

INTISARI

IMPLEMENTASI METODE XGBOOST DAN SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE UNTUK ANALISIS SENTIMEN APLIKASI PEDULILINDUNGI

(Studi Kasus: Ulasan Aplikasi PeduliLindungi pada Google Play)

Malecita Nur Atala Singgih

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

Kasus harian Covid-19 di Indonesia pada 5 Januari 2022 tercatat mengalami penambahan mencapai 404 kasus. Sebelumnya pada akhir Desember 2021 jumlah harian Covid-19 hanya berkisar antara 100 sampai 200 kasus. Kenaikan ini diakibatkan oleh adanya libur menyambut pergantian tahun serta pembatalan Pemberlakuan Pembatasan Kegiatan Masyarakat Level 3 sehingga menambah mobilitas masyarakat kala liburan. Mengatasi hal ini, dalam Instruksi Menteri Dalam Negeri Nomor 66 Tahun 2021 pemerintah mengoptimalkan penggunaan aplikasi PeduliLindungi. Aplikasi dapat diunduh melalui *Google Play*, dalam *Google Play* tersedia fitur ulasan (*review*) dari para pengguna. Fitur ini dapat dimanfaatkan untuk melihat ulasan berisi opini atau komentar yang bersifat positif maupun negatif yang diberikan oleh pengguna. Berdasarkan data ulasan Aplikasi PeduliLindungi, dapat dilakukan analisis sentimen untuk mengetahui bagaimana respon masyarakat terhadap Aplikasi PeduliLindungi. Metode klasifikasi yang digunakan adalah *Extreme Gradient Boosting* dan *Synthetic Minority Over-Sampling Technique* untuk mengatasi ketidakseimbangan kelas data. Analisis ini memberikan hasil bahwa terjadi peningkatan *f1 score*, akurasi, dan AUC dibandingkan sebelum dilakukan penyeimbangan data. Diperoleh *f1 score* sebesar 74,991%, tingkat akurasi sebesar 95,431% dan AUC sebesar 0,841. Hasil ini mengartikan bahwa ketepatan model dalam mengklasifikasi serta memprediksi ulasan positif dan negatif dapat dikatakan baik.

Kata Kunci : Covid-19, PeduliLindungi, Ulasan, XGBoost, SMOTE.

ABSTRACT

IMPLEMENTATION XGBOOST METHOD AND SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE FOR SENTIMENT ANALYSIS PEDULILINDUNGI APPS

(Case Study: Reviews PeduliLindungi Apps in Google Play)

Malecita Nur Atala Singgih

Department of Statistics, Faculty of Mathematics and Natural Sciences
Islamic University of Indonesia

COVID-19 daily cases in Indonesia on January 5, 2022 were recorded increased to 404 cases. Previously, at the end of December 2021 the daily cases only ranged 100 to 200 cases. This increase was caused by the New Year's holiday and the cancellation of Pemberlakuan Pembatasan Kegiatan Masyarakat level 3 increasing the mobility during holidays. To overcome this, in the Instruksi Menteri Dalam Negeri Nomor 66 Tahun 2021 the government optimize the uses of PeduliLindungi apps. The application can be downloaded on Google Play, in Google Play there is a review feature from users. This feature can be used to view the reviews containing positive and negative opinions or comments given by users. Based on the PeduliLindungi reviews data, Sentiment Analysis can be carried out to find out how the public responds to the PeduliLindungi apps. The classification using Extreme Gradient Boosting method and Synthetic Minority Over-Sampling Technique to overcome the imbalance class data. This analysis gives result that there is an increased on f1 score, accuracy, and AUC compared with the data before balancing. Obtained f1 score of 74.991%, an accuracy rate 95.431%, and an AUC 0.841. These results mean that the model is good at classifying and predicting positive and negative reviews.

Keywords: Covid-19, PeduliLindungi, Reviews, XGBoost, SMOTE.