ABSTRACT

DETECTING TWITTER USER DEPRESSION SYMPTOMS USING *SUPPORT VECTOR MACHINE* ALGORITHM

Oleh

RINALDY JABAR SETIA

NIM: 1202160082

The World Health Organization estimates that every 40 seconds suicides occur worldwide due to depression. In Indonesia alone there are around 15.6 million people who experience depression, and continue to multiply. Many cases of depression lead to death due to lack of support from the environment, ignorance is the cause of people - people can not help.

Social media is now no different from real life, user habits related to thoughts, moods, activities, communication, and social life caught on Twitter can indicate users are depressed or not through emotions and the language used. In Indonesia alone, according to We Are Social, social media users reach 150 million users out of a total population of 268 million. One of the most popular social media in Indonesia is Twitter with 78 million users.

Therefore, early detection of depressive symptoms will be carried out using textual data. The first step is crawling process to get the dataset, after getting the data the labeling process will then be performed. Tweets showing depressive symptoms will be given a value of 1 and tweets that do not indicate depressive symptoms will be given a value of 0. Then a text pre-processing process is carried out to clear the data followed by a feature extraction process with the TF-IDF algorithm. After that the data is divided into training data and test data using k-fold cross validation and the classification process is carried out with the Support Vector Machine algorithm. After the classification process is evaluated using the confusion matrix and accuracy testing using k-fold cross validation. In this study, accuracy of 84.81%, precision 79%, recall 58% resulted in an f1-measure of 60%.

Keywords : depression, sentiment analysis, social media, support vector machine, twitter.