Peer Review of the 10th group

1. Summary of the report
   The main analysis framework of the report can be divided into three parts. The first part is variable selection, the second part is logistic regression, and the third part is to use the cubic polynomial feature method to improve the effect of logistic regression. In the comparison of results, weighted average F1-score and Kaggle private score are used to select the appropriate threshold and judge the prediction accuracy of the method used and the improvement of the prediction accuracy.
   The conclusion of the report is that the introduction of cubic polynomial feature appreciably improves model performance in weighted average F1-score, accuracy and the area beneath the ROC curve. But when it comes to Features Analysis, it's hard for them to explain the relationship with economics meanings since that they don't know what external source exactly is.

2. Describe the strengths of the report.
   1. Use the correlation between the variable and TARGET to select the variable for the original logistic regression model and the logistic regression model that introduced cubic polynomial feature
   2. Use cubic polynomial feature to improve the prediction effect of the logistic regression model
   3. It clearly shows the threshold selection process of The Confusion Matrix with the threshold change.
   4. The content of the overall writing is concise and logical

3. Describe the weaknesses of the report
   1. The position of the chart is a bit messy, it should be either all put at the end of the report, or it should be properly interspersed with the text in combination with the content of the article. Some diagrams placed in the text are incorrectly selected, which brings some inconvenience to readers.
   2. When selecting variables, this report directly selects the four most correlated variables. Perhaps it can show the logistic regression results from more to less related variables, so as to determine a more reasonable choice of the number of correlated variables.

4. Evaluation on Clarity and quality of writing (1-5): 4
   The content is concise and clear, but the position of the chart needs to be reformatted.

5. Evaluation on Technical Quality (1-5): 5
   The analysis process and results are fully supported by data, and the results can be replicated. The author also clearly assessed their strengths and weaknesses and gave ideas for future improvements. Perhaps the author can add some citations to related documents.

6. Overall rating: 4- A good report.
   I think the overall logic of the article is clear and accurate, and it also gives ideas for future improvements. There is a slight flaw in the details of some data presentation and the layout of the chart.

7. Confidence on your assessment (1-3): 3- I have carefully read the paper and checked the results.
   I read the article carefully and can understand the overall analysis logic.