

Peer Review

1. The summary of the report

Classmate Chen Liu used both the main dataset `application_train|test.csv` and the supplementary dataset `bureau.csv` to complete his project. And the method he chose to use is the method Gradient Boosting Decision Tree, which is a machine learning technique for regression and classification. His project came out with 2 separate outputs, one with the supplementary dataset, while the other one without. The public scores are 0.75221 and 0.74264 respectively, which are similar to our team's final score. From my point of view, the score is quite acceptable, and his report is clear.

2. Describe the strengths of the report

- Neat looking
- Clear problem-solving logic
- Reliable machine learning technique

3. Describe the weakness of the report

- As for the Methods part, I suggest Chen Liu could write more specific. Maybe it is better to briefly introduce the definition and structure for the machine learning technique he used in this project.
- As for the Experiments part, I suggest Chen Liu adding some clear figures with a detailed comparison about his two outcomes, with supplementary and without supplementary data.
- Maybe he could tell us his guess about why the public score for the 2 outcomes are quite similar? Does the supplementary data unnecessary?

4. Evaluation on Clarity and Quality of Writing:

3.5. This report is clearly written, well organized, and references cited at the end. It would be better if more figures are added.

5. Evaluation on Technical Quality:

4.0. Both the public scores and techniques are sound and acceptable. It would be better if more details were written, like the weaknesses and strengths of his method.

6. Overall Rating:

4.0. Overall, I think it is a good report.

7. Confidence on your assessment: 3

I have carefully read the whole report.