

INTISARI

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

(Studi Kasus : Tes PCR di Indonesia)

Ahmad Zainul Lutfi

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

Perkembangan pengguna *platform* media sosial *Twitter* semakin semakin tinggi seiring dengan perkembangan dunia digital. pada kuartal pertama tahun 2019, ada 330 juta orang yang menjadi pengguna aktif jejaring sosial *Twitter*, menjadikan *Twitter* sebagai aplikasi jejaring sosial yang menempati urutan ke-12 dengan pengguna terbanyak di dunia. dengan pengguna dari berbagai kalangan dan lapisan masyarakat, *Twitter* ialah jejaring sosial tempat para netizen berbagi banyak sekali pendapat tentang setiap isu yang muncul. Melihat berita yang sedang menghangat dari data percakapan di *Twitter* pada Oktober sampai November 2021, berita yang menjadi perbincangan hangat netizen terkait topik tes PCR. oleh sebab itu, penulis melakukan analisis pada jejaring media sosial *Twitter* terkait informasi tes PCR yang sempat menjadi topik perbincangan hangat pada platform jejaring sosial tersebut menggunakan algoritma *Naive Bayes Classifier*. Metode ini dianggap cocok untuk dikaitkan dengan implementasi *text mining* dan *analisis sentimen* karena algoritma ini bertujuan sebagai metode klasifikasi ke dalam kategori positif dan negatif untuk opini publik mengenai tes PCR di Indonesia. hasil klasifikasi *Sentiment Analysis* dihasilkan sebanyak 3642 tweet untuk sentimen negatif, 2285 tweet untuk sentimen positif serta 4073 tweet untuk sentimen netral. berdasarkan hasil klasifikasi menggunakan *Naive Bayes* dihasilkan nilai akurasi sebesar 94%, nilai recall sebesar 96%, serta nilai presisi sebesar 95%. berdasarkan nilai akurasi, recall, dan presisi yang tinggi maka dapat dikatakan klasifikasi tersebut benar. bisa juga dilihat dari nilai lain seperti spesifisitas, FPR, serta AUC yang juga menghasilkan nilai yang besar. untuk nilai AUC didapatkan hasil sebesar 94% yang artinya nilai sangat baik atau klasifikasi dikatakan sangat baik.

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

ABSTRACT

Implementation of Text Mining and Sentiment Analysis on the Twitter Social Media Network using the Naive Bayes Classifier Method

(Case Study : PCR test in Indonesia)

Ahmad Zainul Lutfi

Department of Statistics, Faculty of Mathematics and Natural Sciences
Universitas Islam Indonesia

The development of users of the Twitter social media platform is getting higher along with the development of the digital world. In the first quarter of 2019, there were 330 million people who were active users of the social network Twitter, making Twitter the social networking application that ranks 12th with the most users in the world. With users from all walks of life and walks of life, Twitter is a social network where netizens share various opinions on every issue that arises. Seeing the news that is currently heating up from conversation data on Twitter from October to November 2021, the news that has become a hot topic for netizens is related to the topic of PCR tests. Therefore, the author conducted an analysis on the Twitter social media network regarding PCR test information which had become a hot topic of discussion on the social networking platform using the Naive Bayes Classifier algorithm. This method is considered suitable to be associated with the implementation of text mining and sentiment analysis because this algorithm aims as a classification method into positive and negative categories for public opinion regarding PCR tests in Indonesia. Sentiment Analysis classification results resulted in 3642 tweets for negative sentiment, 2285 tweets for positive sentiment and 4073 tweets for neutral sentiment. Based on the results of the classification using Naive Bayes, the accuracy value is 94%, the recall value is 96%, and the precision value is 95%. Based on the value of accuracy, recall, and high precision, it can be said that the classification is correct. It can also be seen from other values such as specificity, FPR, and AUC which also produce large values. For the AUC value, the result is 94%, which means the value is very good or the classification is said to be very good.

Keywords: *Twitter, PCR Test, Naive Bayes Classifier, Text Mining, Sentiment Analysis*