Group1

• Summary of the report.

The report first introduces the background and dataset, and then introduces the principle of the

algorithm LGBM used. In the specific method and experiment parts, the report first analyzes the

data and selects the features that may be helpful for prediction. In terms of model, on the basis

of LGBM, the problems of memory capacity reduction and data set preprocessing are considered,

and a lag feature quantity and a rolling mean feature are added to further improve the

performance of the model. Finally, the model in the report get 0.32097 public score and 4.07945

private score. In the face of such achievements, the report also puts forward the follow-up

improvement direction.

• Describe the strengths of the report.

I think the report has the advantages of neat format and smooth writing. It clearly shows the

working process from the background, theory to the specific experimental process, and then to

the final result. Secondly, before the formal start of training model and prediction, the report

analyzes and studies the data, which is helpful for the subsequent model construction. Finally, in

terms of model design, LGBM model is selected and added a lag feature quantity and a rolling

mean feature in the report to make the model more optimized.

• Describe the weaknesses of the report.

I think there is some confusion in the layout of the report. Sometimes when the text involves a

picture or a table, I can't find it quickly. In addition, I think a abstract part can be added to the

report, which can help readers master the general content of the report as soon as possible.

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

3

I think report do well on clarity and quality of writing on the whole. But there is some confusion

in the layout of the report as well. Sometimes when the text involves a picture or a table, I can't

find it quickly.

• Evaluation on Technical Quality (1-5):

4

I think the report do well on technical quality. The model is designed well and the public score is

good. Exploratory Data Analysis, a lag feature quantity and a rolling mean feature all proved the

capability of the model. Our group also used LGBM model. The report on the optimization of

LGBM model gives us ideas for future work.

• 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

I think the report is a good one.

• Confidence on your assessment (1-3) (3- I have carefully read the paper and checked the

results, 2- I just browse the paper without checking the details, 1- My assessment can be

wrong)

2

I read the report, but I don’t check the detail of their codes carefully.