

# IMPLEMENTASI METODE *BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS* UNTUK ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI ACCESS

<sup>1</sup>Tri Buwono Bagus Wicaksono, <sup>2\*</sup>Rama Dian Syah

<sup>1</sup>Fakultas Teknologi Industri Universitas Gunadarma, <sup>2</sup>Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Gunadarma

Jl. Margonda Raya No. 100, Depok 16424, Jawa Barat

<sup>1</sup>baguswicaksono48@gmail.com, <sup>2\*</sup>rama\_ds@staff.gunadarma.ac.id

\*) Penulis Korespondensi

## Abstrak

Perkembangan teknologi di era digital ini berkembang pesat di berbagai bidang, salah satunya adalah bidang transportasi umum. Tujuan dari penelitian ini adalah untuk melakukan analisis sentimen terhadap pengguna aplikasi Access by KAI di Google Play Store sehingga dapat digunakan sebagai saran untuk meningkatkan kualitas aplikasi. Penulisan ini menggunakan metode *Bidirectional Encoding Representations from Transformers (BERT)* dengan pretrained model IndoBERT untuk melatih dataset bahasa Indonesia. Metode penulisan ini menggunakan metode CRISP-DM dengan 6 tahapan, yaitu Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment. Dataset yang digunakan sebanyak 10.000 ulasan dan setelah di-processing menjadi 9260. Model yang dibangun berhasil memprediksi sentimen dengan cukup baik dengan persentase sebesar 85%. Namun pada sentimen netral mendapatkan jumlah prediksi salah lebih banyak dari jumlah prediksi benar yaitu 22 ulasan dan jumlah prediksi salah yaitu 150 ulasan. Jumlah prediksi tepat untuk sentimen negatif yaitu 2.822 ulasan dan jumlah prediksi salah yaitu 345 ulasan. Jumlah prediksi tepat untuk sentimen positif yaitu 234 ulasan dan jumlah prediksi salah yaitu 131 ulasan. Model juga berhasil di-deploy dalam bentuk prototipe website dan dapat melakukan prediksi sentimen dengan cukup baik.

**Kata Kunci:** KAI, BERT, Analisis Sentimen, Klasifikasi, Ulasan

## Abstract

Technological developments in this digital era are growing rapidly in various fields, one of which is the field of public transportation. The purpose of this study is to conduct a sentiment analysis of Access by KAI application users on the Google Play Store so that it can be used as a suggestion to improve the quality of the application. This paper uses the *Bidirectional Encoding Representations from Transformers (BERT)* method with the pretrained IndoBERT model to train the Indonesian dataset. This writing method uses the CRISP-DM method with 6 stages, namely Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. The dataset used was 10,000 reviews and after being processed into 9260. The model that was built managed to predict sentiment quite well with a percentage of 85%. However, in neutral sentiment, the number of wrong predictions was more than the number of correct predictions, which was 22 reviews, and the number of wrong predictions, which was 150 reviews. The number of correct predictions for negative sentiment is 2,822 reviews and the number of wrong predictions is 345 reviews. The number of correct predictions for positive sentiment was 234 reviews and the number of wrong predictions was 131 reviews. The model has also been successfully deployed in the form of a website prototype and can strengthen sentiment predictions quite well.

**Keywords:** KAI, BERT, Sentiment analysis, Classification, Reviews

## PENDAHULUAN

Perkembangan teknologi di era digital ini berkembang pesat di berbagai bidang, salah satunya adalah bidang transportasi umum. Perkembangan teknologi dalam bidang transportasi umum membantu memudahkan para pengguna transportasi umum untuk mengakses layanan seperti melihat jadwal perjalanan, memesan tiket, dan lain sebagainya. Salah satu aplikasi yang melayani hal tersebut adalah Access by KAI.

Aplikasi Access by KAI merupakan aplikasi yang dibuat oleh PT. Kereta Api Indonesia (KAI) untuk memudahkan pengguna kereta api dalam mengakses layanan umum dari PT. KAI, seperti pembelian tiket, melihat jadwal perjalanan, ketersediaan kursi, pembatalan tiket, dan pergantian jadwal perjalanan. Kemudian, pengguna juga mendapatkan bukti pembayaran dan e-tiket. Dengan ini, aplikasi Access by KAI telah meningkatkan fleksibilitas terhadap pengguna layanan transportasi umum Kereta Api. Berdasarkan jumlah unduhan di Google Play Store sampai dengan bulan Mei 2024, aplikasi Access by KAI telah diunduh lebih dari 10 juta pengguna dan memiliki 211 ribu ulasan.

Ulasan dari pengguna aplikasi berfungsi untuk mengetahui kualitas aplikasi bagi para pengguna lain dan sebagai masukan bagi para pengembang untuk meningkatkan layanan pada aplikasi [1]. Untuk mengetahui dan memahami ulasan pengguna pada aplikasi Access by KAI, maka dibutuhkan analisis

sentimen. Analisis sentimen berfungsi untuk mengekstrak data opini, memahami, dan mengolah data tekstual secara otomatis untuk melihat sentimen yang terkandung dalam sebuah opini.

Proses klasifikasi sentimen dapat dilakukan dengan berbagai metode, salah satunya adalah *Bidirectional Encoder Representations from Transformers* (BERT). BERT adalah teknologi berbasis jaringan saraf yang bersifat *open source untuk pretraining Natural Language Processing* (NLP). BERT dapat membantu sistem komputer memahami bahasa sebagaimana manusia memahami bahasa dan mengetahui konteks suatu kata dalam pencarian dengan hasil yang relevan. Ada banyak pretrained model yang tersedia dalam metode BERT, salah satu model dari BERT yang dilatih dengan dataset berbahasa Indonesia adalah IndoBERT [2].

Penelitian terkait analisis sentimen menggunakan metode BERT dilakukan oleh beberapa peneliti terdahulu. Penelitian analisis sentimen oleh Atmaja dan Yustanti pada aplikasi Ruangguru menggunakan metode BERT menghasilkan nilai akurasi sebesar 99% [3]. Penelitian analisis sentimen oleh Uno dan Asep Juarna pada aplikasi JOOX menggunakan metode BERT menghasilkan nilai akurasi sebesar 80% [4]. Penelitian analisis sentimen oleh Kusnadi pada aplikasi Genshin Impact menggunakan metode BERT menghasilkan nilai akurasi sebesar 74% [5].

Berdasarkan uraian di atas, maka dibutuhkan analisis sentimen terhadap aplikasi

Access by KAI dari ulasan pengguna di Google Play Store dengan menggunakan model IndoBERT untuk mengetahui sentimen pengguna aplikasi Access by KAI sehingga dapat digunakan sebagai saran untuk meningkatkan kualitas dari aplikasi Access by KAI.

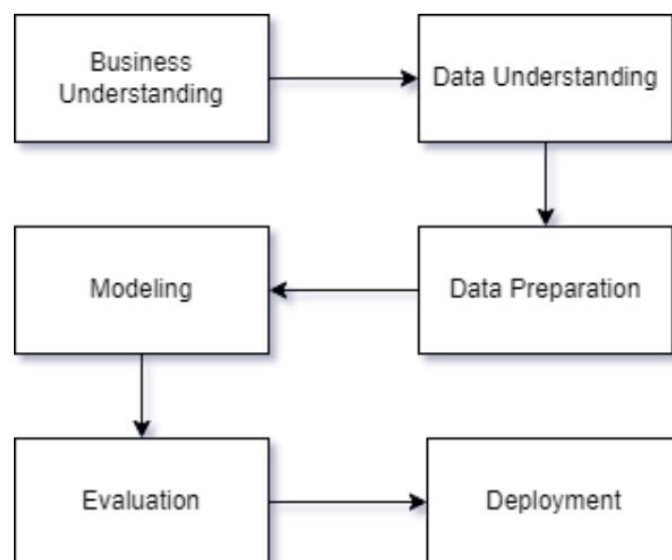
## METODE PENELITIAN

Metode penelitian yang digunakan adalah CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Langkah-langkah penelitian menggunakan metode CRISP-DM dapat dilihat pada gambar 1.

### *Business Understanding*

Pengembangan model IndoBERT untuk aplikasi Access by KAI bertujuan untuk mengoptimalkan analisis sentimen dan

persepsi publik terhadap aplikasi ini, yang dapat diakses melalui ulasan di Google Play Store. Dengan lebih dari 10 juta unduhan dan 211 ribu ulasan, ulasan pengguna mencerminkan pengalaman mereka terkait fungsionalitas, antarmuka pengguna, dan layanan yang disediakan. Memantau sentimen ini sangat penting bagi PT. Kereta Api Indonesia (KAI) untuk terus meningkatkan kualitas layanan mereka. Model IndoBERT yang dibangun akan menganalisis ulasan-ulasan tersebut secara mendalam untuk memberikan wawasan berharga kepada tim pengembang aplikasi, membantu mengidentifikasi isu-isu yang memerlukan perbaikan, serta merancang strategi pengembangan yang lebih efektif. Prototipe website yang di-*deploy* akan memungkinkan tim untuk mengakses hasil analisis ini secara *realtime*.



**Gambar 1. Metode Penelitian**

### Data Understanding

Penelitian ini menggunakan data yang diperoleh dari Google Play Store berupa ulasan pengguna pada aplikasi Access by KAI. *Scraping* dilakukan dengan memanfaatkan library dari *google-play-scraper* yang disediakan oleh pip. *Google-play-scraper* digunakan untuk mengekstrak berbagai atribut pada ulasan aplikasi di Google Play Store, seperti teks ulasan, username pemberi ulasan, skor yang diberikan, dan tanggal dibuatnya ulasan.

Dataset yang berhasil diekstraksi menggunakan *google-play-scraper* sebanyak 10.000 data. Hasil *scraping* kemudian disimpan dalam format *Comma Separated Value* (.csv). Hasil ulasan diurutkan

berdasarkan yang terbaru, dengan target pengambilan hingga 10.000 ulasan. Tidak ada filter skor yang diterapkan, sehingga semua ulasan akan diambil tanpa memandang skornya.

### Data Preparation

Persiapan data dilakukan melalui pelabelan berdasarkan tanggapan pengguna yang diterima [6]. Pada tahap ini, pelabelan dataset dilakukan secara manual untuk mengidentifikasi ulasan sentimen sebagai positif, netral, atau negatif. Tanggapan positif diberikan label angka 2, tanggapan netral diberikan label angka 1, dan tanggapan negatif diberikan label angka 0. Contoh hasil pelabelan data dapat ditemukan pada Tabel 1.

userName	score	at	content
Pengguna Google	3	2024-07-29 18:55:46	Woy, daftar pemesanan tiket nya benerin lagi. Masa, ku cari tiket ku gk ada. Padahal udah ku bayar kok
Pengguna Google	1	2024-07-29 18:42:54	Update bukan makin baik tapi makin kacau
Pengguna Google	1	2024-07-29 17:59:02	Min knpa aplksinya susah digunakan hari ini udh batas limit pesan selalu gagal di proses pembayaran
Pengguna Google	1	2024-07-29 17:47:55	Apk gagna malah jadi beban
Pengguna Google	5	2024-07-29 16:33:00	Kami sebagai masyarakat ekonomi bawah berharap KAI dapat menambah jumlah Kursi Ekonomi KA Erla
Pengguna Google	1	2024-07-29 15:30:01	Mau cancel 1 dari 5 orang, dateng ke loket malah di cancel semua
Pengguna Google	1	2024-07-29 14:20:11	Ini kenapa setelah di update,yang bisa di gunakan adalah kartu Axis,padahal sebelumnya enggak.mohon
Pengguna Google	5	2024-07-29 13:54:24	Sangat membantu sekali disaat kita sibuk untuk mendapatkan tiket lebih mudah,terimakasih kai
Pengguna Google	5	2024-07-29 12:21:28	Sangat membantu dalam pembelian tiket kereta api lokal dan antar kota
Pengguna Google	1	2024-07-29 11:16:36	Di saat urgent malah ga bisa melakukan pembayaran Opsi pembayaran terbatas Tolong perbaiki
Pengguna Google	1	2024-07-29 10:37:13	sering error mau masuk aplikasi gk bisa
Pengguna Google	1	2024-07-29 09:28:20	kirim otp gagal terus
Pengguna Google	1	2024-07-29 09:09:50	Makin ngak jelas kullihat kai access
Pengguna Google	1	2024-07-29 08:16:55	Udh di klik buat pendaftaran basic malah gk bisa" tetep di halaman yg sama berkali" gk bisa"
Pengguna Google	1	2024-07-29 08:16:02	Tiba" Logout sndiri dan skrng ga bisa buat login,tulisannya failed request apalah
Pengguna Google	1	2024-07-29 08:14:16	Aplikasi sampah, mau daftar aja error terus

Gambar 2. Hasil Scraping Data

Tabel 1. Labelling Data

Tanggapan	Label
Aplikasi KAI Access sangat membantu dalam memesan tiket kereta. Prosesnya cepat dan mudah digunakan. Sangat puas dengan layanannya!	2
Fitur dalam aplikasi KAI Access cukup lengkap, tetapi ada beberapa kali mengalami lag saat digunakan	1
Aplikasi sering crash dan tidak bisa digunakan untuk memesan tiket. Sangat mengecewakan	0

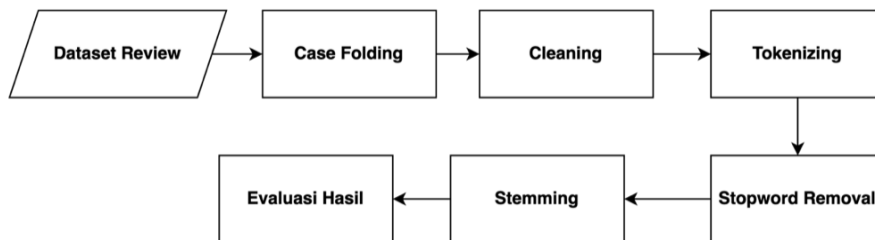
### Data Preprocessing

Tahapan berikutnya yang diterapkan pada dokumen adalah preprocessing teks. Proses ini dilakukan untuk mengubah data yang tidak terstruktur menjadi lebih terstruktur, sehingga mempermudah tahap-tahap selanjutnya. Alur proses preprocessing ditunjukkan pada Gambar 3.

*Case Folding* merupakan proses mengubah huruf pada ulasan pengguna yang mengandung huruf kapital (*uppercase*)

menjadi huruf kecil (*lowercase*) [7]. Proses case folding dilakukan dengan mengambil data ulasan. Kemudian, ulasan tersebut diproses, dan apabila terdapat huruf kapital, huruf tersebut akan diubah menjadi huruf kecil.

*Cleaning* merupakan tahapan dalam menghapus atribut yang tidak digunakan dalam proses pengklasifikasian, seperti angka, emoji, spasi berlebih, tanda baca, atau karakter unik [8]. Contoh hasil proses *cleaning* dapat dilihat pada Tabel 3



Gambar 3. Tahapan Preprocessing

Tabel 2. Hasil Case Folding

No	Tanggapan	Hasil Case Folding
1	Aplikasi KAI Access sangat membantu dalam memesan tiket kereta. Prosesnya cepat dan mudah digunakan. Sangat puas dengan layanannya!	aplikasi kai access sangat membantu dalam memesan tiket kereta. prosesnya cepat dan mudah digunakan. sangat puas dengan layanannya!
2	Fitur dalam aplikasi KAI Access cukup lengkap, tetapi ada beberapa kali mengalami lag saat digunakan.	fitur dalam aplikasi kai access cukup lengkap, tetapi ada beberapa kali mengalami lag saat digunakan.
3	Aplikasi sering crash dan tidak bisa digunakan untuk memesan tiket. Sangat mengecewakan.	aplikasi sering crash dan tidak bisa digunakan untuk memesan tiket. sangat mengecewakan.

Tabel 3. Hasil Cleaning

No	Hasil Case Folding	Hasil Cleaning
1	aplikasi kai access sangat membantu dalam memesan tiket kereta. prosesnya cepat dan mudah digunakan. sangat puas dengan layanannya!	aplikasi kai access sangat membantu dalam memesan tiket kereta prosesnya cepat dan mudah digunakan sangat puas dengan layanannya
2	fitur dalam aplikasi kai access cukup lengkap, tetapi ada beberapa kali mengalami lag saat digunakan.	fitur dalam aplikasi kai access cukup lengkap tetapi ada beberapa kali mengalami lag saat digunakan
3	aplikasi sering crash dan tidak bisa digunakan untuk memesan tiket. sangat mengecewakan.	aplikasi sering crash dan tidak bisa digunakan untuk memesan tiket sangat mengecewakan

*Tokenizing* merupakan proses memecah kata demi kata menjadi beberapa bagian. Kata yang sudah melalui beberapa proses akan dipecah menjadi kata per kata [9]. Proses *tokenizing* dilakukan dengan menggunakan data hasil *cleaning* [10]. Hasil proses *tokenizing* dapat dilihat pada tabel 4.

*Stopword removal* adalah proses penghapusan kata-kata yang tidak memiliki

makna atau tidak memiliki pengaruh terhadap analisis teks [11]. Hasil dari *stopword removal* dapat dilihat pada tabel 5. Proses *stemming* dilakukan dengan mengubah kata-kata menjadi bentuk dasar tanpa menggunakan imbuhan, seperti ke-, ber-, -nya-, di-, dan lainnya [12]. *Stemming* ini dilakukan dengan memanfaatkan library Sastrawi [13]. Hasil *stemming* ditampilkan pada Tabel 6.

**Tabel 4. Hasil *Tokenizing***

No	Hasil <i>Cleaning</i>	Hasil <i>Tokenizing</i>
1	aplikasi kai access sangat membantu dalam memesan tiket kereta prosesnya cepat dan mudah digunakan sangat puas dengan layanannya	[aplikasi, kai, access, sangat, membantu, dalam, memesan, tiket, kereta, prosesnya, cepat, dan, mudah, digunakan, sangat, puas, dengan, layanannya]
2	fitur dalam aplikasi kai access cukup lengkap tetapi ada beberapa kali mengalami lag saat digunakan	[fitur, dalam, aplikasi, kai, access, cukup, lengkap, tetapi, ada, beberapa, kali, mengalami, lag, saat, digunakan,]
3	aplikasi sering crash dan tidak bisa digunakan untuk memesan tiket sangat mengecewakan	[aplikasi, sering, crash, dan, tidak, bisa, digunakan, untuk, memesan, tiket, sangat, mengecewakan]

**Tabel 5. Hasil *Stopword Removal***

No	Hasil <i>Tokenizing</i>	Hasil <i>Stopword Removal</i>
1	[aplikasi, kai, access, sangat, membantu, dalam, memesan, tiket, kereta, prosesnya, cepat, dan, mudah, digunakan, sangat, puas, dengan, layanannya]	[aplikasi, kai, access, sangat, membantu, memesan, tiket, kereta, prosesnya, cepat, mudah, digunakan, sangat, puas, layanannya]
2	[fitur, dalam, aplikasi, kai, access, cukup, lengkap, tetapi, ada, beberapa, kali, mengalami, lag, saat, digunakan,]	[fitur, aplikasi, kai, access, cukup, lengkap, beberapa, kali, mengalami, lag, digunakan]
3	[aplikasi, sering, crash, dan, tidak, bisa, digunakan, untuk, memesan, tiket, sangat, mengecewakan]	[aplikasi, sering, crash, tidak, digunakan, memesan, tiket, sangat, mengecewakan]

**Tabel 6. Hasil *Stemming***

No	Hasil <i>Stopword Removal</i>	Hasil <i>Stemming</i>
1	[aplikasi, kai, access, sangat, membantu, memesan, tiket, kereta, prosesnya, cepat, mudah, digunakan, sangat, puas, layanannya]	[aplikasi, kai, access, sangat, bantu, pesan, tiket, kereta, proses, cepat, mudah, guna, sangat, puas, layan]
2	[fitur, aplikasi, kai, access, cukup, lengkap, beberapa, kali, mengalami, lag, digunakan]	[fitur, aplikasi, kai, access, cukup, lengkap, kali, alami, lag, guna]
3	[aplikasi, sering, crash, tidak, digunakan, memesan, tiket, sangat, mengecewakan]	[aplikasi, sering, crash, tidak, guna, pesan, tiket, sangat, kecewa]

Data yang telah melalui tahap *preprocessing data* akan dibagi menjadi data *training*, *validation*, dan *testing*. Sebanyak 10.000 dataset akan dilakukan pembagian sebanyak tiga kali percobaan, yaitu 80:20, 70:30 dan 60:40. Kemudian data latih hasil pembagian di awal dibagi lagi dengan 10% menjadi data latih dan data validasi.

### Modelling BERT

Pada tahap ini dilakukan implementasi model *pretrained* IndoBERT dari *Indo Benchmark* dan telah dilatih dengan *dataset* besar berbahasa Indonesia. Tahap modelling menggunakan *library transformers* dari *Hugging Face IndoBenchmark* dengan fungsi *Bert Tokenizer* dan *BertForSequence-Classification*. Modelling juga disertai dengan pengaturan *hyperparameter* untuk mengoptimalkan performa pada proses *fine-tuning* yaitu ukuran *batch size*, *learning rate* dan jumlah *epoch* [14]. Berikut penggunaan *hyperparameter* pada penelitian ini.

### Evaluation

Evaluasi ini dilakukan dengan menggunakan library sklearn (scikit-learn) dalam menentukan hasil kinerja dari analisis sentimen.

Library sklearn berfungsi memberikan laporan yang sangat informatif tentang seberapa baik model klasifikasi bekerja dengan menghasilkan metrik evaluasi yang mencakup presisi (precision), recall, f1-score, dan akurasi [15].

### Deployment

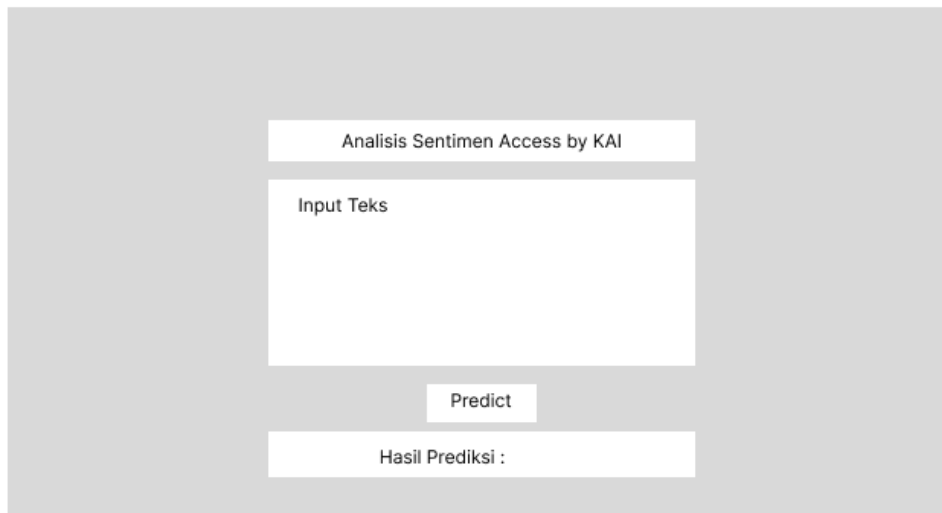
Deployment dilakukan untuk melakukan uji coba model dengan menggunakan platform lain. Pada penelitian kali ini, model akan di-deploy ke dalam bentuk aplikasi website dengan bantuan library Flask yang merupakan sebuah web framework yang ditulis dengan bahasa pemrograman Python. Pada tahap ini, akan dijabarkan rancangan halaman web yang dapat dilihat pada gambar 4.

**Tabel 7. Hasil Pembagian Dataset**

	Rasio 60:40	Rasio 70:30	Rasio 80:20
<b>Data Latih</b>	5000	5833	6667
<b>Data Validasi</b>	556	649	741
<b>Data Uji</b>	3704	2778	1852
<b>Total</b>		9260	

**Tabel 8. Hyperparameter untuk Skenario Pelatihan**

Hyperparameter	Nilai Hyperparameter 1	Nilai Hyperparameter 2
<i>Batch Size</i>	32	64
<i>Learning Rate</i>	2e-5	3e-5
<i>Epoch</i>	10	10



**Gambar 4. Rancangan Tampilan Website**

**Tabel 9. Hasil Pelatihan Skenario Pertama**

Rasio Pembagian <i>Dataset</i>	Hasil Pelatihan		
	Akurasi Latihan (Rata-rata)	Loss Latihan (Rata-rata)	Waktu Pelatihan
60:40	0,85	0,284	10m21d
70:30	0,83	0,113	14m12d
80:20	0,84	0,024	16m44d

**Tabel 10. Hasil Pelatihan Skenario Kedua**

Rasio Pembagian <i>Dataset</i>	Hasil Pelatihan		
	Akurasi Latihan (Rata-rata)	Loss Latihan (Rata-rata)	Waktu Pelatihan
60:40	0,852	0,027	10m10d
70:30	0,83	0,311	12m
80:20	0,84	0,075	13m49d

## HASIL DAN PEMBAHASAN

Pelatihan dilakukan dengan kombinasi tiga skenario pembagian dataset dan dua perbedaan penggunaan hyperparameter. Pelatihan dibagi menjadi dua percobaan skenario pelatihan berdasarkan hyperparameter pada tabel 8. Hasil pelatihan dengan percobaan pertama menggunakan hyperparameter batch size 16, learning rate 2e-5, dan epoch 10 terlihat pada Tabel 9. Hasil pelatihan dengan percobaan kedua

menggunakan hyperparameter batch size 32, learning rate 3e-5, dan epoch 10 terlihat pada Tabel 10.

Berdasarkan hasil dua metode percobaan, pembagian dataset 60:40 dengan hyperparameter (batch size 32, learning rate 3e-5, dan epoch 10) dengan nilai akurasi sebesar 85,2% dengan waktu pelatihan selama 10 menit 10 detik.

Pengujian dilakukan setelah model dilatih. Pengujian menggunakan hasil pelatihan terbaik, yaitu pada percobaan kedua



dengan penggunaan hyperparameter dan rasio pembagian 60:40. Hasil klasifikasi sentimen oleh model dibandingkan dengan sentiment aktual. Pada Tabel 11 terlihat hasil uji coba model yang menunjukkan model mampu melakukan prediksi sentimen berdasarkan ulasan dibandingkan dengan sentimen aktualnya.

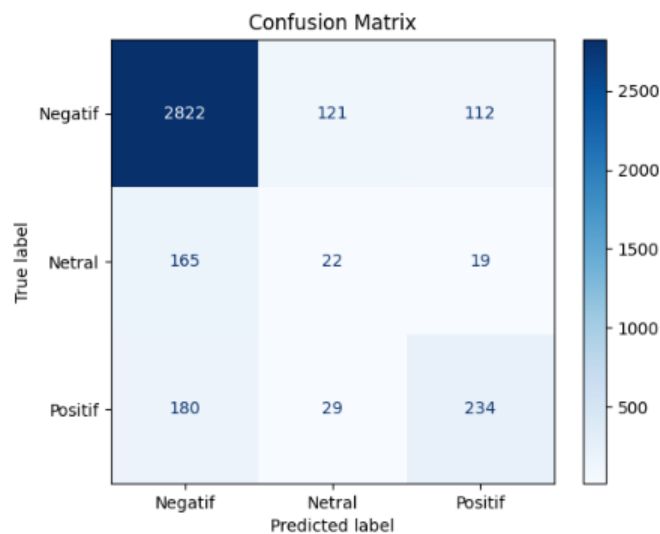
Model dinilai untuk menentukan performanya dalam memprediksi klasifikasi sentimen dengan menggunakan matriks kebingungan (Confusion Matrix). Penilaian ini dilakukan dengan data uji yang telah dipisahkan sejak awal proses analisis. Empat kategori yang diidentifikasi adalah True Positive (TP), yang menunjukkan sentimen positif diprediksi dengan benar sebagai positif, False Negative (FN), yang terjadi ketika

sentimen positif salah diprediksi sebagai negatif, False Positive (FP), yang menunjukkan sentimen negatif salah diprediksi sebagai positif, dan True Negative (TN), yang menunjukkan sentimen negatif diprediksi dengan benar sebagai negatif. Hasil penilaian dengan matriks kebingungan ditunjukkan pada Gambar 5.

Pada gambar 5 dapat disimpulkan bahwa model dapat mengklasifikasi sentimen dengan cukup baik karena jumlah prediksi tepat lebih banyak dibanding dengan jumlah prediksi salah, namun pada klasifikasi sentimen netral mendapatkan hasil jumlah prediksi salah lebih banyak dibanding dengan jumlah prediksi tepat. Jumlah prediksi tepat untuk sentimen netral yaitu 22 ulasan dan jumlah prediksi salah yaitu 150 ulasan.

**Tabel 11. Hasil Uji Coba**

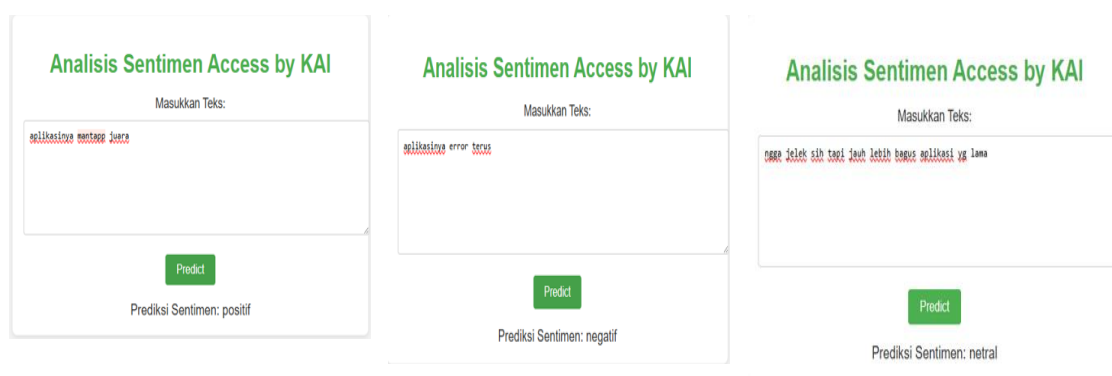
No	Ulasan	Aktual	Prediksi
1	gajelas timeout mulu	Negatif	Negatif
2	mantap	Positif	Positif
3	ada pilihan untuk daftar mapclub tapi setelah dipencet malah gabisa daftar	Netral	Negatif



**Gambar 5. Hasil Evaluasi Confusion Matrix**

**Tabel 12. Hasil Classification Report**

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
<i>Negative</i>	0,89	0,92	0,91	3055 ulasan
<i>Neutral</i>	0,13	0,11	0,12	206 ulasan
<i>Positive</i>	0,64	0,53	0,58	443 ulasan
<i>Accuracy</i>			<b>0,85</b>	3704 ulasan
<i>Macro Avarage</i>	0,55	0,52	0,53	3704 ulasan
<i>Weighted Avarage</i>	0,82	0,83	0,82	3704 ulasan



**Gambar 6. Hasil Deployment Model**

Jumlah prediksi tepat untuk sentimen negatif yaitu 2.822 ulasan dan jumlah prediksi salah yaitu 345 ulasan. Jumlah prediksi tepat untuk sentimen positif yaitu 234 ulasan dan jumlah prediksi salah yaitu 131 ulasan. Pada Gambar 8 terlihat bahwa model menghasilkan nilai accuracy sebesar 85%. Nilai precision yaitu pada sentimen negatif sebesar 89%, netral sebesar 13%, dan positif sebesar 64%. Nilai recall pada sentimen negatif sebesar 92%, netral sebesar 11%, dan positif sebesar 53%. Nilai F1-Score pada sentimen negatif sebesar 91%, netral sebesar 12%, dan positif sebesar 58%. Pada data netral memiliki nilai akurasi yang sangat rendah dikarenakan data support yang digunakan jumlahnya sedikit sehingga mempengaruhi tingkat akurasi analisis sentimen

### Hasil Deployment

Hasil *deployment* model IndoBERT ini berbasis website dengan menggunakan Flask untuk melakukan pengujian model. Pengujian dilakukan dengan menginput ulasan baru yang akan dideteksi oleh model dan menampilkan *output* hasil sentimen positif, negatif atau netral. Hasil dari *deployment* dapat dilihat pada gambar 6.

### KESIMPULAN DAN SARAN

Hasil evaluasi performa model dengan confusion matrix menunjukkan bahwa model dapat mengklasifikasikan sentimen dengan cukup baik namun pada klasifikasi sentimen netral mendapatkan hasil jumlah prediksi salah lebih banyak dibanding dengan jumlah

prediksi tepat. Jumlah prediksi tepat untuk sentimen netral yaitu 22 ulasan dan jumlah prediksi salah yaitu 150 ulasan. Jumlah prediksi tepat untuk sentimen negatif yaitu 2.822 ulasan dan jumlah prediksi salah yaitu 345 ulasan. Jumlah prediksi tepat untuk sentimen positif yaitu 234 ulasan dan jumlah prediksi salah yaitu 131 ulasan. Model menghasilkan accuracy sebesar 85% dengan accuracy untuk sentimen negatif sebesar 89%, untuk sentimen netral sebesar 13%, dan untuk sentimen positif sebesar 64%.

Pengembangan lebih lanjut terkait dengan penelitian ini dapat dilakukan dengan menambahkan tahap normalisasi saat *preprocessing* data, menggunakan *dataset* dengan jumlah ketiga kategori yang seimbang agar model dapat mempelajari dengan lebih baik, dan dapat memproses *input file .csv* pada prototipe *website* hasil dari *deployment*.

## DAFTAR PUSTAKA

- [1] A. Nurian, “Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes,” *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.
- [2] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, “Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN,” *Ilk. J. Ilm.*, vol. 14, no. 3, pp. 348–354, 2022, doi: 10.33096/ilkom.v14i3.1505.348-354.
- [3] R. Mas, R. W. Panca, K. Atmaja1, and W. Yustanti2, “Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers),” *Jeisbi*, vol. 02, no. 03, p. 2021, 2021.
- [4] J. U. S. Lazuardi and A. Juarna, “Analisis Sentimen Ulasan Pengguna Aplikasi Joox Pada Android Menggunakan Metode Bidirectional Encoder Representation From Transformer (Bert),” *J. Ilm. Inform. Komput.*, vol. 28, no. 3, pp. 251–260, 2023, doi:10.35760/ik.2023.v28i3.10090.
- [5] R. Kusnadi, Y. Yusuf, A. Andriantony, R. Ardian Yaputra, and M. Caintan, “Analisis Sentimen Terhadap Game Genshin Impact Menggunakan Bert,” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 6, no. 2, pp. 122–129, 2021, doi: 10.36341/rabit.v6i2.1765.
- [6] N. Z. Rania and R. D. Syah, “Analisis Sentimen Terhadap Aplikasi Gojek Pada Play Store Menggunakan Metode Random Forest Classifier,” *J. Ilm. Inform. Komput.*, vol. 29, no. 2, pp. 144–153, Jul. 2024, doi: 10.35760/ik.2024.v29i2.11877.
- [7] S. M. Fani, R. Santoso, and S. Suparti, “Penerapan Text Mining Untuk Melakukan Clustering Data Tweet Akun Blibli Pada Media Sosial Twitter

- Menggunakan K-Means Clustering,” *J. Gaussian*, vol. 10, no. 4, pp. 583–593, 2021, doi:10.14710/j.gauss.v10i4.30409
- [8] R. Nurul Ikhsani and F. Fauzi Abdulloh, “Optimasi SVM dan Decision Tree Menggunakan SMOTE Untuk Mengklasifikasi Sentimen Masyarakat Mengenai Pinjaman Online,” *J. Media Inform. Budidarma*, vol. 7, no. 4, pp. 1667–1677, 2023, doi:10.30865/mib.v7i4.6809
- [9] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, “Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021, doi: 10.57152/malcom.v1i1.20.
- [10] R. R. Salam, M. F. Jamil, Y. Ibrahim, R. Rahmaddeni, S. Soni, and H. Herianto, “Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggunakan Support Vector Machine,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 27–35, 2023, doi: 10.57152/malcom.v3i1.590.
- [11] A. Santosa, I. Purnamasari, and Mayasari Rini, “Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM,” *J. Sains Komput. Inform.*, vol. 6, no. 1, pp. 81–93, 2022.
- [12] A. S. dan N. Surojudin, “Analisis Dan Perbandingan Stemming Algoritma Porter Dengan Algoritma Ahmad Yusoff Sembok Dalam Dokumen Teks Bahasa Indonesia,” *Pros. Semin. SeNTIK*, vol. 4, no. 1, pp. 347–357, 2020, [Online]. Available: <https://ejournal.jakstik.ac.id/index.php/sentik/article/view/3304>.
- [13] A. E. Budiman and A. Widjaja, “Analisis Pengaruh Teks Preprocessing Terhadap Deteksi Plagiarisme Pada Dokumen Tugas Akhir,” *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 3, pp. 475–488, 2020, doi: 10.28932/jutisi.v6i3.2892.
- [14] A. Awalina, F. A. Bachtiar, and F. Utaminigrum, “Perbandingan Pretrained Model Transformer pada Deteksi Ulasan Palsu,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 597–604, 2022, doi:10.25126/jtiik.2022935696.
- [15] M. Fadli and R. A. Saputra, “Klasifikasi Dan Evaluasi Performa Model Random Forest Untuk Prediksi Stroke,” *JT J. Tek.*, vol. 12, no. 2, pp. 72–80, 2023, [Online]. Available: <http://jurnal.umt.ac.id/index.php/jt/index>.