Three basic Machine Learning Models’ Suitability for Predicting Bank Churn

Yibin Li1\*, Mingze Gao2, Yanfu Zhang3, Changbin Feng4, Muxi Chen5, Luyun Zhang6

1Rutgers Business School, Rutgers University-NB, Piscataway,08854, United States, [yl1857@rutgers.edu](mailto:yl1857@rutgers.edu)

2College of Art & Science, University of Virginia, Charlottesville, 22903, United States mg9aam@virginia.edu

3 School of Economics and Management, Xidian University, Xi 'an, 710126, China, [22061300036@stu.xidian.edu.cn](mailto:22061300036@stu.xidian.edu.cn)

4 Department of Economic and Management, North China Electric Power University, Baoding, 071000, China, [1493056395@qq.com](mailto:1493056395@qq.com)

5 Questrom School of Business, Boston University, Boston, 02215, USA, [chenmuxi@bu.edu](mailto:chenmuxi@bu.edu)

6 University of Edinburgh Business School, The University of Edinburgh, Edinburgh, EH89JS, UK, s2163495@ed.ac.uk

**Abstract.** Credit card customer churn is defined as the situation in which a bank’s customer stops using the bank's service and it leads to a potential loss of profit for the bank. Therefore, developing a customer churn prediction model for predicting customer likelihood of churning is essential for banks. This study aims to find out the key factor that influences credit card customer churn and to build a well performing model with high predicting accuracy. Finally, we can give some advice for banks to decrease customer churn. To achieve the goal, three models are employed, including neural network, logistic regression and decision tree model. The results indicate that all three models can be employed in different situations. However, the decision tree model has the best performance in both training and testing phases. The study further reveals that the top three factors influencing the likelihood of customer churn are total transaction count, total transaction amount, and changes in total transaction count. Based on the strengths and weaknesses of each model, this paper will also illustrate the different scenarios in which each model is most suitable.

**Keywords:** Machine learning; customer churn; predicting model; bank churn

1. Introduction

The cost of acquiring new customers is generally higher than that of retaining existing ones [3]. Furthermore, long-term customers often yield greater profits, making customer retention a critical factor in enhancing overall profitability [3]. In today's increasingly competitive market, retaining existing customers is crucial for any business. According to Nie's research, a bank can increase its profits by 85% by improving the retention rate of existing customers by just 5% [3]. Therefore, customer churn is a critical issue for the banking industry. Customer churn define as “the loss of a customer to a competitor, which leads to losses in profits” [2]. Identifying customers who are likely to terminate their banking services and understanding the patterns associated with these customers can significantly enhance service quality and increase retention rates. In the modern era, machine learning methods are widely employed by banks to predict customer churn. Using big data, these methods classify customers as either churners or non-churners [6-8] .Numerous studies in the literature have developed various predictive models using statistical and data mining techniques, including random forests, logistic regression, neural networks, Bayesian networks, the C5 tree, the chi-square automatic interaction detection (CHAID) tree, and the classification and regression (CR) tree.Using machine learning and statistical methods to predict customer churn is a well established field that has been studied by many researchers. A bunch of well trained models have been developed by researchers in the field. In contrast, this paper will not focus on developing the best model for predicting customer churn. The author will focus on developing models, find each model's strength and match the situation the models fit.

1. Literature Review

Churn management is a part of Customer Relationship Management (CRM), it is important for firms to maintain long-term customer relationships and optimize profitability [1]. The phenomenon of customer churn will affect firms in multiple dimensions. First and foremost, the loss of existing customers is one of the biggest factors that diminished profits. This loss not only results in an immediate decline in product sales but also will increase customer acquisition costs. Furthermore, if the services are under contractual agreements, customer churn represents more than just a reduction in sales - it also means the end of the customer relationship [1]. Under today’s intensely competitive market, the competition between banks is particularly fierce [5]. Therefore, the ability to predict customer churn and identify the causes of customer churn is crucial for banks seeking to sustain and enhance their competitive ability.

Accordingly, banks are closely monitoring their customers for warning signals that may indicate potential churn. Machine learning becomes a common tool for banks to predict churn by analyzing the data collected from their customers. These methods can make banks uncover hidden patterns behind datasets to build predictive models for predicting customer churn [4].

The prediction of credit card customer churn is a well-established area of research, with numerous studies contributing to the field. One remarkable researcher is Dana, her studies is focusing on developing credit card customer churn prediction models using a feature selection method running by five machine learning models: the Bayesian network, the C5 tree, the chi-square automatic interaction detection (CHAID) tree, the classification and regression (CR) tree, and a neural network [2]. Above this basic idea, Dana constructed three variations of each prediction model, utilizing different independent variable selection methods, including the use of all independent variables, two-step clustering, k-nearest neighbor, and feature selection [2]. According to her research, the C5 tree machine learning model performs the best. It has a test recall value of 0.861. The output of the model reveals that the top three variables influencing customer churn were the total transaction count, the total revolving balance on the credit card, and the change in transaction count [2].

Although Dana's research indicates that the C5 tree model delivers the best performance, different models offer unique advantages and are suited to different scenarios. Recognizing this, the author of this study identifies two research gaps:

Exploring the use of various models, including decision trees, logistic regression, and neural networks, to develop a best predictive model for credit card customer churn.

Uncover the strengths of each model and describe the situations in which each model is most effectively applied.

Using the model's results, provide recommendations for the bank to reduce customer churn.

1. Exploratory Data Analysis (EDA)

The dataset used for this research was collected from <https://www.kaggle.com/datasets/teralasowmya/bankchurner> in Karggle websit. The dataset comprises 21 features, including both categorical and numerical variables. It contains records of 10,000 customers. The EDA revealed several insights into the factors influencing customer attrition. The distribution of customer age revealed a higher concentration of customers within the 40-60 age group. Attrition rates varied significantly by gender, with female customers exhibiting a slightly higher attrition rate. The credit limit distribution indicated that customers who churned tended to have lower credit limits. Additionally, the scatter plot of total transaction amount versus customer age showed that younger customers generally had higher transaction amounts, with a noticeable distinction between existing and churned customers.

Correlation analysis identified strong correlations between features such as total transaction count, total transaction amount, and average utilization ratio with the target variable. Furthermore, we observed a significant imbalance in the data between churners and non-churners. As a result, we employed undersampling and oversampling methods to balance the data in subsequent model-building steps.

1. Data Preprocessing
   1. Data Splitting

In the data splitting section, the common approach is to divide the existing dataset into a training set and a test set. However, in this study, we divided the dataset into three parts: the training set, validation set, and test set. The training set is used to train the model, the test set is used to evaluate the model's performance after training and validation, and the newly added validation set is used for fine-tuning the model's performance and parameters after training. The data in the validation set accounts for only 5% of the total data, which is less than the amount in the test set. This approach has two advantages. First, validating the model after training can improve its accuracy during testing. Second, with a smaller amount of data, the time required to optimize the model is reduced, saving time.

* 1. Handling Missing Values

To handle missing values in the dataset, we employed a strategy of filling in the gaps with the most frequent value for each feature. This approach offers several advantages. First, it helps maintain data consistency, as the most frequent value, having already appeared multiple times in the dataset, reflects the common characteristics of the data. Second, it reduces data bias, particularly when compared to methods such as filling with the mean or median, which can be influenced by outliers or extreme values. This is especially advantageous for categorical data, where the most frequent value often represents the predominant category. Additionally, this method is straightforward and requires no complex calculations, making it an efficient and effective way to address missing data. Finally, by using the most frequent value, the original data distribution is better preserved, ensuring that the filled data remains representative of the dataset's overall characteristics. Overall, this approach enhances data quality and, consequently, improves the performance of machine learning models.

1. Model Building
   1. Model 1: Logistic Regression

Logistic regression is an important method used for classification tasks. It finally obtains the probability of each class in the data set or event existing. Unlike linear regression which predicts a continuous value, logistic regression predicts the probability of a binary outcome. It's commonly used in various fields such as healthcare, finance, and marketing.

### 5.1.1 Data Processing

Before doing logistic regression, we can find that the imbalance problem in the data set. Undersampling and oversampling are two methods to address this issue. However, among these two resampling strategies, undersampling has been shown to be a better choice than oversampling [12]. Undersampling is an effective data processing method, particularly suitable in two scenarios. First, when the number of majority class samples is very large, undersampling can reduce these samples, thereby alleviating data imbalance. Additionally, for large datasets, undersampling can decrease the computational load. The dataset in this problem contains over 10,000 entries, making it necessary to preprocess the data using undersampling. After applying undersampling and feeding the processed data back into the logistic regression model, there was a significant improvement in the recall value, as shown in Table 1.

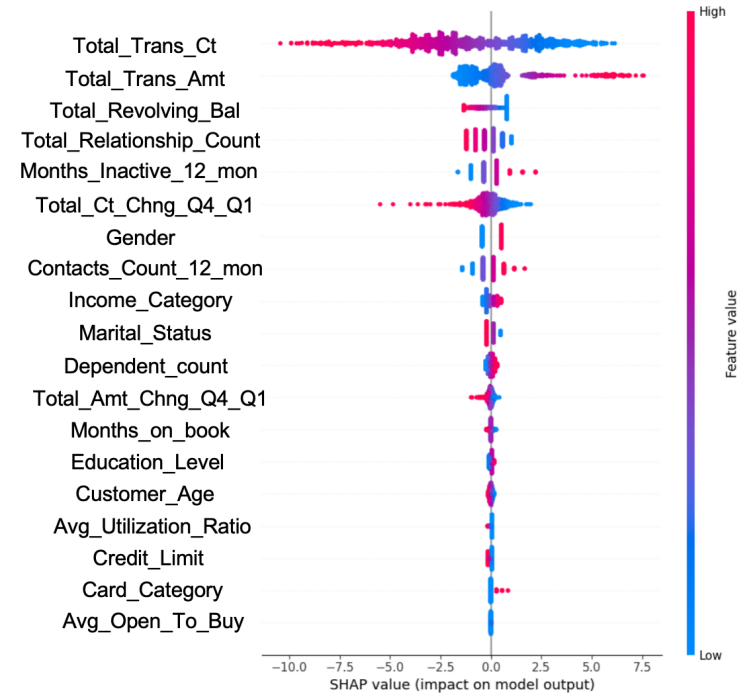
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| 0 | 0.97 | 0.83 | 0.89 | 993 |
| 1 | 0.51 | 0.87 | 0.64 | 203 |
| accuracy |  |  | 0.84 | 1196 |
| macro avg | 0.74 | 0.85 | 0.77 | 1196 |
| weighted avg | 0.89 | 0.84 | 0.85 | 1196 |

**Table 1.** Recall value after undersampling

The improved recall value indicates that the model is highly fitted to the existing data and performs well in predictions, making the results both feasible and highly reliable.

### 5.1.2 Model Results.

The final results of the logistic regression model are generated by using the SHAP package. The SHAP values method is a feature attribution technique that assigns a value to each feature for a specific prediction, which is useful for interpreting the prediction results [17]. SHAP can generate feature importance plots and other global interpretation charts which can help us to understand the overall behavior of the model and the impact of each feature on customer churn. Additionally, SHAP provides explanations for individual predictions, showing the contribution of each feature in specific cases. This is significant for understanding the model's decision-making process at the individual level. The model analyzed the impact of various user features on the possibility of customer churn, and the results displayed in Figure 1.



**Fig. 1.** SHAP result

This figure illustrates the impact of each feature on the model's output using SHAP (SHapley Additive exPlanations) values, which highlight how features contribute to prediction results. The color gradient, ranging from blue (lower values) to red (higher values), indicates the magnitude of feature values in each sample. The X-axis represents SHAP values, where positive values indicate a feature's positive impact on the prediction, and negative values indicate a negative impact. The spread of dots along the X-axis reflects the variation in each feature's impact, while the vertical density of dots shows how frequently similar SHAP values occur across samples. Analyzing these relationships helps identify the most influential features in predicting customer churn and whether they increase or decrease the likelihood of churn, providing valuable insights for effectively addressing customer retention.

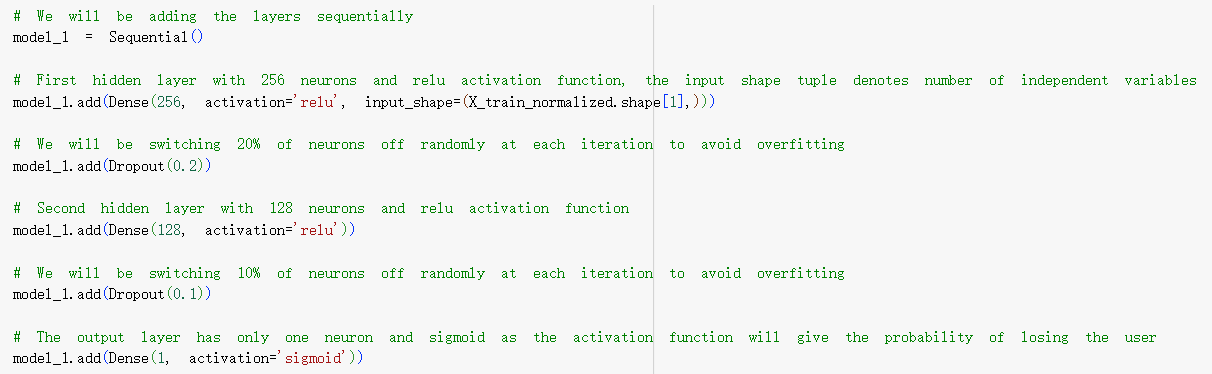
From the above figure, we can easily find that Total transaction count in the last 12 months and Total transaction amount in the last 12 months are two main reasons which infect the churning of the customers. They can also be considered as two key parameters which can predict the future churning of the customers in the neural network.

* 1. Model 2: Neural Network

Muneer et al. (2022) provide a thorough examination of various machine learning techniques for churn prediction in the banking industry, highlighting the significance of feature importance and model performance.[9] This has some inspiration for us to build models. Rahman and Kumar (2021) explore the application of machine learning, including neural networks, for customer churn prediction in banking, emphasizing the relevance of this approach to our study.[10] In this case, the model is designed to predict whether customers will remain with the bank or stop using its services. It does this by analyzing input data that includes 20 features, such as CLIENTNUM, Customer\_Age, Gender, and others, to produce an output indicating true or false.

### 5.2.1 Step one: Determine and define the number of Layers and Nodes.

#### Model 1



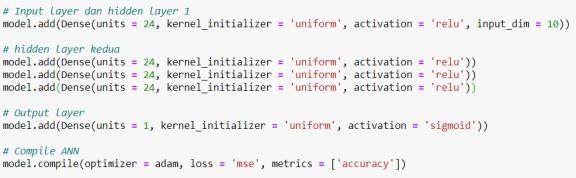
**Fig. 2.** Layers and nodes

#### Model 2



**Fig. 3.** Layers and nodes

#### Model 3



**Fig. 4.** Layers and nodes

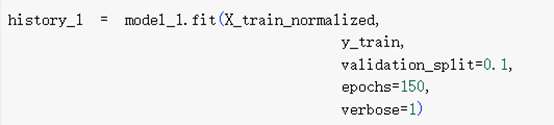
The three neural network models increase in complexity and depth, while sharing core features. Each model has an input layer, multiple hidden layers, and a single-output neuron with a sigmoid function for binary classification.

Model 1 has two hidden layers (256 and 128 neurons) with ReLU activation and dropout layers. Model 2 adds a third hidden layer (64 neurons) and uses tanh activation in all hidden layers, with dropout after the first two. Model 3 includes a fourth hidden layer (32 neurons), maintaining tanh activations and dropout after the first three layers.

All models use weight variables initialized with a uniform function, adjusted during training based on the Mean Squared Error loss function. The Adam optimizer fine-tunes weights to minimize loss and maximize accuracy, using a 0.01 learning rate. Activation functions transform inputs into outputs for each layer, driving the network's predictions through sequential matrix calculations.

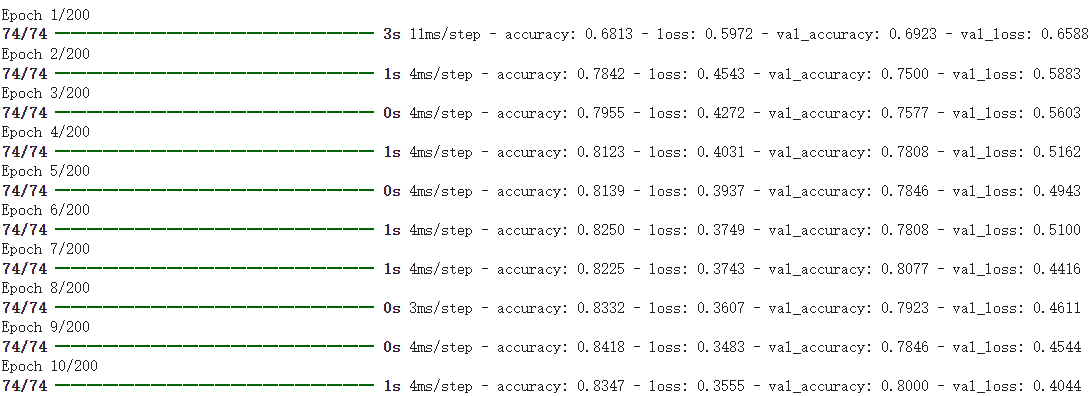
### 5.2.2 Step two: Training Model

Creating three different models allows the system to study the correlation between feature data and labels through mathematical calculations. This approach ensures that the models can be effectively applied to new data in the future.



**Fig. 5.**

By writing the model\_1.fit function and providing certain arguments, the model is trained using normalized feature data (X\_train\_normalized) and corresponding labels (y\_train). The validation\_split=0.1 argument is used to allocate 10% of the training data for validation purposes, helping to monitor the model's performance during training. The epochs=150 argument sets the number of times the model will go through the entire training dataset during the training process, allowing it to learn and adjust its parameters through forward and backpropagation. The verbose=1 argument ensures that detailed information about the training process is displayed, including the loss and accuracy for each epoch.

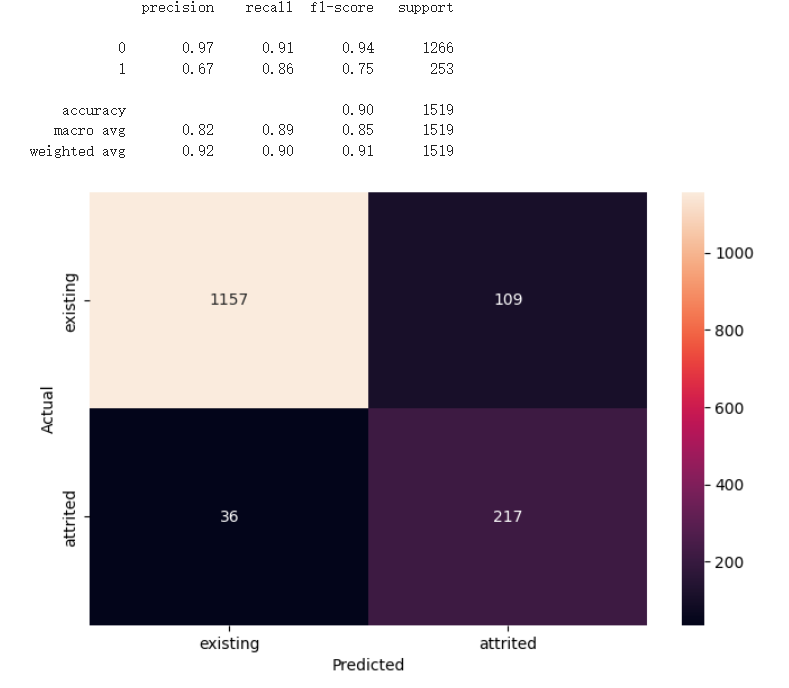


**Fig. 6.** Model training process

Figure 6 shows the model 3 training process (one sample taken from the three model designs) with an epoch of 200, with a loss of 0.0491 and an accuracy of 98%.

### 5.2.3 Step 3: Testing the model.

The data used in this stage are X\_test and Y\_test from the dataset to test whether the model can predict from new data, and not only remember the data from the model training process.

****

**Fig. 7.** Accuracy obtained for testing data

Figure 7 shows that the model achieved a recall of 0.86 in testing, compared to 0.98 in training. This indicates that the model can effectively predict new data, rather than simply memorizing the training data.

|  |  |  |
| --- | --- | --- |
| **recall value** | **Recall of**  **Training** | **Validation**  **Recall** |
| **model 1** | 98% | 72% |
| **model 2** | 98% | 80% |
| **model 3** | 98% | 86% |

**Table 2.**

Table 2 is the result of designing a model from 3 layer and node models that have been created

In conclusion, deep learning models based on artificial neural networks (ANN) are used to address classification problems and enhance predictive services in the banking industry. With input customer data, the model achieves a recall value of 86%. Designed using supervised machine learning and a dataset of existing bank customer churn data, this model demonstrates the ability to make accurate churn predictions. This accuracy can help the company make informed decisions and mitigate potential losses.

* 1. Model 3: Decision Tree

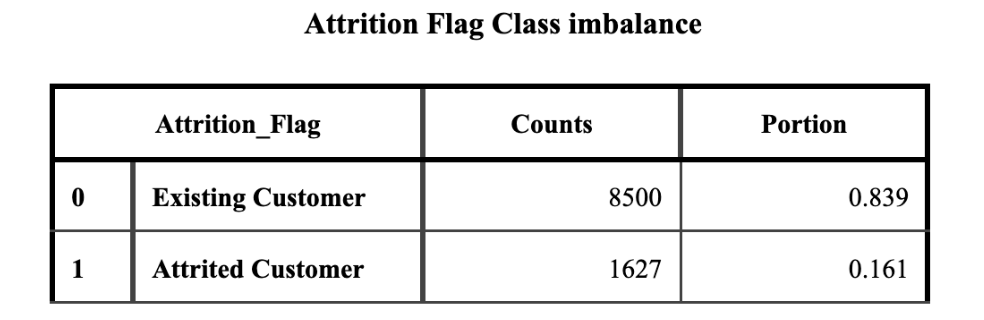
### 5.3.1 Model Construction

Decision tree is one of the predictive modeling approaches used in statistics, data mining and machine learning.[11] The target variable of our decision tree model is Attrition\_Flag, indicating whether a customer is an existing or attrited customer. Our primitive tries of decision tree utilized the Decision Tree Classifier with the criterion set to ‘gini’. This decision tree classifier was trained on the training data without any constraints on its growth. However, the mismatched accuracy of the training set and test set indicates signs of overfitting, especially with our imbalanced datasets (Table.3).

We further refined our model by pruning which reduces the complexity of the decision tree by removing sections that provide minimal classification power. We constrained the decision tree's depth by setting the maximum depth of 3, and required a minimum sample of 20 to split the internal node and at least 10 samples to be at a leaf node. After pruning, the model showed balanced performance with a high accuracy compared to our initial model, but the recall score still indicates high potential to improve. Even though we further constructed our model using the decision tree classifier with ccp\_alpha set to 0.0016994637735378475, there were only slight improvements on the recall score.

By re-examining the zero and ones in our data set referring the attrition flag value, the class imbalance problem is extremely noticeable, therefore we refined our model by employing oversampling and undersampling techniques to address this issue (Table.3).

**Table 3.** Attrition Flag Class imbalance



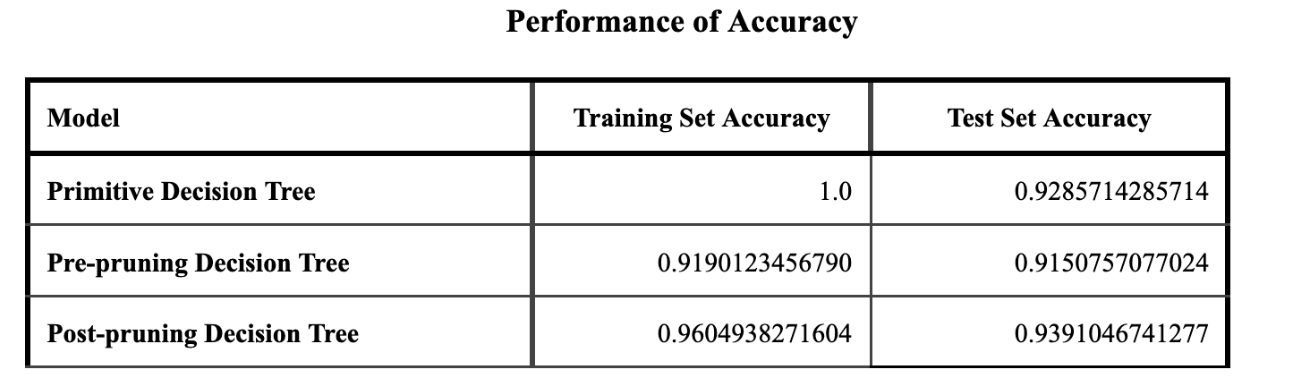
Oversampling increases the number of minority class instances to balance the dataset using the Synthetic Minority Over-sampling Technique (SMOTE). It is based on sampling data from the minority class by simply generating data points on the line segment connecting a randomly selected data point and one of its K-nearest neighbors.[16] The training data was resampled using SMOTE to balance the class distribution followed by a decision tree classifier trained on this oversampled data.

Our final model is undersampled decision tree model. Under sampling reduces the number of majority class instances to balance the dataset by randomly removing samples from the majority class. The training data was resampled using random under sampling to balance the class distribution. A decision tree classifier was then trained on this under sampled data, achieving the overall best recall score.

1. Model evaluation

Our primitive decision tree model was trained without any constraints on its growth, therefore, there was an asymmetry between a perfect train set accuracy and test set accuracy. By pre-pruning to reduce overfit, the accuracy of training and test sets are 0.919 and 0.915. (Table.6) Similarly, the post-pruning achieved an accuracy score of 0.960 and 0.939, (Table.4) both models solved the overfitting problem, but there was still room for improvement of the recall score.

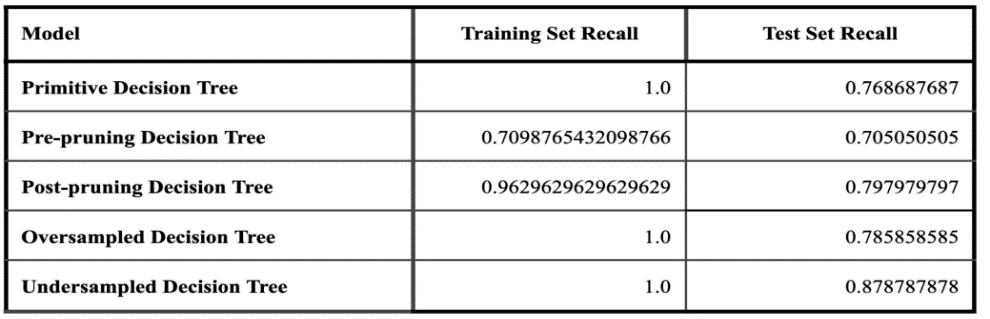
**Table 4.** Performance of Accuracy



The recall scores are the key element when we compare the effectiveness of each model. Recall score was chosen as the evaluation metric because it focuses on the model's ability to correctly identify the crucial minority class (attrited customers).

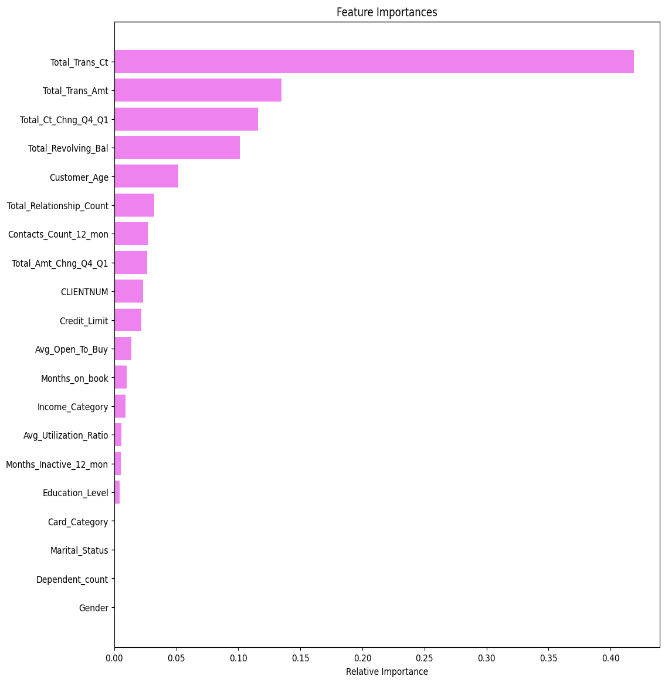
The primitive decision tree achieved a recall score of 0.769, the pre-pruning model 0.705, the post-pruning model 0.798, the oversampled model 0.786, and the undersampled model 0.879(Table.5). By reducing complexity, pruning significantly improved the model's performance, however, the undersampled decision tree model provided the highest recall score, effectively solving the class imbalance and over fitting issue, obtaining an optimal balance between predictive power and interpretability.

**Table 5.** Performance of Decision Tree model



1. Conclusion

Based on the experience of building decision trees, logistic regression, and neural network models, certain characteristics and the suitability of each model for bank churn prediction can be identified. The primary advantage of decision trees is their highly interpretable results, which are presented in an intuitive, tree-like structure that enhances both understanding and the communication of findings [15]. Moreover, the decision tree model has the highest recall value during this study. The weakness of decision tree is prone to overfitting and can be mitigated through pruning and sampling adjustments [14]. Therefore, the decision tree is well-suited for situations where a bank requires highly accurate predictions of churners and a model that can be clearly and effectively presented. Compared to decision trees, logistic regression offers a simpler way to present prediction results. However, its ability to accurately predict churners is generally weaker than that of a decision tree model [13]. Therefore, if a bank seeks to build a straightforward model with reasonable accuracy for predicting churners, a logistic regression model may be appropriate. In contrast to decision trees and logistic regression, neural networks are more complex models. Additionally, the internal workings of a neural network during processing are not visible, making the model less interpretable. The output of a neural network is simply a prediction of whether a customer will churn or not. However, once properly tuned, a neural network can deliver a clear and straightforward prediction. This makes it a valuable tool for banks in identifying potential churners.



**Fig. 8.** Feature importances to Customer churning from Decision Tree

By combining the results of the logistic regression and decision tree models (Figure.1, Figure.8), we identified the top four features influencing customer churn: total transaction count and amount, total revolving balance, customer age, and total relationship count. Moreover, we identified the suitability of each model for different needs within the banking sector when predicting customer churn.

1. Recommendations for Banks

To improve customer retention and reduce churn, banks should consider implementing several strategic initiatives. First, enhancing customer engagement is crucial. This can be achieved by developing and promoting rewards programs tailored to various customer segments, crafting marketing strategies that appeal to different age demographics, and implementing personalized communication strategies to build stronger relationships with customers. Additionally, banks should focus on improving financial support services by offering financial planning and advisory services, as well as providing balance transfer options and lower interest rates for customers with high revolving balances. Strengthening product cross-selling efforts is also important; banks can encourage customers to use multiple banking products by offering bundled deals and discounts for those who hold multiple accounts. Finally, attention should be given to high-risk segments by closely monitoring transaction activity and engaging low-activity customers with targeted incentives to increase their usage of banking services.

Acknowledgement

Yibin Li, Mingze Gao, Yanfu Zhang and Changbin Feng contributed equally to this work and should be considered co-first authors.

Muxi Chen, Luyun Zhang contributed equally to this work and should be considered co-second authors.

Reference

1. Risselada, H., Verhoef, P. C., & Bijmolt, T. H. (2010). Staying power of churn prediction models. *Journal of Interactive Marketing*, *24*(3), 198-208.
2. Al-Najjar, D., Al-Rousan, N., & Al-Najjar, H. (2022). Machine learning to develop credit card customer churn prediction. *Journal of Theoretical and applied electronic commerce research*, *17*(4), 1529-1542.
3. Nie, G.; Rowe, W.; Zhang, L.; Tian, Y.; Shi, Y. Credit card churn forecasting by logistic regression and decision tree. *Expert Syst.Appl.* **2011**, *38*, 15273–15285. [CrossRef]
4. Anil Kumar, D., & Ravi, V. (2008). Predicting credit card customer churn in banks using data mining. *International Journal of Data Analysis Techniques and Strategies*, *1*(1), 4-28.
5. Kumar, D.A.; Ravi, V. Predicting credit card customer churn in banks using data mining. *Int. J. Data Anal. Tech. Strateg.* **2008**, *1*, 4–28. [CrossRef]
6. Xia, G.; He, Q. The research of online shopping customer churn prediction based on integrated learning. In Proceedings of the 2018 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2018), Qingdao, China,30–31 March 2018; pp. 30–31.
7. Rahman, M.; Kumar, V. Machine learning based customer churn prediction in banking. In Proceedings of the 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 5–7 November 2020; IEEE:Piscataway, NJ, USA; pp. 1196–1201.
8. Khodabandehlou, S.; Rahman, M.Z. Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behaviour. *J. Syst. Inf. Technol.* **2017**, *19*, 65–93. [CrossRef]
9. Muneer, A., et al. (2022). Predicting customers churning in the banking industry: A machine learning approach. *Indian Journal of Electrical Engineering and Computer Science, 26*(1), 539-549.
10. Rahman, M., & Kumar, V. (2021). Machine learning-based customer churn prediction in banking. In the 4th *International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1196-1201).
11. Xindong Wu et al. (2007), “Top 10 Algorithms in Data Mining - Knowledge and Information Systems,” SpringerLink, December 4, 2007, <https://link.springer.com/article/10.1007/s10115-007-0114-2>.
12. Lin, W. C., Tsai, C. F., Hu, Y. H., & Jhang, J. S. (2017). Clustering-based undersampling in class-imbalanced data. Information Sciences, 409, 17-26.
13. Long, W. J., Griffith, J. L., Selker, H. P., & D'agostino, R. B. (1993). A comparison of logistic regression to decision-tree induction in a medical domain. *Computers and Biomedical Research*, *26*(1), 74-97.
14. Esposito, F., Malerba, D., Semeraro, G., & Kay, J. (1997). A comparative analysis of methods for pruning decision trees. *IEEE transactions on pattern analysis and machine intelligence*, *19*(5), 476-491.
15. De Ville, B. (2013). Decision trees. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*(6), 448-455.
16. Elreedy, D., & Atiya, A. F. (2019). A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance. *Information Sciences, 505*, 32–64.
17. Meng, Y., Yang, N., Qian, Z., & Zhang, G. (2020). What makes an online review more helpful: an interpretation framework using XGBoost and SHAP values. Journal of Theoretical and Applied Electronic Commerce Research, 16(3), 466-490.

Appendix

## 1.Data Overview

The dataset comprises 21 features, including both categorical and numerical variables. It contains records of 10,000 customers with the following key attributes:

* CLIENTNUM: Unique identifier for the customer.
* Attrition\_Flag: Indicates whether the customer is active or has churned.
* Customer\_Age: Age of the customer.
* Gender: Gender of the customer.
* Dependent\_count: Number of dependents.
* Education\_Level: Educational qualification of the customer.
* Marital\_Status: Marital status of the customer.
* Income\_Category: Income category of the customer.
* Card\_Category: Type of credit card held by the customer.
* Months\_on\_book: Duration (in months) of the customer’s relationship with the bank.
* Total\_Relationship\_Count: Total number of products held by the customer.
* Months\_Inactive\_12\_mon: Number of months inactive in the last 12 months.
* Contacts\_Count\_12\_mon: Number of contacts in the last 12 months.
* Credit\_Limit: Credit limit of the customer.
* Total\_Revolving\_Bal: Total revolving balance on the credit card.
* Avg\_Open\_To\_Buy: Average open-to-buy credit line.
* Total\_Amt\_Chng\_Q4\_Q1: Change in transaction amount (Q4 over Q1).
* Total\_Trans\_Amt: Total transaction amount in the last 12 months.
* Total\_Trans\_Ct: Total transaction count in the last 12 months.
* Total\_Ct\_Chng\_Q4\_Q1: Change in transaction count (Q4 over Q1).
* Avg\_Utilization\_Ratio: Average card utilization ratio