Abstract

The government made a policy of moving the capital of Indonesia on August 26, 2019. This made many people have an opinion on the policy. In this study sentiment analysis will be conducted on public opinion regarding the relocation policy of the Indonesian capital, which is expressed through social media Twitter. Sentiment analysis in this study is a classification process into three classes, namely positive, negative and neutral sentiments. Opinion data was obtained from Twitter social media by scraping tweets that discussed moving the capital. This study aims to establish a text classification system for analyzing capital city relocation sentiments and to determine the effect of preprocessing and feature selection. There are several stages to do sentiment analysis, including data collection, preprocessing, split data into training data and testing data, feature extraction, feature selection, classification and evaluation. The method used in this research is classification using Support Vector Machine with Information Gain and chi square feature selection. Testing uses 3 SVM kernels which are linear, polynomial and radial basis function (RBF). The system built is able to produce the highest accuracy of 70.89% in the rough dataset, 72.01% in the dataset that has only been through the stemming stage of the existing preprocessing circuit, 71.14% in the dataset preprocessed without stemming and 71.64% for the dataset that has gone through all stages preprocessing. Every highest accuracy is obtained with SVM-RBF without applying feature selection.

Keywords: Twitter, Capital City Relocating, Information Gain, Support Vector Machine