

2. Group 2: “Pawpularity” Prediction Using Machine Learning with Tabular Metadata and Images (Chow Hau Cheung Japser)

a. Summary of the report

The topic of this group is about the Pawpularity Prediction. Pawpularity is a composite metric designed by Petfinder.my organization, which measures the popularity of pet photos, estimated using traffic data across the web. He included exploratory data analysis of the data, experimented with simple models, and built complicated models.

b. Describe the strengths of the report

This group introduced the machine learning project in an illustrative way, with informative presentation slides and clear presentation. The flow of presentation is easy to follow, from the introduction of the topic, and exploratory data analysis and data modelling.

c. Describe the weaknesses of the report

Instead of including the numbers in a table, perhaps it is a better idea to turn it into a more illustrative visuals like a bar chart.

d. Evaluation on quality of writing (1-5): 3

The report and presentation slides are easy to follow, with illustrative graphics, and logical thinking skills demonstrated during the whole presentation. The speaker starts by introducing the intuition and baseline of each model, then gradually includes several improvements to the model.

e. Evaluation on presentation (1-5): 3

The presentation style is clear and organized. Numerous graphics and codes are attached so that readers understand the reasons for adopting different models, e.g. CNN, and the effects on the final result. Main points are highlighted in the presentation slides, which makes readers easy to follow the major ideas in the presentation.

f. Evaluation on creativity (1-5): 2

The models adopted in this project are common in the machine learning community and are already developed by other Kagglers in the competition. This group is expected to extend the models to a deeper extent, including CNN ensembling at the beginning for example, followed by adopting self-attention models to examine the possibilities of these state-of-art deep learning models.

g. Confidence on your assessment (1-3): 3

I have carefully read the paper and checked the results