Credit Card Default Prediction using various ML Algorithms

MANAJIT MAHAPATRA : 21MDT0108

MUNJ BHAVESH PATEL : 21MDT0116

VISHAL PRASANNA.A : 21MDT0147

OBJECTIVE

- The use of data analytics has allowed banks to mitigate their market imperfections as a result of their myriad challenges. By using analytics, banks are able to control the real-time information generated by their customers in a healthier manner.
- Using predictive and prescriptive analytics, banks are exploring a whole new horizon that was not possible with descriptive analytics in the past. Machine learning techniques are nowadays widely used in the banking sector.
- Our project aims to identify the correlation and prediction power of factors that contribute to default on credit cards and to make recommendations to the banking industry as a result. To reflect a client's credit rating and attributes, we have developed a credit-score model.

ABSTRACT

- In the banking industry, credit risk plays a major role. The main activities of banks are to provide loans, credit cards, investment products, and mortgages. Throughout the past few years, banks' credit card services have been booming. As credit card usage rises, banks have faced an escalating credit card default rate. By tapping into data analytics, we can tackle the current phenomenon and manage credit risk.

INTRODUCTION

- Credit cards have existed for decades, and they are a product of both changed consumer habits and higher earnings.
- There has been a significant increase in the number of card issuers, issuing banks, as well as transaction volumes.
- Even so, the increase in credit-card transactions has led to problems related to overdue amounts and delinquent rates of credit-card loans.
- This issue is critical to the future success of the industry.

LITERATURE REVIEW

- In a report aimed towards predicting credit card defaults among banks, logistic regression, decision trees and random forests were used. It was found that random forest provided the best performance. The accuracy score surpassed 80% and its AUC score was 77%.
- Other papers examining KNN, logistic regression, DA, NB, NNs, and classification trees compare classification and predictive accuracy between them. It was found that there was little variation among the six methods in terms of error rates. In contrast, the lift curve area ratio of the six techniques differs significantly.

FILLING THE GAP

- Banks can use machine learning to assess credit risk of customers before granting them credit card. In order to stay competitive with their competitors, banks must stay innovative and creative in order to offer their clients valuable products and services. Banks can better cater to their customer base using machine learning techniques.
- Previously, the classification performance of algorithms such as logistic regression, KNN, ANN, DA, NB, NNs, and classification trees has been studied.

PROBLEM STATEMENT

- In this work, six different ML methods will be compared for the prediction of default probability among customers' default payments in Taiwan.
- As part of the current study, we will apply algorithms such as XgRBF, AdaBoost, LGBM Classifier, etc. and use metrics such as ROC-AUC curves and cross validation to compare the results obtained from the models with their results in previous studies.

REFERENCES

- DOI : 10.1109/ICACCAF.2018.8776802
- DOI : https://doi.org/10.1016/j.eswa.2007.12.020
- DATASET : https://www.kaggle.com/enpingzhao/credit-card-default-prediction/data