

Group 3 **Ho Chuen Ho, Chun Lok Him Brian**

The topic of this group is about the Limitations of Translation: How much does translation affect the analysis of Chinese text in different models. They include sentiment analysis, comparison between Chinese Text and translated text, and customized models. The dataset is `weibo_senti_100k.csv`, from a microblogging website.

Strength: They introduce the topic of translation in an illustrative way, with interactive presentation slides and clear presentation. The flow of presentation is easy to follow, from the introduction of the topic, and data preprocessing and data loading,

Weakness: Instead of including the confusion matrix or other scores in plain code, turning it into a more illustrative graphics can enhance the readability of the presentation.

Quality of writing: 5

Their report and presentation slides are clear, easy to follow, with illustrative graphics, and logical thinking skills demonstrated during the whole presentation. Before introducing each model, they included the underlying assumptions, so that the reader can know about what assumptions they made.

Presentation: 5

Their presentation style is clear and organized. Numerous graphics and codes are attached so that readers understand the reasons for specific preprocessing techniques, e.g. Laplace smoothing, and the effects to the final result. They made use of the blackboard and drew my attention to the modelling details of Contextual Word Embedding.

Creativity: 5

The problem statement of their group is niche, and they adopted different natural language processing models including Naive Bayes Classifier, K-Nearest Neighbour Classifier, and BERT models. They did not only apply the models directly, but changed the model architecture and details based on their understanding of the scenario cases.

Confidence on your assessment:3

I have carefully read the paper and checked the results