

PENGEMBANGAN APLIKASI ANALISIS SENTIMEN APLIKASI PEDULILINDUNGI MENGGUNAKAN METODE NAÏVE BAYES

ABSTRAK

Abstract: This study aims to build sentiment analysis application using Naïve Bayes method to analyze public view about an application called PeduliLindungi, and measure it's accuracy by using twitter dataset. PeduliLindungi is an application developed by the government in order to track and stop Coronavirus Disease (COVID-19). The dataset used in this study is collected by using crawling method with Tweepy. Collected dataset will then go through data pre-processing and labeled by using VADER in order to separate it into positive and negative sentiments. The data will be weighted based on the frequency of its occurrence in all tweets using the TF-IDF method. The weighted data will then be classified using the Naïve Bayes method. This research used the incremental method both in model and software development. The results obtained in this study is a model with an accuracy score of 85% and an average precision, recall, and f1-score of 85%.

Keywords: Naïve Bayes, Sentiment Analysis, Incremental, TF-IDF, Twitter

Abstrak: Penelitian ini bertujuan untuk membangun aplikasi analisis sentimen menggunakan metode Naïve Bayes untuk menganalisis pandangan masyarakat terhadap sebuah aplikasi bernama PeduliLindungi, dan mengukur akurasi dengan menggunakan dataset twitter. PeduliLindungi adalah aplikasi yang dikembangkan oleh pemerintah untuk melacak dan menghentikan Coronavirus Disease (COVID-19). Dataset yang digunakan dalam penelitian ini dikumpulkan dengan menggunakan metode crawling dengan API bernama Tweepy. Dataset yang terkumpul kemudian akan melalui data pre-processing dan diberi label menggunakan VADER untuk memisahkannya menjadi sentimen positif dan negatif. Data tersebut akan dihitung bobotnya berdasarkan frekuensi kemunculannya pada semua tweet menggunakan metode TF-IDF. Data yang sudah dihitung bobotnya kemudian akan diklasifikasikan menggunakan metode Naïve Bayes. Penelitian ini menggunakan metode inkremental baik dalam pengembangan modelnya maupun perangkat lunaknya. Hasil yang diperoleh dalam penelitian ini adalah model dengan akurasi skor 85% dan rata-rata presisi, recall, dan f1-score 85%.

Kata kunci : Naïve Bayes, Analisis Sentimen, Inkremental, TF-IDF, Twitter