ABSTRACT

Customers who are loyal to their provider are among the expectations of all companies. The problem that often happens in the company is churn. Churn is a transfer of customers to other providers according to the type of reason. To solve the problem is required churn prediction system at the company to help companies not to lose big. Therefore, in this final project, we designed a system for churn prediction using Neighborhood Cleaning Rule (NCL) and K-Nearest Neighbor (KNN) method. NCL is a sampling technique to handle data imbalance problem in Telecommunication company data. This technique will reduce the major data by identifying the noise data. Noise that misclassified will not be added to the training set, thus maintaining good classification accuracy. K-nearest neighbor is a method to classify objects based on learning data closest to the object. In this final project, the distance calculation used in the KNN method is euclidean, manhattan, chebyshev, mahalanobis and parameter k used k = 3, 5, 7, 51 and 101. The test results obtained from this system test show that this system can increase the value of f1-measure of 7.29% and have optimal performance with accuracy value 95.94% and f1-measure 42.46% at k = 3 manhattan distance.

Keywords: churn prediction, imbalance class, Neighborhood Cleaning Rule, K-Nearest Neighbor