

Implementasi Algoritma BERT pada Sistem Analisis Sentimen Berbasis Web terhadap Kebijakan Elektrifikasi Sepeda Listrik

Implementation of the BERT Algorithm in a Web-Based Sentiment Analysis System regarding the Electric Bicycle Electrification Policy in Indonesia

Winarti Ulandari¹, Aries Dwi Indriyanti²

^{1,2}Sistem Informasi, Universitas Hasyim Asy'ari

E-mail: ¹winartiulandari@mhs.unhasy.ac.id, ²ariesdwiindriyanti@unhasy.ac.id

Abstrak

Pesatnya perkembangan media sosial mendorong lonjakan opini publik di berbagai platform digital, salah satunya *YouTube*. Kebijakan elektrifikasi transportasi, khususnya penggunaan sepeda listrik di Indonesia, menuai beragam respons dari masyarakat. Penelitian ini bertujuan membangun sistem analisis sentimen berbasis web yang mengklasifikasikan komentar *YouTube* terkait kebijakan tersebut menggunakan algoritma BERT (*Bidirectional Encoder Representations from Transformers*). Sistem dikembangkan dengan model *waterfall* dan mencakup *preprocessing* teks (pembersihan teks, *case folding*, normalisasi, dan tokenisasi), klasifikasi sentimen ke dalam kategori positif, netral, dan negatif, serta visualisasi hasil secara interaktif. *Dataset* terdiri atas 1.000 komentar dari 8 video *YouTube* yang dikumpulkan melalui *web scraping* pada periode 2022, 2023 dan 2024. Hasil pengujian menunjukkan model BERT mencapai akurasi 91,50%, dengan *macro precision* 88,34%, *macro recall* 90,80%, dan *macro F1-Score* 89,55%. Distribusi sentimen hasil klasifikasi terdiri atas 201 komentar positif (20,1%), 645 netral (64,5%), dan 154 negatif (15,4%). Sistem yang dibangun mengintegrasikan alur kerja mulai dari unggah *dataset*, *preprocessing*, analisis sentimen, hingga visualisasi secara otomatis, sehingga dapat mendukung pengambilan keputusan berbasis data bagi pemangku kebijakan.

Kata kunci: Analisis Sentimen, BERT, *YouTube*, Sepeda Listrik, Klasifikasi Teks

Abstract

The rapid growth of social media has driven a surge in public opinion across digital platforms, including *YouTube*. The transportation electrification policy, particularly the adoption of electric bicycles in Indonesia, has elicited a wide range of public responses. This study aims to build a web-based sentiment analysis system that classifies *YouTube* comments related to this policy using the BERT (*Bidirectional Encoder Representations from Transformers*) algorithm. The system was developed using the *waterfall* model and covers text preprocessing (text cleaning, *case folding*, using the *waterfall* model and covers text preprocessing (text cleaning, *case folding*,

normalization, and tokenization), sentiment classification into positive, neutral, and negative categories, and interactive visualization of the results. The dataset consists of 1,000 comments from 8 YouTube videos collected through web scraping period of 2022, 2023 and 2024. The BERT model achieved an accuracy of 91.50%, with a macro precision of 88.34%, a macro recall of 90.80%, and a macro F1-Score of 89.55%. The classified sentiment distribution comprises 201 positive comments (20.1%), 645 neutral comments (64.5%), and 154 negative comments (15.4%). The web-based system integrates the entire workflow, from dataset upload and preprocessing to sentiment analysis and visualization, and can therefore support data-driven decision-making for policymakers.

Keywords: Sentiment Analysis, BERT, YouTube, Electric Bicycle, Text Classification

1. PENDAHULUAN

Dalam era digital yang terus berkembang, media sosial telah bertransformasi menjadi kanal utama bagi masyarakat dalam mengekspresikan pendapat serta merespons berbagai isu, termasuk kebijakan yang dikeluarkan oleh pemerintah. *YouTube* merupakan salah satu jejaring sosial berbasis video terpopuler di dunia, dengan lebih dari dua miliar pengguna aktif bulanan, menjadi ruang diskusi yang sangat dinamis di Indonesia. Setiap konten video yang diunggah berpotensi memicu ribuan bahkan jutaan komentar yang mencerminkan beragam sikap dan pandangan masyarakat terhadap suatu topik [1].

Salah satu kebijakan yang tengah mendapat perhatian luas adalah elektrifikasi transportasi, khususnya terkait penggunaan sepeda listrik di Indonesia. Kementerian Energi dan Sumber Daya Mineral Republik Indonesia telah mendorong percepatan elektrifikasi transportasi sebagai komponen strategis dalam transisi energi nasional sekaligus upaya penurunan emisi gas rumah kaca menuju target *net zero emission* pada tahun 2060 sebagaimana tertuang dalam Rencana Strategis Kementerian ESDM. Kebijakan ini menimbulkan beragam tanggapan dari masyarakat, mulai dari dukungan atas aspek ramah lingkungan dan efisiensi biaya operasional, hingga kritik terhadap tingginya harga kendaraan dan terbatasnya infrastruktur pengisian daya.

Analisis sentimen merupakan salah satu teknik dalam bidang *Natural Language Processing* (NLP) yang digunakan untuk mengenali, mengekstrak, dan mengkuantifikasi informasi subjektif secara otomatis dari teks. Dengan memanfaatkan analisis sentimen, opini publik yang tersebar dalam jumlah besar di media sosial dapat diproses secara efisien guna memahami persepsi masyarakat terhadap suatu isu atau kebijakan [2]. Informasi ini sangat bernilai bagi pemangku kebijakan dalam merumuskan strategi komunikasi dan evaluasi yang lebih tepat sasaran.

Pendekatan klasifikasi teks berbasis metode tradisional seperti *Naïve Bayes* dan *Support Vector Machine* (SVM) memiliki keterbatasan dalam memahami konteks bahasa secara menyeluruh, terutama pada teks informal dari media sosial yang kaya singkatan dan variasi bahasa nonstandar. Seiring dengan perkembangan teknologi

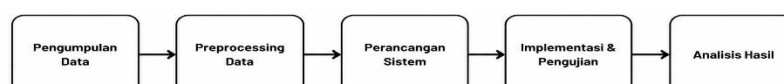
deep learning, model berbasis *Transformer* seperti BERT (*Bidirectional Encoder Representations from Transformers*) [3] hadir sebagai alternatif yang lebih andal. BERT mampu memahami konteks suatu kata secara dua arah sehingga menghasilkan representasi teks yang lebih kontekstual dan akurat, menjadikannya unggul dalam berbagai tugas NLP termasuk klasifikasi sentimen [4].

Beberapa penelitian terdahulu telah membuktikan efektivitas BERT dalam analisis sentimen, baik pada skala internasional maupun dalam konteks bahasa Indonesia. Wu *et al.* [5] meneliti penerapan BERT untuk analisis sentimen di media sosial dan menunjukkan bahwa BERT mengungguli model *machine learning* tradisional maupun pendekatan *deep learning* lainnya dalam memahami nuansa bahasa di platform media sosial. Hidayat dkk. [6] menunjukkan bahwa algoritma BERT mengungguli metode konvensional dalam analisis sentimen data media sosial berbahasa Indonesia. Sholihah dkk. [7] membuktikan efektivitas BERT dalam menganalisis sentimen komentar *YouTube* terkait kinerja KPU pasca-Pemilu 2024 dengan akurasi yang tinggi. Bastian dan Perdana [8] membandingkan metode VADER dan BERT untuk analisis sentimen komentar *YouTube* terkait pidato presiden dan menemukan bahwa BERT menghasilkan klasifikasi yang lebih akurat.

Pada topik sepeda listrik, Lestari dan Isnain [9] telah menganalisis sentimen masyarakat terhadap penggunaan sepeda listrik pada anak-anak di media sosial X menggunakan metode SVM. Meskipun demikian, penelitian mengenai analisis sentimen masyarakat terhadap kebijakan elektrifikasi transportasi sepeda listrik di Indonesia yang menggunakan BERT dan diimplementasikan dalam sistem berbasis web terintegrasi masih sangat terbatas. Oleh karena itu, penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem analisis sentimen berbasis *web* menggunakan algoritma BERT guna mengklasifikasikan komentar *YouTube* terkait kebijakan tersebut ke dalam kategori positif, netral, dan negatif. Nilai lebih penelitian ini terletak pada objek kajian yang spesifik, yaitu kebijakan elektrifikasi sepeda listrik di Indonesia, serta pada implementasi seluruh alur kerja (unggah *dataset*, *preprocessing*, analisis sentimen, visualisasi, dan manajemen data hasil analisis) dalam satu sistem terintegrasi, sehingga dapat mendukung pengambilan keputusan berbasis data bagi pemangku kebijakan di bidang transportasi.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan model pengembangan sistem berbasis metode *waterfall*. Metode *waterfall* dipilih karena pengembangannya bersifat terstruktur dan berurutan, di mana setiap tahapan harus diselesaikan secara tuntas sebelum dapat melanjutkan ke tahapan berikutnya, sehingga meminimalkan risiko kesalahan di tengah proses pengembangan akibat perubahan kebutuhan sistem yang mendadak. Metode ini sangat sesuai penelitian ini karena seluruh spesifikasi kebutuhan fungsional sistem dan alur algoritma BERT telah diidentifikasi secara matang sejak awal. Alur penelitian yang ditempuh terdiri atas lima tahapan utama, yaitu: (1) Pengumpulan Data, (2) *Preprocessing* Data, (3) Perancangan Sistem, (4) Implementasi dan Pengujian, serta (5) Analisis Hasil. Gambar 2.1 mengilustrasikan alur tahapan penelitian secara keseluruhan.



Gambar 2.1 Metodologi Penelitian

2.1 Pengumpulan Data

Data penelitian dikumpulkan dari platform *YouTube* menggunakan teknik *web scraping* pada periode 2022, 2023 dan 2024. Data berupa komentar pengguna pada video *YouTube* yang membahas kebijakan elektrifikasi transportasi sepeda listrik di Indonesia. Proses *scraping* dilakukan menggunakan layanan Outscraper, yaitu layanan *scraping* berbasis *cloud* yang menerima URL video *YouTube* sebagai masukan dan mengembalikan komentar dalam bentuk berkas tabular. Parameter pengumpulan data dikonfigurasi dengan mengurutkan komentar berdasarkan komentar terpopuler dan membatasi pengambilan maksimal sebanyak 150 komentar per video untuk menjaga keseimbangan distribusi data. Sebanyak 8 video *YouTube* digunakan sebagai sumber data. Video diperoleh melalui pencarian dengan kata kunci terkait kebijakan sepeda listrik di Indonesia, seperti “sepeda listrik Indonesia”, “kebijakan sepeda listrik”, dan “subsidi sepeda listrik”. Video dipilih berdasarkan relevansi isi video dengan topik kebijakan elektrifikasi transportasi sepeda listrik. Setiap video yang ditemukan dan dinilai memiliki kesesuaian konten secara langsung dengan topik kebijakan sepeda listrik serta memiliki kolom komentar yang aktif, langsung diambil dengan cara menyalin tautan (*link*) video tersebut. Melalui pendekatan ini terpilih 8 tautan video spesifik yang diunggah dalam rentang tahun 2022, 2023 dan 2024 untuk dijadikan sebagai sumber data.

Komentar yang diperoleh dari kedelapan video tersebut dikumpulkan dan diperiksa sebelum digunakan sebagai *dataset* penelitian, sehingga diperoleh sebanyak 1.000 komentar. Proses seleksi dilakukan dengan memfilter dan menghapus komentar duplikat, berupa *spam*, ataupun komentar yang berisi tautan (*link*) dan emoji tanpa teks. Data komentar disimpan dalam format Microsoft Excel (.xlsx) dan dikonversi ke format CSV agar dapat diunggah ke sistem untuk diproses lebih lanjut menggunakan algoritma BERT. Proses penentuan label sentimen (*ground truth*) dilakukan secara manual oleh peneliti secara langsung di dalam halaman manajemen data pada sistem antarmuka berbasis *web* yang telah dibangun. Setiap komentar diberi label sentimen ke dalam tiga kategori, yaitu positif, netral, dan negatif berdasarkan pedoman penilaian kontekstual terhadap isi opini komentar. Total *dataset* berjumlah 1.000 komentar dengan distribusi awal: 215 komentar positif, 638 komentar netral, dan 147 komentar negatif (Tabel 2.1).

Tabel 2.1 Distribusi Data Berdasarkan Kategori Sentimen

No	Kategori Sentimen	Jumlah Data	Persentase (%)
1	Positif	215	21,5%
2	Netral	638	63,8%
3	Negatif	147	14,7%
	Total	1.000	100%

2.2 Preprocessing Data

Tahap *preprocessing* dilakukan untuk menyiapkan data teks agar memiliki struktur yang bersih dan seragam sebelum diproses lebih lanjut oleh model BERT [10], [11]. Tabel 2.2 mengilustrasikan contoh hasil dari setiap tahapan *preprocessing* yang diterapkan.

Tabel 2.2 Contoh Hasil Tahapan Preprocessing

Tahapan	Input	Output
Teks Asli	Harga sepeda listrik MAHAL bgt 🚫!! @user123	-
<i>Cleaning Text</i>	-	Harga sepeda listrik MAHAL bgt
<i>Case Folding</i>	Harga sepeda listrik MAHAL bgt	harga sepeda listrik mahal bgt
Normalisasi	harga sepeda listrik mahal bgt	harga sepeda listrik mahal banget
Tokenisasi	harga sepeda listrik mahal banget	[harga, sepeda, listrik, mahal, banget]

2.3 Algoritma BERT

Algoritma BERT (*Bidirectional Encoder Representations from Transformers*) merupakan model bahasa pra-latih yang dikembangkan dan dibangun di atas arsitektur *Transformer* [3]. Keunggulan utama BERT terletak pada kemampuannya memproses teks secara dua arah (*bidirectional*), sehingga mampu memahami konteks suatu kata berdasarkan hubungannya dengan kata-kata yang mendahului maupun mengikutinya secara bersamaan dalam satu kalimat. Tahapan kerja BERT pada penelitian ini adalah sebagai berikut.

1. *Token Classification & Embedding*: Setiap kata pada teks yang telah melalui *preprocessing* dikonversikan menjadi token khusus. Token [CLS] ditambahkan di awal kalimat sebagai representasi agregat seluruh teks, dan token [SEP] sebagai pemisah. Selanjutnya setiap token direpresentasikan sebagai vektor numerik (*embedding*) berdimensi tinggi.
2. *Contextual Processing (Self-Attention)*: Lapisan *Self-Attention* pada arsitektur *Transformer* menghitung bobot hubungan semantik antar-kata secara simultan. Melalui mekanisme ini, model BERT dapat menangkap makna implisit dan konteks kalimat secara lebih holistik dibandingkan model sekuensial konvensional.
3. *Output Layer (Softmax)*: Representasi fitur dari token [CLS] diteruskan ke lapisan *Fully Connected (Dense Layer)* dengan fungsi aktivasi *Softmax* untuk menghasilkan distribusi probabilitas ke dalam tiga kelas sentimen target: Positif, Netral, dan Negatif. Kelas dengan nilai probabilitas tertinggi ditetapkan sebagai hasil klasifikasi.

Model BERT yang digunakan pada penelitian ini Adalah varian *IndoBERT-base* (*indobenchmark/indobert-base-p1*). Model tersebut dipilih karena telah dilatih pada korpus berskala besar dalam bahasa Inonesia, sehingga memiliki pemahaman bahasa konseptual dan kontekstual yang baik pada data berbahasa Indonesia. Konfigurasi model dan skema validasi dirangkum pada Tabel 2.3, sedangkan prosedur klasifikasi sentimen secara keseluruhan disajikan pada Algoritma 1.

Tabel 2.3 Konfigurasi Model dan Skema Validasi

Parameter	Nilai
Model pra-latih	Indobenchmark/indobert-base-p1
Sumber model	Hugging Face Model Hub
Kelas keluaran	Positif, Netral, Negatif (3 kelas)
Panjang urutan maksimum (<i>max sequence length</i>)	128
Ukuran <i>batch</i>	32
Laju pembelajaran (<i>learning rate</i>)	2e-5
Jumlah <i>epoch</i>	3
<i>Optimizer</i>	AdamW
Skema validasi	<i>Train-test split</i>

Algoritma 1 Prosedur Klasifikasi Sentimen Komentar YouTube

```

Masukan : D = himpunan komentar YouTube (berkas CSV)
Keluaran: label sentimen tiap komentar  $\in$  {positif, netral, negatif}
1. muat D dari berkas CSV dan simpan ke basis data MySQL
2. untuk setiap komentar  $c \in D$  lakukan
3.    $c \leftarrow$  hapus URL, mention, emoji, dan tanda baca berlebihan
4.    $c \leftarrow$  ubah semua huruf menjadi huruf kecil (case folding)
5.    $c \leftarrow$  ganti kata tidak baku dengan bentuk bakunya (normalisasi)
6.    $T \leftarrow$  tokenisasi( $c$ )
7. selesai
8. bagi D menjadi data latih dan data uji dengan metode train-test split
9. muat BERT pra-latih + lapisan Dense + Softmax (3 kelas)
10. latih (fine-tuning) pada data latih
11. untuk setiap komentar uji lakukan
12.  $y \leftarrow \text{argmax Softmax BERT}([\text{CLS}] \ T \ [\text{SEP}])$ 
13. selesai
14. hitung Accuracy, Precision, Recall, F1-Score (macro)
15. tampilkan distribusi sentimen dan simpan hasil ke basis data
    
```

2.4 Perancangan Sistem

Sistem analisis sentimen berbasis web dirancang dengan menggunakan pemodelan *Unified Modeling Language* (UML) yang mencakup *Use Case Diagram*, *Activity Diagram*, dan *Class Diagram*. Basis data dirancang menggunakan MySQL dengan enam tabel utama, yakni: Admin, Dataset, Komentar, Preprocessing, SentimenAnalysis, dan VisualisasiExport.

Stack teknologi yang dimanfaatkan dalam pengembangan sistem meliputi PHP dan JavaScript untuk pengembangan antarmuka web (*front-end* dan *back-end*); MySQL sebagai sistem manajemen basis data; serta CSS untuk tampilan antarmuka yang responsif dan ramah pengguna.

2.5 Evaluasi Model

Evaluasi model dilakukan menggunakan empat metrik, yaitu Precision, Recall, F1-Score, dan Accuracy. Perhitungan didasarkan pada matriks konfusi (*Confusion Matrix*) yang mencatat nilai True Positive (TP), *False Positive* (FP), dan *False Negative* (FN) untuk setiap kelas sentimen. Formula yang digunakan adalah sebagai berikut:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{Accuracy} = (\text{Total Prediksi Benar}) / (\text{Total Data}) \times 100\%$$

Nilai akhir dilaporkan menggunakan *Macro Average*, yaitu rata-rata aritmatika dari nilai masing-masing kelas tanpa mempertimbangkan ketidakseimbangan distribusi data antarkelas. Pengujian fungsional sistem memakai metode *black-box testing* untuk memastikan bahwa setiap fitur berfungsi sesuai dengan spesifikasi yang telah ditetapkan.

Validasi model dilakukan dengan metode *train-test split* dengan memisahkan dataset ke dalam porsi data latih (*training data*) dan data uji (*testing data*). Seluruh metrik dihitung dari prediksi model pada 1.000 data uji.

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Sistem

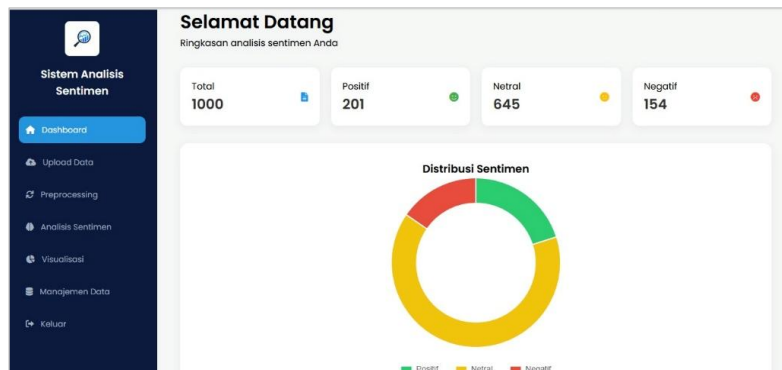
Sistem analisis sentimen berbasis web berhasil dikembangkan menggunakan PHP, JavaScript, CSS, dan MySQL. Sistem mencakup tujuh modul utama yang saling terintegrasi: *Login*, *Dashboard*, *Upload Dataset*, *Preprocessing*, Analisis Sentimen, Visualisasi, dan Manajemen Data. Ringkasan fungsi setiap modul dan contoh tampilan antarmuka utama disajikan pada Subbab 3.2.

3.2 Tampilan Antarmuka Sistem

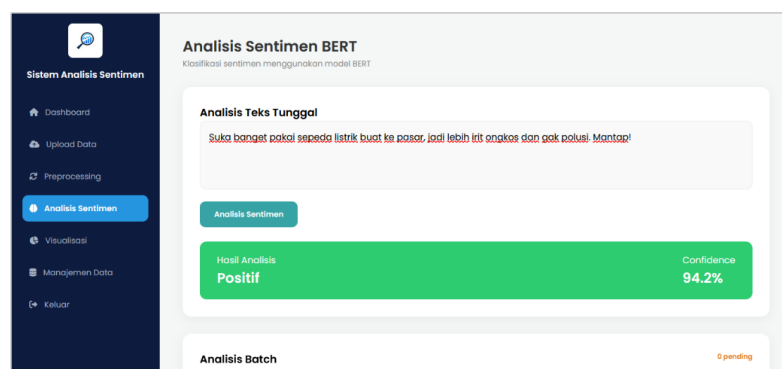
Antarmuka sistem terdiri atas tujuh modul. Modul *Login* memvalidasi kredensial admin terhadap basis data sehingga hanya pengguna terdaftar yang dapat mengakses sistem. Modul *Upload Dataset* menerima berkas komentar berformat CSV, menampilkan pratinjau data, dan menolak berkas non-CSV atau *dataset* kosong. Modul *Preprocessing* menjalankan pembersihan teks, *case folding*, normalisasi, dan tokenisasi secara berurutan serta menampilkan perbandingan teks asli dan teks hasil pemrosesan. Modul Analisis Sentimen mengklasifikasikan komentar dengan BERT, baik pada mode analisis tunggal maupun *batch*, disertai *confidence score*.

Modul Visualisasi menyajikan diagram batang dan diagram lingkaran, sedangkan modul Manajemen Data mendukung peninjauan, penghapusan, dan ekspor hasil analisis. Modul *Dashboard* merangkum jumlah komentar dan distribusi sentimen.

Agar pembahasan berfokus pada hasil penelitian, hanya tiga tampilan yang paling relevan yang disajikan. Gambar 3.1 menunjukkan bahwa *Dashboard* merangkum jumlah komentar dan distribusi sentimen dalam bentuk angka dan grafik sehingga pengguna memperoleh gambaran umum tanpa membuka data mentah. Gambar 3.2 memperlihatkan hasil klasifikasi per komentar beserta *confidence score* pada mode analisis tunggal dan *batch*, sedangkan Gambar 3.3 menampilkan proporsi sentimen dalam diagram batang dan diagram lingkaran yang memudahkan pemangku kepentingan membandingkan proporsi antarkelas.



Gambar 3.1 Halaman Dashboard



Gambar 3.2 Halaman Hasil Analisis Sentimen



Gambar 3.3 Halaman Visualisasi

3.3 Hasil Analisis Sentimen

Berdasarkan proses klasifikasi terhadap 1.000 komentar *YouTube* menggunakan model BERT, diperoleh distribusi sentimen sebagai berikut: 201 komentar positif (20,1%), 645 komentar netral (64,5%), dan 154 komentar negatif (15,4%). Rincian hasil klasifikasi disajikan pada Tabel 3.1.

Tabel 3.1 Distribusi Hasil Klasifikasi Sentimen

Kelas Sentimen	Jumlah Komentar	Persentase	Deskripsi Umum
Positif	201	20,1%	Dukungan terhadap manfaat lingkungan dan efisiensi
Netral	645	64,5%	Pernyataan fakta, pertanyaan, opini umum
Negatif	154	15,4%	Keluhan harga mahal dan keterbatasan infrastruktur
Total	1.000	100%	

Distribusi pada Tabel 3.1 menunjukkan dominasi komentar netral (64,5%), yang umumnya berupa pernyataan fakta, pertanyaan, atau opini umum tanpa keberpihakan yang tegas terhadap kebijakan. Komentar positif (20,1%) berisi dukungan terhadap manfaat lingkungan dan efisiensi, sedangkan komentar negatif (15,4%) berisi keluhan mengenai harga yang mahal dan keterbatasan infrastruktur; jumlah komentar positif sekitar 1,3 kali jumlah komentar negatif. Temuan ini sejalan dengan dua sisi tanggapan masyarakat yang diuraikan pada Pendahuluan. Namun, distribusi tersebut berasal dari 8 video pada periode Januari hingga Maret 2024 sehingga belum tentu mewakili seluruh opini publik terhadap kebijakan ini.

3.4 Evaluasi Model BERT

Evaluasi kinerja model BERT dilakukan menggunakan perhitungan *Precision*, *Recall*, dan *F1-Score* berdasarkan *confusion matrix* dengan skema validasi yang dijelaskan pada Subbab 2.5 menggunakan metode *train-test split* dengan total 1.000 data uji. Tabel 3.2 menampilkan nilai TP, FP, dan FN beserta metrik evaluasi untuk setiap kelas sentimen.

Tabel 3.2 Hasil Evaluasi Model BERT per Kelas Sentimen

Kelas Sentimen	TP	FP	FN	Precision	Recall	F1-Score
Positif	182	36	19	83,5%	90,5%	86,8%
Netral	595	26	50	95,8%	92,2%	94,0%
Negatif	138	23	16	85,7%	89,6%	87,6%
Macro Avg	—	—	—	88,34%	90,80%	89,55%

Berdasarkan Tabel 3.2, model BERT menghasilkan performa terbaik pada kelas netral dengan F1-Score 94,0%, disusul kelas negatif (87,6%) dan kelas positif (86,8%). Kelas netral memiliki *Precision* (95,8%) dan *Recall* (92,2%) tertinggi, yang kemungkinan dipengaruhi oleh proporsinya yang dominan dalam *dataset* (sekitar 64%). Kelas positif memiliki *Precision* terendah (83,5%) karena jumlah *false positive*-nya paling banyak (36 komentar), yang berarti sebagian komentar dari kelas lain kemungkinan besar netral, diprediksi sebagai positif. Sebaliknya, kelas netral memiliki *false negative* terbanyak (50 komentar), yaitu komentar netral yang salah diklasifikasikan ke kelas positif atau negatif. Pola ini wajar pada klasifikasi tiga kelas karena batas antara opini netral dan opini yang bernada mendukung atau mengkritik sering kali tipis. Akurasi keseluruhan model mencapai 91,50% (915 dari 1.000 data diklasifikasikan dengan benar), dengan *Macro Precision* 88,34%, *Macro Recall* 90,80%, dan *Macro F1-Score* 89,55%. Karena distribusi kelas tidak seimbang, *macro average* digunakan agar setiap kelas berkontribusi setara terhadap nilai akhir. Sebagai pembandingan, memprediksi seluruh komentar sebagai kelas mayoritas (netral) hanya akan menghasilkan akurasi sekitar 64%, sehingga akurasi model berada jauh di atas *baseline* tersebut. Hasil ini menunjukkan bahwa kemampuan BERT dalam memproses teks secara *bidirectional* memberikan kinerja yang baik pada komentar media sosial berbahasa Indonesia yang beragam dalam *dataset* ini.

Hasil ini sejalan dengan temuan penelitian sebelumnya. Hidayat dkk. [6] melaporkan keunggulan BERT dibandingkan metode konvensional dalam analisis sentimen berbahasa Indonesia, Sholihah dkk. [7] melaporkan akurasi yang tinggi pada komentar *YouTube* berbahasa Indonesia, dan Bastian dan Perdana [8] menemukan BERT lebih akurat daripada VADER. Perbandingan angka antarpenelitian perlu dilakukan dengan hati-hati karena perbedaan *dataset*, topik, dan skema validasi.

Penelitian ini memiliki beberapa keterbatasan. Pertama, *dataset* hanya berjumlah 1.000 komentar dari 8 video pada satu periode pengambilan data sehingga generalisasi ke opini publik yang lebih luas masih terbatas. Kedua, proporsi kelas tidak seimbang dengan dominasi kelas netral. Ketiga, kinerja BERT belum dibandingkan dengan model pembandingan (misalnya SVM atau LSTM) pada *dataset* yang sama sehingga keunggulannya pada kasus ini belum diuji secara langsung.

3.5 Pengujian Sistem (Black-Box Testing)

Pengujian fungsional sistem dilakukan menggunakan metode *black-box testing* yang berfokus pada verifikasi fungsionalitas sistem tanpa menelaah struktur kode internal. Pengujian dilaksanakan terhadap sembilan skenario utama yang mencakup seluruh modul sistem.

Tabel 3.3 Rekapitulasi Hasil Black-box Testing

No	Modul	Skenario Pengujian	Hasil yang Diharapkan	Hasil
1	Login	Masukkan username & password benar	Berhasil diarahkan ke dashboard	Valid
2	Login	Masukkan username/password salah	Tampil error autentikasi	Valid
3	Upload Dataset	Unggah file CSV valid	Dataset tersimpan ke database	Valid
4	Upload Dataset	Unggah file non-CSV	Tampil pesan error format file	Valid
5	Preprocessing	Jalankan preprocessing dataset	Data bersih ditampilkan di tabel	Valid
6	Analisis Sentimen	Analisis tunggal komentar	Klasifikasi & confidence score tampil	Valid
7	Analisis Sentimen	Analisis batch seluruh dataset	Semua data terklasifikasi otomatis	Valid
8	Visualisasi	Tampilkan grafik distribusi sentimen	Grafik batang & lingkaran tampil	Valid
9	Manajemen Data	Kelola & ekspor data hasil analisis	Data berhasil dikelola & diunduh	Valid

Seluruh sembilan skenario pengujian menghasilkan status Valid, yang berarti sistem berfungsi sesuai dengan spesifikasi yang telah ditetapkan. Hasil *black-box testing* ini mengonfirmasi bahwa sistem analisis sentimen berbasis web yang dikembangkan layak digunakan sebagai alat bantu dalam menganalisis opini publik terhadap kebijakan elektrifikasi transportasi sepeda listrik di Indonesia.

KESIMPULAN

Penelitian ini bertujuan merancang dan mengimplementasikan sistem analisis sentimen berbasis web menggunakan algoritma BERT untuk mengklasifikasikan komentar *YouTube* terkait kebijakan elektrifikasi transportasi sepeda listrik di Indonesia. Berdasarkan hasil dan pembahasan, tujuan tersebut telah tercapai. Pertama, sistem berbasis web berhasil dibangun dan mengintegrasikan unggah *dataset*, *preprocessing* teks, klasifikasi sentimen dengan BERT, visualisasi, serta manajemen data dalam satu platform; seluruh sembilan skenario *black-box testing* dinyatakan valid. Kedua, model BERT *IndoBERT-base* yang divalidasi dengan metode *train-test split* menghasilkan akurasi 91,50%, *Macro Precision* 88,34%,

Macro Recall 90,80%, dan Macro F1-Score 89,55%, dengan kinerja terbaik pada kelas netral (F1-Score 94,0%) dan terendah pada kelas positif (86,8%). Ketiga, dari 1.000 komentar yang dianalisis, sentimen netral mendominasi (64,5%), diikuti sentimen positif (20,1%) dan negatif (15,4%), yang mengindikasikan bahwa sebagian besar komentar bersifat informatif atau tidak menunjukkan keberpihakan yang tegas terhadap kebijakan. Dominasi sentimen netral ini dapat menjadi masukan bagi pemangku kebijakan untuk mengkaji efektivitas komunikasi kebijakan, meskipun kesimpulan sebab-akibat memerlukan analisis lebih lanjut.

Hasil penelitian ini terbatas pada 1.000 komentar dari 8 video *YouTube* yang dikumpulkan pada periode 2022, 2023 dan 2024 dan belum dibandingkan dengan model lain. Untuk pengembangan selanjutnya disarankan: (1) penggunaan *dataset* yang lebih besar dan beragam dari berbagai platform media sosial dan periode; (2) perbandingan kinerja dengan algoritma lain seperti LSTM atau RoBERTa; (3) penambahan fitur *real-time scraping*; (4) pengembangan sistem agar dapat menangani data berbahasa daerah; dan (5) analisis aspek atau topik untuk mengetahui isu yang mendorong sentimen positif dan negatif.

DAFTAR PUSTAKA

- [1]. Y. K. Dwivedi *et al.*, "Setting the Future of Digital and Social Media Marketing Research: Perspectives and Research Propositions," *International Journal of Information Management*, vol. 59, p. 102168, 2022.
- [2]. A. Pratama and R. Nugroho, "Pemanfaatan Analisis Sentimen Media Sosial dalam Evaluasi Kebijakan Publik," *Jurnal Administrasi Publik*, vol. 18, no. 1, pp. 45–58, 2022.
- [3]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. 2019 Conf. North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, vol. 1, Minneapolis, MN, USA, 2019, pp. 4171–4186.
- [4]. Y. Putra and A. Wibowo, "Implementasi Self-Attention pada Model BERT untuk Klasifikasi Sentimen Teks Media Sosial," *Jurnal Teknologi Informasi dan Komputer*, vol. 10, no. 1, pp. 33–44, 2023.
- [5]. J. Wu, P. Qu, B. Zhang, and Z. Zhou, "Sentiment Analysis in Social Media: Leveraging BERT for Enhanced Accuracy," *Journal of Industrial Engineering and Applied Science*, vol. 2, no. 4, pp. 143–149, 2024.
- [6]. R. Hidayat, A. Pratama, and D. Lestari, "Penerapan Algoritma BERT untuk Analisis Sentimen Media Sosial Berbahasa Indonesia," *Jurnal Informatika dan Sistem Informasi*, vol. 10, no. 2, pp. 120–130, 2024.
- [7]. N. Sholihah, F. F. Abdulloh, and M. Rahardi, "Sentiment Analysis on KPU Performance Post-2024 Election via *YouTube* Comments Using BERT," *Sinkron*, vol. 8, no. 4, pp. 2222–2232, 2024.
- [8]. A. A. Bastian and A. Perdana, "Sentiment Analysis of Public Comments on *YouTube* Regarding the Inaugural Speech of the 8th President of Indonesia Using VADER and BERT Methods," *International Journal Software Engineering and Computer Science (IJSECS)*, vol. 5, no. 1, pp. 126–140, 2025.

-
- [9]. R. D. Lestari and A. R. Isnain, "Analisis Sentimen Masyarakat terhadap Penggunaan Sepeda Listrik pada Anak-Anak di Media Sosial X Menggunakan Metode SVM," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 1954–1965, 2024.
- [10]. R. Anugrah and S. Amalia, "Penerapan Text Preprocessing pada Analisis Sentimen Media Sosial Berbahasa Indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, pp. 145–154, 2024.
- [11]. N. Sari, D. Wibowo, and R. Kurniawan, "Pengaruh Preprocessing terhadap Akurasi Analisis Sentimen Teks Media Sosial," *Jurnal Informatika dan Komputer*, vol. 12, no. 2, pp. 88–97, 2023.