

Analisis Sentimen Mengenai Kuliah Daring Di Twitter Menggunakan Metode Naïve Bayes dan Pembentukan Stopword dengan Term Based Random Sampling

Raditya Rinandyaswara, Yuita Arum Sari, Muhammad Tanzil Furqon

Abstrak

Pemerintah Indonesia mengadakan sistem *New Normal* untuk mempercepat penanganan Covid-19. Kuliah daring menjadi salah satu efek dari sistem *New Normal*. Namun kuliah daring ini menjadi perbincangan masyarakat di sosial media. Sentimen Analisis diperlukan untuk mengatasi masalah tersebut dan mengetahui pendapat masyarakat. Dalam penelitian ini metode yang digunakan adalah *Naïve Bayes* serta algoritme *Term Based Random Sampling* untuk pembuatan *stopword*. Data yang digunakan terdiri dari 100 tweet negatif, 100 tweet netral, dan 100 tweet positif. Hasil evaluasi dengan *stopword Term Based Random Sampling* mendapatkan akurasi tertinggi dengan kombinasi X sebesar 10, Y sebesar 10, dan L sebesar 40 *macroaverage accuracy* sebesar 75,8%, *macroaverage precision* sebesar 65,8%, *macroaverage recall* sebesar 63,6%, dan *macroaverage f-measure* sebesar 64,7%. Berdasarkan hasil pengujian parameter disimpulkan bahwa semakin besar nilai X , Y , dan L maka semakin tinggi kemungkinannya untuk *accuracy*, *precision*, *recall*, dan *f-measure* turun. Hal ini dibuktikan ketika X bernilai 10, garis *accuracy* lebih tinggi dan lebih stabil jika dibandingkan dengan X ketika bernilai 50, hal ini berlaku juga untuk kedua parameter lainnya yaitu Y dan L . Hasil evaluasi sistem membuktikan bahwa analisis sentimen dengan *stopword Term Based Random Sampling* lebih tinggi dibandingkan dengan hasil evaluasi yang menggunakan *stopword Tala* maupun yang tanpa menggunakan proses *stopword removal*.

Kata kunci: *sentimen analisis, kuliah daring, twitter, Naïve Bayes, Term Based Random Sampling, stopwords*

Abstract

The Indonesian government established a *New Normal* system to accelerate the handling of Covid-19. Online lectures are one of the effects of the *New Normal* system. However, this online lecture is the talk of the public on social media. Sentiment analysis is needed to solve these problems and find out the public opinion. In this research, the method used is *Naïve Bayes* and the *Term Based Random Sampling* algorithm for making *stopwords*. The data used consisted of 100 negative tweets, 100 neutral tweets, and 100 positive tweets. The results of the evaluation with the *Term Based Random Sampling* *stopword* get the highest accuracy with a combination of X of 10, Y of 10, and L of 40 *macroaverage accuracy* of 75.8%, *macroaverage precision* of 65.8%, *macroaverage recall* of 63.6%, and *macroaverage f-measure* of 64.7%. Based on the results of parameter testing, it can be concluded that the greater the X , Y , and L values, the higher the probability for decreasing *accuracy*, *precision*, *recall*, and *f-measure*. This is evidenced when X is 10, the *accuracy* line is higher and more stable when compared to X when it is 50, this also applies to the other two parameters, namely Y and L . The results of the system evaluation prove that sentiment analysis with *stopword Term Based Random Sampling* is higher than the evaluation results using *stopword Tala* or without using the *stopword removal* process.

Keywords: *sentiment analysis, online school, twitter, Naïve Bayes, Term Based Random Sampling, stopwords*