

Abstract

A lot of information is scattered in various online social networks, including twitter. On twitter, many people share information around them. However, Twitter can only send tweets of up to 280 characters. Therefore, many users cut/shorten words and also use variations of words in tweets so that one tweet includes a lot of information. The use of word variations such as emoticons, slang, and abbreviations in tweets makes the mismatch of words and sentences that are conveyed so difficult to find difficult. In this case, the author expands the feature to classify topics on Twitter in order to overcome these problems. The method used for expansion is FastText and the method used in classification is Nave Bayes-Support Vector Machine (NBSVM). The results of this study indicate that the topic classification system with FastText feature expansion using the NBSVM method has an accuracy of 82.01%.

Keywords: *classification, NBSVM, fasttext, feature expansion*