CHURN MODEL FOR LOAN DISBURSAL

An Internship Report submitted to

DR. VISHWANATH KARAD
MIT WORLD PEACE UNIVERSITY

Submitted by,

Shriya Padhi (1032181174)

Under the supervision of

Mr. Arpan Mondal

And

Prof. Abhishek Chunawale

School of Computer Engineering and Technology

MIT World Peace University, Pune
Kothrud, Pune-411038

(Period from 18.11.2021 to 17.02.2022)

ACKNOWLEDGEMENT

Yabx has achieved such great heights under the leadership of Rajat Dayal. I would like to express my gratitude to Arpan Mondal who gave me a golden opportunity to be a part of the ‘Churn model for loan disbursal’ project at the esteemed organization and helped me in the completion of my project. I thank Barsa Panda and Soumalya Banerjee whose guidance and instructions have been a major contribution towards my project completion. I would like to extend my gratitude to Mohit Thakur to get me on board seamlessly and Puneeth Chandel who ensured my internship progressed smoothly throughout my time at Yabx. I also thank Ajay Kumar and the IT department to help me with timely access of my project related resources. Lastly, I would like to acknowledge the crucial role of MIT World Peace University and Professor Abhishek Chunawale, my university supervisor who guided me and supported me during my internship.

ABSTRACT

Yabx's aim is to enable the underserved with limited credit history to get fair access to financial services tailored to their needs. It has a lot of experience in handset financing due to its credit scoring, customer lifecycle management and E2E service delivery pipelines. Through sub-processes such as feature selection, feature creation, data modeling and interpretation, a loan disbursal mechanism is created using churn model to ensure efficient handset financing. All the customer information, purchase history, SaaS metrics, and previous churn data are used here and turned into a statistical prediction of when certain types of customers might churn in the future. It helps to analyze retention and see the probability of specific customer segments churning more clearly. To help maximize retention, the information can be used to formulate a plan that directly targets each of the company's cohorts and makes prioritizing its actions easy. For the same we perform Exploratory Data Analysis, create models, and evaluate them. Comprehensive customer profiles help the company see what types of customers are canceling their accounts and why they're churning by learning more about the customer pain points causing them to churn.

TABLE OF CONTENTS

CHAPTER I- INTRODUCTION

Yabx's aim is to enable the underserved with limited credit history to get fair access to financial services tailored to their needs [1]. As a result, Yabx creates a credible financial identity for unbanked and underbanked population based on alternate data along with management of the entire customer's credit lifecycle starting from acquisition to settlement. Yabx is the fintech venture of Mahindra Comviva, a global leader in Mobile Financial Services and is a part of USD 20 Billion Mahindra Group.

Smartphone financing makes it easier to pay for smartphone purchases over time rather than paying cash up front [2]. Options for financing a smartphone include promotional offers at retail stores that sell cellphones, financing through cellphone providers, and point-of-sale installment loans.

As a result, Yabx has a lot of experience in handset financing due to its credit scoring, customer lifecycle management and E2E service delivery pipelines. Through sub-processes such as feature selection, feature creation, data modeling and interpretation, a loan disbursal mechanism is created using churn model to ensure efficient handset financing.

CHAPTER II- REVIEW OF LITERATURE

Aji et.al proposed a data mining approach on an Indonesian case study involving mortgage loan instead of conventional statistics like logistic regression and discriminant analysis as those techniques produce a good accuracy but some of the assumptions aren't accomplished by the data [3].

Hassan et.al proposed a scoring approach and categorized loans-based scoring level to be good loans or bad loans [4]. Various data mining techniques have been used to analyze and evaluate these loans to grant them with minimum risk.

Abakarim et.al proposed a Real-Time Binary classification model to deal with loan approval which is based on a deep neural network, and it permits to classify loan applicant as good or bad risk [5]. Experimental results prove that the proposed Real-Time model, based on deep neural network, outperforms typical binary classifiers, in terms of precision recall and accuracy.

Han et. al proposed a multi-objective decision-making model to maximizing bank benefits, minimizing risks, and minimizing customer churn [6]. Statistical data verification and genetic algorithm is also used to solve the credit problem.

CHAPTER III- METHODOLOGY/APPROACH

Building a predictive churn model for a business starts with categorizing everything one knows about their customers into customer profiles as it helps create a more targeted churn model [7]. With this data, one can easily spot patterns in churned customers related to their demographics and segment them into cohorts for more granular analysis. Then, expand on customer profiles by including information about their purchase and billing history. Knowing their overall lifetime value (LTV) helps build a clear picture of how billing processes impact the churn. It is also essential to include a customer's chosen pricing tier in the churn data to see how its pricing decisions affect how customers churn.

One of the most significant contributors to voluntary churn is the customer experience. Tracking every customer's interaction with the company team and its product is essential. Including this information in customer profiles help to see the impact of the product and the customer experience on churn rates. Tracking past interactions can also be valuable for surfacing points along the customer journey where churn is more likely to occur. Customer profiles are the basis for more in-depth churn analysis. With this data, one can start looking for how and why different types of customers leave the company service.

It is essential to segment the churned customers into the two buckets of voluntary and involuntary churn first. The model for voluntary churn will rely on more customer-focused data points than the model for involuntary churn, which will deal mainly with the internal mechanisms of the business.

Building an effective churn model requires a solid grasp of why customers churn at certain times instead of others. The underlying reasons for a 30-day-old customer leaving differ from those for a 90-day or 6-month-old customer. When someone cancels, the company can always send an anonymous exit survey to find out why. Each response is a potential indicator of how the service failed the customer. Seasonality can impact even the most established SaaS business. Budgets are usually reviewed every quarter, so churned customers might result from economic changes in the new year. These questions provide insight into the many reasons a customer might cancel their service, and the answers act as data points the company can use when building the churn model.

The more data the company has at its disposal, the more specific its churn model will be. All data should be reviewed for accuracy and validity before modeling. Mathematical modeling for churn is built on a statistical process called logistic regression, and this process determines the relationships between data points based on a formula and limits the outcome to between 0 and 1.

All the customer information, purchase history, SaaS metrics, and previous churn data are used here and turned into a statistical prediction of when certain types of customers might churn in the future. It helps to analyze retention and see the probability of specific customer segments churning more clearly. To help maximize retention, the information can be used to formulate a plan that directly targets each of the company's cohorts and makes prioritizing its actions easy. Then the company can implement its retention strategy while keeping track of how it impacts its churn rate over the next few months by gathering data to see the business impact and making additional changes in the plan based on its analysis.

CHAPTER IV- ANALYSIS/RESULT

The parameters in the dataset included :

- Count of outflow transactions in complete week.
- Average amount of outflow transactions in complete week
- Count of secondary parties involved in complete week for outflow transactions.
- Count of unique service types involved in outflow transactions in a complete week.
- Count of person-to-person outflow transactions in the complete week.
- Count of cash out outflow transactions in the complete week.
- Count of merchant payments outflow transactions in the complete week.
- Count of recharge outflow transactions in the complete week.
- Count of other outflow transactions in the complete week.
- Total amount of person-to-person outflow transactions in the complete week.
- Total amount of cash out outflow transactions in the complete week.
- Total amount of merchant payments outflow transactions in the complete week.
- Total amount of recharge outflow transactions in the complete week.
- Total amount of other outflow transactions in the complete week.

The same parameters were included in inflow analysis too.

First, we carry out Exploratory Data Analysis on our data for handling missing values. Second, we perform data preprocessing. Categorical features need to be converted to numbers to be included in calculations done by a machine learning model. Then, we perform encoding and continuous variables can also be scaled. When target variables have imbalanced class distribution which is undesirable for machine learning models, up-sampling is used where we increase the number of samples of the class with fewer samples by randomly selecting rows from it. Next, we create the model and evaluate it. Methods like Random Forest, Ridge Classifier are used. Lastly, for improving the model, methods like GridSearchCV can be used for parameter tuning.

CHAPTER V- CONCLUSION

Churn analysis evaluates a company's customer loss rate to reduce it [8]. The company should analyze churn frequently and accurately to keep track of its KPI-oriented goals and engage with customers effectively.

Comprehensive customer profiles help the company see what types of customers are canceling their accounts and why they're churning by learning more about the customer pain points causing them to churn.

To avoid customer churn, the company should ensure that they identify the correct set of customers in the first place. The company can retain customers effectively if its services help its customers achieve their goals. The company should prioritize quality customer support from the beginning and keep an eye on its competitors to focus on what sets it apart. Effective pricing and eliminating delinquent churn also play a huge role in customer retention.

CHAPTER VI- SUGGESTIONS FOR FUTURE WORK

The future scope of this project can be a more intelligent predictive churn analytics model which is fed with both traditional and non-traditional data to achieve holistic customer understanding. Thus, we can combine company data from social media and customer feedback to reduce churn rate.

BIBLIOGRAPHY

[1] https://yabx.co/

[2] https://www.investopedia.com/smartphone-financing-what-you-need-to-know-5190144

[3] N. A. Aji and A. Dhini, "Credit Scoring Through Data Mining Approach: A Case Study of Mortgage Loan In Indonesia," 2019 16th International Conference on Service Systems and Service Management (ICSSSM), 2019, pp. 1-5, doi: 10.1109/ICSSSM.2019.8887731.

[4] A. Hassan and R. Jayousi, "Financial Services Credit Scoring System Using Data Mining," 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), 2020, pp. 1-7, doi: 10.1109/AICT50176.2020.9368572.

[5] Y. Abakarim, M. Lahby and A. Attioui, "Towards An Efficient Real-time Approach To Loan Credit Approval Using Deep Learning," 2018 9th International Symposium on Signal, Image, Video and Communications (ISIVC), 2018, pp. 306-313, doi: 10.1109/ISIVC.2018.8709173.

[6] Y. Han, B. He and J. Zhao, "Credit decision algorithm for SMEs based on data mining," 2020 2nd International Conference on Information Technology and Computer Application (ITCA), 2020, pp. 178-181, doi: 10.1109/ITCA52113.2020.00044.

[7] https://www.profitwell.com/customer-churn/

[8] https://towardsdatascience.com/