

Abstract

Churn prediction on telecommunication customer case is an effort to predict/classify telecommunication service customers who quit or switch their subscribe from one operator to another. However, the dataset on its case generally have imbalanced class which instances from one class (active or negative or majority class) have a greater number from the other (churn or positive or minority class). As a consequence, mostly classifier skewed to classify the majority class and neglect the minority class, so that its accuracy become very low.

One approach for addressing the problem is modifying the dataset instances distribution which known as sampling-based approach. The resampling technique including over-sampling, under-sampling and combine-sampling.

With Ensemble of SVM (EnSVM) method hopefully can minimalize the majority and minority class misclassification which provided by just one classifier (SVM). It uses Combined Sampling technique to inherit the strength of both sampling technique (SMOTE and under-sampling) thus increases majority and minority class accuracy.

Through the empirical study, the result show that CombinedSampling-SVM method only can provide Gini Index performance better than SMOTE-SVM and pure-SVM method on Telecommucation Data generally.

Keywords: *churn prediction, imbalance, combined sampling, ensemble, SVM, performance.*