

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Di era digital ini, media sosial menjadi platform utama bagi masyarakat untuk menyampaikan pendapat dan berbagi informasi. Twitter, sebagai salah satu platform media sosial yang sangat dinamis, memberikan ruang bagi pengguna untuk mengungkapkan opini mereka, termasuk mengenai aplikasi-aplikasi yang mereka gunakan, seperti Aplikasi KAI Access. Sebagai aplikasi resmi milik PT Kereta Api Indonesia, Aplikasi KAI Access memungkinkan pengguna untuk memberikan opini, pengalaman, dan pandangan mereka terhadap layanan yang disediakan.

Analisis sentimen terhadap komentar publik mengenai Aplikasi KAI Access sangat penting untuk memahami persepsi masyarakat terhadap layanan tersebut. Namun, salah satu tantangan utama dalam analisis sentimen adalah rendahnya akurasi yang muncul ketika model tidak mampu memahami konteks kalimat secara dua arah. Model tradisional biasanya memproses kata-kata berdasarkan konteks satu arah [1], sehingga sering kali mengabaikan makna ganda atau konteks yang lebih luas. Selain itu, model-model konvensional kesulitan mengatasi bahasa informal, termasuk singkatan, slang, atau variasi bahasa yang sering ditemukan di Twitter.

Untuk mengatasi masalah ini, BERT (Bidirectional Encoder Representations from Transformers) menjadi solusi yang sangat relevan. BERT, dengan kemampuannya untuk memahami konteks kalimat secara dua arah, mampu menangkap makna kata dalam konteks yang lebih luas, yang sangat penting untuk meningkatkan akurasi analisis sentimen. Khusus untuk bahasa Indonesia, penggunaan Multilingual BERT menjadi pilihan yang tepat karena mampu menangani berbagai bahasa, termasuk kombinasi antara bahasa formal dan informal yang sering ditemukan pada komentar publik di Twitter [2]. Dengan mengoptimasi BERT Multilingual melalui teknik fine-tuning pada dataset yang relevan, model ini

dapat lebih akurat dalam mengenali nuansa bahasa Indonesia serta variasi unik seperti singkatan, kode-switching, atau slang.

Selain fokus pada peningkatan akurasi, optimasi Multilingual BERT juga bertujuan untuk meningkatkan efisiensi model, yang diukur berdasarkan waktu pelatihan. Dengan menggunakan teknik seperti distilasi model dan optimasi hiperparameter, penelitian ini berusaha menemukan konfigurasi terbaik yang memungkinkan model tetap berkinerja tinggi tanpa mengorbankan efisiensi. Efek dari optimasi ini akan dievaluasi dengan membandingkan model sebelum dan sesudah optimasi, dengan mempertimbangkan metrik seperti akurasi.

Penggunaan BERT Multilingual yang telah dioptimasi diharapkan dapat memberikan hasil analisis sentimen yang lebih akurat dan efisien. Hal ini tidak hanya membantu memahami opini masyarakat terhadap layanan KAI Access, tetapi juga menawarkan solusi yang dapat diterapkan secara luas di lingkungan operasional. Penelitian ini bertujuan untuk menunjukkan bagaimana optimasi BERT dapat meningkatkan akurasi dan efisiensi sekaligus memberikan wawasan yang bermanfaat untuk pengembangan layanan.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang di atas, masalah-masalah yang ada dapat dirumuskan sebagai berikut:

1. Bagaimana kinerja model BERT multilingual dalam mengklasifikasikan komentar terkait Aplikasi KAI Access, serta seberapa efektif optimasi model tersebut dalam meningkatkan akurasi?

## **1.3 Batasan Masalah**

Untuk menghindari permasalahan yang lebih luas dilakukan pembatasan sebagai berikut:

1. Penelitian dilakukan dengan data tweet dari user yang memuat komentar ataupun opini terkait kai access pada platform X
2. Data yang akan diproses hanya data tweet dari bahasa Indonesia yang memuat kata kunci "kai access".

3. Pengambilan data komentar dilakukan pada rentan waktu November 2023-April 2024.
4. Data yang diperoleh dari hasil crawling sebanyak 12.477.
5. Labeling menggunakan metode pipeline dengan model Bahasa Indonesia. Komentar hanya diklasifikasi menjadi komentar *positive* dan *negative*

#### **1.4 Tujuan Penelitian**

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah sebagai berikut :

1. Menghasilkan model BERT yang digunakan untuk melakukan klasifikasi aplikasi kai access beserta evaluasinya.
2. Mengetahui pengukuran efek optimasi terhadap efisiensi model BERT

#### **1.5 Manfaat Penelitian**

Hasil penelitian ini diharapkan dapat memberi manfaat sebagai berikut :

1. Mengembangkan kemampuan tentang sentiment analisis dengan metode BERT
2. Mendapatkan masukan yang berharga untuk meningkatkan efektivitas algoritma BERT
3. Hasil penelitian ini dapat memberikan pemahaman yang lebih mendalam tentang persepsi dan opini pengguna X terhadap aplikasi kai access.

#### **1.6 Sistematika Penulisan**

Berikut ini merupakan sistematika penulisan yang ada dalam penelitian ini

### **BAB I PENDAHULUAN**

Berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian serta sistematika penulisan skripsi.

### **BAB II TINJAUAN PUSTAKA**

Bab ini berisikan studi literatur yang mencakup penelitian-penelitian terdahulu dan dasar teori yang digunakan dalam penelitian ini.

### BAB III METODE PENELITIAN

Di dalamnya membahas metode yang dilakukan dalam penelitian serta Langkah-langkah yang harus dilakukan untuk mencapai tujuan dalam menyelesaikan penelitian tugas akhir.

### BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi penjelasan mengenai tahap-tahap yang telah dilakukan pada penelitian ini yang mencakup penumpulan data, pra-pemrosesan; pelatihan, hingga evaluasi. Bab ini juga membahas hasil temuan yang didapat setelah proses evaluasi dan perbandingan model dilaksanakan.

### BAB V PENUTUP

Bab ini berisikan kesimpulan dari penelitian yang dilakukan serta saran-saran yang berguna untuk penelitian selanjutnya. Berikut ini merupakan sistematika penulisan yang ada dalam penelitian ini