

Abstract

Most of the telecommunications company's revenue comes from customers. Therefore each company will strive to provide the best service to retain customers. However, many telecommunication companies make increasingly fierce competition to get customers so that customers will potentially churn. Churn prediction is a way to detect the movement of customers from one provider to the other provider. Problems encountered when predicting churn is imbalanced data between classes churn and not churn. In this study used methods Underbagging and logistic regression. Underbagging methods used to handle cases of imbalanced data which allows deleting data usage by not churn class so that the gap proportion of not churn and churn classes reduced. Logistic regression is binary classification techniques to maximize the probability function in matching the data to be processed. This study produced a useful system for identifying customers who wish to churn. System testing uses data Telecommunication Company generates prediction accuracy of 88.23% and is able to increase the F1-Measure of 0.3022.

KEYWORD : Churn, Underbagging, Logistic regression, F1-Measure, Imbalanced Data