

## CHAPTER V

### CONCLUSION

#### 5.1 Conclusion

1. Typographical errors can affect the performance of a model.
2. Typographical errors can give the probability of the Likelihood being 0; eventually, the model can be classified incorrectly and reduce the model performance's accuracy.
3. By correcting typographical errors with edit distances 1 and 2, the model performance has increased.
4. The best model performance obtained using edit distance 2 with an average increase of **5.9%**, and the best word frequency list is by using the **built-in** frequency list or with a **gwicks** frequency list that provides an average accuracy of **79.7%**.

#### 5.2 Future Works

The accuracy of pspellchecker cannot be the leading benchmark for correcting typographical errors because it has not been able to correct more than two distances such as "goooooood" which should be "good", and also it has not been able to recognize the actual word wanted, such as given the word "god". However, it will return the corrected word into "god" whereas the wanted word is "good". Hence, further research can be done by improving the pspellchecker with other methods to overcome typo words more than 2 distances away, slang, abbreviation, et cetera, and improve recognizing the actual desired word. Future research can also be conducted to find another result with another classification algorithm combined with pspellchecker to overcome the typographical error.