BUAN 6341.003 Applied Machine Learning

Group 7

Project Proposal

Credit Score Classification

Group Member

Abiodun Bello

Ying-Tzu Chen

Mrudula Jethe Bhanushali

Shruthi Ramesh

Arwin Kumar Ravi

Rohith Thiagarajan

Business Problem

Credit scores are crucial for companies for several reasons like:
Risk Assessment, Loan Approval, Pricing Strategies, Fraud Prevention, Market Analysis, Regulatory Compliance.
Overall, credit scores are integral to managing financial risk and optimizing business operations. Thus, it is extremely important for a company to attain credit scores for various customers with the at most accuracy. This is exactly what we are trying to address by building this particular prediction model.

Type of Problem

The dataset we possess contains labeled data that specifies the Credit score of the customer along with all other variables. So, in this problem, we will be using Supervised Learning algorithms specifically deploying Classification Algorithms for predicting the 'Credit Score' variable as Good, Standard or Poor.

Dataset

- This particular dataset was acquired from Kaggle. The link to the same is: https://www.kaggle.com/datasets/ayushsharma0812/dataset-for-credit-score-classification/data?select=credit_score.csv

- The dataset has 100,000 records providing the customer data relating to various financial prospects.

- Attributes: No. of Attributes = 28
  → Credit_Score: Whether customer has a Good, Standard or Poor credit score. (Target Variable)
  → Other Dependent Variables: Month, Age, Occupation, Annual_Income, Monthly_Inhand_Salary, Num_Bank_Accounts, Num_Credit_Card, Interest_Rate, Num_of_Loan, Delay_from_Due_Date, Num_of_Delayed_Payment, etc.

- Number of Categorical Features =12
  Number of Numerical Features = 16

Method

In this problem, we need to predict the 'Credit_Score' variable as Good, Standard or Poor. Using the labeled Customer Churn dataset, we plan to deploy the following classification algorithms so we can choose one with the best performance: KNN, Decision Tree, SVM, EDA.