

ABSTRAK

ANALISIS SENTIMEN SEBAGAI PENILAIAN OPINI TERHADAP POLEMIK BANTUAN SOSIAL COVID-19 MENGGUNAKAN ALGORITMA RANDOM FOREST

PRAMADITA SIELDA DEWI – NIM 1177050089

Upaya pemerintah dalam melaksanakan program Bantuan Sosial COVID-19 kerap kali menemui permasalahan dan menggiring opini banyak pihak. Opini yang diutarakan salah satunya pada media sosial *twitter*. Sentimen dari opini tersebut dapat dianalisis untuk mengetahui penilaian dan bahasan dari setiap sentimen yang dapat dijadikan bahan evaluasi dari program Bantuan Sosial telah dilakukan. Pada penelitian ini, sentimen setiap teks yang telah di-*preprocessing* didapatkan dari proses pelabelan dengan penilaian *polarity* dan *subjectivity* yang memanfaatkan *library TextBlob*. Hasil dari penilaian sentimen netral, positif, dan negatif dibobotkan menggunakan *TF-IDF* agar dimengerti oleh mesin. Kata yang telah diformat menjadi numerik kemudian diklasifikasikan menggunakan algoritma Random Forest. Parameter yang digunakan pada kasus ini disesuaikan dengan dokumentasi pada *sklearn*. Dilakukan pula evaluasi algoritma dengan metode *10 k-fold cross validation* sebagai validasi performa dari hasil *testing* setiap bagian data. Didapatkan performa *accuracy* sebesar 76%, *precision* 71,4%, *recall* 60,67%, dan *f1-score* 63,67.

Persentase klasifikasi teks menunjukkan bahwa penilaian topik yang dibahas adalah positif. Hanya saja, dari persentase tersebut ditemukan bahwa metode penilaian *polarity* dan *subjectivity* masih memiliki kekurangan dalam mendeteksi kalimat sarkasme. Namun secara keseluruhan masih bisa menghasilkan topik utama sesuai dengan masing-masing sentimen. Topik yang dibahas pada sentimen positif salah satunya adalah upaya pemerintah membuat fitur untuk memonitoring penyaluran bantuan sosial. Pembahasan kata sebutan yang digunakan pada kasus korupsi yang terjadi sebagai sentimen negatif. Adapun sentimen netral yang menunjukkan urgensi untuk membahas isu Bantuan Sosial.

Kata Kunci: Analisis Sentimen, Bansos, *Random Forest*, *Twitter*, *Polarity*, *Subjectivity*, *Cross Validation*.

ABSTRACT

SENTIMENT ANALYSIS AS ASSESSMENT OF THE COVID-19 SOCIAL ASSISTANCE POLLEMIC USING RANDOM FOREST ALGORITHM

PRAMADITA SIELDA DEWI – NIM 1177050089

The government's endeavors in organizing the COVID-19 Social Assistance program often encounter problems and lead to the opinion of many parties. One of the opinions expressed on social media is twitter. Sentiments from these opinions can be analyzed to find out the assessment and discussion of each sentiment that can be used as evaluation material for the Social Assistance program. In this study, the sentiment of each preprocessed text were obtained using a labeling process with an assessment of polarity and subjectivity from TextBlob library. The results of the neutral, positive, and negative sentiment assessments were then weighted using the TF-IDF for machine understanding. Words that have been formatted into numeric are then classified using the Random Forest algorithm. The parameters used in this case were in accordance with the documentation on sklearn. An evaluation of the algorithm was also carried out using the 10 k-fold cross validation method as a performance validation of the results of testing each piece of data. Accuracy performance was obtained at 76%, precision 71,4%, recall 60,67%, and f1-score 63,67.

The percentage of text classification indicates that the assessment of the topics discussed is positive. However, from this percentage it was found that the polarity and subjectivity assessment methods still have shortcomings in detecting sarcasm sentences. But overall it can still produce the main topic according to each sentiment. One of the topics discussed on positive sentiment was the government's endeavor to create a feature to monitor the distribution of social assistance. Discussion of the terms used in corruption cases that occur as negative sentiments. The neutral sentiment shows the urgency to discuss the issue of Social Assistance.

Keywords: *Sentiment Analysis, Bansos, Random Forest, Twitter, Polarity, Subjectivity, Cross Validation.*