A Dynamic Classification Approach to Delinquency Prediction in the Credit Card Issuing Industry

*Emergent Research Forum (ERF) Papers*

**Wingyan Chung Atit Acharya**

Dept. of Computer Science, UT Tyler University of Texas at Tyler

# Abstract

Credit card delinquency is a significant problem for banks and financial institutions. Predicting credit card delinquency is a challenging task due to the complex nature of the data. In this study, we developed predictive models for credit card delinquency using three machine-learning algorithms: random forest, XGBOOST, XGBoost with DART, and DMatrix. We used 100 balanced datasets with 2240 values and having equal distribution of negative and positive classes of target features in each dataset. Each dataset contains eight input features and one target feature. These features contain financial and behavioral factors. We evaluated the performance of each algorithm using several performance metrics, including accuracy, precision, and recall score. Our results showed that XGBoost with DART outperformed the other algorithms in terms of performance metrics. We also identified the most significant predictors of credit card delinquency, which included outstanding balance, user’s card external status, etc. Our study provides insights into the most effective ensemble of machine learning algorithms for predicting credit card delinquency and the most significant predictors of credit card delinquency.

## Keywords (Required)

Random forest classifier (RFC), XGBoost with DART and DMatrix, delinquency, hyperparameters

# Introduction

Credit Card Delinquency is the process where credit card users fail to make timely payments on their outstanding balances. This is a severe problem for banks and financial institutions that can lead to financial losses for the bank and negatively impact the user’s credit score. Predicting credit card delinquency can help banks and financial institutions to identify high-risk customers and take proactive measures to prevent delinquency. This paper proposes different classification techniques such as the Random Forest Classifier model, XGBOOST with DMatrix, and XGBOOST with DART to find the binary logistics classification report and the accuracy to predict credit card delinquency using a balanced dataset. To date, several studies have been conducted on credit card delinquency prediction using various machine-learning techniques. For instance, (Zhang et al. 2019) used a logistic regression model to predict credit card delinquency, and (Li et al. 2020) used a support vector machine (SVM) classifier to predict credit card default. However, these studies did not address the issue of dynamic and multiple datasets with multiple features affecting the prediction of credit card delinquency.

This research helps to develop and validate a dynamic classification approach to predict the financial behavior of credit card users for delinquency prediction. These multiple classification approaches are applied to delinquency prediction on 100 balanced datasets consisting of 2240 values each with equal distribution of positive and negative values selected randomly from 1,236,973 values. The banking, credit card issuing companies, and credit unions need delinquency prediction to protect against potential risk and minimize losses when possible. According to the datasets used for this research, the delinquency queues are considered zero or reset once the credit card users make a required minimum payment whereas those whose payments become outstanding for 70-89 days are delinquent (University of Central Florida, 2022-23).

**Literature Review**

Delinquency prediction allows banks, credit card companies, and credit unions to detect potential users who are not paying their outstanding balance amount at the end of payment cycles to reduce monetary loss. Previous work related to delinquency prediction was based on feature engineering, data mining models, rare event prediction, and modeling approaches Their strengths and weaknesses are reviewed below:

**Modeling Approaches**

The work done in credit card default prediction in the past includes logistic regression, decision tree algorithms, support vector machine (SVM), gradient boosting machine (GBM), Bayesian network model, random forest model, etc. (Shen and Huang 2018) The logistic regression model is used for credit card delinquency prediction consisting a simple and interpretable model that can handle both categorical and continuous variables however, have a drawback of not capturing complex non-linear relationship. (Yang and Kim 2019) The hybrid approach uses a logistic and gradient boosting machine (GBM) for the prediction of credit card delinquency that can capture both linear and non-linear relationships between variables but the hybrid model may be more complex and harder to interpret than a single model. (Yang et al. 2018) A decision tree approach for the prediction of credit card delinquency was used in China that can identify non-linear relationships between variables and are easy to interpret; however, decision trees can be prone to overfitting and may not be generalized to new datasets. (Khandani et al. 2010) Bayesian networks were used to handle complex relationships between variables and that can be used to make probabilistic predictions but such networks can be computationally intensive and may require large datasets to estimate parameters. Hence, there is a need to provide an efficient and clear approach to processing large, complex, imbalanced, non-linear data.

**Delinquency Related Features**

Delinquency prediction may depend on a number of features such as credit score, payment history, account status, etc. The study is related to the prediction of credit card delinquency using logistic regression (Lu and Zhang 2017) using different features such as credit limit, outstanding balance, and payment history. The author used a real-world dataset which added external validity and generalizability and also provides insights into the most significant predictors of credit card delinquency. However, the authors did not provide a comprehensive list of features used in the analysis, making it difficult to assess the relevance of each feature and their work did not compare the performance of logistic regression with other machine learning algorithms, limiting the scope of the analysis. We need to identify the most important features in delinquency prediction that can assist credit card issuing companies to prevent loss and delinquencies.

**Rare Event Prediction**

The challenge in delinquency prediction is the imbalanced dataset that can significantly affect the learning performance of machine learning models. The study for rare-event prediction uses an extended over-sampling approach to address the issue of imbalanced data in credit card delinquency prediction(Chen & Li 2019). The strengths of this study are that it addressed the issue of imbalanced data and evaluated the performance of the proposed approach using various metrics, including AUC-ROC and recall. However, it only used an ensemble of decision trees as the prediction model and did not compare the performance of the proposed approach to other models, and didn’t use any feature selection techniques, which could potentially improve the performance of the prediction model. Also, (Wang et al. 2020) a study based on hybrid feature selection and ensemble learning which was effective in predicting rare events in credit card delinquency, and the use of mutual information and maximum relevance minimum redundancy criteria in the feature selection process helped to identify the most relevant features for prediction. However, the study did not compare the performance of the proposed approach to other methods, making it difficult to assess its effectiveness relative to other approaches and it only used logistic regression and decision tree as the base classifier, which may not be as effective in capturing complex relationships in the data compared to other models (Wang et al. 2020).

**A Dynamic Classification Approach to Delinquency Prediction**

This study aims to predict delinquent credit card users by developing and validating dynamic classification approaches through the optimized model specifications. These approaches contain multiple classification techniques such as Random Forest Classifier, XGBOOST with DMatrix, and XGBOOST with DART for dynamic and balanced datasets.

**Experiments and Preliminary Results**

The datasets used in this study include over 2240 values with balanced target features having an equal number of values for positive and negative classes and 100 such samples are used. The input and target features selected for this study are listed in Table 1.

|  |  |
| --- | --- |
| Features | Feature Description |
| LastStatementMinimumPaymentDueAmount | Amount of the minimum payment due that appeared on the last statement |
| CardExternalStatus | Status inputted by staff on the card account |
| LastStatementBalanceAmount | Amount of the ending balance that appeared on the last statement |
| CardType | Credit Card BIN type |
| CreditLine | The maximum amount of credit that can be used |
| LastStatementPurchaseAmount | Amount of purchases that were posted on the last statement |
| LastStatementPaymentTotalAmount | Amount of payments posted on the last statement |
| LastStatementPurchaseReturnAmount | Amount of returned transactions during the statement cycle |
| Target | Delinquent days count between 70 and 89 days |

**Table 1. Selected Features related to payment activities**

We used a dataset from Addition Financial’s Collections department containing information on credit card users, including demographic information, credit limits, card types, status, payment information, credit history, (University of Central Florida, 2022-23), etc. From the large dataset, 100 samples with an equal number of negative and positive target features in each are taken and fetched to multiple machine learning models to predict delinquency using dynamic hyperparameters. This study tries to predict a generalized delinquency prediction with the following specifications:

* Training, validating, and testing 100 datasets using RFC, DMatrix, and XGBoost with DART with dynamic hyperparameters using 80%, 10%, and 10% samples from each dataset respectively.
* Finding the average performance metrics for the 100 datasets using dynamic classification approaches as mentioned above.

A series of trials are conducted to find the key features and optimal specifications from those 100 datasets.

***Hyperparameters***

Hyperparameters play a crucial role in the performance of classification models, and their tuning is essential for achieving optimal results. As stated by Goodfellow et al. (2016), “Hyperparameters are settings for the learning algorithm that determine how much emphasis to place on different aspects of the model and training procedure” (p. 401). For this study, the hyperparameters settings for training, validating, and testing are objective as binary logistics, evaluation metrics as logarithmic loss (logloss), 80% of the observations are sampled for each tree, 80% of columns samples for each tree, the maximum depth for decision tree is wet to 10, L1 regularization term on weights is set to 0.1 and minimum loss reduction required is set to 0.1

***Feature Selection***

This process of feature selection was carried out to find the best and optimal features that can help to predict delinquency from the original datasets that contain 37 features. Out of these 37 features, only 8 of them with the following accuracy score when each of them are subjected to the model individually as listed in Table 2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Features | RFC | XGBClassifier | DMATRIX | XGBDART |
| LastStatementMinimumPaymentDueAmount | 0.9062 | 0.9118 | 0.9106 | 0.9126 |
| CardExternalStatus | 0.8102 | 0.8091 | 0.8095 | 0.8093 |
| LastStatementBalanceAmount | 0.7288 | 0.7472 | 0.7407 | 0.7433 |
| LastStatementPaymentTotalAmount | 0.7137 | 0.7152 | 0.7138 | 0.7156 |
| LastStatementPurchaseAmount | 0.6744 | 0.6745 | 0.6746 | 0.6745 |
| LastStatementPurchaseReturnAmount | 0.6187 | 0.6241 | 0.6208 | 0.6245 |
| CreditLine | 0.6187 | 0.6241 | 0.6208 | 0.6245 |
| CardType | 0.5689 | 0.5687 | 0.5690 | 0.5690 |

**Table 2. Accuracy Score for each classification approaches**

From Table 2, further three key features have high accuracy scores to predict delinquency namely: LastStatementBalanceAmount, LastStatementMinimumPaymentDueAmount, and CardExternalStatus. Accuracy scores for all features are shown in Table 2.

***Predictive Modeling***

This study uses datasets chosen randomly and most of the values (nearly 80%) are used as training samples whereas 10% are used for validation and the remaining 10% are used to test the model for prediction. Accuracy, Precision, recall, and confusion matrix are used to measure model performance. The following formulae can be used to calculate the performance measure. TP is true positive, TN is true negative, FN is false negative, and FP is false positive. As we have taken the balanced dataset, the performance measures can be high according to the importance of the features that can be chosen from the given datasets whereas recall and precision are essential to measuring how well the model in predicting the target. i.e., delinquency.

**Results**

After, fetching the 100 datasets to the prediction models and hyperparameters, as mentioned above the performance metrics such as accuracy, precision, and recall score, are obtained as shown in Table 3 using XGBOOST (almost all models give the same result but XGBOOST with DART outperformed other approaches). Also, when the top features are dropped, we still got a better performance score which can be possible from the combination of other features. For example, if the user is ‘active’ with a good credit line and still has zero outstanding balance, then the user can fall into the non-delinquent class.

|  |  |  |  |
| --- | --- | --- | --- |
| Performance Measures (Average) | Results (including all features) | Results (when dropped top 2 features) | Results (when dropped top 3 features) |
| Accuracy Score | 0.9763 | 0.9372 | 0.8069 |
| Precision Score | 0.9664 | 0.9409 | 0.7815 |
| Recall Score | 0.9867 | 0.9322 | 0.8481 |

**Table 3. Results using all features vs. dropping top 3 features**

**Feature Importance**

Our preliminary results indicate and display that “LastStatementMinimumPaymentDueAmount”, “CardExternalStatus”, and “LastStatementBalanceAmount” are the most influential delinquency predictors. Credit Card users who pay the minimum payment due tend to be non-delinquent as compared to those who don’t pay at the end of the payment cycle and those who have more LastStatementBalanceAmount at the end of the payment cycle are more vulnerable to delinquency. However, when these top 3 features are dropped then the accuracy, precision, and recall scores decrease drastically.

# Conclusion

Our results showed that XGBOOST using DART (XGBDART) outperformed random forest and DMatrix in terms of accuracy, precision, recall, and other performance measures. We can see how the average performance measures of the model in Table 3 change when the top-ranked features according to their accuracy score as shown in Table 2 are excluded. The most significant predictors of credit card delinquency were LastStatementMinimumPaymentDueAmount, CardExternalStatus, LastStatementBalanceAmount, etc. However, the combination of features can also influence the change in values of the target feature.

# Acknowledgments

The authors acknowledge the data provider’s kind assistance as well as the conference editors and reviewers for their valuable comments.

# REFERENCES

Li, Y., Xie, Y., Zhou, W., & Feng, J. (2020). “Predicting credit card default with machine learning methods,” *International Journal of Machine Learning and Cybernetics*, 11(11), 2427-2441.

Zhang, Z., Huang, W., & Gao, R. (2019). “Predicting credit card delinquency: A comparative study of logistic regression and decision tree,” *Journal of Business Research,* 98, 80-89.

Goodfellow, I., Bengio, Y., & Courville, A. 2016. “Deep Learning,” *MIT Press,* p. 401.

Yang, Y., Zhao, J., & Zhang, H. (2018). “A decision tree approach for predicting credit card delinquency in China,” *Journal of Financial Services Research*, 54(1), 53-70.

Shen, C., & Huang, C. (2018). “Logistic regression model for credit card delinquency prediction,”*Journal of Credit Risk*, 14(1), 25-44.

Yang, H., Kim, Y. (2019). “A hybrid approach to predict credit card delinquency using logistic regression and gradient boosting machine,” *Journal of Financial Services Research*, 55(2), 163-186.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). “Consumer credit-risk models via machine-learning algorithms,” *Journal of Banking & Finance*, 34(11), 2767-2787.

Lu, Y., & Zhang, L. (2017). “Predicting Credit Card Delinquency Using Logistic Regression,” *Journal of Financial Risk Management* 6(3), 63-70.

Chen, C. H., & Li, M. (2019). “Rare-event prediction for credit card delinquency using an extended over-sampling approach,” *Expert Systems with Applications*, 120, 254-265.

Wang, J., Yang, X., & Yin, L. (2020). “A novel rare event prediction model for credit card delinquency based on hybrid feature selection and ensemble learning,” *Expert Systems with Applications*, 143, 113063.

University of Central Florida. (2022-23). “Statistics and Data Science” from [Addition Financial Competition 2022-23 - Statistics and Data Science (ucf.edu)](https://sciences.ucf.edu/statistics/addition-financial-competition-2022-23/)