ABSTRAK

MENDETEKSI GEJALA DEPRESI PENGGUNA *TWITTER* MENGGUNAKAN *ALGORITMA NAÏVE BAYES CLASSIFIER*

Oleh

TUBAGUS RAHMAN RAMADAN

NIM : 1202162337

Dengan seiringnya perkembangan teknologi informasi (TI) yang semakin pesat. Pertukaran informasi menjadi sangat mudah dan cepat dengan melalui berbagai macam media, salah satunya adalah media sosial. Masyarakat Indonesia banyak yang tertarik dengan salah satu media sosial, yaitu *twitter*. *Twitter* dapat menjadi sumber data sentimen dan pendapat masyarakat untuk penelitian karena digunakan banyak orang untuk mengekspresikan emosi dan pendapat seseorang dengan bentuk teks. Pada penelitian ini dilakukan deteksi gejala depresi pengguna *twitter*, dikarenakan depresi menjadi salah satu penyakit serius dalam dunia kesehatan.

Pada penelitian ini bertujuan untuk mengetahui bagaimana analisis sentimen mendeteksi gejala depresi pengguna *twitter* dengan menggunakan algoritma *naïve bayes classifier*. Seseorang atau pengguna *twitter* yang mengalami gejala depresi dapat di lihat dari *tweets* yang berisi kata-kata dan kalimat yang menandakan gejala depresi, seperti kesedihan atau lekas marah, gangguan dalam tidur, menangis, dan pikiran untuk bunuh diri. Untuk dapat mendeteksi gejala depresi dilakukan melalui beberapa tahap, yaitu pengumpulan data, memberi label pada data, *preprocessing data,* ekstraksi fitur, dan klasifikasi data.

Pengumpulan data dilakukan dengan *web crawling* menggunakan API *twitter*. Data *tweet* yang mengandung gejala depresi ditandai dengan 1 dan yang tidak mengandung gejala depresi ditandai dengan 0. Terdapat enam proses yang dilakukan pada *preprocessing data*, yaitu *case folding, slang word removal, tokenization, stopword removal,* dan *stemming.* Ekstraksi fitur dilakukan dengan TF-IDF, dan klasifikasi data menggunakan algoritma *naïve bayes classifier* dengan validasi menggunakan *k-fold cross validation* dengan k =10. Setelah dilakukan klasifikasi maka selanjutnya dievaluasi menggunakan *confusion matrix*. Didapatkan hasil akurasi sebesar 83,68%, *precision* 61,51%, *recall* 50,23%, dan *f1-score* 46,20%.

Kata kunci: analisis sentimen, gejala depresi, media sosial, *naïve bayes classifier, twitter*.