Summary of the report

The team analyzed the Home Credit Default Risk dataset to predict the repayment abilities of clients. By conducting the feature engineering, attempting different machine learning models, and tuning parameters, they got AUC score of 0.748 from Kaggle. They found that compared to other machine learning methods, Light Gradient Boosting Machine performed best for not only in-sample data set, but also out-sample prediction.

Describe the strengths of the report

1. They mined as many new features as possible and tried different algorithms as well as combination of parameters to look for the best model.
2. The team utilized figures to support the report.
3. The report follows NIPS conference style and is well organized; besides, there are no grammar and typo issues.

Describe the weaknesses of the report

The score from Kaggle is not the top.

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

Is the report clearly written? 5

Is there a good use of examples and figures? 5

Is it well organized? 5

Are there problems with style and grammar? 5

Are there issues with typos, formatting, references, etc.? 5

Please make suggestions to improve the clarity of the paper and provide details of typos: Change the background of the screenshot to light.

Evaluation on Technical Quality (1-5):

Are the results technically sound? 5

Are there obvious flaws in the reasoning? 5

Are claims well-supported by theoretical analysis or experimental results? 5

Are the experiments well thought out and convincing? 5

Will it be possible for other researchers to replicate these results? 5

Is the evaluation appropriate? 5

Did the authors clearly assess both the strengths and weaknesses of their approach? 2

Are relevant papers cited, discussed, and compared to the presented work? 2

Overall rating: (5- My vote as the best-report. 4- A good report. 3- An average one. 2- below

average. 1- a poorly written one): 4

Confidence on your assessment (1-3): 3