Abstract

Lately the telecommunications companies are competing in terms of attracting consumers to use their services. This results in a churn. Churn is the transfer of customers for some reason. Displacement of the customer in the presence of a customer's inconvenience to use the service. Therefore, a company must be prepared a specific strategy to predict the transfer of customers. Strategies to retain customers can be done by changing marketing strategies. To fit the number of shifting customers, a model or system that can predict churn is required. In this final project, we designed a model that can predict churn, using the method of Synthetic Minority Over-Sampling Technique (SMOTE) and Pseudo K-Nearest Neighbor (KNN). SMOTE is a technique for handling data imbalance in telecommunication company data. To facilitate the classification of telecommunication company's customer data is used Pseudo K-Nearest Neighbor (KNN) method. Pseudo KNN is an algorithm that works according to the shortest distance from query instance to training sampling to determine ketetanggannya. Based on the test results using the SMOTE and Pseudo KNN models to predict churn, the most optimal result with parameters N is the minor data increase of N = 103 and the parameter k is the nearest neighbor number with k = 3 with the F1-Measure result of 38.56% And accuracy of 95.70%.

Keywords: churn prediction, imbalance, Synthetic Minority Over-samplingTechnique, Pseudo K-Nearest Neighbor, query instance, training.