

Summary of the report:

This group used application {train / test}.csv and 3 extra datasets for training and prediction. The group has clearly described the data aggregation method, handled missing values and removed outliers. Feature engineering and feature selection have been done to add more diversity and select the most important features respectively. Generalized Linear Models and Tree Models have been experimented. LightGBM yielded best results.

Describe the strengths of the report:

Extensive data preprocessing, feature engineering and feature selection have been applied. Multiple models have been tested. Good visualization on feature importance.

Describe the weaknesses of the report:

The group did not visualize the AUC scores of the tested models. No hyperparameters tuning on LightGBM is mentioned and experimented.

Evaluation on Clarity and quality of writing (1-5):

5

Evaluation on Technical Quality (1-5):

4

Overall rating (1-5):

4.5

Confidence on your assessment (1-3):

3