Abstract

Churn prediction is a prediction to the telecommunication cellular customers that potentially churn. Another issue is that some data is not particularly useable or affecting the process in reaching the expected purposes. Because of that, to perform the mining process is impossible to use all the available features.

In order to solve it, it will perform feature selection to the available data. The feature selection will perform by using tool and human analysis. The tool analysis is using Fast Correlation Based Filter and Correlation Coefficient method also Clementine's feature selection. Meanwhile, the human analysis is performed by the field expert. For the classification process is using C5.0 method that integrated in the Clementine 10.1. But for the accuration calculation is using Gini Coefficient, Top Decile Lift and Lift Curve calculation.

The result from this research show that feature selection method can increase accuration. For telecomunication campany data, the Based feature selection method is collaboration between human analysis and tools analysis. Meanwhile for turnament data, the best feature selection method is tools analysis. Human analysis will give the optimal result in high dimensionality data if number of human analyst more many than human analyst to low dimensionality data. But if many human analyst will analysis the data then much expense will be outputed by company.

Keyword : feature selection, churn prediction, tools analysis, human analysis.