

BAB I

PENDAHULUAN

A. Latar Belakang

Transformasi digital yang berlangsung pada teknologi informasi dan komunikasi kini merambah berbagai sector kehidupan, salah satunya sektor dalam pelayanan transportasi. Peran transportasi sendiri memiliki peran yang sangat penting karena mendukung pergerakan masyarakat sehari-hari. Dengan perkembangan ini dapat mempermudah masyarakat melakukan aktivitasnya sehari-hari karena memiliki kesetaraan nilai baik melalui transportasi laut, udara maupun darat. Kondisi inilah yang menjadikan peningkatan prasarana dan sarana transportasi menjadi perhatian utama pada masyarakat [1].

Aplikasi Ferizy diluncurkan pada Mei tahun 2020 oleh PT ASDP sebagai salah satu inovasi layanan transportasi digital. Dengan aplikasi ini, salah satunya layanan transportasi laut dapat digunakan untuk memudahkan masyarakat dalam memesan tiket keberangkatan yang sebelumnya masyarakat harus membeli tiket secara langsung di loket. Namun dengan aplikasi Ferizy proses pemesanan tiket berbasis online dengan mengunduh aplikasi melalui ponsel kapan saja dan dimana saja menjadi lebih mudah.



Gambar 1. Visualisasi Data Jumlah Pengguna Aplikasi Ferizy

Berdasarkan Gambar 1. terkait visualisasi data jumlah pengguna pada aplikasi Ferizy pada tahun 2020 didapatkan 438.105 pengguna. Pada tahun 2021, pengguna aplikasi Ferizy menurun dengan jumlah yang didapatkan sekitar 354.700 pengguna. Hal ini juga dikarenakan pada tahun 2021 terkait kasus Covid-19, pemerintah mengambil keputusan yaitu adanya pemberlakuan pembatasan kegiatan masyarakat. Sedangkan, pada tahun 2022 mengalami peningkatan dengan jumlah 527.730 pengguna dan tahun 2023 berjumlah 655.951 pengguna. Pada tahun 2024 setelah meredanya kasus Covid-19 tersebut, pengguna pada aplikasi Ferizy mengalami peningkatan drastis yaitu berjumlah 2.500.000 pengguna (ASDP, 2024). Meskipun telah memberikan berbagai kemudahan dalam upaya meningkatkan sarana transportasi pengalaman pengguna yang optimal, terdapat kekurangan dalam aplikasi mobile-nya. Hal ini berpotensi memunculkan berbagai tanggapan dalam bentuk ulasan atau komentar di *Google Play Store*.

Ulasan yang diberikan pengguna Ferizy memiliki kedudukan penting untuk dijadikan acuan dalam meningkatkan layanan di masa mendatang bagi pihak pembuat aplikasi. Didalamnya tersimpan informasi bernilai yang memungkinkan pemilik aplikasi mengetahui kekurangan sekaligus kelebihan aplikasinya [2]. Upaya ini dikenal dengan istilah analisis sentimen dalam memahami penilaian pengguna terhadap aplikasi Ferizy.

Analisis sentimen merupakan proses untuk mengidentifikasi emosi serta memahami sentimen atau opini dari seseorang dalam bentuk sebuah teks terhadap suatu produk atau merek, bahkan layanan. Tujuan analisis sentimen ini dapat mengetahui dan mengkategorikan apakah sentimen yang terkandung sebagai sentimen positif atau sentimen negatif dengan memanfaatkan model [3].

Efektivitas model *IndoBERT* juga dibuktikan pada penelitian yang berhasil mengklasifikasikan sentimen ulasan aplikasi *mobile banking BRI* dengan akurasi mencapai 90% yang menunjukkan performa tinggi pada domain ulasan aplikasi yang sejenis dengan objek penelitian ini [4]. Selain itu, penelitian yang menerapkan model *BERT* pada komentar video YouTube mengenai *Paylater* justru hanya memperoleh akurasi 58%, akibat penggunaan dataset yang kecil (326 komentar) dan tidak seimbang antar kelas [5]. Perbedaan hasil ini menegaskan bahwa performa tinggi model berbasis *BERT/IndoBERT* sangat bergantung pada ukuran

dataset yang memadai serta penanganan ketidakseimbangan kelas yang tepat, bukan semata-mata pada arsitektur model itu sendiri. Hal ini diperkuat pada penelitian yang menunjukkan bahwa performa *IndoBERT* dapat dioptimalkan lebih lanjut melalui eksplorasi *hyperparameter* menggunakan *Grid Search*, yang berhasil menaikkan akurasi hingga 90,55% pada klasifikasi sentimen ulasan aplikasi *E-Government* [6].

Berdasarkan uraian sebelumnya, penelitian ini mengusulkan klasifikasi analisis sentimen dengan menerapkan metode *IndoBERT*. Pemilihan model *IndoBERT* ini bersandar pada bukti empiris sejumlah penelitian terdahulu yang memperlihatkan keunggulan model berbasis *transformer* dalam tugas analisis sentimen berbahasa Indonesia. *IndoBERT* sendiri merupakan salah satu adaptasi dari model *BERT* (*Bidirectional Encoder Representation From Transformers*) yang di rancang khusus untuk bahasa Indonesia dan mampu menangani berbagai tugas pemrosesan bahasa alami, seperti analisis sentimen [7]. Dibanding model pemrosesan bahasa alami umumnya, kinerja model *IndoBERT* lebih unggul dalam mengolah teks bahasa Indonesia. Fokus penelitian ini diarahkan pada klasifikasi sentimen ulasan pengguna Ferizy ke dalam dua kategori, yaitu positif dan negatif. Implementasi model *IndoBERT*, diharapkan mampu menghasilkan pemahaman yang lebih mendalam mengenai sentimen pengguna terhadap aplikasi Ferizy, yang selanjutnya dapat digunakan untuk perbaikan layanan berkelanjutan.

B. Perumusan Masalah

Merujuk pada pemaparan latar belakang diatas, rumusan masalah yang diangkat dalam penelitian ini adalah :

1. Bagaimana tingkat akurasi dan performa model *IndoBERT* (*Indonesian Bidirectional Encoder Representation From Transformers*) ketika diterapkan untuk mengklasifikasikan sentimen positif dan negatif pada ulasan aplikasi Ferizy di *google play store*?

C. Batasan Penelitian

Penelitian ini akan membatasi cakupannya agar pembahasan tidak melebar terlalu luas. Penelitian ini menetapkan sejumlah batasan sebagai berikut :

1. Data penelitian bersumber dari komentar-komentar pengguna aplikasi Ferizy di *Google Play Store*, dengan total sebanyak 8.000 ulasan terhitung sejak 25 Juni 2020 – 05 Juli 2025.
2. Tahapan *preprocessing* data mencakup tiga proses yaitu, *cleaning*, *case folding*, dan normalisasi.
3. Klasifikasi sentimen pada penelitian ini fokus pada dua kategori, yaitu positif dan negatif.
4. Penanganan ketidakseimbangan kelas melalui metode *Random Undersampling* (RUS) diterapkan pada keseluruhan dataset hasil pelabelan sebelum tahap split data, sehingga turut memengaruhi jumlah data latih (*training*), data validasi (*validation*), dan data uji (*testing*) pada skenario data seimbang. Penelitian ini tidak membandingkan RUS dengan teknik penyeimbangan data lain seperti *oversampling* atau kombinasi keduanya.
5. Pada tahap *split* data, baik dataset tidak seimbang maupun dataset hasil RUS dibagi dengan rasio yang sama, yaitu 70:10:20, sehingga menghasilkan jumlah data latih, validasi, dan testing yang berbeda pada masing-masing skenario.
6. Klasifikasi sentimen pada penelitian ini menerapkan algoritma *IndoBERT* (*Indonesian Bidirectional Encoder Representations from Transformers*) yang dikombinasikan dengan *hyperparameters tuning* melalui *Grid Search*.
7. Proses pelatihan model *IndoBERT* memanfaatkan dua jenis data yaitu data latih untuk melatih model serta data validasi sebagai pengawasan dan evaluasi dalam mencari model terbaik.
8. Fokus utama penelitian ini berupa hasil akurasi analisis sentimen yang diperoleh dari penerapan metode *IndoBERT*.

D. Tujuan Penelitian

Penelitian ini bertujuan untuk memperoleh hasil akurasi dan menguji performa model *IndoBERT* (*Indonesian Bidirectional Encoder Representation From Transformers*) dalam mengklasifikasikan sentimen positif dan sentimen negatif pada ulasan pengguna aplikasi Ferizy di *Google Play Store*.

E. Manfaat Penelitian

Penelitian ini diharapkan memberikan kontribusi baik dari sisi praktis maupun teoritis, dengan beberapa temuan yang dapat diterapkan. Manfaat yang diharapkan dari penelitian ini antara lain :

1. Bagi peneliti, memperoleh pemahaman yang lebih mendalam terkait penerapan metode *IndoBERT* dalam mengklasifikasikan sentimen.
2. Bagi peneliti lain maupun akademisi, temuan penelitian ini dapat dijadikan acuan atau referensi bagi penelitian mendatang serta menjadi dasar bagi penelitian lanjutan yang lebih akurat.

