

CHAPTER 6

CONCLUSION

6.1 Conclusion

Based on the result from Chapter 5, it can be concluded:

1. **K-Means** algorithm can't check whether it's a fake review or not, but the upvotes can be played. There are 2 factors, firstly Steam Curators and secondly Developers.
2. From 500 data, it can be seen that the reviews that have a high **weight** value don't have a high **helpful** value.
3. No, from 500 data, the reviewers who has hour of played around 500-1000 hours get a fairly small amount of **helpful** value. It means the quality of a review is more considered by the customers.
4. The algorithm accuracy score for 100 sample data is **84%** and **100%** for 500 sample data. This means the algorithm accuracy is great.

6.2 Suggestions

There are some suggestions to improve this code in future:

1. To classify bigger data, or much diverse data need bigger class, for now this program can only hold 5 classes, changes to the classification process must be made in order to accommodate more classes.
2. TF-IDF need to be simplify, TF-IDF process still very slow, it needs more multiprocessor work.
3. The program needs hardware that have a better processor and bigger RAM to work faster.