Group 10

* Summary

Group 10 also choose the topic Home Credit Default. They compared four method KNN, LightGBM, Xgboost, RondomForest and come with the highest score of 77%. The report contains both EDA, methods and reflection of this project.

* Strengths

Compared different methods with explicit experiment results written in a table.

High accuracy.

Code is clear with notes.

Report with a lot of reflections.

* Weakness

Some questions:

Report written that used coefficient to subtract the feature dimensions, it seems not in the code. And it seems not to talk about the numerical feature and categorical feature separately.

Besides, during the “pre-processing” part in the code, it seems there is no reason for creating some feature.

To conclude, the feature extraction is a bit vague.

Maybe you can add some explanations on the results like auc.

Maybe you can add more visualization.

* Evaluation

|  |  |
| --- | --- |
| Evaluation Direction | Score |
| Clarity and quality of writing | 4 |
| Technical quality | 4 |
| Overall Rating | 4 |
| Confidence | 2.5 |

Explicit Reason:

Report written is very clear with a lot of real reflections. They showed the result of different models in table which is clear and showed the workload they have done. But I thought there might be some problems: how to choose tables, feature extraction is vague, analysis the result, visualization.