

BAB 1 PENDAHULUAN

Media sosial telah menjadi platform utama bagi individu untuk mengekspresikan perasaan dan berinteraksi dengan komunitas. Salah satu platform media sosial yang digunakan oleh komunitas akademik di Telkom University adalah My Tel-U. Aplikasi ini dikembangkan oleh Direktorat Pusat Teknologi Informasi Telkom University untuk memberikan akses mudah terhadap informasi akademik bagi dosen, karyawan, dan mahasiswa [1]. Selain menyediakan informasi utama seperti jadwal kuliah dan informasi akademik lainnya, My Tel-U juga memiliki fitur *timeline* yang memungkinkan pengguna untuk berinteraksi melalui unggahan teks, gambar, atau kombinasi keduanya. Melalui fitur ini, pengguna sering mengekspresikan perasaan yang memberikan wawasan berharga tentang persepsi mereka terhadap lingkungan sekitar [2]. Memahami sentimen dari unggahan ini melalui *sentiment analysis* dapat menjadi langkah penting dalam meningkatkan pengalaman pengguna sekaligus membantu universitas mengevaluasi kinerjanya [3].

Penelitian sebelumnya dalam *sentiment analysis* telah mengidentifikasi berbagai keterbatasan, termasuk penggunaan dataset yang spesifik domain atau berskala kecil, *overfitting* pada model *machine learning*, tantangan dalam memproses teks informal atau campuran bahasa, dan keterbatasan pemahaman kontekstual. Selain itu, belum ada penelitian yang secara sistematis mengevaluasi efektivitas Large Language Models (LLMs) untuk *sentiment analysis* pada platform sosial privat dalam komunitas tertutup seperti institusi akademik. Penelitian ini bertujuan untuk mengisi kesenjangan tersebut.

Untuk mengatasinya, penelitian ini menggunakan *dataset* yang lebih besar dan beragam yang dikumpulkan dari platform My Tel-U serta menerapkan model berbasis LLM seperti BERT, RoBERTa, dan IndoBERT.

Model-model ini dikenal memiliki kemampuan pemahaman konteks yang kuat. IndoBERT, yang secara khusus dilatih pada korpus bahasa Indonesia, efektif dalam memproses teks monolingual maupun yang *code-switched*, serta mengungguli model seperti mBERT dalam tugas *sentiment analysis* [4], [5]. RoBERTa juga menunjukkan generalisasi yang kuat dan kinerja yang lebih unggul dibandingkan BERT dalam berbagai tugas klasifikasi [6].

Selain itu, penelitian ini juga membandingkan kinerja model LLM dengan metode tradisional seperti Support Vector Machine (SVM), untuk mengevaluasi efektivitas kedua pendekatan pada *dataset* My Tel-U, yang memiliki karakteristik bahasa informal dan distribusi data yang tidak seimbang. Penelitian ini diharapkan dapat berkontribusi dalam pengembangan metode *sentiment analysis* yang lebih akurat, khususnya pada platform media sosial privat di Indonesia.