Abstract

Churn prediction is a method to predict behaviour of customers who potentially for being *churn*. This prediction is needed by telecommunication company to prevent its customer being *churn* so that the company be able to maintain its revenue. Data mining especially classification is expected to be one of alternative solution to build accurate *churn* prediction model. However, output from classification become not accurate because data of *churn* have the *imbalance* characteristic. Class data become not stable because data will be inclined to part data which has greater composition side.

One method to handle this problem is modify dataset or known as *sampling* method. *Sampling* method including *oversampling*, *undersampling*, and combine.

The analysis in this final exam is knowing how the effect of applying *sampling* method to accuration of *churn* prediction by counting *lift curve*, top decile, and *gini coefficient* as accuration of *churn* prediction model, and also f-measure as accuration of *imbalance* case .

The result from this research show that *sampling* method can increase accuration in *churn* evaluation. The application of *sampling* method causing *churn* data can be classified without sacrificing minor class which is focused in this research. Using *sampling* method make a different accuration to dataset as *churn* data or *imbalance* data.

Keywords: *churn prediction, imbalance, sampling, accuration, evaluation*.