INTISARI

SUPPORT VECTOR MACHINE DENGAN *STOCHASTIC GRADIENT DESCENT TRAINING* (SVM-SGD) PADA ANALISIS SENTIMEN

(Studi kasus: Ulasan Aplikasi PeduliLindungi di Google Play)

Oleh
Nur Isfandiarie Puteri
18/427706/PA/18666

Salah satu upaya pemerintah Indonesia untuk menanggulangi Covid-19 adalah dengan memperkuat 3T (*Testing, Tracing*, dan Treatment). Pelaksanaan 3T secara digital didukung dengan penggunaan aplikasi buatan Kemenkominfo yaitu PeduliLindungi. Tempat umum (pusat perbelanjaan, restoran, dsb.) kini mewajibkan seluruh pengunjung untuk scan QR code PeduliLindungi. Selain itu, banyak fitur lainnya yang mendorong masyarakat Indonesia menggunakan PeduliLindungi. Penting bagi pihak pengembang PeduliLindungi untuk menganalisis opini pengguna terhadap PeduliLindungi melalui ulasan yang ditulis langsung oleh pengguna salah satunya yang dapat diakses pada Google Play. Diperlukan metode untuk mengklasifikasikan opini pengguna (positif dan negatif) secara otomatis mengingat banyaknya ulasan yang masuk setiap harinya. Metode klasifikasi yang dapat digunakan adalah *Support Vector Machine* dan metode hybrid *Support Vector Machine with Stochastic Gradient Descent Training*. Karena data yang digunakan berbentuk teks, maka penting untuk mempertimbangkan teknik representasi vektor yang sesuai dengan metode klasifikasi yang digunakan, pada penelitian ini diujikan dua teknik yaitu BoW dan TF-IDF. Diperoleh metode SVM dengan vektor TF-IDF adalah yang paling baik diantara model lainnya, karena memiliki nilai akurasi, recall, presisi, dan f1-score paling tinggi yaitu nilai akurasi 95.68%, presisi 95.93%, recall 95.40%, dan f1-score 95.67%.

Kata kunci: analisis sentimen, *Support Vector Machine, Support Vector Machine with Stochastic Gradient Descent Training*, BoW, TF-IDF

ABSTRACT

SUPPORT VECTOR MACHINE WITH STOCHASTIC GRADIENT DESCENT TRAINING (SVM-SGD) ON SENTIMENT ANALYSIS

(Case study: PeduliLindungi App Review on Google Play)

By

Nur Isfandiarie Puteri

18/427706/PA/18666

One of the Indonesian government's efforts to tackle Covid-19 is to strengthen 3T (Testing, Tracing, and Treatment). The implementation of 3T digitally is supported by the use of an application made by the Ministry of Communication and Information, namely PeduliLindungi. Public places (shopping centers, restaurants, etc.) now require all visitors to scan the PeduliLindungi QR code. In addition, there are many other features that encourage Indonesian people to use PeduliLindungi. It is important for PeduliLindungi developers to analyze user opinions on PeduliLindungi through reviews written directly by users, one of which can be accessed on Google Play. A method is needed to classify user opinions (positive and negative) automatically considering the number of reviews that come in every day. The classification methods that can be used are the Support Vector Machine and the hybrid Support Vector Machine with Stochastic Gradient Descent Training method. Because the data used is in the form of text, it is important to consider vector representation techniques that are in accordance with the classification method used. In this study, two techniques were tested, namely BoW and TF-IDF. The SVM method with the TF-IDF vector is the best among other models, because it has the highest accuracy, recall, precision, and f1-score values, with 95.68% accuracy, 95.93% precision, 95.40% recall, and 95.67% f1-score.

Keywords: *sentiment analysis, Support Vector Machine, Support Vector Machine with Stochastic Gradient Descent Training,* BoW, TF-IDF