

ABSTRAK

COVID-19 kembali meningkat, di Indonesia pada bulan Januari 2021 sampai Februari 2021 tercatat 1.217.468 orang yang terkonfirmasi positif virus corona. Karena peningkatan angka tersebut, pemerintah melakukan upaya pencegahan yang salah satunya adalah dengan pendistribusian vaksin. Upaya Vaksinasi Covid-19 yang dilakukan oleh pemerintah memberikan pengaruh luas pada kalangan masyarakat melalui media sosial (khususnya *Twitter*) yang kemudian memunculkan pro dan kontra. Oleh karena itu, dibutuhkan analisis sentimen untuk memprediksi kecenderungan opini masyarakat terkait kebijakan vaksinasi Covid-19 yang diklasifikasikan menjadi opini positif, netral, dan negatif. *Random Forest Classifier* lemah dalam tingkat keakuratan akurasi, sehingga membutuhkan fitur seleksi yaitu dengan menerapkan *Information Gain* yang dapat meningkatkan akurasi. Hasil penelitian menunjukkan penerapan *Information Gain* dapat meningkatkan akurasi dengan akurasi tertinggi didapatkan pada percobaan 1 sebesar 0.00747 yaitu 0.94776 dari 0.94029 dengan nilai *precision* sebesar 65%, *recall* 43% dan *f1-score* 47% serta memiliki kecenderungan opini netral pada cuitan masyarakat mengenai vaksinasi Covid-19 pada *Twitter*.

Kata Kunci: Analisis Sentimen, COVID-19, *Information Gain*, *Random Forest Classifier*, *Twitter*, Vaksinasi.

ABSTRACT

COVID-19 in Indonesia has increased from January 2021 until February 2021 there were 1,217,468 people who were confirmed positive for the corona virus. As a result of the increase in the number, the government has made prevention efforts, one of which is the distribution of vaccines. The government's Covid-19 vaccination efforts have had a broad influence on the community through social media (especially Twitter) which then led to pros and cons. Therefore, sentiment analysis is needed to predict the tendency of public opinion regarding the Covid-19 vaccination policy which is classified into positive, neutral, and negative opinions. The Random Forest Classifier is weak in accuracy, so it requires a selection feature by applying Information Gain which can improve accuracy. The results show that the application of Information Gain can improve accuracy with the highest accuracy obtained in experiment 1 of 0.00747 that is 0.94776 from 0.94029 with a precision value of 65%, recall 43% and f1-score 47% and have a tendency to have a neutral opinion on public tweets about the Covid-19 vaccination on Twitter.

Keyword: *COVID-19, Information Gain, Random Forest Classifier, Sentiment Analysis, Twitter, Vaccination.*