VOL. 03. NO. 01 JANUARI 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

ANALISIS SENTIMEN DI MEDIA SOSIAL TERHADAP KEBIJAKAN JAM MASUK SEKOLAH JAWA BARAT MENGGUNAKAN INDOBERT

Irvan Maulana[1], Rudi Kurniawan[2], Bani Nurhakim[3], Nining Rahaningsih[4], Willy Prihartono[5]

Program Studi Teknik Informatika[12]
Program Studi Manajemen Informatika[3]
Program Studi Komputerisasi Akuntansi[45]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
irvanmaul4321@gmail.com

(*) Corresponding Author : irvanmaul4321@gmail.com
Published: 30 Januari 2026

Abstract- *Advances in information technology have led toincreased use of sentiment analysis in identifying public opinionon social media. The school start time policy implemented by the West Java Provincial Governmenthas sparked widespread debate, requiring an analytical approachto objectively determine public perception. This studyaims to classify public sentiment towards the policyusing the IndoBERT model. Data was collected through web scraping methods from Twitter (X) and YouTube, resulting in 1,103 comments thatwere then processed through preprocessing, manual labeling, anddata division. The IndoBERT Base-P1 model was fine-tuned using WeightedCross-Entropy Loss to handle class imbalance. The results of the studyshow that IndoBERT is capable of achieving an accuracy of 75% and a weighted F1-score of0.745. This study proves that IndoBERT is effective as asentiment analysis model for Indonesian-based public policy issues.*

Keyword: IndoBERT,Sentiment Analysis, Social Media, Public Policy, NLP

Abstrak- Perkembangan teknologi informasi mendorong meningkatnya penggunaan analisis sentimen dalam mengidentifikasi opini publik di media sosial. Kebijakan jam masuk sekolah yang diberlakukan oleh Pemerintah Provinsi Jawa Barat memicu perdebatan luas, sehingga diperlukan pendekatan analitis untuk mengetahui persepsi masyarakat secara objektif. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap kebijakan tersebut menggunakan model IndoBERT. Data dikumpulkan melalui metode web scraping terhadap Twitter (X) dan YouTube, menghasilkan 1.103 komentar yang selanjutnya diproses melalui tahapan preprocessing, pelabelan manual, dan pembagian data. Model IndoBERT Base-P1 di-fine-tune menggunakan Weighted Cross-Entropy Loss untuk menangani ketidakseimbangan kelas. Hasil penelitian menunjukkan bahwa IndoBERT mampu mencapai akurasi 75% dan F1-score tertimbang 0.745. Penelitian ini membuktikan bahwa IndoBERT efektif digunakan sebagai model analisis sentimen untuk isu kebijakan publik berbasis Bahasa Indonesia.

Kata Kunci : IndoBERT, Analisis Sentimen, Media Sosial, Kebijakan Publik, NLP

INTRODUCTION

Kebijakan jam masuk sekolah di Provinsi Jawa Barat menimbulkan perdebatan luas di ruang digital karena dianggap membawa konsekuensi pada kesiapan fisik siswa, kondisi sosial, serta efektivitas proses belajar. Permasalahan utama dalam penelitian ini adalah bagaimana mengidentifikasi dan memetakan persepsi masyarakat secara objektif berdasarkan opini yang diungkapkan di media sosial, mengingat karakteristik teks digital bersifat tidak terstruktur, mengandung *slang*, dan sangat dinamis. Solusi yang ditawarkan pada penelitian ini adalah pemanfaatan teknologi *Natural Language Processing* (NLP) berbasis model *transformer*, khususnya IndoBERT, untuk melakukan klasifikasi sentimen publik terhadap kebijakan tersebut secara otomatis, terukur, dan dapat direplikasi.

Media sosial seperti Twitter dan YouTube telah digunakan sebagai sumber opini publik dalam berbagai penelitian karena mampu merepresentasikan reaksi masyarakat secara langsung dan real-time. Dalam lima tahun terakhir, sejumlah penelitian yang relevan menunjukkan bahwa IndoBERT memiliki performa unggul dalam tugas analisis sentimen Bahasa Indonesia. Studi oleh [1] membuktikan bahwa IndoBERT mampu mencapai akurasi tinggi untuk ulasan aplikasi

kesehatan. Penelitian[2] juga menunjukkan efektivitas IndoBERT pada dataset ulasan hotel meskipun mengalami kendala ketidakseimbangan kelas.

Penelitian[3] menggunakan IndoBERT pada opini publik di platform X mengenai kendaraan listrik dan menunjukkan bahwa model ini mampu menangani teks pendek dan informal. [4] mengusulkan teknik *Confident Learning* untuk memperbaiki kualitas label pada dataset sentimen berbasis IndoBERT dan mendapatkan peningkatan signifikan pada akurasi. Sementara itu, [5]menerapkan pendekatan *aspect-based sentiment analysis* (ABSA) menggunakan IndoBERT pada ulasan mahasiswa dan menunjukkan bahwa model ini efektif pada konteks multi-aspek. Studi pendukung lainnya seperti [6] dan [7]menegaskan pentingnya analisis sentimen untuk memahami polarisasi masyarakat terhadap isu kebijakan publik.

Penelitian lain yang relevan juga memperlihatkan bahwa pendekatan berbasis *transformer* efektif digunakan untuk menganalisis opini publik pada isu kebijakan dan pelayanan pemerintah. [8] menunjukkan bahwa analisis sentimen mampu mengungkap persepsi publik terhadap strategi komunikasi lembaga pemerintah melalui pemanfaatan model bahasa Indonesia modern. [9] menemukan bahwa metode pembelajaran mesin dapat mendeteksi kecenderungan sentimen masyarakat terhadap kebijakan lingkungan nasional. Pada konteks layanan digital, [10] membuktikan bahwa kombinasi leksikon dan IndoBERT mampu menangkap sentimen pengguna aplikasi secara lebih akurat. [11] menegaskan pentingnya *post-training* berbasis domain untuk meningkatkan performa IndoBERT dalam klasifikasi sentimen.

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa IndoBERT efektif dalam memproses teks Bahasa Indonesia. Namun, terdapat *gap* penelitian berupa minimnya kajian yang mengaplikasikan IndoBERT pada isu kebijakan pendidikan tingkat daerah, terutama kebijakan jam masuk sekolah. Selain itu, sebagian besar studi masih terbatas pada satu platform media sosial, sedangkan opini publik tersebar di berbagai platform dengan karakteristik wacana yang berbeda. Tantangan tambahan muncul karena teks media sosial di Indonesia sarat dengan bahasa informal, sehingga membutuhkan strategi preprocessing yang tepat agar model dapat memahami konteks dengan baik.

Berdasarkan *gap* tersebut, penelitian ini bertujuan untuk: (1) menganalisis sentimen masyarakat terhadap kebijakan jam masuk sekolah di Jawa Barat berdasarkan data Twitter dan YouTube, (2) mengevaluasi performa model IndoBERT dalam klasifikasi sentimen multi-kelas pada teks informal media sosial, dan (3) menyajikan gambaran sentimen publik sebagai masukan bagi pembuat kebijakan agar dapat merumuskan strategi yang lebih responsif dan berbasis data. Harapannya, penelitian ini memberikan kontribusi pada pengembangan NLP Bahasa Indonesia dalam konteks kebijakan publik serta menjadi referensi untuk penelitian lanjutan pada isu sosial-politik lokal.

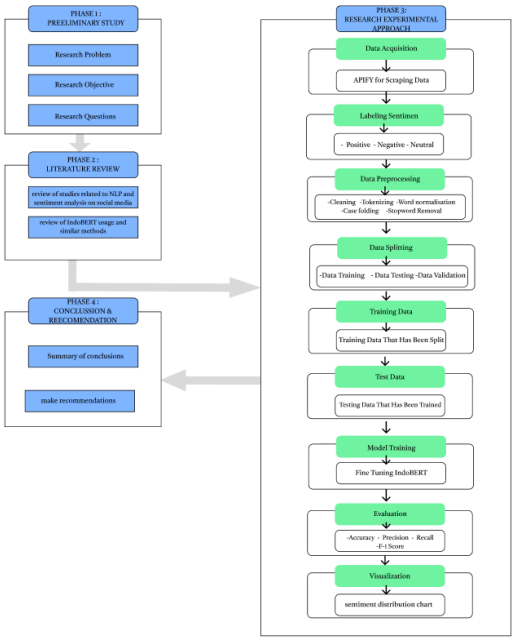
membutuhkan.

MATERIALS AND METHODS

1. Tahapan Penelitian

Tahapan penelitian ini disusun untuk menggambarkan alur kerja secara sistematis mulai dari pengumpulan data hingga evaluasi model IndoBERT. Secara umum, penelitian terdiri dari enam tahapan utama: (1) pengumpulan data, (2) pelabelan sentimen, (3) preprocessing teks, (4) pembagian data, (5) pelatihan model IndoBERT, dan (6) evaluasi performa.

Tahapan ini memastikan bahwa proses penerapan metode dilakukan dengan terstruktur sehingga menghasilkan model analisis sentimen yang optimal dan dapat direplikasi.



Gambar 1. Tahapan Penelitian

Gambar 1. menunjukkan proses lengkap penelitian mulai dari akuisisi data hingga evaluasi model. Setiap tahap saling terhubung dan berperan dalam menghasilkan model analisis sentimen berbasis *transformer* yang akurat.

Tabel 1. Tahapan dan Fungsi Proses Penelitian

Tahap	Fungsi
Akuisisi Data	Mengambil komentar dari Twitter & YouTube
Labeling	Memberikan label sentimen manual
Preprocessing	Membersihkan dan menormalisasi teks
Splitting	Membagi data menjadi train/valid/test
Modeling	Fine-tuning IndoBERT
Evaluasi	Mengukur akurasi dan F1-score

Tabel 1 digunakan untuk memperjelas tahapan utama penelitian dan dirujuk dalam teks sebelumnya.

2. Data Acquisition

Data dikumpulkan dari dua platform media sosial, yaitu Twitter (X) dan YouTube, menggunakan teknik *web scraping* dengan bantuan APIFY. Kata kunci yang digunakan antara lain: "jam masuk sekolah", "Jawa Barat sekolah pagi", dan frase terkait.

Kriteria inklusi:

1. komentar/unggahan publik,
2. berbahasa Indonesia,
3. relevan dengan isu kebijakan.

Total data yang diperoleh sebanyak 1.103 komentar. Seluruh data tidak mengandung informasi pribadi dan mengikuti etika pengumpulan data digital.

3. Sentimen Labeling

Pelabelan dilakukan secara manual oleh peneliti untuk memastikan kualitas label. Kategori sentimen meliputi:

1. Positif
2. Netral
3. Negatif

Rincian kuantitas data untuk setiap kelas sentimen pasca-pelabelan dapat dilihat pada Tabel 2 berikut.

Tabel 2. Distribusi Kelas Sentimen

No	Kelas Sentimen	Jumlah
1.	Negatif	630
2.	Positif	248
3.	Netral	225

Berdasarkan Tabel 2, terlihat bahwa distribusi data antarkelas memiliki proporsi yang tidak seimbang (*imbalanced dataset*). Kelas sentimen Negatif mendominasi dataset dengan jumlah terbanyak, yaitu 630 data. Jumlah ini jauh lebih besar dibandingkan dengan kelas sentimen Positif yang berjumlah 248 data dan kelas Netral sebanyak 225 data. Ketimpangan jumlah data ini mengindikasikan bahwa respons masyarakat terhadap topik yang diangkat cenderung lebih banyak bermuatan opini kontra atau negatif.

4. Preprocessing

Preprocessing dilakukan untuk membersihkan dan menstandarkan teks sehingga siap diproses oleh model IndoBERT. Tahapan yang diterapkan adalah:

1. Case Folding
   Mengubah seluruh teks menjadi huruf kecil agar variasi kata dapat diseragamkan.
2. Cleaning
   Menghapus URL, *hashtag*, *mention*, angka, emotikon, karakter non-alfabet, dan duplikasi.
3. Normalisasi
   Mengubah *slang* dan singkatan ke bentuk baku, misalnya: "bgt" → "banget".
4. Tokenization
   Memecah teks menggunakan tokenizer IndoBERT yang berbasis *WordPiece*.
5. Stopword Removal
   Menghapus kata-kata umum yang tidak relevan dengan konteks sentimen.

Tahapan ini memperbaiki kualitas data sehingga model lebih mudah memahami konteks linguistik.

5. Data Spliting

Untuk mengukur performa model secara objektif, dataset dibagi ke dalam tiga himpunan bagian (*subset*) dengan skema rasio 65:15:20. Rincian distribusi data untuk *training*, *validation*, dan *testing* disajikan pada Tabel 3 berikut.

Tabel 3. Pembagian Data

Jenis Data	Presentase	Jumlah Data	Fungsi
Training	65%	717	Melatih bobot model IndoBERT
Validation	15%	165	Evaluasi saat training dan tuning
Testing	20%	221	Pengujian performa akhir
Total	100%	1103	

Berdasarkan Tabel 3, mayoritas data dialokasikan sebagai data latih (training set), yaitu sebanyak 717 data (65%). Jumlah yang besar ini diperlukan agar model pre-trained IndoBERT dapat

mempelajari pola kebahasaan dan fitur sentimen secara optimal.

Selanjutnya, sebanyak 165 data (15%) digunakan sebagai data validasi untuk memantau performa model di setiap epoch dan mencegah terjadinya overfitting. Sisa 221 data (20%) dialokasikan secara eksklusif sebagai data uji (testing set). Data uji ini merupakan data asing (unseen data) yang tidak dilibatkan dalam proses pelatihan, sehingga hasil evaluasi yang diperoleh dapat merepresentasikan kemampuan generalisasi model yang sebenarnya di dunia nyata."

6. Modeling

Tahap pemodelan dilakukan dengan menerapkan teknik *fine-tuning* pada model *pre-trained* IndoBERT Base-P1. Implementasi model dibangun menggunakan pustaka *Hugging Face Transformers* berbasis *framework* PyTorch. Untuk mendapatkan performa optimal, konfigurasi *hyperparameter* utama disajikaan pada Tabel 4.

Tabel 4. Konfigurasi Parameter Fine-Tuning IndoBERT

No	Parameter	Keterangan
1.	Base Model	IndoBERT Base-P1
2.	Framework	PyTorch & Transformers
3.	Learning Rate	$2\ x\ 10^{-5}(2e-5)$
4.	Batch Size	16
5.	Epoch	4
6.	Loss Function	Weighted Cross-Entropy
7.	GPU	T4 GPU Google Colab

Berdasarkan Tabel 4, *learning rate* ditetapkan pada nilai rendah ($2\ x\ 10^{-5}$) sesuai rekomendasi standar untuk model berbasis BERT. Nilai yang kecil ini dipilih untuk menjaga stabilitas bobot model *pre-trained* agar tidak rusak saat mempelajari fitur baru (*catastrophic forgetting*). Durasi pelatihan dibatasi sebanyak 4 *epoch* dengan *batch size* 16 untuk menyeimbangkan efisiensi komputasi memori GPU dengan konvergensi model.

7. Data Training

Proses pelatihan (*training*) dijalankan dengan memasukkan Data Latih (*Training Set*) ke dalam model secara iteratif. Pada tahap ini, model mempelajari pola fitur linguistik dari data latih untuk meminimalkan nilai *error*. Mekanisme pelatihan dilakukan dengan ketentuan sebagai berikut:

1. Input Data**:** Model menerima masukan berupa token dari data latih yang telah melalui proses *tokenisasi*.
2. Validasi Berkala: Pada setiap akhir *epoch*, model dievaluasi menggunakan Data Validasi (*Validation Set*). Hal ini bertujuan untuk memantau apakah model mampu mengenali data di luar data latih dan untuk mendeteksi indikasi *overfitting* sedini mungkin.
3. Optimalisasi Bobot**:** Berdasarkan umpan balik dari fungsi kerugian (*loss function*), *optimizer* akan memperbarui bobot internal model agar prediksi sentimen semakin akurat seiring bertambahnya iterasi.

Hal yang krusial dalam konfigurasi pelatihan ini adalah penggunaan *Loss Function* jenis *Weighted Cross-Entropy.* Mengingat distribusi data latih yang tidak seimbang (*imbalanced dataset*) yang didominasi oleh kelas negatif, fungsi ini diterapkan untuk memberikan bobot penalti lebih besar pada kesalahan prediksi di kelas minoritas. Strategi ini bertujuan agar model tidak bias dan mampu mengenali kelas dengan jumlah data sedikit sebaik mengenali kelas mayoritas.

8. Evaluation

Evaluasi performa model dilakukan secara komprehensif pada data uji (*testing set*) yang merupakan data asing (*unseen data*) bagi model. Hal ini bertujuan untuk mengukur kemampuan generalisasi model secara objektif. Pengukuran kinerja didasarkan pada metrik evaluasi standar klasifikasi, yaitu:

a. *Accuracy*
b. *Precision*
c. *Recall*
d. *F1-Score*
e. *Confusion Matrix*

Penggunaan kombinasi metrik ini sangat penting mengingat karakteristik dataset yang tidak seimbang (*imbalanced dataset*). Metrik Akurasi memberikan gambaran performa secara global, namun sering kali bias terhadap kelas mayoritas. Oleh karena itu, metrik **Presisi** (*Precision*), Recall, dan F1-Score digunakan sebagai penentu utama untuk memastikan model mampu mengklasifikasikan kelas minoritas (Positif dan Netral) dengan baik, bukan hanya unggul memprediksi kelas Negatif.

Sebagai pelengkap, Confusion Matrix dihadirkan untuk memvisualisasikan detail persebaran prediksi yang benar dan salah (*True Positive, False Positive, dst.*) pada setiap kelas. Analisis matriks ini memungkinkan identifikasi

spesifik terhadap kesalahan klasifikasi yang dilakukan oleh model IndoBERT.

RESULTS AND DISCUSSION

1. Data Acquisition

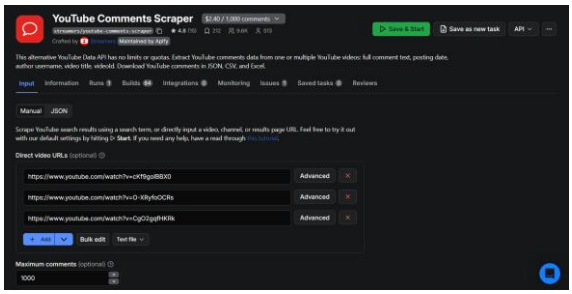
Tahap awal penelitian ini melibatkan proses akuisisi data (*data acquisition*) dari dua platform media sosial utama, yaitu Twitter (X) dan YouTube. Kedua platform ini dipilih karena memiliki intensitas diskusi publik yang tinggi dan responsif terhadap isu terkini, khususnya terkait kebijakan jam masuk sekolah pukul 06.00 WIB di Provinsi Jawa Barat. Proses pengambilan data dilakukan dengan memanfaatkan layanan Apify, sebuah platform ekstraksi data web yang memungkinkan penarikan data secara otomatis, terstruktur, dan efisien.

Pada platform YouTube, data dikumpulkan dari kolom komentar tiga video berita yang diunggah oleh saluran berita nasional. Video-video ini dipilih karena secara spesifik membahas pro dan kontra implementasi kebijakan tersebut. Ekstraksi komentar dilakukan menggunakan modul *YouTube Comments Scraper* yang mampu mengambil teks komentar beserta metadata interaksinya. Adapun tiga video yang menjadi sumber data utama adalah:

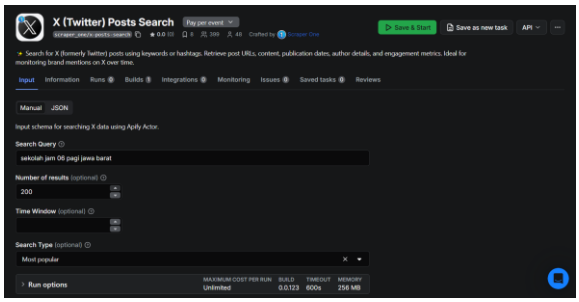
1. https://www.youtube.com/watch?v=cKf9goIBBX0
2. https://www.youtube.com/watch?v=O-XRyfoOCRs
3. https://www.youtube.com/watch?v=CgO2gqfHKRk

Dari ketiga sumber video tersebut, berhasil diakuisisi sebanyak 533 komentar yang merepresentasikan persepsi, argumen, dan respons masyarakat terhadap wacana kebijakan yang bergulir.

Sementara itu, pengumpulan data dari platform Twitter dilakukan menggunakan modul *X (Twitter) Post Search* dengan kueri pencarian spesifik *"sekolah jam 06 pagi Jawa Barat"* untuk menjamin relevansi topik. Mengingat adanya batasan teknis pada API yang membatasi pengambilan data maksimal sekitar 300 entri per eksekusi, proses *scraping* dilakukan dalam dua tahapan siklus. Strategi ini diterapkan untuk mendapatkan volume data yang lebih representatif. Melalui proses tersebut, terkumpul sebanyak **570** cuitan (*tweets*) yang mencerminkan opini spontan, kritik, serta dukungan publik.



Gambar 2. Youtube Comments Scraper



Gambar 3. X (Twitter) Posts Search

Secara keseluruhan, integrasi data dari kedua platform menghasilkan dataset mentah (*raw data*) sebanyak **1.103** entri teks. Seluruh data tersebut disimpan dalam format *Comma Separated Values* (.CSV) untuk memfasilitasi tahapan pengolahan selanjutnya. Sebagai langkah awal persiapan data sebelum pemodelan, setiap entri teks diberikan label secara manual ke dalam tiga kelas polaritas sentimen, yaitu **Negatif (-1)**, **Netral (0)**, dan **Positif (1)**. Dataset terstruktur inilah yang kemudian menjadi landasan empiris dalam proses pra-pemrosesan teks, eksplorasi data, serta pelatihan model IndoBERT.

2. Sentimen Labeling

Sebelum memasuki tahapan pra-pemrosesan (preprocessing), seluruh data mentah melalui proses pelabelan sentimen (annotation) secara manual. Proses ini dilakukan oleh peneliti yang bertindak sebagai anotator tunggal untuk menjaga konsistensi penilaian terhadap setiap data. Pelabelan ini bertujuan untuk menyusun ground truth (kebenaran dasar) yang valid, yang nantinya akan digunakan oleh model IndoBERT untuk mempelajari pola bahasa dalam paradigma supervised learning. Setiap entri data diklasifikasikan ke dalam tiga kelas polaritas sentimen dengan definisi operasional sebagai berikut:

a. Negatif (-1): Diberikan pada komentar yang mengandung ekspresi penolakan, kritik tajam, keluhan, atau ketidaksetujuan

terhadap kebijakan jam masuk sekolah pukul 06.00 pagi.

b. Netral (0): Diberikan pada komentar yang bersifat objektif, informatif, berupa pertanyaan, atau pernyataan yang tidak memihak (tidak mendukung maupun menolak).
c. Positif (1): Diberikan pada komentar yang secara eksplisit menunjukkan dukungan, persetujuan, atau apresiasi terhadap kebijakan tersebut (misalnya terkait aspek kedisiplinan).

Proses pelabelan manual ini dilakukan secara teliti pada keseluruhan dataset untuk meminimalisir bias. Contoh tampilan data yang telah melalui proses pelabelan dapat dilihat pada Gambar 4.

Gambar 4. Pelabelan Sentimen

3. Preprocessing

Tahap pra-pemrosesan data merupakan fase fundamental dalam penelitian ini yang bertujuan untuk menjamin kualitas data masukan (*input quality*). Mengingat data yang bersumber dari media sosial cenderung memiliki karakteristik tidak terstruktur (*unstructured text*) dan mengandung banyak *noise*, proses ini diperlukan untuk mentransformasi teks mentah menjadi format yang bersih, konsisten, dan terstandarisasi. Transformasi ini krusial agar model IndoBERT dapat mengenali pola fitur linguistik dan sentimen secara akurat tanpa terganggu oleh elemen data yang tidak relevan.

Rangkaian tahapan pra-pemrosesan yang diterapkan dalam penelitian ini meliputi:

a. Case Folding

Proses *case folding* merupakan langkah awal standarisasi teks. Pada tahap ini, seluruh karakter huruf dikonversi menjadi huruf kecil (*lowercase*). Tujuannya adalah untuk menyeragamkan format teks dan mengurangi dimensi variasi kata yang harus dipelajari model. Sebagai contoh, kata "Sekolah", "SEKOLAH", dan "sekolah" akan dianggap sebagai token yang identik ("sekolah"), sehingga konsistensi representasi leksikal terjaga.

```
def clean_text(text):
    # 1. Case Folding: Mengubah semua teks menjadi huruf kecil
    text = text.lower()
```

Gambar 5. Case Folding

b. Data Cleaning

Tahap pembersihan data (*cleaning*) difokuskan pada eliminasi elemen non-tekstual yang tidak memiliki kontribusi semantik terhadap sentimen. Menggunakan teknik *Regular Expression* (Regex), proses ini menghapus komponen *noise* seperti:

1. Tautan URL (misal: *http://...*)
2. *Mention* pengguna (*@username*)
3. Tagar (*#hashtag*)
4. Karakter non-alfanumerik dan simbol baca yang berlebihan. Hasil dari tahap ini adalah teks bersih yang hanya memuat konten linguistik relevan untuk dianalisis.

```
def clean_text(text):
    text = re.sub(r'http\S+|www\S+|https\S+', '', text, flags=re.MULTILINE)
    text = re.sub(r'@\w+', '', text)
    text = re.sub(r'#\w+', '', text)
    text = re.sub(r'[^a-z0-9\s]', ' ', text)
    text = re.sub(r'\s+', ' ', text).strip()
    return text
```

Gambar 6. Cleaning

c. Text Normalization

Media sosial sarat dengan penggunaan bahasa tidak baku, singkatan, dan istilah gaul (*slang*). Tahap normalisasi bertujuan mengatasi ketidakteraturan ini dengan mengonversi kata-kata *alay* ke dalam bentuk baku sesuai kaidah Bahasa Indonesia. Proses ini memanfaatkan kamus normalisasi khusus (alay_dict) yang memetakan kata tidak baku ke padanan resminya (contoh: "yg" "yang", "bgt" "banget"). Fungsi normalize_text() bekerja dengan memindai setiap token dan mensubstitusinya jika ditemukan kecocokan dalam kamus. Langkah ini krusial untuk meminimalisir munculnya token [UNK] (*unknown*) pada saat tokenisasi IndoBERT dan memperkaya pemahaman konteks model.

Gambar 7. Text Normalization

d. Tokenization

Tokenisasi adalah proses memecah kalimat menjadi unit-unit kata diskrit yang disebut token. Dalam penelitian ini, fungsi tokenize_text() digunakan untuk memisahkan teks berdasarkan spasi (*white space*), mengubah struktur *string* menjadi *list of tokens*. Struktur data ini memudahkan analisis lebih lanjut dan menjadi format dasar sebelum teks diproses oleh *Sub-word Tokenizer* milik IndoBERT.

```
def tokenize_text(text):
    # Tokenizing sederhana (pemecahan string menjadi list kata)
    return text.split()

def detokenize_text(tokens):
    # Menggabungkan kembali list kata menjadi string
    return " ".join(tokens)

# Tokenization ke dalam list (list of tokens)
df['tokenized_text'] = df['normalized_text'].apply(tokenize_text)
```

Gambar 8. Tokenization

e. Stopword Removal

Stopword removal atau penghilangan kata umum (seperti "dan", "yang", "di") diterapkan hanya untuk keperluan analisis eksploratif dan visualisasi *Word Cloud*. Tujuannya adalah agar kata-kata kunci (*keywords*) yang membawa makna topik (tematik) dapat muncul dominan tanpa tertutup oleh kata sambung yang frekuensinya tinggi. **Catatan Penting:** Proses ini tidak diterapkan pada data yang digunakan untuk pelatihan model IndoBERT. Hal ini dikarenakan model berbasis BERT membutuhkan struktur kalimat utuh (termasuk kata sambung) untuk memahami konteks sintaksis dan semantik antar-kata secara komprehensif.

```
# Inisialisasi Stopword Remover
factory = StopWordRemoverFactory()
stopword_remover = factory.create_stop_word_remover()


def remove_stopwords(tokens):
    # Gabungkan token menjadi string, hapus stopword, lalu tokenisasi lagi
    text = detokenize_text(tokens)
    text_no_stopwords = stopword_remover.remove(text)
    return tokenize_text(text_no_stopwords)

# Stopword Removal
df['final_tokens'] = df['tokenized_text'].apply(remove_stopwords)
df['final_text'] = df['final_tokens'].apply(detokenize_text)
```

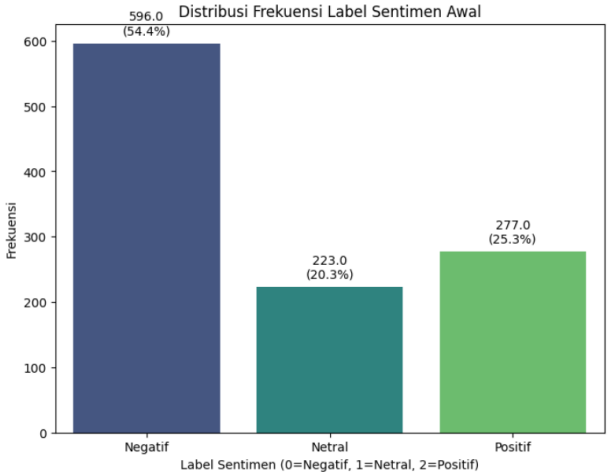
Gambar 9. Stopword Removal

f. Label Mapping & Weighted Loss

Sebelum pemodelan, label kategori dikonversi menjadi format numerik: Negatif (-1) 0, Netral (0) 1, dan Positif (1) 2. Selanjutnya, untuk menangani masalah ketidakseimbangan kelas (*imbalanced dataset*), diterapkan mekanisme *Class Weighting*. Bobot untuk setiap kelas dihitung secara invers proporsional terhadap frekuensinya menggunakan rumus:

$$w_c = \frac{N_{samples}}{N_{classes}\ x\ N_{samples_c}} \qquad (1)$$

Hasil perhitungan bobot ini diimplementasikan ke dalam fungsi kerugian *Weighted Cross-Entropy Loss*. Dengan metode ini, model akan menerima penalti (hukuman) yang lebih besar jika salah memprediksi kelas minoritas, sehingga mendorong model untuk belajar lebih adil dan tidak bias ke kelas mayoritas.



Gambar 10. Distribusi Frekuensi Label & Class Weights

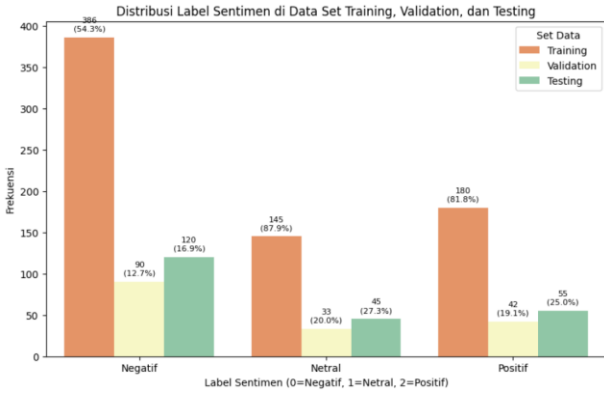
4. Data Spliting

Untuk menjamin validitas evaluasi, dataset dibagi menjadi tiga himpunan bagian (*subset*) menggunakan teknik *Stratified Random Sampling*. Teknik ini dipilih untuk memastikan proporsi kelas sentimen (Positif, Negatif, Netral) pada data *Training*, *Validation*, dan *Testing* tetap seimbang dan merepresentasikan distribusi populasi data asli.

Tabel 5. Distribusi Pembagian Data

Set Data	Proporsi
Training	65%
Validation	15%
Testing	20%

Visualisasi distribusi pada Gambar 10 mengonfirmasi bahwa rasio kelas tetap konsisten di ketiga *subset*, sehingga bias distribusi dapat dihindari.

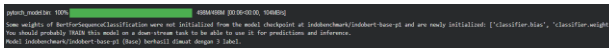


Gambar 11. Label Training, Validation, Testing

5. Modeling

Penelitian ini menerapkan teknik *Fine-Tuning* pada model *IndoBERT Base-P1*, sebuah model *pre-trained* Transformer yang telah dilatih pada korpus Bahasa Indonesia berskala besar.

a. Arsitektur: Model dimodifikasi dengan menambahkan lapisan klasifikasi (*classification layer*) pada bagian output yang memiliki 3 neuron (sesuai jumlah kelas sentimen).
b. Regularisasi**:** *Dropout* sebesar 0.1 diterapkan untuk mencegah *overfitting*.
c. Optimasi: Menggunakan algoritma *AdamW* dengan *learning rate* (2e-5), yang merupakan standar rekomendasi untuk menjaga stabilitas bobot *pre-trained* BERT.



Gambar 12. Inisialisasi Model

6. Data Training

Proses pelatihan berlangsung selama 4 *epoch* menggunakan GPU Google Colab. Selama pelatihan, *Weighted Cross-Entropy Loss* digunakan sebagai fungsi objektif untuk mengoptimalkan performa pada data yang tidak seimbang. Ringkasan performa pelatihan disajikan pada Tabel 6.

Tabel 6. Hasil Pelatihan Model per Epoch

Epoch	Training Loss	Validation Loss	Validation Accuracy	Time (s)
1	0.985833	0.8878	0.578409	4.11005
2	0.673384	0.698425	0.743182	4.58660
3	0.400616	0.764242	0.747727	4.27812
4	0.279555	0.740161	0.782955	10.71720

Analisis Tabel 6 menunjukkan tren konvergensi yang positif. *Training Loss* menurun drastis dari 0.9858 ke 0.2796, menandakan model berhasil mempelajari pola data latih dengan sangat baik. *Validation Accuracy* mencapai puncaknya pada Epoch ke-4 sebesar 78.30%. Meskipun *Validation Loss* sedikit fluktuatif (naik di epoch 3 lalu turun di epoch 4), nilainya masih dalam batas wajar, mengindikasikan model memiliki kemampuan generalisasi yang cukup baik.

7. Evaluation

Model terbaik (Epoch 4) dievaluasi menggunakan data uji (*Testing Set*) sebanyak 221 data. Hasil evaluasi komprehensif disajikan dalam *Classification Report* pada Tabel 4.3.

Tabel 7. Laporan Klasifikasi pada Data Uji

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif (-1)	0.76087	0.875000	0.813953	120
Netral (0)	0.619048	0.577778	0.597701	45
Positif (1)	0.85	0.618182	0.715789	55
Akurasi	**0.75**	0.75	**0.75**	0.75
Macro Avg	0.743306	0.690032	0.709148	220
Weighted Avg	0.754143	0.75	0.745179	220

Berdasarkan tabel di atas, model mencatat akurasi global sebesar **75%**. Secara spesifik:

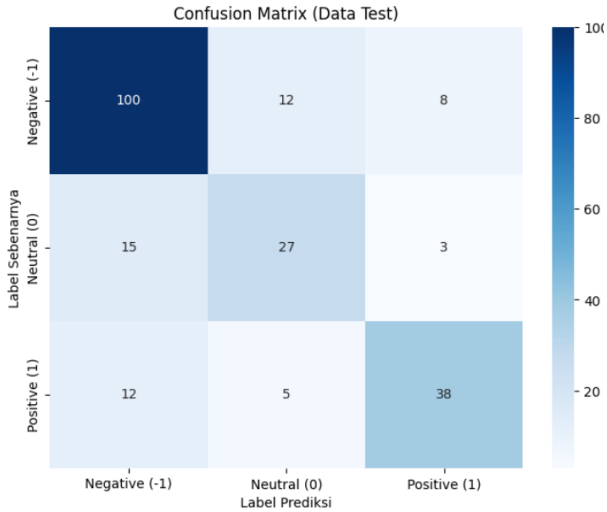
a. Kelas Negatif memiliki performa terbaik (F1-Score 0.81), didukung oleh jumlah data latih yang dominan.
b. Kelas Positif memiliki *Precision* tertinggi (0.85), artinya ketika model memprediksi Positif, kemungkinannya benar sangat tinggi, meskipun *Recall*-nya (0.62) masih perlu ditingkatkan.
c. Kelas Netral memiliki performa terendah, yang wajar terjadi karena seringnya ambiguitas linguistik antara sentimen netral dengan sentimen lainnya.

8. Visualisation dan Analisis Result

Untuk memperdalam interpretasi, dilakukan analisis visual sebagai berikut:

a. Confusion Matrix

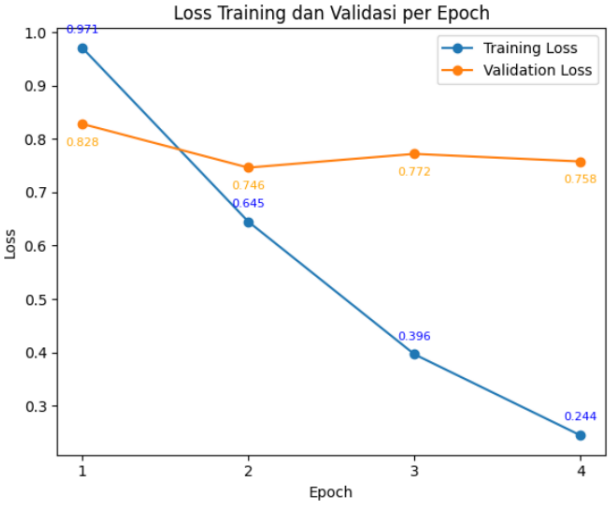
Matriks kebingungan menunjukkan bahwa model sangat andal dalam mengklasifikasikan sentimen Negatif (True Positive tinggi). Kesalahan prediksi (*misclassification*) paling sering terjadi pada kelas Netral yang terprediksi sebagai Positif atau Negatif, yang mengindikasikan adanya irisan fitur bahasa yang tipis pada data kelas tersebut.



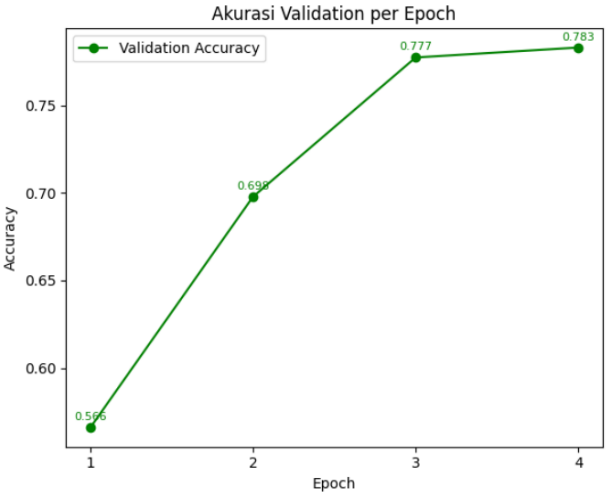
Gambar 13. Confusion Matrix

2. Grafik Loss dan Akurasi

Grafik memperlihatkan *Training Loss* (garis menurun tajam) dan *Validation Loss* (garis landai). Jarak (*gap*) antara keduanya menunjukkan model belajar agresif pada data latih. Peningkatan *Validation Accuracy* yang konsisten hingga Epoch 4 menegaskan bahwa keputusan untuk menghentikan pelatihan di epoch ini sudah tepat untuk mencegah *overfitting* lebih lanjut.



Gambar 14. Training Loss dan Validation Loss



Gambar 15. Validation Accuracy

3. Word Cloud

Visualisasi kata yang paling sering muncul pasca-*stopword removal* menampilkan istilah-istilah dominan yang menjadi fokus publik. Kata-kata ini merepresentasikan inti diskusi terkait penolakan atau dukungan terhadap kebijakan jam masuk sekolah, memvalidasi bahwa model memang mempelajari fitur kontekstual yang relevan dengan topik penelitian.

Gambar 16. Most Common Word

CONCLUSION

Bagian ini berisi kesimpulan yang menjawab hal segala permasalahan yang terdapat didalam penelitian. Isi kesimpulan tidak berupa point-point, namun berupa paragraph. Banyaknya kata pada bagian ini berkisar.

Penelitian ini menyimpulkan bahwa analisis sentimen berbasis IndoBERT mampu memberikan gambaran yang objektif mengenai persepsi masyarakat terhadap kebijakan jam masuk sekolah di Jawa Barat. Berdasarkan data 1.103 komentar dari Twitter dan YouTube, ditemukan bahwa sentimen negatif mendominasi opini publik, menunjukkan adanya resistensi dan ketidaksetujuan terhadap kebijakan tersebut.

Proses *fine-tuning* IndoBERT menghasilkan akurasi 75% dan F1-score tertimbang 0.745, yang menunjukkan bahwa model mampu mengklasifikasikan sentimen secara cukup efektif pada teks Bahasa Indonesia yang bersifat informal. Kinerja model yang sangat baik pada kelas Negatif menjadi indikator bahwa isu ini memunculkan reaksi emosional yang kuat di masyarakat.

Sementara itu, performa kelas Netral yang relatif rendah menunjukkan bahwa komentar bernuansa informatif atau ambigu masih menantang bagi model, terutama karena keterbatasan jumlah data dan keragaman bahasa informal media sosial. Secara keseluruhan, penelitian ini membuktikan bahwa IndoBERT dapat digunakan sebagai alat bantu analitis untuk memahami opini publik terhadap kebijakan pendidikan. Hasil penelitian ini diharapkan menjadi masukan bagi pembuat kebijakan untuk mengevaluasi kebijakan yang telah diterapkan dan meningkatkan strategi komunikasi publik agar lebih responsif terhadap suara masyarakat.

REFERENCE

[1] M. I. Imaduddin, M. R. A'la, and E. Nugroho, "Implementation of IndoBERT for sentiment analysis of health service applications using TF-IDF and word2vec," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 610–618, 2023, doi: 10.14569/IJACSA.2023.0140813.

[2] Y. A. Singgalen, "Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data," *Article in Journal of Information System Research*, vol. 6, no. 2, 2025, doi: 10.47065/josh.v6i2.6505.

[3] R. Merdiansah and A. Ali Ridha, "Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI*, vol. 7, no. 1, pp. 221–228, 2024.

[4] D. Al Akhdaan, Taufik Edy Sutanto, and Muhaza Liebenlito, "Confident Learning pada IndoBERT: Peningkatan Kinerja Klasifikasi Sentimen," *The Indonesian Journal of Computer Science*, vol. 13, no. 5, Oct. 2024, doi: 10.33022/ijcs.v13i5.4401.

[5] A. Jazuli, Widowati, and R. Kusumaningrum, "Optimizing Aspect-Based Sentiment Analysis Using BERT for Comprehensive Analysis of Indonesian Student Feedback," *Applied Sciences (Switzerland)*, vol. 15, no. 1, Jan. 2025, doi: 10.3390/app15010172.

[6] C. Juditha, "The Phenomenon of 'Indonesia Gelap' on Social Media: Sentiment Analysis and Public Opinion Polarization," *Jurnal Komunikasi*, vol. 17, no. 1, pp. 157–170, Jul. 2025, doi: 10.24912/jk.v17i1.33846.

[7] G. P. Rizky and W. Alrasyid, "AI-Based Sentiment Analysis of Social Media to Detect Public Opinion on Government Policies," 2025.

[8] B. Tri Nugroho and A. Noorlistyo Adi, "Sinergi International Journal of Communication Sciences Sentiment Analysis to Know Public Perception Regarding to Public Communication of Indonesian Customs and Excise," vol. 1, no. 2, 2024.

[9] A. Fontanella, H. Ihsan, and Z.-H. Wang, "Sentiment Analysis of Indonesian Government Policy in the Environmental Sector Using Machine Learning Method," 2021. [Online]. Available: https://akuntansi.pnp.ac.id/rafgo

[10] R. Asri, R. A. Kuswardani, D. M. Suliyanti, N. H. Manullang, and F. Ansyari, "Sentiment classification on online customer reviews using IndoBERT pretrained language model," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 3, pp. 1554–1561, 2025, doi: 10.11591/ijeecs.v30.i3.pp1554-1561.

[11] N. P. I. Maharani, Y. Yustiawan, F. C. Rochim, and A. Purwarianti, "Domain-Specific

Language Model Post-Training for Indonesian Financial NLP," Oct. 2023, [Online]. Available: http://arxiv.org/abs/2310.09736