

ABSTRAK

My Tel-U merupakan aplikasi media sosial privat yang digunakan oleh sivitas akademika Universitas Telkom untuk mengakses informasi akademik dan mengekspresikan opini melalui fitur *timeline*-nya. Postingan yang dibagikan di *timeline* sering kali mencerminkan sentimen pengguna terhadap layanan kampus atau lingkungan sekitarnya. Namun, pendekatan analisis sentimen sebelumnya masih bergantung pada *dataset* berukuran kecil dan metode tradisional yang kurang mampu menangani teks Bahasa Indonesia yang informal dan bercampur bahasa. Penelitian ini bertujuan untuk mengatasi keterbatasan tersebut dengan mengevaluasi efektivitas Large Language Models (LLMs), termasuk BERT, RoBERTa, dan IndoBERT, dibandingkan dengan model tradisional seperti Support Vector Machine (SVM), dalam analisis sentimen pada data My Tel-U yang tidak seimbang dan penuh gangguan (*noisy*). Penelitian dilakukan melalui tiga skenario eksperimen, yaitu mengamati pengaruh teknik *oversampling* yang berbeda (*SMOTE*, *random oversampling*, dan tanpa *oversampling*) terhadap IndoBERT, perbandingan performa antar LLM, serta perbandingan antara LLM dan SVM dengan representasi TF-IDF dan Word2Vec. Studi ini merupakan salah satu yang pertama dalam mengeksplorasi secara sistematis penggunaan LLM berbahasa Indonesia untuk klasifikasi sentimen pada platform media sosial pendidikan privat. Hasil penelitian menunjukkan bahwa IndoBERT mengungguli semua model lainnya dengan skor F1-Macro sebesar 0,8423, serta mampu memberikan performa tinggi tanpa memerlukan teknik *oversampling*. Hal ini menunjukkan bahwa pelatihan awal (*pretraining*) pada korpus berbahasa Indonesia secara signifikan meningkatkan pemahaman kontekstual. Sebaliknya, LLM berbasis bahasa Inggris dan model SVM tradisional menunjukkan performa yang lebih rendah karena keterbatasannya dalam memahami struktur linguistik Bahasa Indonesia. Temuan ini memperlihatkan potensi praktis integrasi LLM spesifik bahasa ke dalam sistem umpan balik dan pemantauan institusi, sehingga memungkinkan universitas untuk lebih memahami dan merespons sentimen pengguna.

Kata Kunci: analisis sentimen, *large language models*, indobert, data tidak seimbang, my tel-u.