specific process is as follows: Assuming we have nbase models with respective prediction outcomes denoted as: $y_1,y_2,...,y_n$,our goal is to ascertain the weight: $w_1,w_2,...,w_n$,for each base model such that the final prediction result is $y_{final}=w_1\cdot y_1+w_2\cdot y_2,\cdots w_n\cdot y_n$.
First, calculate the weights of the base models: Based on the training dataset, we can obtain the prediction error for each base model (such as mean squared error, log loss, etc.) through methods like cross-validation, denoted as $e_1,e_2,...,e_n$,Typically, the reciprocal of the error can serve as a candidate value for the weight, denoted as $w_i=\frac{1}{e_i}$.Then, normalize the weights: To ensure that the sum of the weights of all base models equals 1, we need to normalize the weights. Specifically, divide each candidate weight w_i by the sum of all candidate weights, resulting in $w_i=\frac{w_i}{\sum_{j=1}^{n}w_j}$. Finally, apply the weights: The final prediction result, denoted as y_{final}, is the weighted sum of the prediction results from each base model, where the weights are denoted as $w_1,w_2,...,w_n$.Through these steps, we can determine the weight of each base

model within the bagging model, thereby obtaining the final prediction result. This method effectively combines the predictive power of multiple base models, enhancing the overall model's generalization ability and stability.

Experimental Result Analysis

Train and predict using the ensemble model, and calculate its Kappa value. Formulate the comparison results by contrasting it with algorithm models such as LightGBM, XGBoost, and CatBoost, as shown in Table 3 below:

Table 3
Comparison of Kappa score

Model	Kappa score
XGBoost	0.47368633
LightBoost	0.46795530
CatBoost	0.47211799
Fusion Model(average)	0.47822253

Conduct a comparative analysis of the Kappa values for both the single models and the ensemble model to evaluate the performance improvement of the ensemble model relative to the single models. From the experimental results, we can observe that the individual models XGBoost, LightBoost, and CatBoost have Kappa scores of 0.47368633, 0.46795530, and 0.47211799, respectively. Meanwhile, the hybrid model obtained by averaging these models using the bagging method achieved a Kappa score of 0.47822253.

(1) Performance of Individual Models: XGBoost performs slightly better than CatBoost but is somewhat inferior to LightBoost. LightBoost exhibits the best performance in this set of experiments, boasting the lowest performance metric. CatBoost's performance is marginally lower than both XGBoost and LightBoost.

(2) Performance of the Hybrid Model: The Kappa score of the hybrid model, at 0.47822253, is significantly higher than all individual models, showing superior performance. The hybrid model, by averaging the prediction results of each individual model, integrates their strengths, thereby achieving better performance in this experiment.

(3) Advantages of the Hybrid Model: The hybrid model can reduce the risk of overfitting and enhance the model's generalization capability by combining the prediction results of multiple individual models. As the hybrid model integrates the strengths of various models, it often achieves more accurate prediction results. The hybrid model can reduce the random errors of individual models, improving the overall model's stability and reliability. Different models have their own advantages in various datasets or problems, and the hybrid model can better adapt to different situations.

In summary, by integrating the strengths of multiple individual models, the hybrid model can improve prediction accuracy while reducing the risk of overfitting. It possesses better stability and adaptability, therefore, in many cases, it can achieve better performance than individual models alone.

Findings and Conclusion

This study proposed a novel multiclass ensemble learning approach for predicting customer churn in commercial banks. By categorizing customers into different churn risk levels instead of binary classification, it provided a more comprehensive understanding of customer loyalty trends.

An ensemble model combining XGBoost, LightGBM and CatBoost using bagging was developed. Experimental results on real customer data from a large Chinese bank demonstrated that the ensemble model achieved superior predictive performance over individual algorithms, with a Kappa score of 0.47822253, outperforming other models.

Analysis of customer behavior patterns across various attributes using the ensemble model offered actionable insights into different customer segments. Banks can leverage these findings to design targeted retention programs and optimize resource allocation.

This research contributes to advancing the theoretical study of customer churn prediction through a multiclass perspective and application of ensemble learning techniques. The proposed framework also provides practical guidance for banks to strategically minimize customer outflows through precision marketing interventions.

Some limitations include the focus on a single bank's data. Further validation using multi-source customer data could boost the generalizability of findings. Future work will explore incorporating deep learning models to handle more complex customer characteristics in big data contexts.

Overall, the multiclass ensemble learning approach developed in this study presents an effective solution for predicting customer churn levels and enhancing customer retention capabilities in commercial banks. The model offers continuous value as customer profiles evolve with technological advancement.

Recommendation

In light of the insights garnered from our investigation, we propose a strategic framework for commercial banks aiming to refine their customer churn prediction and management methodologies. This framework is centered on the integration of a multiclass ensemble learning approach, which marks a departure from traditional binary classification models. The recommendations are detailed as follows:

1. Implementation of Multiclass Customer Segmentation: Banks are encouraged to adopt a nuanced segmentation approach by classifying customers into multiple levels of churn risk. This strategy facilitates a deeper understanding of customer loyalty dynamics, enabling the development of more personalized customer engagement and retention strategies.
2. Adoption of an Advanced Ensemble Learning Model: We recommend the deployment of an ensemble learning model that synergizes the strengths of advanced algorithms such as XGBoost, LightBoost, and CatBoost through a bagging technique. This composite model is demonstrated to surpass the predictive accuracy of its constituent algorithms when applied to churn prediction.
3. Strategic Analysis and Targeted Interventions: Utilizing the ensemble model to dissect and interpret customer behavior across diverse attributes will unveil critical insights. These insights should guide the creation of bespoke retention initiatives and the judicious distribution of resources tailored to the unique needs of distinct customer segments.

4. Ongoing Monitoring and Adaptive Model Refinement: To maintain the relevance and efficacy of the churn prediction model, we advise establishing a regimen for continuous monitoring and periodic recalibration of the model. This approach ensures that the model adapts to evolving customer profiles and behaviors, thereby enabling banks to preemptively address churn risks with timely and appropriate measures.
5. Enhanced Strategic Positioning through Precision Marketing: By embracing the outlined recommendations, banks can significantly reduce customer attrition and augment the lifetime value of their clientele. This strategic pivot not only facilitates precision marketing efforts but also bolsters risk management practices in a competitive banking landscape.

Acknowledgments
Our project is supported by Guangxi First-class Discipline Applied Economics Construction Project Fund (2022GSXKB05).

References
Bashir, M. A., Ali, M. H., Wai, L. M., Hossain, M. I., & Rahaman, M. S. (2020). Mediating Effect of Customer Perceived Value on the Relationship between Service Quality and Customer Satisfaction of E-Banking in Bangladesh. *International Journal of Advanced Science and Technology*. 29(2), 3590 – 3606

Bhattacharjee, A., Jahanshahi, A. A., Polas, M. R. H., Hossain, M. I., & Asheq, A. S. (2019). Customer Care Service Management is Moving Forward to Achieve Sustainable Customer Retention in Every Industry. Does it play a Role to Increase Brand Retention. *International Journal of Management and Sustainability*, 8(2), 88-97.

Ben-lan, H. (2014). A Study of the Application of SVM in Prediction about Decrease in Bank's Customers. *Financial Forum*, 19(9), 70-74.

Cai-xian, Y. U., & Zhi-rong, Z. (2013). Mathematical modeling analysis of customer churn prediction in banks [Mathematical modeling and analysis on bank customer churn prediction]. *Journal of Changchun University of Technology (Natural Science Edition)*, 34(1), 5-8. http://doi.org/10.3969/j.issn.1674-1374.2013.01.002

Dalvi, P. K., Khandge, S. K., Deomore, A., Bankar, A., & Kanade, V. A. (2016, 2016/1/1). *Analysis of customer churn prediction in telecom industry using decision trees and logistic regression*. Paper presented at the.

De Caigny, A., Coussement, K., & De Bock. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*(269 (2)), 760-772.

Ganesh, J., Arnold, M. J., & Reynolds, K. E. (2000). Understanding the customer base of service providers: An examination of the differences between switchers and stayers. *Journal of Marketing*, 64(3), 65-87.

Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B. O., Rocha, A. M. A. C., Tarantino, E., & Torre, C. M. (2020). Machine Learning for Customer Churn Prediction in Retail Banking (12251, pp. 576-589). Springer International Publishing AG. http://doi.org/10.1007/978-3-030-58808-3_42.

Hossain, M. I., Limon, N., Amin, M. T., & Asheq, A. S. (2018). Work Life Balance Trends: A Study on Malaysian GenerationY Bankers. *IOSR Journal of Business and Management,* 20 (9), 01-09.

Jain, H., Khunteta, A., & Srivastava, S. (2020). Churn prediction in telecommunication using logistic regression and logit boost. *Procedia Computer Science*, 167, 101-112.

Jiyang, D. (2020). Research on Predicting Fund Customer churn Based on Decision Tree Automated Feature Selection: Reflections in the Post Epidemic Era. *Shandong Social Science*(09), 74-80.

John Britto, M., & Gobinath, R. (2020). WITHDRAWN: REHC: Reduction through exclusive homogeneous clusters for imbalance dataset. *Materials Today: Proceedings* http://doi.org/10.1016/j.matpr.2020.11.093

Jones, K. I., & Sah, S. (2023). The Implementation of Machine Learning In The Insurance Industry With Big Data Analytics. *International Journal of Data Informatics and Intelligent Computing*, 2(2), 21-38.

Khaled, A. S., Ahmed, S., Tabash, M. I., Al-Homaidi, E. A., & Hossain, M. I. (2019). The Impact of Technological and Marketing Innovations on Retailing Industry: Evidence of India. *Journal of Reviews on Global Economics,* 8, 948-957

Li, B., & Xie, J. (2020). Study on the Prediction of Imbalanced Bank Customer Churn Based on Generative Adversarial Network. *Journal of Physics. Conference Series*, 1624(3), 32054. http://doi.org/10.1088/1742-6596/1624/3/032054

Lin Rui, C. X. (2012). A Customer churn analysis model for banks based on artificial. *Computer Knowledge and Technology*(3)

Maldonado, S. (2015). Churn prediction via support vector classification: An empirical comparison. *Intelligent Data Analysis*, 19(s1), S135-S147.

Mavri, M., & Ioannou, G. (2008). Customer switching behaviour in Greek banking services using survival analysis. *Managerial Finance*, 34(3), 186-197.

Parvez, M. O. (2020). Use of machine learning technology for tourist and organizational services: high-tech innovation in the hospitality industry. *Journal of Tourism Futures*, 7(2), 240-244.

Polas, M. R. H., Juman ,M. K., Karim, A. M., Tabash, M. I., Hossain, M. I., (2020). Do Service Quality Dimensions Increase the Customer Brand Relationship among Gen Z? The Mediation Role of Customer Perception between the Service Quality Dimensions (SERVQUAL) and Brand Satisfaction. *International Journal of Advanced Science and Technology*. 29(4), 1050-1070

Reinartz, W. J., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, 67(1), 77-99.

Shehab, M., Abualigah, L., Shambour, Q., Abu-Hashem, M. A., Shambour, M. K. Y., Alsalibi, A. I., & Gandomi, A. H. (2022). Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine*, 145, 105458.

Shirazi, F., & Mohammadi, M. (2019). A big data analytics model for customer churn prediction in the retiree segment. *International Journal of Information Management*, 48, 238-253.

Tékouabou, S. C. K., Gherghina, Ș. C., Toulni, H., Mata, P. N., & Martins, J. M. (2022). Towards Explainable Machine Learning for Bank Churn Prediction Using Data Balancing and Ensemble-Based Methods. *Mathematics*, 10(14), 2379. http://doi.org/10.3390/math10142379

Tran, H., Le, N., & Nguyen, V. (2023). Customer Churn Prediction In The Banking Sector Using Machine Learning-Based Classification Models. *Interdisciplinary Journal of Information, Knowledge & Management*, 18

Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., & Kim, S. W. (2019). A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector. *Ieee Access*, 7, 60134-60149.

Verma, P. (2020). Churn prediction for savings bank customers: A machine learning approach. *Journal of Statistics Applications & Probability*, 9(3), 535-547.

Wang, Y., & Yu, W. (2021, December). Research on the Influencing Factors of Consumer Experience Under the New Retail Mode of Fresh Food. In *Proceedings of the 2021 5th International Conference on Computer Science and Artificial Intelligence* (pp. 337-346).

Weiqing, W., Rao, Y., & Cheng, L. (2014). The influencing factors of customer churn in commercial banks: a study based on survival analysis method. *Financial Forum*, 19(01), 73-79.

Zhu, B., Baesens, B., Backiel, A., & Vanden Broucke, S. K. (2018). Benchmarking sampling techniques for imbalance learning in churn prediction. *Journal of the Operational Research Society*, 69(1), 49-65.

Zobair, K. M., Sanzogni, L., Houghton, L., & Islam, M. Z. (2021). Forecasting care seekers satisfaction with telemedicine using machine learning and structural equation modeling. *Plos One*, 16(9), e257300.