

Name : Yang, Po-Yen
Student ID : 20561878

Group 3

1. Summary of the report

Ho Chuen Ho and Chun Lok Him Brian are working on a Natural Language Processing (NLP) project, entitled "Limitations of Translation: How much translation affect the analysis of Chinese text in different models?" They collected their data from weibo, and it is quite a huge dataset as it contains 100k data. The data preprocessing they performed on the data are as follows, first, they created machine translation on the data as a comparison using Google translate, then they utilized frequency-inverse document frequency (TF-IDF) to reflect how important a word is in the dataset. The methods that they used are the self-created Naive Bayes model, self-created KNN model, and BERT model. For the self-created Naive Bayes model, the token is used instead of a character for English text, and the Chinese test set accuracy is 0.7941, Chinese train set accuracy is 0.7913, English test set accuracy is 0.8692, and the English train set accuracy is 0.8294. This is the only model with prediction results better for English text compared to that of Chinese text. As for the self-created KNN model, they have tried different norms. For the first KNN model, the Chinese test set accuracy is 0.8624, Chinese train set accuracy is 0.8604, English test set accuracy is 0.6523, and the English train set accuracy is 0.6497. For the second KNN model, the Chinese test set accuracy is 0.7462, Chinese train set accuracy is 0.7426, English test set accuracy is 0.5271, and the English train set accuracy is 0.5272. Lastly, for the BERT model, they compared three types of variant models, which are full, reduced, and emoji. For the full BERT model, the Chinese test set accuracy is 0.829, Chinese train set accuracy is 0.943, English test set accuracy is 0.852, and the English train set accuracy is 0.918. For the reduced BERT model, the Chinese test set accuracy is 0.836, Chinese train set accuracy is 0.946, English test set accuracy is 0.966, and the English train set accuracy is 0.974. For the emoji BERT model, the Chinese test set accuracy is 0.980, Chinese train set accuracy is 0.983, English test set accuracy is 0.971, and the English train set accuracy is 0.981. They then gave their conclusion, contributions, and references at the end.

2. Strengths of the report

- (1) They performed different methods for their prediction, also, they have created their own models, which is quite impressive.
- (2) They performed predictions on both Chinese and English training dataset and test dataset.
- (3) They drew some graphs, tables, and put definitions for the models which are great for visualization and understanding.

3. Weaknesses of the report

- (1) They only provide a short conclusion but didn't give some analysis or which part can be improved in the future.

- (2) The color for the table is the same as the background, although they have highlighted the highest accuracy for a column, it is sometimes a little hard to read at first sight.

4. Evaluation on quality of writing (4)

I think the report is clearly written as it did express their thoughts clearly, and there were figures for visualization and there were no typos. I think the parts that this report can improve are as follows :

- (1) For the reduced BERT model, the English text has higher accuracy, however, they only mentioned that their self-created Naive Bayes model resulted in a higher English text accuracy.
- (2) They could add some analysis or how it can be improved in the future at the very last to make the report more complete.

5. Evaluation on presentation (5)

The presentation was clear and well organized, and the language flow was fluent. The slides were also clear and well prepared as he used different colors for highlight points and also screenshots of the code in corresponding slides. Additionally, they provided clear details of the models they used, and they demonstrated their method and models on the whiteboard which made it clearer for the audiences to understand. However, there are some points that I think he can improve, which are as follows :

- (1) Some of the slides have too many background colors, and it might be hard to focus on the words which they want to present.
- (2) There are too many words in some of the slides, and when the word is black and the background color is red, it is hard to see clearly all the words.

6. Evaluation on creativity (5)

I think the work did propose new ideas, as they have created their own models, and NLP is one of the current hot topics for research recently. I do think that BERT model can be counted as one of the state-of-the-art results, additionally, the accuracy for the BERT model and their first self-created KNN model is relatively good and I may want to utilize the parameters of these two modes, 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.