

INTISARI

Penelitian ini bertujuan untuk mengetahui performa penggunaan *stemming* secara empiris sentimen terhadap BPJS Kesehatan pada media sosial *twitter* menggunakan algoritma *Naive Bayes*. Dalam analisa sentimen diwakili oleh dua tahap yaitu *preprocessing*, dan klasifikasi. Penggunaan algoritma *Naive Bayes Classifier* pada penelitian ini karena merupakan pendekatan klasifikasi probabilitas paling sederhana dan memiliki tingkat akurasi yang sangat tinggi. Tahapan penelitian ini yaitu, *crawling* data, *preprocessing*, pelabelan, pembuatan data uji dan latih, klasifikasi *Naive Bayes*, evaluasi, dan visualisasi data. Penelitian ini menggunakan 2000 data *tweet* sebagai sampel penelitian. Berdasarkan penelitian ini, diketahui sentimen negatif yang dihasilkan sebesar 74,2%, dan 25,8% menghasilkan sentimen positif. Hasil penelitian yang menggunakan algoritma *Naive Bayes* ini menunjukkan akurasi sebesar 76,55% saat menggunakan *stemming* dan 72,89% tanpa menggunakan *stemming* yang menandakan bahwa penggunaan *stemming* cukup berpengaruh pada akurasi.

Kata Kunci: Sentimen; BPJS Kesehatan; *Twitter*; Algoritma *Naive Bayes*;

ABSTRACT

This research aims to determine the performance of empirically stemming sentiment towards BPJS Health on Twitter social media using the Naive Bayes algorithm. In sentiment analysis, it is represented by two stages, namely preprocessing and classification. The Naive Bayes Classifier algorithm was used in this research because it is the simplest probability classification approach and has a very high level of accuracy. The stages of this research are, data crawling, preprocessing, labeling, creating test and training data, Naive Bayes classification, evaluation, and data visualization. This research uses 2000 tweet data as a research sample. Based on this research, it is known that negative sentiment generated was 74.2%, and 25.8% generated positive sentiment. The results of research using the Naive Bayes algorithm show an accuracy of 76.55% when using stemming and 72.89% without using stemming, which indicates that the use of stemming has quite an effect on accuracy.

Keywords: *Sentiment; BPJS Kesehatan; Twitter; Naive Bayes Algorithm;*