Name : Yang, Po-Yen
Student ID : 20561878

Group 6

1. Summary of the report
   Tse Justin Chung Heng is working on the Titanic dataset with a comparison between tree-based model and linear model on this dataset. He used the family size instead of isAlone, which he hopes to incorporate more information into the dataset, also, he used one-hot encoding for nominal data, and extracted more title from the Title variable. He also plotted some graphs for visualization. The methods that he compared are tree-based model vs linear model, where there are six tree-based models including Decision Tree, Decision Tree with Bagging, Randomized Decision Trees (extra-trees), AdaBoost, Gradient Boosting Classifier, and Random Forest. On the other hand, there are two linear models, which are Logistic Regression with L1 regularization and Logistic Regression with L2 regularization. He also compared decision tree vs ensemble method. Lastly, he gave a conclusion and put his references at the end.

2. Strengths of the report
   (1) He performed different methods for the model comparison.
   (2) He drew lots of graphs for the decision boundary of each model.
   (3) He gave detailed test accuracy results for the comparison of tree-based model and linear model.

3. Weaknesses of the report
   (1) He only provided a brief conclusion but didn't give some analysis or which part can be improved in the future.
   (2) The comparison table is too small, and I can't see it clearly.
   (3) Though there are many decision boundaries, each graph is a bit too small to see clearly.

4. Evaluation on quality of writing (3)
   I think the report is clearly written as it did express his thoughts clearly, and there were figures for visualization. But there were some typos and grammar mistakes, for example in the data preprocessing and feature engineering part, incooperate should be incorporate, in the introduction part, tree-based model should use has instead of having, and in part four, it should be one of the reasons instead of one of the reason.
   I think the parts that this report can improve are as follows :
   (1) He could add the parameters that he used for some or all of the models.
   (2) He could add some analysis or how it can be improved in the future at the very last to make the report more complete.
   (3) The graphs and tables for the models can be larger so that the readers can see them clearly.

5. Evaluation on presentation (4)

The presentation was clear and the language flow was fine. The slides were also clear and well prepared as he circled the important parts that he want to highlight during the presentation. Also, he put lots of graphs and screenshots and used a large word font which is great for visualization. However, there are some points that I think he can improve, which are as follows :

(1) He didn't say much about the reasons for setting the parameters for the models
(2) Sometimes it's a bit hard to understand the point that he was trying to point out

6. Evaluation on creativity (3)
   I don't think the work proposes any new ideas, as I believe these methods have been proposed for the Titanic dataset before. However, I do think that comparisons of all these models should be new, and some of the tree-based models can be counted as state-of-the-art results, which I think is an extension of existing ideas.

7. Confidence on my assessment (3)
   I have carefully read the paper and checked the results.