

Group 6 **Tse Justin Chung Heng**

The topic of his project is improving classification performance on the Titanic dataset. Titanic competition is the starter project on Kaggle, and he tries to extend the work from project 1, improving the performance with different techniques including discretization, various models, and hyper-parameters tuning.

Strength: Adequate number of graphics are included in the presentation and report. The problem statement is clear and easy to follow.

Weakness: Presentation can be more organized and fluency can be improved. Although the problem statement is clear, the complexity of the project can be enhanced. The feature discretization should have been considered in project 1, as it is a basic step in data processing.

Quality of writing: 1

The poster is somewhat messy and hard to follow. The only graphic presented is the decision boundary of each model, but the model names in the graphics are very small, and it does not have an intuitive comparison of the accuracy between each model. More graphics can be included to enhance the readability of the report. Moreover, he just listed several models he used, without the reasons, assumptions and parameters of each model.

Presentation: 1

The fluency and organization of the presentation can be enhanced. The presentation is hard to follow and makes readers confused about the techniques used in the project. For example, in P.8 of the presentation, he presented 14 graphics in one slide, with just one sentence explaining the effect of binning towards linear and non-linear classifiers. These graphics can be divided into multiple slides, and some graphics units can be used to enhance the readability, for example, adding circles or arrows to highlight the important parts in the presentation.

Creativity: 1

Since this is a starter project in machine learning, many techniques have been covered in others' works. The writer just included several data preprocessing techniques, and common machine learning models. For hyperparameters tuning, it should be included in the first project and more variety of the machine learning techniques can be demonstrated, like ensembling methods or state-of-art deep learning models.

Confidence on your assessment: 3

I have carefully read the paper and checked the results