Data: 2019-09-20

Other peers’ reviews of my paper.

Peer 1:

1. Briefly summarise what the author has done

The author firstly briefly introduced what is blend words and then used statistics in blend word list to set up appropriate parameters for later blend word detecting algorithm. After made a few assumptions, the author had used both global edit distance and Jaro-Winkler similarity to develop the detecting algorithm. After listing some unsuitable situation and potential improvement, the results had been evaluated. Finally, the author wrote a review of related works, drawn a conclusion and stated what knowledge had been gained.

2. Indicate what you think the author has done well, and why

The entire paper is well structured and easy to read. Although the algorithm is somehow complicated using both Global Edit Distance and Jaro-Winkler Similarity, the author had clearly listed all methods used and even a segment of pseudo code for easier understanding. The author also did well in the analyzing part, two metrics were used to evaluate the algorithm. The author also did a research on related work, and all required parts, such as knowledge gained and related work review, are clearly stated as well.

3. Indicate what you think could have been improved, and why

As required by the subject’s instructor, the true blend list is only used to evaluate the algorithm rather than used to develop it. The author may not want to analyze the true blend words statistic at first to develop the algorithm, and the author should use results of other related articles to set up the algorithm parameters instead. Secondly, as a minor one, the words count had exceeded paper requirement, this can be done by refining the language used.

Peer 2:

1. Briefly summarise what the author has done

In this report, the author implements a approximate string matching system to detect blend words from candidates dataset and then evaluate the result of the proposed detection algorithm including precision and recall. There are two approximate string matching algorithms used in this project, which is Jaro-Winkler similarity and Global Edit Distance. First, the author used prefix tree and suffix tree to get the prefix and suffix of the candidates then calculate Jaro-Winkler similarity to find the most similar prefix and suffix. Finally, limit the GED distance to find predict blend words.

2. Indicate what you think the author has done well, and why

The author given a clear hypotheses about his(her) system and showed the knowledge gained from this project. In order to get a clear view of the dataset, he(she) get the statistical of the number of prefix and suffix also the length of blend word in order to simplify further analysis. This report's structure is logical and formal, in line with typical standards in academic writing. What's more, Ideas and arguments are clear and supportive, where the whole report is clear related to the each component.

3. Indicate what you think could have been improved, and why

In finding the prefix and suffix in this system, the author combined the most similar prefix words with the most similar suffix words respectively, it will miss some true blend words. There can be some improvement on finding prefix and suffix. In some parts of the report, there are lack of evidence to support the idea the author implement, it is better to add more evidence and illustrative examples. This system can add other algorithms to make a comparison and find a better algorithm to find blend words.