1. Group 3

1) Summary of the report

This report aims to build a model that accurately predicts an applicant’s

repayment abilities by utilizing various alternative data. The evaluation

criterion is mainly AUC score.

2) Strengths of the report

a) The report is well organized. Firstly, it introduces the datasets and

does the Exploratory Data Analysis. Then it does data processing and

feature engineering. Finally, it describes the model construction and

evaluation.

b) The figures in report is very clear and technical.

c) The analysis and description is very well.

d) It uses statistical knowledge to judge the importance of each

categorical feature and uses correlation heat map to analysis

numerical features.

3) Weakness of the report

a) There are some typos in the report, such as “20.811937”in Table

2.

b) In the Feature Encoding Part, it is not necessary for LightGBM model

to do one-hot encoding for categorical features, since LightGBM can

directly handle this type of features. One-hot methods will increase

the feature dimension and maybe reduce the performance of model.

4) Evaluation on Clarity and Quality of Writing (1-5): 5

5) Evaluation on Technical Quality (1-5): 4

6) Overall rating (1-5): 4

7) Confidence on my assessment (1-3): 3