Abstract

Customers are very precious for companies if we talk about bussiness. One of the examples are in telecommunication based companies, customers will continue using their services if they satisfied with it. But if the services can't meet their satisfaction, they might be unsubscribe the service. This is very impactfull for successabilty of the companies. Therefore, a strategy must be implemented to predict wich customer whose have potential to stop using the service (churn). There is a problem before prediction can be done, which is imbalance data problem. Data mining approach can solve this problem. Then the processed data go to the classifitacion stage. The technique to counter imbalanced data that used in this paper is RUSBoost, and then SVM for the classification. The results obtained from this final task is the result of customer prediction who will do churn from total numbers of customer on indihome service. Then the results are compared with sistem that either use boosting or not. With boosting, the results of accuracy and F1score are mostly better. But, the results obtained are not stable both in accuracy or F1score, this is due to the use of the random undersampling method repeatedly that produce unique data. The best result obtained is in boosting with 2 loops and linear kernel svm, the score of F1score is 0,3195 and the accuracy is 0,8852.

Keywords: imbalance data, classification, RUSBoost, SVM