

Credit Card Defaulter Analysis

ISOM3360 Group 23: Project Progress Report

LAM, Ho Chit
SID: 20607878
hclamao@connect.ust.hk

LEE, Ho Wan Owen
SID: 20604852
hwlee@connect.ust.hk

LEE, Wai Chung
SID: 20702733
wcleeeaj@connect.ust.hk

1. Current Progress

Quick recap of the goals of this project:

- Help credit card issuers identify credit card defaulters
- Take extra precautions on people that are risky of delayed credit repayment

As per the idea report, we have selected a proper dataset and determined the business objective of this group project. The steps we have taken since then mostly revolve around preparation before implementing and evaluating the models themselves. The details are described in the following subsections:

1.1 Data Pre-processing

After preliminary inspection of the dataset, we have selected a few possible features that would be useful to our model training, including family members count, occupation type, annuity, goods price, etc. One-hot encoding is applied to discrete value columns such as occupation, and normalization techniques are used for certain continuous data types. Missing values are either filled by the population mean or mode. Columns of significant unique value counts are dropped.

1.2 Model Selection

Based on the nature of our business objective, our group has selected 4 preliminary classification models to implement: Decision Tree, Logistic Regression, Naïve Bayes, and k-Nearest-Neighbours. The use of Multilayer Perceptron Classifier is also discussed, but its implementation is contingent on our progress. Performance of the models will be further evaluated once we reach the implementation stage.

2 Next Steps

The following details a reference timetable for our future actions:

- Week 8:
 - Apply finishing touches to data pre-processing flow
 - Cherry pick useful features and drop irrelevant ones
 - Complete prediction result analysis function
 - Begin implementing Decision Tree Classifier
 - Complete project progress report
- Week 9:
 - Implement Logistic Regression Classifier
 - Evaluate Decision Tree and Logistic Regression models
 - Complete model evaluation function
- Week 10:
 - Implement Naïve Bayes Classifier and k-Nearest-Neighbours Classifier
 - Evaluate Naïve Bayes and k-Nearest-Neighbours models
- Week 11:
 - Fine-tune models and compare the performance of the models. Overfitting inspection.
- Week 12:
 - Complete Project Final Report