

## CHAPTER 6

### CONCLUSION

Sentiment analysis which is done on tweets in a document has been realized with a fairly high degree of accuracy. The algorithm used also has a very good role. In addition to being easy to implement, k-NN also has high predictive accuracy when combined with Tf-Idf calculations. Python functions and libraries which are used also work well. Several conclusions can be drawn from this project such as:

1. The data was taken from twitter using scrap method. Searching from trending topic based-on hashtag. The hashtags which is taken were the hashtags with suspicious conditions such as the existence of debates on issues of religion, race, and culture.
2. The project often takes Indonesia region as the main source trending, otherwise the U.S and other Asian regions are the sources also.
3. The project needs at least 5000 thousands data. Therefore, each document needs more than 90 data tweets.
4. Final conclusion of this project is the result which has less point has the nearest score on negative text. The document which declared negative is the document which has more than 40% negative texts inside. The accuracy of this calculation is 82,3%.

Despite the project is fully complete, yet needs remedy. Improvement on the algorithm required in the future and data should be added more.