

Customer Churn Analysis Based on User Profile and Activity

Suwandi Dwi S and Zulkarnain
Department of the Industrial Engineering
University of Indonesia
Depok, Indonesia
suwandi.dwi81@ui.ac.id, zulkarnain@ui.ac.id

Abstract

Using datasets from one crypto exchange company in Indonesia, this study aims to predict churn in cryptocurrency exchange and analyze the factor that impacts it. The model developed in this work uses decision tree and random forest. Due to imbalance data, the undersampling method is applied before modelling. As the result, decision tree has 74.59% accuracy and the random forest 74.34%. From the two models, it is also found that the use of Google Authenticator and the frequency of cryptocurrency transactions are important factors to determine whether a customer will experience churn.

Keywords

Cryptocurrency Exchange, Churn Analysis, Decision Tree and Random Forest

1. Introduction

Since the emergence of bitcoin in 2009, the cryptocurrency industry has been overgrown in most countries. In Indonesia, the emergence of a cryptocurrency trading platform in 2013 was the beginning of the entry of the cryptocurrency industry. Cryptocurrency is increasingly recognized by many people in Indonesia, according to a study by Mckinsey & company, Indonesia has outpaced the rest of the world in terms of digital adoption with a score of 99% for the period of 2014-2017 outpacing India, South Korea, the United Kingdom, as well as the United States[1]. because of the massive expansion of cryptocurrencies, it appears that cryptocurrencies have created an entirely new and global industry. Then in 2018, the government regulated cryptocurrency trading through Badan Pengatur Perdagangan Berjangka Komoditi (BAPPEBTI), an agency under the Ministry of Trade.

As the Indonesian cryptocurrency industry grows and there are many cryptocurrency trading platforms, the competition for customers is becoming more and more difficult. Therefore, it is essential for crypto asset trading companies to retain customers already using the platform.

1.1 Objectives

There are many studies that discuss customer churn using data that already has a churn feature, especially in the telecommunication industry. Only a few of them describe the criteria used to determine whether a customer experiences churn or not, some research is contained in Table 1, but none of them specifically address customer churn in the crypto industry.

The primary objective of the study is to analyze the churn problem in cryptocurrency exchange company uses decision tree and random forest and find important factors that lead to churn. In this research, also we introduce a new criterion for determining the churn of a customer by combining the recency of making transactions and the balance amount in his wallet.

Table 1. Previous Researches

Title	Year	Authors	Industry	Churn Criteria	Method
-------	------	---------	----------	----------------	--------

Customer Churn Prediction in Influencer Commerce : An Application of Decision Trees	2021	Sulim Kim Heeseok Lee	E-Commer ce	Number of Purchases : 1 Purchase is churn, 2 or more is not churn	Decision Trees
Churn prediction in digital game-based learning using data mining techniques: Logistic regression, decision tree, and random forest	2022	Mai Kiguchi Waddah Saeed Imran Medi	Educatio nal Technolo gy	recency > avg_between + (2 * std_between)	Decision Tree, Logistic Regression and Random Forest
Customer churn prediction for web browsers	2022	Xing Wu Pan Li Ming Zhao Ying Liu Rubén González Crespo Enrique Herrera-Viedma	Web Browser		Multivariate Behavior Sequence Transformer (MBST)
Churn in the mobile gaming field: Establishing churn definitions and measuring classification similarities	2021	Ana Perisic Dubravka Sisak Jung Marko Pahor	Gaming	4 methods	Random Forest
Applying the CG-logistic Regression Method to Predict the Customer Churn Problem	2018	Lang Wu Menggang Li	Airline	Churn clients: the number of flights in the second year is less than 0.5 over the number of flights in the first year	Conjugate Gradient Logistic Regression
Combining Sequential and Aggregated Data for Churn Prediction in Casual Freemium Games	2019	Jeppe Theiss Kristensen Paolo Burelli	Gaming	30 days inactive	Neural Networks

2. Literature Review

2.1. Cryptocurrency Exchange

As a digital currency, cryptocurrencies can be traded. Cryptocurrency can be traded on a cryptocurrency exchange platform. Cryptocurrency exchanges can be divided into two based on the transaction mechanism, namely centralized exchanges (CEXs) and decentralized exchanges (DEXs). Many cryptocurrency traders trade on centralized exchanges. Currently there are over 200 centralized exchanges in the world.

In Indonesia there are 25 cryptocurrency exchange that registered in BAPPEBTI. BAPPEBTI requires users to carry out the KYC process before trade crypto in a legal crypto exchange.

2.2. Customer Churn

To get a customer, a company usually carries out a series of promotional activities. After a customer chooses to use a product or service from a company, the company still has to retain the customer to continue to use their product or services. Customer churn occurs when a customer decides to no longer use a company's products or services. In this work, the definition of customer churn and not churn is defined below:

- churn : no transaction and the average balance amount is less then \$1 in the last 4 months
- not churn : minimum one transaction in the last 4 months or the average balance more than \$1 in the last 4 months.

3. Methods

3.1. Preparation

The first step is to apply churn criteria to label users who are churn and who are not churn. From this process it is known that more than 90% of users are not churn. Due to unbalanced data, we apply the undersampling process before the modeling process.

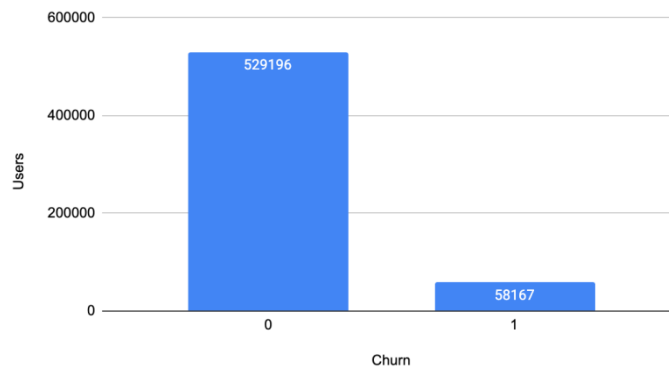


Figure 1. Users by churn

3.2. Modelling

The decision tree and random forests method were used to predict the customer churn model. Apart from being commonly used in classification, both methods have calculations to determine which features have the most importance in classification.

4. Data Collection

The case study is based on data collected from one of the cryptocurrency exchange companies in Indonesia that is officially registered with BAPPEBTI. The data obtained from customer profiles and transaction activity from January - December 2021. The customer profile data consists of country, gender, occupation, age, KYC level, the platform used when registering, the use of google authenticator to secure the account, the referral code and etc. The transaction activity consists of deposit and trade (FIAT or Crypto).

5. Results and Discussion

5.1 Numerical Results

Both methods have almost the same results in terms of accuracy, precision, recall and F1-Score.

Tabel 2. Evaluation Metrics

Method	Accuracy	Precision	Recall	F1-Score
Decision Tree	74.59%	78.76%	72.51%	75.50%

Random Forest	74.34%	77.70%	72.61%	75.07%
----------------------	--------	--------	--------	--------

5.2 Graphical Results

From the calculation of feature importance, it is found that the use of Google Authenticator and trading frequency are important factors that determine whether a customer churns or not.

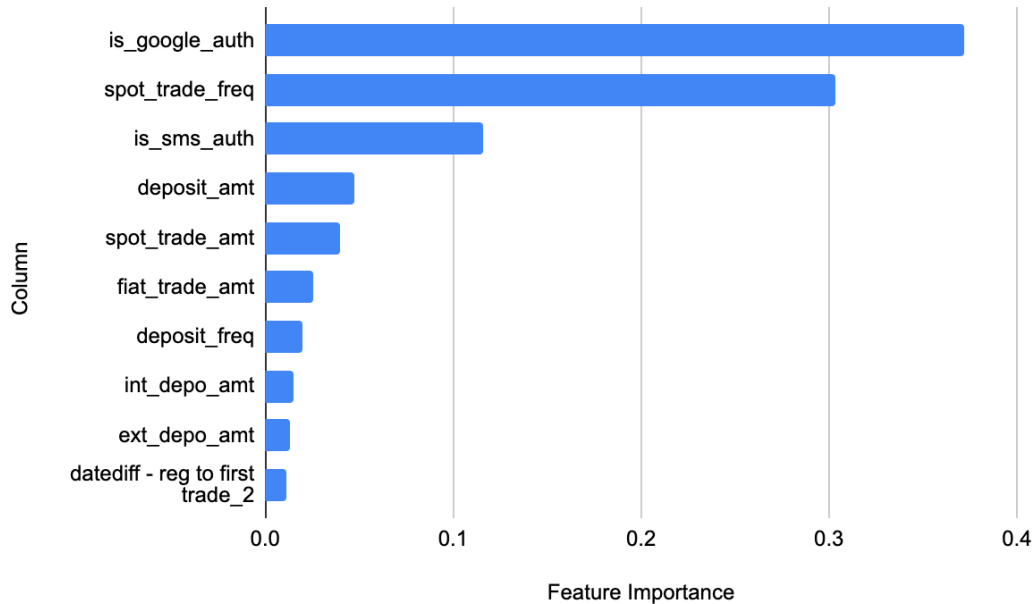


Figure 2. Feature Importance

5.3 Proposed Improvements

There was no common determination of churn and churn prediction in crypto industry because of lack of research. Although in this study, the churn criterion has taken the recency of transactions and the average balance amount but the proposed approach is not suitable for new customer.

6. Conclusion

The competition to get new customers is getting hard, the costs incurred for promotion are also expensive. Retain customers is one of the sustainable methods that can be carried out to continue to generate profits. To increase the effectiveness of retain customer activity, it is important to know the potential churn from customers. Churn analysis become one of method that can be used to predict customer churn and to know the important features that cause customer churn. The result from churn analysis can be used by the marketing team to determine the right strategy to retain customers.

References

- Ding, Y. and Chen, W. Probing the Mystery of Cryptocurrency Exchange: The Case Study Based on Mt.Gox, 2022 International Conference on Service Science (ICSS), Zhuhai, China, May 13-15, 2022.
- Kim, S. and Lee, H., Customer Churn Prediction in Influencer Commerce : An Application of Decision Trees, *The 8th International Conference on Information Technology and Quantitative Management*, pp. 1332-1339, Chengdu, China, July 9-11, 2021.
- Kiguchi, M., Saeed, W. and Medi I., Churn prediction in digital game-based learning using data mining techniques: Logistic regression, decision tree, and random forest, *Applied Soft Computing*, vol. 132, 2022.
- Perisic, A., Jung, D. S. and Pahor, M., Churn in the mobile gaming field: Establishing churn definitions and measuring classification similarities, *Expert Systems with Applications*, vol. 191, 2022.

Wu, L. and Li, M., Applying the CG-logistic Regression Method to Predict the Customer Churn Problem, *2018 5th International Conference on Industrial Economics System and Industrial Security Engineering (IEIS)*, Toronto, Canada, August 3-6, 2018.

Wu, X., Li, P., Zhao, M., Ilu, Y., Crespo, R. G. and Viedma, E. H., Customer churn prediction for web browsers, *Expert Systems with Applications*, vol. 214, 2022.