Explainable Artificial Intelligence for Customer Churning Prediction in Banking*

Nham Cao[0000−0001−5808−9819]

University of Economics, The University of Danang, Danang, Vietnam
nhamct@due.udn.vn

Abstract. In banking industry, customer churn prediction plays important role in business success due to the fact that the cost of attracting new customers is much more than that of retaining ones. Several Machine Learning (ML) models are being employed to make predictions in customer churn and achieve excellent performance. However, the problem with these models is a lack of transparency and interpretability. The goal of this work is developing explanations for customer churn prediction model using Shapley Additive exPlanations (SHAP) method.

Keywords: explainable artificial intelligence · customer churn prediction

1 Introduction

Customer churn also known as customer attrition, customer turnover, or customer defection, is defined as the propensity of customers to cease doing business with a company in a given time period [8]. It has become a significant problem and is one of the biggest challenges that many companies worldwide are facing. Customer churn introduces not only some loss in income but also other negative effects on the operations of companies[8]. According to [20], attrition rates in the banking industry hover around 15%, and the annual churn rates on new customers are roughly in the 20-25% range during the first year.

Churn management is the concept of identifying those customers who are intending to move their custom to a competing service provider[7]. Risselada stated that churn management is becoming part of customer relationship management [17]. It is important for companies to consider it as they try to establish long-term relationships with customers and maximize the value of their customer base.

Recent studies in predicting customer churn has investigated a wide range of algorithms. They are often evaluated on their predictive performance or their ability to discriminate between churners and non-churners. The performance of these machine learning models is remarkable but it is not the only aspect that is utmost importance. For customer churn prediction, understanding of the model and its outputs is important as well. It helps to target incentives to customers who have a high risk of churning to induce them to stay.

eXplainable AI (XAI) provides a suite of ML techniques that produce more explainable models [2]. Several studies have been proposed to explain models in banking industry such as explaining the outputs of credit card fraud detection, interpreting credit scoring models or giving explanations for underwriting loan system[14,21,18]. However, there have been not of focus on explaining customer churn prediction models.

This work focuses on explanations with a state-of-art method, namely Shapley Additive exPlanations (SHAP) to explain customer churn prediction by several ML models.

2 Related works

In financial service field, customer churn prediction has been extensively researched using different machine learning algorithms, particularly in banking industry. Decision trees and logistics regression are widely used because of their good predictive performance and robust results[12,15,9]. Also K-means, Support Vector Machine (SVM), Artificial Neural Network (ANN) have been proven to give excellent predictive performance[13,**?**]. Some advanced tree-based algorithms have been tested on electronic customer data[22] and hybrid methodologies combining logistics regression and decision trees has been applied to solve the problem[3,6]. Researchers also attempted to develop dynamic approach to optimizing customer churn prediction model by using time-series predictors, multiple time periods, and rare event detection[10].

XAI has been applied in several fields such as healthcare, medical, and finance[19]. This is mainly based on the improvement of the overall feature importance [1], Shapley Additive exPlanations (SHAP) value and other methods such as LIME[2]. In banking industry, SHAP and LIME method has been applied successfully to interrupt fraud detection model [14]. In [21], researchers use SHAP to provide global and local interpretation of the credit scoring model predictions to formulate a human-comprehensive approach to understanding the decision-maker. XAI has also applied to give details for the outputs of automating loan underwriting system [18]. However, so far there have not been of focus on explainability of ML models that provide predictions of customer churn.

In this work, we explore explanations of customer churn prediction models with SHAP method.

3 Methodology

In this work, explanations of customer churn prediction models are performed in three steps. As illustrated in Figure 1 the data is first preprocessed and then classified by several machine learning algorithms. Finally, the ML models are explained by SHAP method.

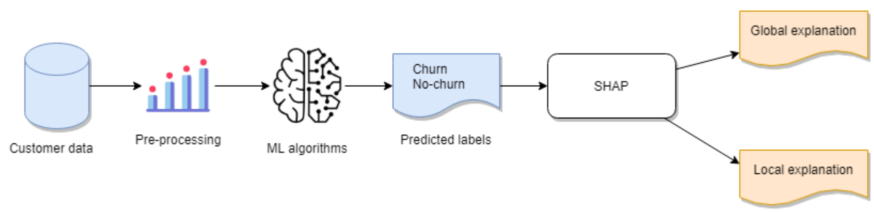


Fig. 1: Pipeline of explaining customer churn prediction models.

3.1 Data pre-processing

Firstly, the data is cleaned by dropping irrelevant columns, handling any special values in the credit card customer dataset, converting text variables to numeric values, removing outliers, converting categorical values. Finally, the dataset is split with 70:30 ratios for training set and testing set.

3.2 Classification

After preprocessing data, most commonly used classification techniques are implemented to build customer churn prediction models. The implemented ML algorithms include Logistics Regression (LR), Gradient Boosting(GB), Random Forest(RF), and Naïve Bayes. We also compare these models in term of their performance.

3.3 Explainable AI

The main goal of this work is to enhance customer churn prediction models by applying model-agnostic techniques. Model-agnostic methods are techniques of explainable artificial intelligence that can be used on any machine learning models and are applied after the model has been trained [2].They usually work by analyzing feature input and output pairs[16].

Explanations can be divided into two types: global explainability or local explainability. The distinction between them is in terms of scope. While global explainability focuses on how the model works from a global view point, local explainability targets specific predictions from the machine learning model [2].

Several methods as widely used to interpret ML models: feature imputation importance for ranking features that contribute to a model's decision, Local Interpretable Model-agnostic Explanations (LIME) which focuses on training local surrogate models to explain individual predictions [11], and Shapley Additive exPlanations (SHAP) that is a game theoretic approach to explain the output of any machine learning model[19]. SHAP method aims to explain the prediction of a specific instance by computing the contribution of each feature to the prediction [11]. SHAP can be also used to make global explanations using the combination or average across all local instances.

In this work, we use SHAP method with KernelSHAP library because it provides both global and local explanations and can be applied for all ML models.

4 Evaluation

4.1 Dataset

The credit card customer dataset used in this work was collected from Kaggle [5]. The original dataset has roughly 10000 customer records. Each record contains 21 features as described in Table 1.

Table 1: Data features

Feature	Description
CLIENTNUM	Unique identifier for the customer holding the account
Attrition_Flag	If the account is closed then 1 else 0
Customer_Age	Customer's Age in Years
Gender	M=Male, F=Female
Dependent_count	Number of dependents
Education_Level	Educational qualification of the account holder
Marital_Status	Married, Single, Divorced, Unknown
Income_Category	Annual income category of the account holder
Card_Category	Type of Card (Blue, Silver, Gold, Platinum)
Months_on_book	Period of relationship with 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 on the credit card
Total_Revolving_Bal	Total revolving balance on the credit card
Avg_Open_To_Buy	Open to buy credit line (Average of last 12 months)
Total_Amt_Chng_Q4_Q1	Change in transaction amount (Q4 over Q1)
Total_Trans_Amt	Total transaction amount (Last 12 months)
Total_Trans_Ct	Total transaction count (Last 12 months)
Total_Ct_Chng_Q4_Q1	Change in transaction count (Q4 over Q1)
Avg_Utilization_Ratio	Average card utilization ratio

The dataset has highly imbalanced classes, with attrition accounting for 16.07% of all customers.

Figure 2 is the correlation matrix between features evaluated in this work. Correlations are assessed using Pearson correlation coefficient.

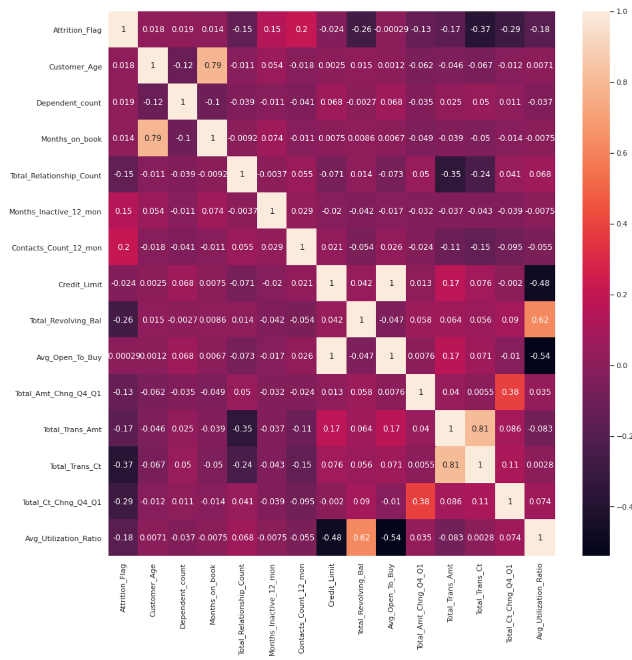


Fig. 2: Pearson correlation of features.

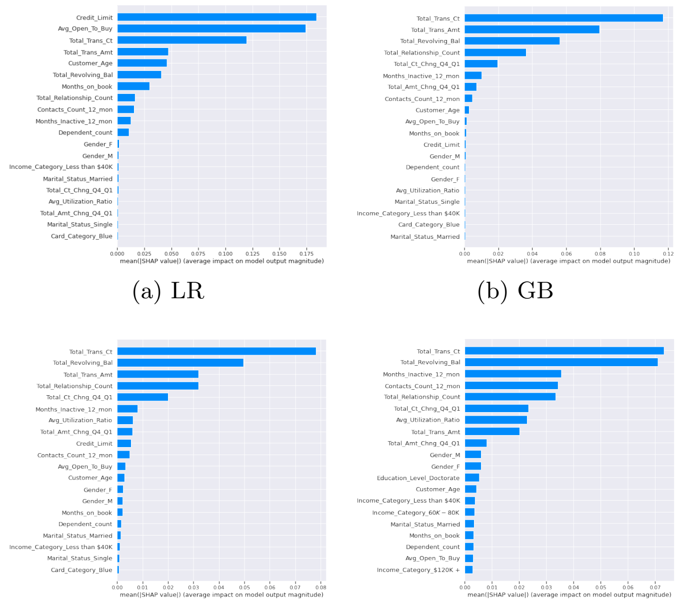
4.2 Experimental setups

In this work, I used Scikit-learn libraries to experiment with several machine learning algorithms, namely Logistics Regression (LR), Gradient Boosting (GB), Random Forest (RF) and Naïve Bayes (NB).

Table 2: Performance results

Algorithms	Precision	Recall	F1 score
Logistics Regression	0.7067	0.4711	0.5654
Gradient boosting	0.9412	0.8560	0.8965
Random forest	0.9301	0.7942	0.8568
Naïve Bayes	0.6983	0.6049	0.6483

Table 2 shows the classification results of each model. The performance of models was evaluated based on precision, recall, and F1 score. As mentioned in the Table 2, Gradient Boosting algorithm achieves the best performance on the chosen dataset.



(a) LR (b) GB

(c) RF (d) NB

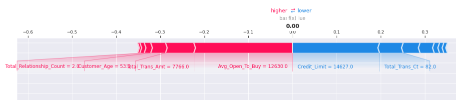
Fig. 3: Top features ranked by SHAP method

4.3 Experimental results

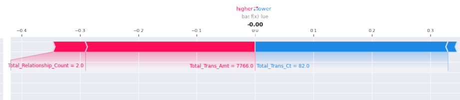
Global explanations Figure 3 illustrates the global explanation results for the implemented ML models on the chosen dataset. Explanations for four models produce similar top features, however the rankings of features vary across all models. For instance, Gradient Boosting, Random Forest and Naïve Bayes models agree that Total_Trans_Ct is the feature that affects the model's outputs the most, while that of Logistics Regression is Credit_Limit. These results are also consistent with top features related to the target variable in the correlation matrix in Figure 2.

Local explanation Local explanations provide a local understanding on how and why a specific prediction was made. With SHAP method, explanations for a specific prediction can be illustrated in a force plot. The red arrows in a force plot represent the features that drive the prediction higher, while the blue arrows represent the features that drive the prediction lower. The size of the arrows is proportional to the magnitude of the pushing force.

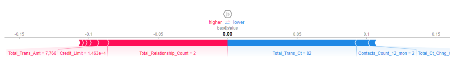
Figure 4 is the force plots for four implemented ML models which give explanations to a same single prediction. From the plots we can see that outputs of all implemented ML models are 0.0. This means that the customer is predicted to be not churn. However, features that impact to the outputs are different from models. In the RF and NB model, Total_Relationship_Count is the feature that major impact on the increase of output value, while that of GB and LR model is Total_Trans_Amt and Avg_Open_To_Buy respectively.



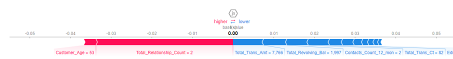
(a) Logistics Regression



(b) Gradient boosting



(c) Random forest



(d) Naïve Bayes

Fig. 4: Local explanation of a sample prediction

5 Discussion

In this work, we perform global and local explainability for four implemented ML models using model-agnostic approach with KernelSHAP library. Given explanations help us understand the factors that result in customer churn. After performing the experiment, we realize that increasing the size of background dataset leads to increase runtime linearly as illustrated in Figure 5. This is consistent with the fact that Shapley values are basically estimated based on random sampling.

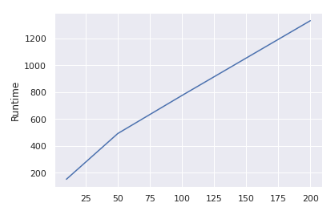


Fig. 5: Runtime for explaining LR model in seconds

Within the scope of the paper, we only performed explaining ML models of customer churn prediction but did not evaluate the given explanations. This limitation of our research is due to the fact that evaluation has proven to be a challenge[4]. In the future, we will focus on evaluation techniques to measure the accuracy of explanations of model-agnostic methods.

6 Conclusion

In order to the success of machine learning algorithms adopted in banking industry, model explainability is necessary to ensure accurate results. The relevant literature proposed several methods to predict customers who are probably to move their custom to other banks. However, we found that exploration of the explainability was limited. In this work, we provide insights for explanations of customer churn prediction by using SHAP - a model-agnostic method. We also found that the runtime of SHAP method with KernalSHAP library increases linearly when increasing the size of background dataset.

References

1. Altmann, A., Laura Toloşi, O.S., Lengauer, T.: Permutation importance: a corrected feature importance measure. Bioinformatics **26**(10), 1340–1347 (2010)
2. Arrieta, A.B., Díaz-Rodríguez, N., Bennetot, J.D.S.A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F.: Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. Information Fusion **58**, 82–115 (2020)
3. Caigny, A.D., Coussement, K., Bock, K.W.D.: 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 (2018)
4. Das, A., Rad, P.: Opportunities and challenges in explainable artificial intelligence (xai): A survey (2020)
5. Goyal, S.: Credit card customers, https://www.kaggle.com/sakshigoyal7/credit-card-customers
6. Özden Gür Ali, Arıtürk, U.: Dynamic churn prediction framework with more effective use of rare event data: The case of private banking. Expert Systems with Applications **41**(7), 7889–7903 (2014)
7. Hadden, J., Tiwari, A., Roy, R., Ruta, D.: Computer assisted customer churn management: State-of-the-art and future trends. Computers Operations Research **34**(3), 2902–2917 (2007)
8. K, C., Y-H, H., Y-C, H.: Predicting customer churn from valuable b2b customers in the logistics industry: a case study. Information Systems and e-Business Management **13**, 475–494 (2015)
9. Karvana, K.G.M., Yazid, S., Syalim, A., Mursanto, P.: Customer churn analysis and prediction using data mining models in banking industry. 2019 International Workshop on Big Data and Information Security (IWBIS) pp. 33–38 (2019)
10. Leung, H.C., Chung, W.: A dynamic classification approach to churn prediction in banking industry. AMCIS 2020 Proceedings (2020)
11. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model. 31st Conference on Neural Information Processing Systems (2017)
12. Mutanen, T., Nousiainen, S., Ahola, J.: Customer churn prediction – a case study in retail banking. Data Mining for Business Applications **218**, 77–83 (2010)
13. Olaniyi, A.S., Olaolu, A.M., Jimada-Ojuolape, B., Kayode, S.Y.: Customer churn prediction in banking industry using k-means and support vector machine algorithms. International Journal of Multidisciplinary Sciences and Advanced Technology **1**(2), 48–54 (2020)

14. Psychoula, I., Gutmann, A., Mainali, P., Lee, S.H., Dunphy, P., Petitcolas, F.A.P.: Explainable machine learning for fraud detection. arXiv (2021)
15. Rahman, M., Kumar, V.: Machine learning based customer churn prediction in banking. 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (2018)
16. Ribeiro, M.T., Singh, S., Guestrin, C.: Why should i trust you?: Explaining the predictions of any classifier. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining pp. 1135–144 (2016)
17. Risselada, H., Verhoef, P.C., Bijmolt, T.H.: Staying power of churn prediction models. Journal of Interactive Marketing **24**(3), 198–208 (2010)
18. Sachan, S., Yang, J.B., Xu, D.L., Benavides, D.E., Li, Y.: An explainable ai decision-support-system to automate loan underwriting. Expert Systems with Applications **144** (2020)
19. Samek, W., Montavon, G., Vedaldi, A., Hansen, L.K., Müller, K.R.: Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Springer (2019)
20. Statist: Customer churn rate by industry u.s. 2020, https://www.statista.com/statistics/816735/customer-churn-rate-by-industry-us
21. Torrent, N., Visani, G., Bagli, E.: Psd2 explainable ai model for credit scoring. arXiv (2020)
22. Xie, Y., Li, X., Ngai, E., Ying, W.: Customer churn prediction using improved balanced random forests. Expert Systems with Applications **36**(3), 5445–5449 (2009)