Abstract

Sentiment analysis on movie review documents is an attempt to classify movie review documents into a positive opinion classes or negative opinion classes. Problems that arise in the classification of documents is the amount of words (features) that will be considered, although very few of those words actually express the sentiment. This causes the classification of documents is slower because more words that must be processed instead of the actual need to be processed. In addition, it can reduce the accuracy because the system must consider the words that are not important when performing classification.

The approach taken to address these issues is to perform feature selection before the classification process. Feature selection method used is Categorical Proportional Difference (CPD). While the methods used for classification is the Support Vector Machine (SVM).

The results showed that application of feature selection using the CPD method can improve the accuracy of SVM classification even though the increase is relatively small.

Keywords: *sentiment analysis, movie review, CPD, SVM.*