

## ABSTRAK

Rendahnya tingkat literasi masyarakat Indonesia berdasarkan survei *PISA* dan terbatasnya akses literasi berbasis digital terkait kata serapan mendorong pengembangan solusi teknologi pembelajaran yang inovatif. Penelitian ini bertujuan mengembangkan aplikasi Android yang dapat mengenali dan mengklasifikasikan kata serapan dari input percakapan pengguna menggunakan algoritma *Support Vector Machine (SVM)*. Metodologi *CRISP-DM* digunakan sebagai kerangka kerja penelitian, dengan dataset teks yang terdiri dari 3.433 kata serapan dan 3.433 kata murni yang dikumpulkan dari Buku Kamus Kata Serapan karya Js. Badudu, KBBI, dan Buku Senarai Kata Serapan. Dataset audio dihasilkan menggunakan platform *text-to-speech LuvVoice* dengan total 27.456 file audio dari empat varian suara. Ekstraksi fitur dilakukan menggunakan *TF-IDF* berbasis karakter untuk data teks dan *Mel-Frequency Cepstral Coefficients (MFCC)* untuk data audio. Hasil evaluasi menunjukkan model berbasis teks mencapai akurasi 88% dengan precision 90% dan recall 86% untuk kata serapan, sementara model berbasis audio mencapai akurasi 73% dengan precision 78% dan recall 66%. Pengujian pada data baru menunjukkan performa yang bervariasi: model teks mencapai akurasi 96% untuk kata serapan dan 55% untuk kata murni, sementara model audio mencapai 62% untuk kata serapan dan 15% untuk kata murni. Aplikasi yang dikembangkan mengintegrasikan model *SVM* dengan arsitektur terdistribusi, dilengkapi fitur autentikasi pengguna, monitoring percakapan, dan pembuatan laporan. Meski masih memerlukan pengembangan untuk meningkatkan akurasi prediksi, terutama pada model audio dan klasifikasi kata murni, aplikasi ini diharapkan menjadi solusi inovatif untuk pembelajaran kata serapan bahasa Indonesia.

Kata kunci: Kata Serapan, *Machine Learning*, *CRISP-DM*, *TF-IDF*, *MFCC*, *SVM*, Aplikasi, Android

