Project proposal

Project title

Credit Card Defaulters Detection using Machine Learning

Team members

- Harshini Kurupati (A20480932) – hkurupati@hawk.iit.edu
- Vismaya Mohanram (A20519405) – vm@hawk.iit.edu
- Yasha Srinivas (A20516915) – ysrinivasa@hawk.iit.edu

Description of the problem

People and financial institutions increasingly rely on internet services to complete transactions, which has resulted in an exponential surge in credit card usage. In present day, the COVID-19 pandemic has caused lockdowns, business closures, and widespread job losses due to which many people have found themselves struggling to make ends meet. This in turn has led to a significant increase in credit card defaulters as people struggle to pay their bills on time. The development of an effective Credit Card Defaulters Detection system is required in order to minimize the damage sustained by customers and financial institutions.

A brief survey of what have been done and how the proposed work is different

Numerous papers used datasets that lacked crucial features needed to properly predict the target variable. Additionally, several of the algorithms used did not generate sufficiently accurate results for the datasets. Reassessing these with the most accurate algorithms based on the essential features may enhance prediction accuracy.

Preliminary plan

We will be working on the below steps of implementation for the project to produce much better outcomes:

- Data Preparation and Mining
  Checking all the dataset for null or missing values, checking correlation between variables, plotting a heatmap. Using a pairplot to understand the various relations.
- Feature Engineering
  Manipulation by adding, deleting and combing data.
- Modelling and Prediction

Using the below Machine Learning algorithms we perform the predictions:

- Logistic regression
- Decision tree
- Support Vector Machine
- K-Nearest Neighbors Algorithm
- AdaBoost algorithm
- Random forest

Implementation of Proposed_model which will compare all the models by combining the classification results of test dataset of each model by majority voting.

Our final goal is to build a model with the best accuracy to predict the credit card defaulters.

Dataset

https://raw.githubusercontent.com/MLWave/Black-Boxxy/master/credit-card-default.csv

References

[1] S. S. H. Padmanabhuni, A. S. Kandukuri, D. Prusti and S. K. Rath, "Detecting Default Payment Fraud in Credit Cards," *2019 IEEE International Conference on Intelligent Systems and Green Technology (ICISGT)*, Visakhapatnam, India, 2019, pp. 15-153, doi: 10.1109/ICISGT44072.2019.00018.

[2] Y. Sayjadah, I. A. T. Hashem, F. Alotaibi and K. A. Kasmiran, "Credit Card Default Prediction using Machine Learning Techniques," 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), Subang Jaya, Malaysia, 2018, pp. 1-4, doi: 10.1109/ICACCAF.2018.8776802.

[3] Ying Chen, Ruirui Zhang, "Research on Credit Card Default Prediction Based on *k*-Means SMOTE and BP Neural Network", *Complexity*, vol. 2021, Article ID 6618841, 13 pages, 2021. https://doi.org/10.1155/2021/6618841

[4] M. Zan, G. Yanrong, and F. Guanlong, "Credit card fraud classification based on GAN-AdaBoost-DT imbalance classification algorithm," *Journal of Computer Applications*, vol. 39, no. 2, pp. 314–318, 2019.

[5] R. Mei, Y. Xu, and G. Wang, "Study on analysis and influence factors of credit card default prediction model," *Statistics and Applications*, vol. 5, no. 3, pp. 263–275, 2016.

[6] Akshay A. Bardiya, "Credit Card Defaulters Using Machine Learning", Volume 10, Issue V, May 2022. DOI Link: https://doi.org/10.22214/ijraset.2022.43255