CHAPTER 6
CONCLUSION

In this research, the author aims to solve fraudulent transaction prevention problem that occurs in E-Commerce. From several algorithm the author used, RF performs the best out of all scenario. In 30x averaged run score, RF almost always net a near-perfect score in both precision and recall that indicates low FP and FN. Both SVM methods fall short when given an imbalanced dataset but may perform better with a balanced dataset. The recall score of both SVM never reached 50% which is the baseline accuracy and only ever achieved a maximum of 20% in recall score. SVm RBF's precision however, can reach a near perfect score with 20% recall in several scenarios which means while some fraudulent transactions may be classified as a legitimate transaction, when the algorithm do classify a transaction fraudulent, it is almost always fraudulent.

RF with PCA applied have a higher precision and recall score than both SVM algorithms but still have relatively low recall score; The average accuracy an algorithm is safe to use for real-world scenarios is at about 80% but vary a lot depends on how important precision is or how important recall is e.g. some predictions needs high precision if it's something so important so there's no false positives allowed, or high recall when some false positives is okay, but false negatives leads to some serious consequences. In this research, the author needs an algorithm with high precision so no legitimate customer is flagged as fraudster while having acceptable recall rate to not allow too high of a fraudulent transaction to pass.

The author will need to gather some more real fraudulent transaction records to balance the dataset more since SVM seems to perform poorly with imbalanced dataset but may perform better with balanced dataset. Since we need a real transaction record, it is understandable to end up with small fraction of

fraudulent transaction, so it may be faster to just use sampling on the dataset, since even with the disadvantages, sampling is still proves to be quite the good measure to solve data imbalance [20]. Another thing the author would like to do is to expand the number of algorithm and combination of technique used to sample the result. Initially the author wanted to use a LOR but seeing as a LOR needs linear dataset and this dataset the author uses is certainly not linear, the author decides against using LOR in this research.