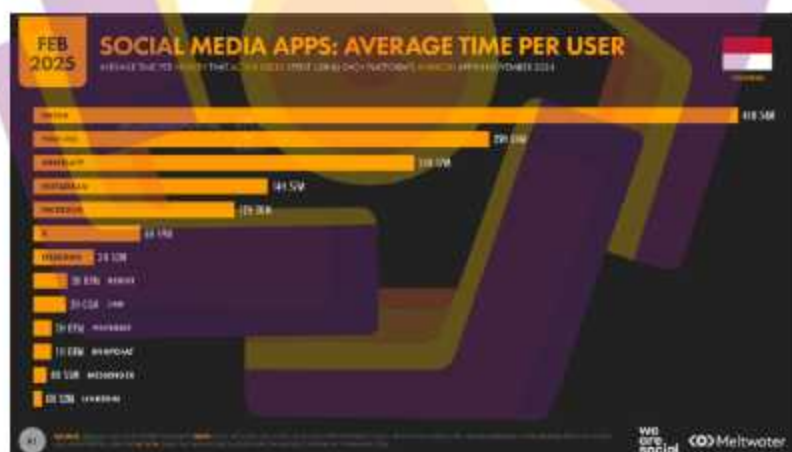


BAB I PENDAHULUAN

1.1 Latar Belakang

Media sosial TikTok merupakan platform berbagi video pendek yang dikembangkan oleh perusahaan teknologi ByteDance asal Tiongkok dan dirilis secara global pada tahun 2018. TikTok berkembang pesat dan memiliki jumlah pengguna aktif yang besar sehingga konten video dapat dengan cepat tersebar dan mudah dikonsumsi oleh pengguna aktif di berbagai kelompok usia [1]. Setiap konten video yang diunggah memungkinkan pengguna untuk memberikan komentar berupa opini, kritik, atau dukungan terhadap berbagai isu populer [2]. Pola interaksi berbasis komentar tersebut sering kali mencerminkan reaksi spontan masyarakat terhadap berbagai isu viral, karena TikTok turut menjadi sarana pertukaran informasi, ide, dan pengalaman antar pengguna [3].



Gambar 1. 1 Grafik pertumbuhan pengguna TikTok di Indonesia

Salah satu topik yang banyak menarik perhatian publik di media sosial adalah Whoosh, yaitu layanan kereta cepat Jakarta-Bandung yang menjadi moda transportasi kereta cepat pertama di Indonesia dan Asia Tenggara. Proyek ini merupakan bagian dari proyek strategi nasional yang diresmikan pada tahun 2023

dan dioperasikan oleh PT Kereta Cepat Indonesia China (KCIC). Kehadiran Whoosh membawa perubahan signifikan dalam sistem transportasi nasional, terutama untuk dalam efisiensi waktu tempuh dan modernisasi layanan transportasi publik [4].

Sejak awal perencanaan hingga masa operasionalnya, Whoosh telah menjadi topik perbincangan publik yang luas dan menimbulkan beragam opini di masyarakat. Layanan ini telah mengangkut jutaan penumpang sejak beroperasi, menunjukkan tingginya antusiasme publik terhadap moda transportasi tersebut [5]. Opini dan diskusi tentang Whoosh tidak hanya disampaikan melalui media konvensional, tetapi juga tersebar secara luas di media sosial seperti TikTok dan Detik.com yang memperlihatkan berbagai respons netizen terhadap berbagai aspek operasionalnya [6]. Selain itu, proyek kereta cepat Whoosh juga mendapat perhatian media internasional sebagai kereta cepat pertama di Asia Tenggara, yang turut memperluas cakupan pemberitaan dan opini publik [7]. Tingginya intensitas komentar serta beragam pandangan masyarakat terhadap Whoosh menjadikan analisis sentimen sebagai pendekatan yang relevan untuk memahami kecenderungan opini publik secara sistematis.

Analisis sentimen merupakan pendekatan komputasi dalam bidang pemrosesan bahasa alami (Natural Language Processing) yang digunakan untuk mengelompokkan opini publik ke dalam kategori umum, seperti sentimen positif, negatif, dan netral. Teknik ini banyak dimanfaatkan untuk memahami persepsi masyarakat terhadap fenomena publik, kebijakan, produk, maupun layanan. Dalam konteks TikTok, analisis sentimen memiliki tingkat relevansi yang tinggi karena mampu mengolah data opini publik dalam jumlah besar secara otomatis dan objektif, sehingga kecenderungan sikap masyarakat terhadap suatu isu dapat diidentifikasi dengan lebih akurat [8].

Perkembangan teknologi deep learning turut mendorong peningkatan performa analisis sentimen, khususnya dengan hadirnya model Bidirectional Encoder Representations from Transformers (BERT). Model ini dirancang untuk memahami konteks bahasa secara dua arah melalui mekanisme transformer,

sehingga mampu menangkap makna kata berdasarkan hubungan dengan kata-kata lain dalam suatu kalimat. Keunggulan tersebut menjadikan BERT lebih efektif dalam mengolah teks pendek dan tidak baku, seperti komentar media sosial. Penelitian-penelitian terkini menunjukkan bahwa varian BERT untuk Bahasa Indonesia, seperti IndoBERT dan IndoBERTweet, memiliki tingkat akurasi yang lebih tinggi dibandingkan pendekatan machine learning konvensional maupun model deep learning berbasis Recurrent Neural Network (RNN) dan Long Short-Term Memory (LSTM)[9].

Karakteristik komentar TikTok mengenai Whoosh yang bersifat informal, ambigu, dan sering kali mengandung ekspresi emosional menuntut penggunaan model analisis sentimen yang mampu memahami konteks bahasa secara mendalam. Oleh karena itu, penerapan model BERT dinilai tepat untuk menganalisis opini publik terhadap Whoosh secara komputasional, karena kemampuannya dalam menangkap makna kontekstual teks secara lebih akurat.

Berdasarkan uraian tersebut, penelitian ini dilakukan untuk menganalisis sentimen komentar pengguna TikTok terhadap video Whoosh menggunakan model BERT. Penelitian ini difokuskan pada proses pengumpulan data komentar, pra pemrosesan teks, pelabelan sentimen, serta penerapan dan evaluasi model BERT, yang juga dibandingkan dengan metode analisis sentimen lainnya. Hasil penelitian diharapkan dapat memberikan gambaran ilmiah mengenai efektivitas model BERT dalam mengklasifikasikan sentimen pada platform TikTok, sekaligus memberikan pemahaman yang lebih objektif terhadap persepsi publik mengenai operasional kereta cepat Whoosh.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka dihasilkan rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kinerja model BERT dalam mengklasifikasikan sentimen positif, negatif, dan netral pada data komentar TikTok mengenai Whoosh yang bersifat heterogen?

2. Bagaimana distribusi dan kecenderungan sentimen publik terhadap operasional kereta cepat Whoosh berdasarkan hasil analisis pada berbagai macam konten video di platform TikTok?
3. Seberapa efektif penggunaan model BERT dalam menangani karakteristik bahasa nonformal dan noise pada data hasil scraping komentar TikTok dibandingkan dengan metode konvensional?

1.3 Batasan Masalah

Untuk menjaga agar penelitian ini tetap fokus pada tujuan yang ditetapkan dan menghindari pembahasan yang terlalu luas, maka ditetapkan batasan masalah sebagai berikut:

1. Data yang digunakan terbatas pada komentar teks yang diperoleh dari hasil *scraping* pada platform TikTok dengan kata kunci pencarian “Whoosh” dan tidak mencakup data dari platform media sosial lainnya.
2. Komentar yang dianalisis hanya menggunakan Bahasa Indonesia. Penggunaan bahasa asing atau bahasa daerah yang dominan akan diabaikan atau dikeluarkan dari *dataset*.
3. Sentimen dibatasi ke dalam tiga kategori utama yaitu positif, negatif, dan netral. Penelitian ini tidak melakukan analisis aspek (*aspect-based*) yang lebih mendalam pada setiap komponen pelayanan Whoosh.
4. Atribut profil pengguna seperti nama akun, jumlah pengikut (*followers*), foto profil, serta lokasi geografis pengguna dianggap sebagai parameter konstanta yang diabaikan karena diperkirakan tidak berpengaruh signifikan terhadap hasil klasifikasi sentimen teks.
5. Penelitian hanya menganalisis teks komentar dan mengabaikan elemen visual maupun audio dari video TikTok yang menjadi sumber data.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah menghasilkan analisis sentimen otomatis yang akurat terhadap opini publik mengenai Whoosh di platform TikTok menggunakan model BERT. Upaya ini dilakukan untuk menangani karakteristik

bahasa non-formal pada data hasil scraping. Selain itu, penelitian ini bertujuan menghitung distribusi sentimen guna mengetahui kecenderungan respons masyarakat terhadap operasional kereta cepat tersebut secara terstruktur dan sistematis.

1.5 Manfaat Penelitian

Manfaat yang dihasilkan dari penelitian adalah sebagai berikut:

1. Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan ilmu pengetahuan di bidang Pemrosesan Bahasa Alami (Natural Language Processing), khususnya dalam membuktikan efektivitas model BERT untuk mengolah data teks non-formal berbahasa Indonesia.
2. Secara praktis, penelitian ini memberikan manfaat bagi pihak pengelola operasional Whoosh sebagai bahan evaluasi untuk memahami persepsi, keluhan, maupun apresiasi masyarakat secara objektif. Bagi organisasi atau pengambil kebijakan, sistem analisis sentimen otomatis ini dapat dimanfaatkan sebagai alat pemantauan dinamika opini publik secara real-time. Secara teknis, kerangka kerja (framework) yang dihasilkan dapat digunakan kembali untuk menganalisis isu strategis lainnya pada platform TikTok dengan lebih efisien.

1.6 Sistematika Penulisan

Penelitian ini menerapkan sistematika penulisan supaya penyusunan penulisan menjadi lebih teratur dan dapat dipahami dengan mudah, sistematika penulisan yang terdapat pada penelitian adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi pengantar mengenai penelitian yang dilakukan. Isi bab ini mencakup latar belakang masalah yang menjelaskan urgensi analisis sentimen publik terhadap operasional Kereta Cepat Whoosh di platform TikTok serta potensi model berbasis Transformer (BERT) untuk mengolah karakteristik bahasanya, Rumusan Masalah, Batasan Masalah yang mendefinisikan ruang lingkup data komentar dan metode, Tujuan Penelitian secara rinci, Manfaat Penelitian (teoretis dan praktis), dan

diakhiri dengan sistematika penulisan ini.

BAB II LANDASAN TEORI

Bab ini menyajikan dua bagian utama, yaitu tinjauan pustaka dan dasar teori. Tinjauan Pustaka akan memaparkan state-of-the-art atau penelitian terdahulu yang relevan terkait analisis sentimen menggunakan model BERT dan metode klasifikasi teks lainnya. Dasar teori akan menjabarkan konsep-konsep kunci yang digunakan, meliputi teori mengenai Natural Language Processing (NLP), karakteristik media sosial TikTok, operasional Kereta Cepat Whoosh, prinsip kerja pemrosesan teks, serta arsitektur teoretis dari algoritma BERT (khususnya IndoBERT), Bidirectional Long Short-Term Memory (BiLSTM), dan Support Vector Machine (SVM).

BAB III METODE PENELITIAN

Bab ini memaparkan secara rinci langkah-langkah teknis penelitian. Di dalamnya terdapat penjelasan mengenai objek penelitian (deskripsi dataset komentar TikTok mengenai Whoosh), alur kerja penelitian (framework), serta metodologi rinci. Metodologi ini mencakup tahapan pengumpulan data (scraping), pra-pemrosesan teks (preprocessing/cleansing), pelabelan kelas sentimen (Positif, Negatif, Netral), pembagian data, desain eksperimen model utama (IndoBERT dengan pendekatan Focal Loss dan layer freezing) beserta model pembanding (BiLSTM dan SVM), hingga tahap evaluasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini merupakan inti dari penelitian yang menyajikan seluruh temuan dari eksperimen komputasi. Bagian pertama berfokus pada hasil analisis distribusi data sentimen publik, yang dilanjutkan dengan hasil analisis komparatif kinerja model melalui tabel evaluasi metrik (Precision, Recall, F1-Score, Macro Average) dari ketiga algoritma yang diuji. Seluruh temuan tersebut, termasuk fenomena pelatihan model dan kendala ketidakseimbangan kelas (class imbalance) akan dibahas dan diinterpretasikan secara mendalam untuk menjawab performa akhir model.

BAB V PENUTUP

Bab terakhir ini berisi Kesimpulan yang dirangkum dari hasil pembahasan pada bab

sebelumnya dan secara lugas menjawab poin-poin pada rumusan masalah mengenai distribusi sentimen Whoosh serta efektivitas arsitektur model. Selain itu, bab ini juga memuat Saran yang berisi rekomendasi konstruktif untuk pengembangan penelitian di masa mendatang, baik dari segi penanganan ketidakseimbangan data (class imbalance) maupun eksplorasi arsitektur model NLP lainnya.

