CHAPTER 6
CONCLUSION

Based on the results of research Comparing the Linear Regression and Random Forest Algorithms in the regression of the Paris housing dataset, it can be concluded that linear regression is accurate works in this study and also it suitable for this dataset, but for the random forest, the answer is no because random forest output has too many errors.

The main steps for the regression of the Linear regression and random forest method are dividing the data based on the probability that training data and testing data appear, with the comparison with different data samplers, the linear regression method has less error than the random forest method, with 60%, 70%, 80%, and 90%, linear regression method have less error, also linear regression also have less error in MSE, RMSE, And MAE. The only thing that cannot make a difference is R2, R2 always has the same output in those two algorithms.

Based on the above statement, the following suggestions are given for future study, use another method such as KNN, Tree, AdaBoost, neural network, or other algorithms that support regression-based data, and also for the random forest author suggests using another dataset aside from this Paris housing sales prediction to gain better result.