

Rebuttal

Review1:

More details regarding the steps of processing the text data are expected. Also, more NLP models could be applied as mentioned by the group, these two methods are not much different but the location of "mask". It would be great to see a variety of models giving varying results.

We could only apply countable models due to the limit of running capacity of our computers and google collab resources. Thus, we only include popular models. The results of around 5 models we tried are very similar, so we feel redundant to present all of them. Instead, we only mentioned the most well-performed two.

Also, it would be great to include hyperparameters tuning to improve the accuracy.

This part is mentioned in our presentation.

d. Evaluation on quality of writing (1-5): 2

Other than the first sentence, this part is irrelevant with writing but concerned with project design and should not be a reason for mark deduction. We appeal for grading only based on writing quality but not irrelevant parts.

e. Evaluation on presentation (1-5): 2

The reviewer did not give his reason for the three-mark deduction.

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

Indeed, BERT is a common way for such problem, that is why we propose models OTHER THAN BERT. Also, in the Q&A session we explained why it is difficult to provide data insight, which is such model usually has tens of thousands of parameters and extract them all would require a running capacity beyond accessible computers. If using popular models should be viewed as zero creativity, then we would have no way to conduct the project.

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

I have carefully read the paper and checked the results.

Given that the reviewer even couldn't summarize the project himself but copied the summary part entirely from our poster and the reviewer seemed to miss some information in our presentation, we doubt whether s/he is actually careful as he stated. From the reviewer's reasons above, we cannot see the reason behind his grading.