

ABSTRAK

Keamanan komunikasi digital di Indonesia menghadapi tantangan besar akibat meningkatnya *spam call*, penipuan telepon, dan kebocoran data pribadi. Hiya Global Call Threat Report mencatat bahwa Indonesia menjadi negara dengan *spam call* tertinggi di Asia Pasifik sejak 2023 hingga 2025, yang turut menurunkan kepercayaan masyarakat terhadap aplikasi pengelola data sensitif. Dalam situasi ini, aplikasi pendeteksi penelepon seperti Getcontact banyak digunakan, namun efektivitasnya dalam memberikan keamanan dan kenyamanan pengguna masih perlu dikaji. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap Getcontact berdasarkan ulasan Google Play Store serta mengevaluasi performa model klasifikasi dalam membedakan ulasan positif dan negatif. Selain analisis, penelitian ini juga menghasilkan implementasi sistem analisis sentimen yang mencakup fitur proses analisis, pencarian sentimen, analisis tren sentimen, dan analisis permasalahan yang berkaitan dengan latar belakang penelitian. Sebanyak 40.376 ulasan diproses melalui tahap normalisasi dan diberi label menggunakan *InSet Lexicon*, menghasilkan 6.678 ulasan positif dan 31.201 ulasan negatif. Proses klasifikasi dilakukan menggunakan algoritma *Support Vector Machine* (SVM), yaitu metode pembelajaran mesin yang bekerja dengan mencari batas pemisah paling optimal untuk membedakan dua kelas data. Klasifikasi berbasis TF-IDF ini kemudian dibandingkan menggunakan dua jenis kernel, yaitu Linear dan *Radial Basis Function* (RBF). Hasil optimasi menunjukkan bahwa kernel Linear memiliki parameter terbaik $C = 1$, sedangkan kernel RBF mencapai performa optimal pada kombinasi $C = 100$ dan γ (Gamma) = 0.1. Dari seluruh rasio data yang diuji, pembagian data 80:20 memberikan hasil terbaik, di mana kernel RBF memperoleh akurasi sebesar 96,20%, mengungguli kernel Linear dengan akurasi 95,90%. Temuan ini menunjukkan bahwa kernel RBF lebih efektif dalam menangkap pola bahasa kompleks, sementara analisis sentimen mengungkapkan dominasi ulasan negatif, terutama terkait kebijakan premium, privasi data, dan ketidakakuratan tagar.

Kata Kunci: Analisis Sentimen, Getcontact, Google Play Store, *InSet Lexicon*, Kernel Linear, Kernel RBF, SVM, TF-IDF.

ABSTRACT

Digital communication security in Indonesia faces major challenges due to the increase in spam calls, telephone fraud, and personal data leaks. The Hiya Global Call Threat Report notes that Indonesia has been the country with the highest number of spam calls in the Asia Pacific region from 2023 to 2025, which has contributed to a decline in public trust in applications that manage sensitive data. In this situation, caller detection apps like Getcontact are widely used, but their effectiveness in providing security and user comfort still needs to be examined. This study aims to analyze user sentiment towards Getcontact based on Google Play Store reviews and evaluate the performance of classification models in distinguishing between positive and negative reviews. In addition to analysis, this study also produced the implementation of a sentiment analysis system that includes analysis process features, sentiment search, sentiment trend analysis, and analysis of issues related to the research background. A total of 40,376 reviews were processed through a normalization stage and labeled using InSet Lexicon, resulting in 6,678 positive reviews and 31,201 negative reviews. The classification process was carried out using the Support Vector Machine (SVM) algorithm, a machine learning method that works by finding the most optimal separating boundary to distinguish two classes of data. This TF-IDF-based classification was then compared using two types of kernels, namely Linear and Radial Basis Function (RBF). The optimization results showed that the Linear kernel had the best parameter $C = 1$, while the RBF kernel achieved optimal performance at a combination of $C = 100$ and γ (Gamma) $= 0.1$. Of all the data ratios tested, the 80:20 data split gave the best results, with the RBF kernel achieving an accuracy of 96.20%, surpassing the Linear kernel with an accuracy of 95.90%. These findings indicate that the RBF kernel is more effective at capturing complex language patterns, while sentiment analysis reveals a predominance of negative reviews, particularly related to premium policies, data privacy, and hashtag inaccuracy.

Keywords: Getcontact, Google Play Store, InSet Lexicon, Linear Kernel, RBF Kernel, Sentiment Analysis, SVM, TF-IDF.