

INTISARI

Implementasi *Text Mining* dan *Sentiment Analysis* pada Jejaring Media Sosial *Twitter* menggunakan Metode *Naive Bayes Classifier*

(Studi Kasus : Vaksin COVID-19 di Indonesia)

Bima Yudha Wicaksono

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia

Melihat perkembangan dunia digital pada saat ini, pengguna *platform* media sosial *Twitter* bertambah seiring berkembangnya dunia digital. Pada kuartal pertama tahun 2019 tercatat sejumlah 330 juta orang menjadi pengguna aktif jejaring sosial *Twitter*, hal ini menjadikan *Twitter* sebagai aplikasi jejaring sosial yang menempati peringkat ke-12 dengan jumlah pengguna terbanyak di dunia. Dengan pengguna yang berasal dari berbagai kalangan dan lapisan masyarakat, opini warganet di jejaring sosial *Twitter* menjadi sangat beragam pada setiap kemunculan berita yang ada. Jika melihat tren isu dari data percakapan di *Twitter* sejak Oktober hingga Desember 2020, beberapa isu yang menjadi puncak pembicaraan warganet adalah terkait topik vaksin *COVID-19*. Oleh karena itu, penulis melakukan analisis pada jejaring media sosial *Twitter* terkait dengan isu vaksin *COVID-19* yang sempat menjadi perbincangan hangat di *platform* jejaring sosial tersebut menggunakan metode algoritma *Naive Bayes Classifier*. Metode ini dirasa cocok dikaitkan dengan penggunaan implementasi *text mining* dan *sentiment analysis* dikarenakan algoritma ini bertujuan sebagai metode klasifikasi ke dalam kategori positif dan negatif terhadap pendapat masyarakat terkait vaksin *COVID-19* di Indonesia. Hasil klasifikasi *Sentiment Analysis* didapatkan sebanyak 2960 tweet masuk dalam kelas sentimen negatif, 1475 masuk dalam kelas sentimen positif dan 1601 tweet masuk dalam kelas *sentiment* netral. Berdasarkan hasil klasifikasi menggunakan *naive bayes* didapatkan nilai akurasi sebesar 92%, untuk nilai *Recall* sebesar 88%, dan untuk nilai *precision* sebesar 97%. Berdasarkan nilai akurasi, *recall*, dan *precision* yang tinggi maka dapat dikatakan bahwa klasifikasi sudah tepat. Dapat juga dilihat dari nilai lain seperti *specificity*, *FPR*, dan *AUC* juga menghasilkan nilai yang besar juga. Untuk nilai *AUC* didapatkan hasil sebesar 0,92 yang artinya bahwa nilai tersebut sudah sangat baik atau klasifikasi dikatakan sudah sangat baik.

Kata Kunci : *Naive Bayes Classifier* , *Sentiment Analysis*, *Text Mining*, *Twitter*