

Abstract

Twitter social media users are usually only interested in *tweets* that fall under certain types of topics. *Tweets* that only contain no more than 140 characters, make classification of *tweets* a lot of challenges, because *tweets* are short, noisy, and less focused on the topic. The solution to solve these challenges in this study uses the expansion feature to enrich the text so that it looks like a large text document. The method chosen in the expansion feature is Word2Vec, to group vectors of similar words together in a vector space, meaning to detect similarity mathematically. The author uses the Random Forest method for the classification of *tweet* data in this study, because it is well known for maintaining data imbalance in different classes, especially very large data sets.

Kata kunci : *Tweet, Expansion Feature, Word2Vec, Random Forest.*