

ABSTRAK

Judul : Analisis Sentimen Kalimat Depresi Pada Pengguna Twitter Dengan Naive Bayes, Support Vector Machine, Random Forest

Nama : Mohammad Fachriza

Program Studi : Teknik Informatika

Pengguna aktif media sosial terus meningkat, dengan Twitter menjadi salah satu platform yang populer termasuk Twitter, menjadi tempat bagi pengguna untuk mengekspresikan pendapat, mengenai masalah kesehatan mental seperti depresi. Penelitian ini bertujuan untuk menganalisis sentimen *tweet* pengguna Twitter terkait depresi menggunakan metode klasifikasi seperti *Naïve Bayes Classifier*, *Support Vector Machine (SVM)*, dan *Random Forest*. Pengumpulan data *tweet* menggunakan metode *crawling* menggunakan API yang disediakan oleh Twitter dengan kata kunci yang berhubungan dengan depresi. Data *tweet* yang digunakan sebanyak 1502 *tweet*, yang selanjutnya dibersihkan pada tahap *preprocessing* dan diberi label dengan validasi oleh pakar terkait depresi, data yang sudah diberi label sebagai data untuk pengujian pada algoritma yang digunakan. Dari hasil pengujian performa pada algoritma yang diuji dapat disimpulkan bahwa algoritma *Random Forest* memiliki hasil performa yang lebih tinggi dibandingkan dengan *Naive Bayes Clasifier* dan *Support Vector Machine* dengan Hasil akurasi: 83.33%, presisi: 83.04%, recall: 83.33%, dan *f1-scores*: 82.62%. Penelitian ini Juga memberikan Word cloud untuk memberikan gambaran visual tentang kata-kata yang paling sering muncul dalam *tweet*. Kata "depresi" memiliki frekuensi lebih tinggi. Kata-kata yang dominan dapat memberikan indikasi tentang topik atau isu yang paling mendominasi terkait kesehatan mental seseorang.

Kata kunci : Analisis Sentimen, Depresi, Naive Bayes Classifier, Random Forest, Twitter.

ABSTRACT

Title : *Sentiment Analysis of Depressive Tweets on Twitter Using Naive Bayes, Support Vector Machine, and Random Forest*
Name : Mohammad Fachriza
Studi Program : Teknik Informatika

Active users of social media continue to increase, with Twitter being one of the popular platforms including Twitter, being a place for users to express opinions, regarding mental health issues such as depression. This research aims to analyze the sentiment of Twitter user tweets related to depression using classification methods such as Naïve Bayes Classifier, Support Vector Machine (SVM), and Random Forest. Collecting tweet data using the crawling method using the API provided by Twitter with keywords related to depression. The tweet data used is 1502 tweets, which are then cleaned at the preprocessing stage and labeled with validation by experts related to depression, data that has been labeled as data for testing the algorithms used. From the results of performance testing on the algorithms tested, it can be concluded that the Random Forest algorithm has higher performance results compared to Naive Bayes Clasifier and Support Vector Machine with accuracy results: 83.33%, precision: 83.04%, recall: 83.33%, and f1-scores: 82.62%. This research also provides a word cloud to provide a visual representation of the most frequently occurring words in tweets. The word "depression" has a higher frequency. Dominant words can give an indication of the most dominating topics or issues related to one's mental health.

Keywords : *Sentiment Analysis, Depression, Naive Bayes Classifier, Random Forest, Twitter.*