Group 1

• Summary of the report.

Group 1's report shows their approaches to the Titanic's problem of investigating what kinds of passengers are more likely to survive in this tragic accident. Specifically, they use Logistics Regression, Decision tree and Random Forest to make predictions based on attributes like age, gender, and some other information, and provide some analyses on their performance with R^2 statistics. Their report begins by briefly introducing about the dataset, how they pre-processed it, and providing some exploratory analysis. Subsequently, some diagnostics for their model are vaguely illustrated and explained before discussing about models' selection and comparison. Finally, the report ends with discussion about differences between models' performance and ways to improve them together with a clear conclusion about what kinds of passengers are more likely to survive.

• Describe the strengths of the report.

There are some positive sparks in Group 1's report that I want to commend:

- They describe the dataset and how they perform data cleansing and pre-processing for it in details. They also have some meaningful observations when exploring the dataset.
- In terms of model diagnostics, they reasonably did multicollinearity check for their variables before using them to fit their model.
- In the end of their report, they make some rational conclusions and mention about ways to improve their work even though it's short. Besides, they provide a clear answer for the topic question of what kinds of passengers are more likely to survive.

• Describe the weaknesses of the report.

There are some issues about the report that I want to point out:

- The report does not explain why they selected Logistics Regression, Decision tree and Random Forest as their model.
- The way they organize their report is somewhat hard to follow. For example, it does not have a clear format and illustration such as figures and tables are placed inappropriately without caption.

• Evaluation on quality of writing (1-5): Is the report clearly written? Is there a good use of examples and figures? Is it well organized? Are there problems with style and grammar? Are there issues with typos, formatting, references, etc.? Please make suggestions to improve the clarity of the paper, and provide details of typos.

4. Evaluation is as follow and Group 1 can improve based on this.

- The report is somewhat clearly written.
- They did not really make good use of illustration. Figures and tables are placed inappropriately without caption. Especially, in the result table, they also mentioned the accuracy of KNN and SVM which are not their used models and explain nothing why they did it.
- It is quite clumsy in terms of organization.
- No problems with style and grammar I suppose

• Evaluation on presentation (1-5): Is the presentation clear and well organized? Are the language flow fluent and persuasive? Are the slides clear and well elaborated? Please make suggestions to improve the presentation.

4. Evaluation is as follow and Group 1 can improve based on this

- The presentation is clear, somewhat organized, and the slides are aligned with the report.
- CAI presented not fluently, making his part quite hard to follow. There is imbalance in terms of time used between CAI and SONG's part

• Evaluation on creativity (1-5): Does the work propose any genuinely new ideas? Is this a work that you are eager to read and cite? Does it contain some state-of-the-art results? As a reviewer you should try to assess whether the ideas are truly new and creative. Novel combinations, adaptations or extensions of existing ideas are also valuable.

3. Evaluation is as follow

- Their work doesn't propose any novel idea

• 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 just browse the paper without checking the details