Credit Score Classification Using Machine Learning Algorithms

Aman Kumar aman20279@iiitd.ac.in 2020279
Pritish Poswal pritish20321@iiitd.ac.in 2020321
Vibhu Jain vibhu20151@iiitd.ac.in 2020151
Karan Prasad Gupta karan20439@iiitd.ac.in 2020439

1. Motivation

We wanted to select something which had significance in Today's world at the same time within the scope of our course. We also had to keep this fact that the data should have sufficient samples.

Credit cards have become an integral part of our lives and a huge fraction of young and medium age people use it. The usage of credit cards has increased over the years and with the emergence of companies like Cred there is more incentive for people to use credit cards.

Classification of credit score is important because credit score acts as a feedback to validate the users, thus a good credit score can be very beneficial as it helps the user to get more favourable loans, credit cards and more.

Credit score classification is a complex problem as it depends upon many parameters and factors, we aim to solve this problem using ML. This is the reason we chose credit score classification as our project.

2. Related Work

1. This Research study uses Decision Tree Classifier and Deep neural networks Techniques which predicts the Credit Score , which is useful for Financial Institute and helping the lender for financial Losses. [3]

2. This Project explores the Machine Learning approaches that are used in Credit Card Scoring which result in Classification of it . They applied Cost Sensitive Learning to the algorithm since the dataset was imbalanced . [2]

3. This Tutorial helps to develop a model for the imbalanced German Credit Score Classification. [1]

3. Timeline

[Week 1] - Initial Analysing and selection of Dataset and planning/discussing about the future use/trimming of the dataset according to real world application.

[Week 2] - Preprocessing the data and visualising the available data using various kinds of graphing techniques

[Week 3-4] - Extracting and selecting useful and relevant features and finding correlation present in the available data

[Week 5-7] - Applying various possible appropriate ML models and techniques to the data through code like Logistic Regression, SVM, Random Forest, Naive .etc

[Week 8-9] - Analysing the performance of each model for accuracy through various parameters

[Week 10] - Further Optimization of the model with best results according to various parameters and finally deploying the model for the project.

[Week 11] - Documentation of the work along with creating PPTs and doing final fixes if any.

4. Individual Tasks

Aman : Dataset selection, Preprocessing and visualisation of data, Applying various ML models, Performance Analysis and testing Models .

Karan : Documentation, Data Extraction and finding correlation and using K-Fold, Finding possible ML models for the data.

Pritish : Initial analysing of the dataset, EDA , Extracting useful features in data , applying multiple ML models to the data.

Vibhu: Dataset selection,Extracting useful features in data, Testing accuracy of all the applied ML models and evaluating their performance, Further Optimization and final deployment.

5. Final Outcome

To develop a ML model with that classifies the credit score of person as Good, Standard, Poor with the help of available creditinal information in the dataset such that the result is highly accurate. At the end after analyzing such a large set of data we would be able to understand and comprehend this problem more closely from the perspective of machine learning. The current classification is more algorithmic and doesn't relies much on the data. The work done can help in quick classification of the real world problem if the machine learning algorithm predicts poor credit score then credit/loans to such individuals can be directly rejected by financial institutions and vice versa.

References

[1] Jason Brownlee. Develop a Model for the Imbalanced Classification of Good and Bad Credit, Feb. 2020.

[2] Bornvalue Chitambira, Christopher Engstrom, and Va Ster. Credit Scoring using Machine Learning Approaches. page 62.

[3] Ashwani Kumar, D. Shanthi, and Pronaya Bhattacharya. Credit Score Prediction System using Deep Learning and K-Means Algorithms. *Journal of Physics: Conference Series*, 1998(1):012027, Aug. 2021.