ABSTRACT

The rapid development of recent technologies demands major companies to maximize their profits and to minimize their loss. Using churn prediction, companies could solve that problem. But with the increasingly high amount of data, to generate an accurate prediction, there are things that need to be done, two of them being imbalanced data and an accurate model to predict the data to its right class. Handling imbalanced data is a necessity as it is reported that imbalanced data would impact negatively to the performance of some of the classification methods. In this research, underbagging will be used to handle the imbalanced data, and to classify the class, one of the supervised learning method support vector machine is going to be used as a classification method. The best F_1 -measure result in this research is 67.69% which belong to the test scenario with 38 attributes, 90% training data using Polynomial kernel.

Keywords : churn prediction, imbalanced data, underbagging, Support Vector Machine.