Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.1, February 2026, pp. 1038~1049
e-ISSN: 2548-6861

Performance Analysis of KNN and BERT Algorithms for Classifying Student Sentiments Towards Campus Services

Mutmainna [1], Heliawati Hamrul [2], Wawan Firgiawan[3]
Teknik Informatika, Universitas Sulawesi Barat
innanurul03@gmail.com [1], heliawatyhamrul@unsulbar.ac.id [2], wawanfirgiawan@unsulbar.ac.id [3]

Article Info

Article history:

Received 2025-09-26
Revised 2026-01-23
Accepted 2026-01-30

Keyword:

Sentiment Analysis,
KNN,
BERT,
IndoBERT,
Student Feedback.

ABSTRACT

This study addresses the limitations of campus service evaluation processes that are still conducted manually and are unable to optimally process students' textual opinions. The objective of this research is to analyze and compare the performance of the K-Nearest Neighbor (KNN) and BERT algorithms in classifying student sentiments toward campus services. The research stages include text preprocessing, the generation of IndoBERT embeddings for the KNN model, and fine-tuning IndoBERT for direct sentiment classification. The dataset consists of student evaluation texts from the Faculty of Engineering at UNSULBAR, labeled as negative, neutral, and positive sentiments. Model evaluation is performed using accuracy, precision, recall, and F1-score metrics. The results show that the basic KNN model achieves an accuracy of 79%, while KNN with hyperparameter tuning improves performance to 86%. The BERT model delivers the best performance, achieving an accuracy of 88.68%, precision of 87.87%, recall of 90.19%, and an F1-score of 88.79%. These findings indicate that transformer-based approaches, particularly IndoBERT, are more effective in understanding the contextual nuances of student language than traditional methods, and are therefore more recommended for sentiment analysis implementation in campus service evaluation.

I. PENDAHULUAN

Perguruan tinggi sebagai lembaga penyedia jasa pendidikan dituntut untuk memberikan pelayanan akademik maupun nonakademik yang berkualitas. Kualitas layanan menjadi salah satu indikator penting dalam menentukan performa dan daya saing sebuah institusi pendidikan tinggi [1]. Kepuasan mahasiswa terhadap layanan yang diterima juga menjadi tolok ukur utama dalam menilai mutu perguruan tinggi [2]. Namun, pada praktiknya masih banyak ditemui berbagai keluhan mahasiswa, seperti lambatnya respon fakultas terhadap permasalahan, komunikasi yang tidak efektif, serta keterbatasan fasilitas pendukung [3]. Kondisi ini berdampak pada kepuasan mahasiswa sekaligus citra perguruan tinggi secara keseluruhan. Sistem evaluasi layanan kampus saat ini masih terkendala oleh banyaknya data teks dari kuesioner yang belum terolah secara otomatis, sehingga analisis sering dilakukan secara manual dan memakan waktu. Selain itu, metode klasifikasi tradisional kurang mampu memahami konteks bahasa mahasiswa yang bersifat panjang, bervariasi, dan ambigu, sehingga akurasi penilaian sentimen sering tidak konsisten. Kondisi ini menunjukkan perlunya pendekatan analisis sentimen yang lebih cerdas dan adaptif untuk menghasilkan evaluasi layanan yang lebih tepat dan informatif.

Berbagai penelitian sebelumnya telah mengkaji kepuasan mahasiswa terhadap layanan kampus dengan memanfaatkan metode survei maupun analisis sentimen. Analisis sentimen sendiri merupakan pendekatan *text mining* yang berfungsi untuk menggali opini atau ulasan yang diberikan mahasiswa dalam bentuk teks, kemudian dikategorikan ke dalam sentimen positif, netral, atau negatif [4]. Salah satu algoritma yang banyak digunakan dalam penelitian terdahulu adalah *K-Nearest Neighbor* (KNN) karena kesederhanaan dan efektivitasnya dalam mengklasifikasi data berbasis kemiripan. Penelitian Lukman Rizky Darmawan (2020) misalnya, menunjukkan bahwa KNN mampu mencapai akurasi sebesar 86% dalam menganalisis sentimen mahasiswa

terhadap sistem informasi akademik [5].Sementara itu, penelitian Putri Nabila (2023) membandingkan algoritma KNN dan Naïve Bayes pada analisis sentimen terkait perkuliahan *online*, di mana hasilnya menunjukkan bahwa KNN memperoleh akurasi 78.52%, lebih tinggi dibandingkan *Naïve Bayes* yang hanya mencapai 77.97% dalam mengklasifikasi opini mahasiswa [6].Penelitian lain oleh Meisi Angraini Korompot dkk. (2024) mengembangkan aplikasi evaluasi berbasis web untuk analisis sentimen layanan akademik menggunakan algoritma KNN dengan dataset sebanyak 500 data. Hasil penelitian ini menunjukkan performa yang tinggi dengan *precision* 93%, *recall* 85%, *f1-score* 88%, dan akurasi sebesar 94% [7]. Namun, penelitian-penelitian tersebut masih terbatas pada penggunaan algoritma tradisional seperti KNN, *Naïve Bayes*, atau SVM tanpa mengevaluasi kelemahannya dalam memahami konteks semantik kalimat panjang maupun kalimat ambigu. Selain itu, sebagian besar penelitian terdahulu hanya menggunakan pendekatan berbasis *bag-of-words* atau TF-IDF, sehingga tidak mampu menangkap hubungan antar kata dalam kalimat. Padahal, studi terbaru menunjukkan bahwa model berbasis *transformer* seperti BERT dan IndoBERT mampu meningkatkan akurasi analisis sentimen berbahasa Indonesia secara signifikan karena dapat memahami konteks secara *bidirectional* [8]. Meskipun KNN cukup efektif, algoritma ini memiliki keterbatasan dalam memahami konteks semantik yang lebih mendalam [9].

Hal ini menyebabkan kesulitan dalam membedakan kalimat yang secara struktur mirip, tetapi memiliki makna berbeda. Seiring dengan perkembangan teknologi, pendekatan modern berbasis deep learning seperti *Bidirectional Encoder Representations from Transformers* (BERT) mulai banyak digunakan. Model ini dirancang untuk memahami hubungan antar kata dalam kalimat secara kontekstual, sehingga lebih akurat dalam menangkap makna sentimen. IndoBERT, salah satu varian BERT yang dilatih khusus dengan data berbahasa Indonesia, terbukti mampu memberikan hasil yang lebih representatif dalam pemrosesan bahasa alami *Natural Language Processing* (NLP) [10].

Beberapa Beberapa penelitian terbaru menunjukkan bahwa BERT memiliki performa lebih tinggi dibandingkan algoritma klasik dalam berbagai tugas NLP, termasuk analisis sentimen. Namun, penelitian yang secara khusus membandingkan efektivitas KNN sebagai metode tradisional dengan BERT/IndoBERT sebagai metode modern dalam konteks analisis sentimen mahasiswa terhadap layanan kampus masih sangat terbatas. Padahal, perbandingan ini penting untuk mengetahui sejauh mana model *deep learning* mampu mengatasi keterbatasan algoritma tradisional dalam mengevaluasi persepsi mahasiswa. Selain membandingkan akurasi, penelitian ini juga bertujuan mengevaluasi waktu komputasi, kebutuhan sumber daya, serta kompleksitas implementasi dari masing-masing algoritma. Tujuan ini memberikan gambaran yang lebih komprehensif terkait efektivitas KNN sebagai metode tradisional dan BERT sebagai model modern dalam pengolahan teks opini mahasiswa.

Berdasarkan hal tersebut, penelitian ini dilakukan untuk mengisi gap dengan membandingkan algoritma KNN dan BERT dalam analisis sentimen teks evaluasi layanan kampus. Perbandingan ini diharapkan memberikan pemahaman yang lebih jelas mengenai keunggulan dan kelemahan masing-masing algoritma serta kontribusi praktis bagi peningkatan kualitas layanan berbasis data. Penelitian sebelumnya umumnya hanya membandingkan KNN dengan algoritma tradisional seperti *Naïve Bayes* atau SVM, sementara studi yang menilai kemampuan BERT dalam mengatasi keterbatasan KNN pada teks mahasiswa yang bersifat campuran dan kontekstual masih sangat terbatas. Oleh karena itu, penelitian ini menjadi pembeda dengan menghadirkan evaluasi langsung antara KNN sebagai baseline klasik dan BERT sebagai model modern yang mampu memahami konteks bahasa lebih mendalam, khususnya varian IndoBERT yang relevan untuk bahasa Indonesia. Kedua algoritma tersebut mewakili pendekatan tradisional dan *deep learning* yang layak dianalisis secara komparatif.

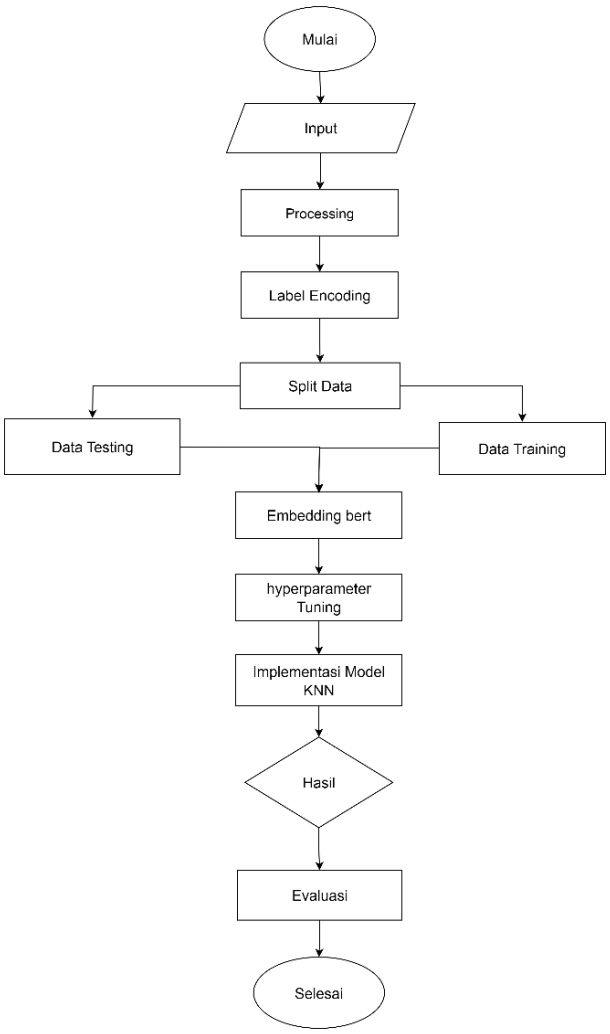
Hingga saat ini belum ada penelitian yang secara langsung membandingkan performa KNN sebagai metode tradisional dengan BERT/IndoBERT dalam analisis sentimen layanan kampus, khususnya pada opini mahasiswa Fakultas Teknik UNSULBAR yang memiliki karakteristik bahasa campuran, panjang bervariasi, dan sering ambigu. Penelitian terdahulu menggunakan KNN umumnya hanya menyoroti akurasi tanpa membahas keterbatasannya dalam memahami konteks kalimat kompleks, sementara studi berbasis BERT lebih banyak diterapkan pada domain umum seperti ulasan produk atau media sosial, bukan pada layanan kampus. Selain itu, belum ada penelitian yang menilai kemampuan kedua algoritma dalam mengklasifikasikan sentimen netral, yang dikenal lebih menantang. Oleh karena itu, penelitian ini mengisi gap tersebut dengan menghadirkan perbandingan komprehensif antara KNN dan BERT pada dataset evaluasi mahasiswa, sekaligus menunjukkan bagaimana kedua pendekatan bekerja pada konteks bahasa yang bersifat domain-spesifik. Kontribusi ilmiah dari penelitian ini tidak hanya terbatas pada perbandingan performa algoritma, tetapi juga pada penerapan *fine-tuning* IndoBERT menggunakan data opini mahasiswa Fakultas Teknik. Selain itu, penelitian ini menyajikan rancangan alur pemrosesan khusus untuk menangani karakteristik teks kuesioner terbuka yang pendek, bervariasi, serta sering menggunakan bahasa tidak baku. Dengan demikian, penelitian ini memberikan kontribusi metodologis berupa adaptasi konteks lokal terhadap model transformer modern dan evaluasi mendalam terhadap efektivitas *baseline* klasik seperti KNN.

II. METODE

Pada tahap ini, data yang dikumpulkan berupa teks hasil evaluasi mahasiswa terhadap layanan kampus, yang diperoleh melalui penyebaran kuesioner secara online maupun langsung kepada mahasiswa aktif di Fakultas Teknik. Kuesioner

tersebut dirancang untuk mengumpulkan pendapat, keluhan, serta masukan mahasiswa mengenai berbagai aspek layanan kampus, seperti layanan administrasi (pengurusan UKT,surat akademik),layanan akademik (jadwal kuliah dosen,nilai, proses akademik), serta fasilitas fisik (kelas,laboratorium ,internet, AC).

Data yang terkumpul kemudian dijadikan sebagai sumber utama untuk menganalisis tingkat kepuasan mahasiswa serta mengidentifikasi area yang memerlukan peningkatan. Proses pengumpulan ini juga bertujuan untuk menangkap variasi pengalaman mahasiswa, baik yang bersifat positif maupun negatif, sehingga diperoleh gambaran yang lebih komprehensif mengenai kualitas layanan yang ada. Dengan adanya masukan langsung dari mahasiswa, pihak fakultas diharapkan dapat merumuskan strategi perbaikan yang lebih efektif, terukur, dan sesuai dengan kebutuhan nyata, sehingga kualitas layanan kampus dapat terus berkembang seiring dengan harapan aktivitas akademika.



Gambar 1. *Flowchart* tahapan penelitian analisis sentimen menggunakan *preprocessing, label encoding, embedding* BERT, *hyperparameter tuning*, dan implementasi model KNN hingga evaluasi hasil[11].

A. Input Data

Dalam penelitian ini, data yang digunakan berupa teks evaluasi atau komentar mahasiswa terkait layanan kampus, yang diperoleh dari kuesioner terbuka, survei daring, maupun media sosial internal fakultas. Data yang di peroleh sebanyak 527 responden,terkumpul total 589 kalimat opini karena sebagian responden memberikan lebih dari satu tanggapan. Data pada penelitian ini berasal dari survei dengan format pertanyaan terbuka *open-ended*. Responden diberikan satu pertanyaan utama dan diminta memberikan opini secara bebas tanpa batasan kata maupun pilihan jawaban. Sebelum dilakukan proses klasifikasi, seluruh data diberi label sentimen oleh peneliti. Distribusi sentimen awal adalah negatif sebanyak 254 data (48,2%), netral sebanyak 517 data (29,8%), dan positif sebanyak 116 data (22,0%). Informasi distribusi sentimen tersebut berfungsi sebagai input utama untuk proses analisis sentimen dengan dua pendekatan berbeda, yaitu *K-Nearest Neighbors* (KNN) dan BERT. Karena data diperoleh melalui kuesioner terbuka, teks yang dihasilkan cenderung pendek, tidak baku, dan bervariasi, sehingga berpotensi menyulitkan model dalam memahami konteks. Untuk mengatasi hal tersebut, penelitian ini melakukan beberapa langkah normalisasi, seperti pembersihan teks, penyeragaman format penulisan, serta penghapusan elemen yang tidak relevan. Selain itu, model BERT digunakan karena kemampuannya membaca konteks secara *bidirectional*, sehingga tetap efektif meskipun input berupa kalimat pendek atau beragam. Dengan tahapan ini, kualitas representasi teks dapat ditingkatkan sebelum masuk ke proses klasifikasi.

B. Preprocessing

Setelah menginput data selanjutnya yaitu Preprocessing. Pada tahap ini dilakukan pembersihan teks, dimana teks mahasiswa mengandung bahasa campuran antar indonesia baku dan tidak baku, penghapusan tanda baca atau karakter khusus, konversi ke huruf kecil [12]. Penelitian ini hanya menerapkan tahap data *cleaning* dan *case folding* pada proses *preprocessing*. Pendekatan ini dipilih karena model BERT telah dirancang untuk menangani teks mentah melalui mekanisme tokenisasi berbasis *subword*. selain itu, penggunaan *preprocessing* minimal bertujuan untuk menjaga keutuhan makna kalimat opini mahasiswa yang umumnya bersifat pendek, kontekstual, dan ambigu. Dengan demikian, tahapan yang diterapkan dinilai sudah memadai untuk mendukung kinerja kedua algoritma yang digunakan.

C. Data Cleaning

Pada tahap ini, data teks dibersihkan dari unsur-unsur yang tidak relevan untuk analisis sentimen, seperti simbol-simbol khusus, angka, emoji, tautan (URL), mention (@username), dan tagar (#)[13]. Pembersihan dilakukan menggunakan *library regular expression* (re) di Python untuk memastikan data siap masuk ke tahap pemrosesan berikutnya. Tujuan utama dari proses ini adalah menghilangkan *noise* agar hasil analisis algoritma KNN dan BERT lebih akurat.

D. Case Folding

Case folding merupakan proses mengubah seluruh huruf dalam teks evaluasi menjadi huruf kecil (*lowercase*). Langkah ini penting untuk mengurangi redundansi kata akibat perbedaan kapitalisasi. Misalnya, kata "Kampus" dan "kampus" akan dianggap sama[14].Proses ini dilakukan dengan fungsi *Python* yang menerapkan *metode .lower*() pada setiap string. Ini memberikan keseragaman bagi algoritma KNN dan BERT dalam membaca teks.

Teks kuesioner sering mengandung bahasa tidak baku, campuran bahasa, dan kesalahan penulisan (*typo*) yang dapat memengaruhi hasil klasifikasi sentimen. Oleh karena itu, pada tahap prapemrosesan data penelitian ini hanya diterapkan proses *case folding* dan *cleaning*. Proses *case folding* bertujuan untuk mengubah seluruh huruf menjadi huruf kecil, sedangkan cleaning dilakukan untuk menghapus karakter khusus, tanda baca, dan elemen yang tidak relavan. Selanjutnya, analisis sentimen dilakukan menggunakan model IndoBERT yang dinilai mampu menangani variasi bahasa Indonesia. Penentuan label sentimen pada dataset dilakukan melalui proses penilaian berbasis aturan yang disusun oleh peneliti. Aturan penilaian ini mempertimbangkan konteks evaluasi layanan kampus dan diterapkan secara otomatis, sehingga setiap data teks memiliki label sentimen positif, negatif, atau netral sebelum digunakan dalam proses analisis [15].

E. Label Encoding

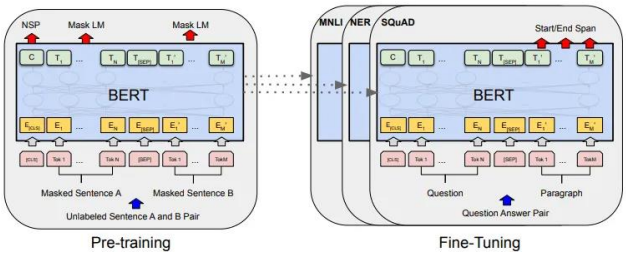
Label encoding merupakan proses mengubah label sentimen yang semula berbentuk teks menjadi format numerik agar dapat diproses oleh algoritma klasifikasi. Dalam penelitian ini menggunakan tiga kelas sentiment positif, negatif, netral, dimana label positif dikonversi menjadi 1, negatif menjadi 0, dan netral menjadi 2[16].Dengan cara ini, data sentimen dapat diproses secara numerik sehingga memudahkan algoritma KNN maupun BERT dalam melakukan klasifikasi.

F. Split Data

Split data merupakan proses membagikan dataset menjadi dua kelompok yaitu data *testing* dan data *training.* Pembagian data menggunakan *library sklearn* dengan fungsi *train test split* untuk membagi data. Rasio pembagian data pada umumnya adalah 80:20 (80% adalah data *training* dan 20% adalah data *testing*) [17]. Pembagian data tersebut menggunakan metode *hold-out* yaitu 20% bagian dari dataset digunakan untuk data testing dan sisanya digunakan untuk data *training*. Dalam penelitian ini, data dibagi menjadi dua bagian menggunakan metode *train–test split* dengan proporsi 80% data latih dan 20% data uji. Proporsi ini dipilih karena umum digunakan dalam penelitian analisis sentimen dan memberikan keseimbangan antara jumlah data untuk pelatihan model dan data untuk evaluasi performa. Pembagian ini memastikan bahwa model diuji dengan data yang belum pernah dilihat sebelumnya sehingga validitas hasil evaluasi tetap terjaga [18].

G. Embedding Bert

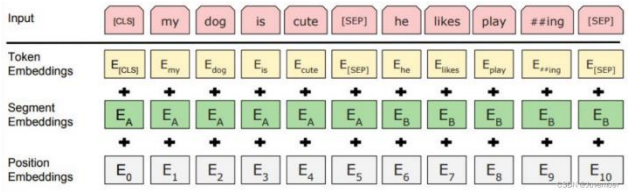
Setelah pembagian data tahap selanjutnya yaitu *Embedding* bert. Dimana Tahap pertama yang dilakukan adalah *set-up pre-trained* BERT, pada tahap ini mencakup instalasi BERT *pre-trainedmodel* sebelum dapat digunakan untuk keperluan sentimen analisis. Penelitian ini menggunakan model IndoBERT Base, yaitu versi BERT yang telah dilatih secara khusus menggunakan korpus besar berbahasa *Indonesia.* Pemilihan IndoBERT dilakukan karena kemampuan model ini lebih sesuai untuk memahami karakteristik bahasa Indonesia dibandingkan *Multilingual* BERT. Penggunaan versi model yang spesifik ini juga penting untuk memastikan replikasi penelitian dapat dilakukan secara konsisten oleh studi lain.Model yang digunakan adalah bert-*base-uncase*.Tahap kedua adalah *Tokenizer and token embedding with* BERT (*Pre-trained* BERT), pada tahap ini, teks inputakan diproses menggunakan *tokenizer* BERT, yang berfungsi untuk memecah kalimat atau teks menjadi potongan kecil yang disebut token. Token-token ini kemudian diubah menjadi representasi numerik melalui *embedding* yang sudah dipelajari di *pre-trained* model BERT [19].
BERT mempunyai dua tahap kerja yaitu *pre- training* dan *fine-tuning.*



Gambar 2. *Prosedur Pre-training dan Fine- tuning* BERT [20]

Gambar 2 tersebut menggambarkan arsitektur dan alur kerja model BERT yang terdiri dari dua tahap utama, yaitu *pre- training* dan *fine-tuning.* Pada tahap *pre-training* (bagian kiri), BERT dilatih menggunakan korpus teks besar tanpa label melalui dua tugas utama, yaitu *Masked Language Model* (MLM) dan *Next Sentence Prediction* (NSP). Pada MLM, beberapa token dalam kalimat sengaja disembunyikan (*masked*) dan *model* dilatih untuk memprediksi token yang hilang, sedangkan NSP digunakan untuk mempelajari hubungan antar kalimat dengan menentukan apakah dua kalimat saling berurutan atau tidak. Proses ini memungkinkan BERT memahami konteks kata secara dua arah (*bidirectional*). Selanjutnya, pada tahap *fine-tuning* (bagian kanan), model BERT yang telah dipralatih disesuaikan dengan tugas spesifik menggunakan data berlabel, seperti klasifikasi teks, pengenalan entitas bernama (NER), atau tanya jawab (*Question Answering*). Pada tahap ini, arsitektur dasar BERT tetap dipertahankan, namun ditambahkan lapisan keluaran sesuai kebutuhan tugas. Dengan mekanisme ini, BERT mampu memanfaatkan pengetahuan bahasa yang telah dipelajari pada tahap pre-training untuk menghasilkan

performa yang lebih baik pada berbagai tugas pemrosesan bahasa alami, termasuk analisis sentimen.



Gambar 3. Representasi *Input* BERT [21]

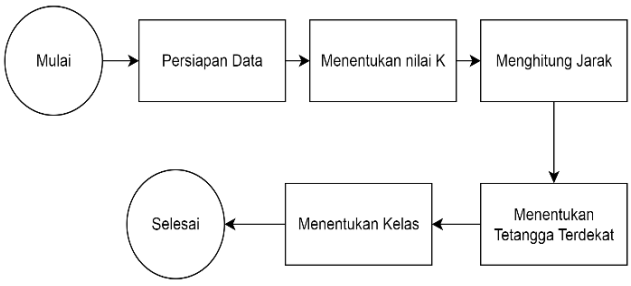
Gambar 3 menunjukkan representasi *input* dari BERT. Setiap urutan dimulai dengan token khusus [CLS]. Token khusus [SEP] ditambahkan untuk memisahkan setiap kalimat. Setiap token memiliki *segment embedding* yang membedakan antara kalimat A dan B. Positional *embedding* untuk memberikan tanda posisi setiap token dalam urutan. *Input encoder* BERT dihasilkan dengan menggabungkan Token *Embeddings*, Segment *Embedding*, dan *Positional Embedding*.

H. Hyperparameter Tuning

Hyperparameter tuning adalah proses mengoptimalkan performa algoritma dengan mencari konfigurasi parameter terbaik. Penelitian ini menggunakan metode *Grid Search* yang secara sistematis menguji berbagai kombinasi *hyperparameter tuning*, seperti nilai k pada KNN serta learning rate, epoch, batch size, dan dropout pada BERT [22]. Setiap kombinasi dievaluasi dengan metrik akurasi, *precision, recall*, atau *F1-score*, lalu dipilih yang menghasilkan performa terbaik meskipun membutuhkan waktu komputasi lebih besar.

I. Implementasi Model KNN

Setelah *Hyperparameter tuning* selanjutnya tahap model KNN. Algoritma KNN melibatkan beberapa tahap yaitu menenentukan nilai K, menentukan nilai K maksimum, mengurutkan nilai K dari yang tertinggi hingga terendah, dan memprediksi kategori objek menggunakan kategori mayoritas tetangga terdekat [23].



Gambar 4. merupakan langkah-langkah metode KNearest Neighbor.

1. Persiapan Data

Pada proses persiapan data ini, telah melakukan crawling data pada komentar Instagram sebagai dataset yang akan dianalisis dalam penelitian ini.

2. Menentukan Nilai K

Mencari jumlah tetangga terdekat untuk membuat keputusan klasifikasi dan menemukan nilai K yang optimal menggunakan cross-validation. Nilai K biasanya digunakan sebagai bilangan ganjil untuk menghindari hasil klasifikasi seri.

3. Menghitung Jarak

Menghitung jarak antara titik data uji dan pelatihan menggunakan Euclidean untuk melakukan analisis kalimat baru dengan setiap kalimat dalam dataset pelatihan.

$$d(xy) = \sqrt{\sum_{i=0}^{n} (x_{1i} - y_{j1i})^2 + (x_{2i} - y_{2i})^2 + \cdots} \quad (1)$$

Keterangan :

- d(x,y)adalah jarak antara dua vektor x dan y
- xi dan yi adalah nilai fitur ke-1 dari masing -masing vektor
- n adalah jumlah dimensi aau fitur dalam representasi vektor

4. Menentukan Tetangga Terdekat

Berdasarkan jarak yang dihitung pada langkah sebelumnya, pilih k pada kalimat terdekat yang akan digunakan untuk menentukan sentiment kalimat baru.

5. Menentukan Kelas

Menentukan kelas yang paling sering muncul (voting mayoritas) dari K tetangga terdekat digunakan untuk mengklasifikasikan data baru.

J. Evaluasi

Setelah pelatihan selesai, model dievaluasi untuk mengukur efektivitasnya dalam melakukan klasifikasi sentimen. Evaluasi yang digunakan yaitu *Confusion Matrix* .*Confusion Matrix* adalah pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih. *Confusion Matrix* adalah tabel dengan 4 kombinasi berbeda dari nilai prediksi dan nilai aktual[24].Berikut adalah struktur *confusion matrix* 2x2 :

TABEL I
CONFUSION MATRIX [25]

Predicted Actual	Positive(1)	Negative(0)
Positive(1)	True Positive(TP)	False Negative(FN)
Negative(0)	False Negative(FN)	True Negative(TN)

Penjelasan tabel I *confusion matrix* di atas adalah sebagai berikut:[26]

1. *True Positive (TP)* : jumlah prediksi yang benar,dimana nilai kelas aktual positif dan diprediksi sebagai positif
2. *False Negative(FN)* : jumlah prediksi yang salah,dimana nilai kelas aktual positif namun diprediksi sebagai negatif.
3. *False Positive(FP)* : jumlah prediksi yang salah, dimana nilai kelas aktual negatif namun diprediksi sebagai positif.

4. *True Negative(TN)* : jumlah prediksi yang benar, dimana nilai kelas aktual negatif dan diprediksi sebagai negatif.

Berikut adalah *performance measure* yang dapat diukur oleh *Confusion Matrix* :

a. *Accuracy*

Accuracy adalah perbandingan antara data yang terdeteksi benar dengan seluruh data hasil prediksi[27].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

b. *Precision*

Precision adalah perbandingan antara data yang terdeteksi benar dengan seluruh data prediksi pada suatu kelas atau tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem [27].

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

c. *Recall*

Recall adalah perbandingan hasil klasifikasi dengan kelas sesungguhnya atau tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi[27].

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

d. *F1-Score*

F1-Score menggambarkan perbandingan rata-rata precision dan recall yang dibobotkan[27].

$$F1 - score = \frac{(2*Recall*Precision)}{Recall+Precision} \quad (5)$$

Selain akurasi, penelitian ini juga menggunakan beberapa metrik evaluasi lain yang umum dipakai dalam analisis performa model NLP, yaitu *precision, recall, F1-score*, dan *confusion matrix. Precision* digunakan untuk mengukur ketepatan prediksi pada masing-masing kelas, *recall* mengukur kemampuan model dalam mengenali seluruh data pada kelas tertentu, sementara *F1-score* memberikan gambaran keseimbangan antara *precision* dan *recall. Confusion matrix* digunakan untuk melihat distribusi prediksi benar maupun salah pada setiap kelas sentimen. Penggunaan metrik-metrik ini penting agar evaluasi performa model tidak hanya bergantung pada akurasi, tetapi juga mencerminkan kemampuan klasifikasi secara lebih komprehensif.

III. HASIL DAN PEMBAHASAN

A. Preprocessing

Pada tahap *preprocessing*, data teks yang semula masih bercampur dengan tanda baca, angka, emotikon, serta variasi penulisan berhasil dibersihkan sehingga lebih seragam dan mudah diolah.untuk menyeleksi data teks agar lebih terstruktur dengan melakukan *cleaning* dan *case folding*. Proses *cleaning* dilakukan untuk menghilangkan karakter atau simbol seperti link URL, emoji, simbol, angka, baris kosong, kolom kosong, missing value, dan data duplikat. Pada Tabel 2 menunjukkan hasil data setelah proses *cleaning*.

TABEL II
HASIL PEMBERSIHAN DATA(CLEANING)

No	Sebelum melakukan *Clening*	Sesudah melakukan *Clening*
1.	1. Peningkatan komunikasi antara fakultas, dosen, dan mahasiswa. 2. Implementasi sistem *feedback* dan evaluasi untuk meningkatkan kualitas layanan. 3. Peningkatan kesadaran dan komitmen seluruh civitas akademika terhadap kualitas layanan yang telah tersedia.	peningkatan komunikasi antara fakultas dosen dan mahasiswa implementasi sistem *feedback* dan evaluasi untuk meningkatkan kualitas layanan peningkatan kesadaran dan komitmen seluruh civitas akademika terhadap kualitas layanan yang telah tersedia
2.	Tambah dan perbaiki ruang kelas dengan peralatan dasar seperti proyektor, *whiteboard*, AC/kipas, dan koneksi internet yang stabil☺	tambah dan perbaiki ruang kelas dengan peralatan dasar seperti proyektor *whiteboard* ackipas dan koneksi internet yang stabil
3.	Perlu perawatan rutin ruang kelas dan laboratorium, terutama kebersihan, pendingin ruangan:)	perlu perawatan rutin ruang kelas dan laboratorium terutama kebersihan pendingin ruangan
4.	Kalau dari saya lebih di tingkatkan lagi apa yang sudah di laksanakan di perbaiki dan untuk fasilitas penunjang kegiatan belajar mohon di perhatikan untuk pihak fakultas agar* lebih baik dan membuat nyaman mahasiswa ☺□☺□	kalau dari saya lebih di tingkatkan lagi apa yang sudah di laksanakan di perbaiki dan untuk fasilitas penunjang kegiatan belajar mohon di perhatikan untuk pihak fakultas agar lebih baik dan membuat nyaman mahasiswa
5.	Perbanyak sosialisasi prosedur akademik, misalnya terkait pengajuan cuti kuliah, pengunduran diri, pengajuan beasiswa, dll	perbanyak sosialisasi prosedur akademik misalnya terkait pengajuan cuti kuliah pengunduran diri pengajuan beasiswa dll
6.	bagusnya libur sesuai jadwal akademik , jangan fakultas lain sudah libur kita fakultas teknik masih ada UAS	bagusnya libur sesuai jadwal akademik jangan fakultas lain sudah libur kita fakultas teknik masih ada uas
7.	Memperjelas prosedur pengajuan keringanan UKT dengan panduan langkah-langkah yang detail.	memperjelas prosedur pengajuan keringanan ukt dengan panduan langkahlangkah yang detail
8.	Untuk meningkatkan kualitas layanan kampus UNSULBAR Fakultas Teknik, beberapa saran dapat diimplementasikan, seperti sistem online untuk pengurusan UKT dan KRS, evaluasi rutin dan pengembangan profesional untuk dosen, sistem administrasi digital, serta penyediaan fasilitas modern dan stabil.	untuk meningkatkan kualitas layanan kampus unsulbar fakultas teknik beberapa saran dapat diimplementasikan seperti sistem online untuk pengurusan ukt dan krs evaluasi rutin dan pengembangan profesional untuk dosen sistem administrasi digital serta penyediaan fasilitas modern dan *internet* stabil
9.	Fasilitas belajar seperti laboratorium dan ruang kelas sudah cukup baik dan memadai untuk kegiatan pembelajaran yang efektif.	fasilitas modern dan internet stabil fasilitas belajar seperti laboratorium dan ruang kelas sudah

		cukup baik dan memadai untuk kegiatan pembelajaran yang efektif
10.	menurut saya pembayaran ukt yang belum jelas kapan tglnya dan berapa jumlah minimal uktnya mungkin bisa ditambahkan disiakad jika sudah selesai semester lansung ada jumlah ukt yang jelas dan pembayarannya atau membuat *portal* informasi yang bisa diakses oleh mahasiswa untuk memperoleh informasi tentang pembayaran ukt termasuk batas wktu pembayaran dan metode pembayaran, jgan disampaikan ketika sudah ingin pengisian krs agar tidak terburu2 dan tetap bisa dikondisikan, dn fasilitas menurut saya mungkin sudah lebih baik dari kemarin2 jadi itu saja yang bisa saya tambahkan:)))	menurut saya pembayaran ukt yang belum jelas kapan tglnya dan berapa jumlah minimal uktnya mungkin bisa ditambahkan disiakad jika sudah selesai semester lansung ada jumlah ukt yang jelas dan pembayarannya atau membuat *portal* informasi yang bisa diakses oleh mahasiswa untuk memperoleh informasi tentang pembayaran ukt termasuk batas wktu pembayaran dan metode pembayaran jgan disampaikan ketika sudah ingin pengisian krs agar tidak terburu dan tetap bisa dikondisikan dn fasilitas menurut saya mungkin sudah lebih baik dari kemarin jadi itu saja yang bisa saya tambahkan
. . .	. . .	. . .
527	Tingkatkan kedisiplinan dosen dalam mengajar sesuai jadwal	tingkatkan kedisiplinan dosen dalam mengajar sesuai jadwal

TABEL III
HASIL CASE FOLDING

NO	Sebelum melakukan *Case Folding*	Sesudah melakukan *Case Folding*
1.	1. Peningkatan komunikasi antara fakultas, dosen, dan mahasiswa. 2. Implementasi sistem *feedback* dan evaluasi untuk meningkatkan kualitas layanan. 3. Peningkatan kesadaran dan komitmen seluruh civitas akademika terhadap kualitas layanan yang telah tersedia.	1. peningkatan komunikasi antara fakultas, dosen, dan mahasiswa. 2. implementasi sistem *feedback* dan evaluasi untuk meningkatkan kualitas layanan. 3. peningkatan kesadaran dan komitmen seluruh civitas akademika terhadap kualitas layanan yang telah tersedia
2.	Tambah dan perbaiki ruang kelas dengan peralatan dasar seperti proyektor, *whiteboard*, AC/kipas, dan koneksi internet yang stabil☺	tambah dan perbaiki ruang kelas dengan peralatan dasar seperti proyektor, *whiteboard*, ac/kipas, dan koneksi internet yang stabil☺
3.	Perlu perawatan rutin ruang kelas dan laboratorium, terutama kebersihan, pendingin ruangan:)	perlu perawatan rutin ruang kelas dan laboratorium, terutama kebersihan, pendingin ruangan:)
4.	Kalau dari saya lebih di tingkatkan lagi apa yang sudah di laksanakan di perbaiki dan untuk fasilitas penunjang kegiatan belajar mohon di perhatikan untuk pihak fakultas agar* lebih baik dan membuat nyaman mahasiswa ☺□☺□	kalau dari saya lebih di tingkatkan lagi apa yang sudah di laksanakan di perbaiki dan untuk fasilitas penunjang kegiatan belajar mohon di perhatikan untuk pihak fakultas agar* lebih baik dan membuat nyaman mahasiswa ☺ ☺
5.	Perbanyak sosialisasi prosedur akademik, misalnya terkait pengajuan cuti kuliah, pengunduran diri, pengajuan beasiswa, dll	perbanyak sosialisasi prosedur akademik, misalnya terkait pengajuan cuti kuliah, pengunduran diri, pengajuan beasiswa, dll.
6.	bagusnya libur sesuai jadwal akademik , jangan fakultas lain sudah libur kita fakultas teknik masih ada UAS	bagusnya libur sesuai jadwal akademik , jangan fakultas lain sudah libur kita fakultas teknik masih ada uas
7.	Memperjelas prosedur pengajuan keringanan UKT dengan panduan langkah-langkah yang detail.	memperjelas prosedur pengajuan keringanan ukt dengan panduan langkahlangkah yang detail
8.	Untuk meningkatkan kualitas layanan kampus UNSULBAR Fakultas Teknik, beberapa saran dapat diimplementasikan, seperti sistem online untuk pengurusan UKT dan KRS, evaluasi rutin dan pengembangan profesional untuk dosen, sistem administrasi digital, serta penyediaan fasilitas modern dan *internet* stabil.	untuk meningkatkan kualitas layanan kampus unsulbar fakultas teknik beberapa saran dapat diimplementasikan seperti sistem online untuk pengurusan ukt dan krs evaluasi rutin dan pengembangan profesional untuk dosen sistem administrasi digital, serta penyediaan fasilitas modern dan *internet* stabil.
9.	Fasilitas belajar seperti laboratorium dan ruang kelas sudah cukup baik dan memadai untuk kegiatan pembelajaran yang efektif.	fasilitas modern dan internet stabil fasilitas belajar seperti laboratorium dan ruang kelas sudah cukup baik dan memadai untuk kegiatan pembelajaran yang efektif.
10.	menurut saya pembayaran ukt yang belum jelas kapan tglnya dan berapa jumlah minimal uktnya mungkin bisa ditambahkan disiakad jika sudah selesai semester lansung ada jumlah ukt yang jelas dan pembayarannya atau membuat *portal* informasi yang bisa diakses oleh mahasiswa untuk memperoleh informasi tentang pembayaran ukt termasuk batas wktu pembayaran dan metode pembayaran, jgan disampaikan ketika sudah	menurut saya pembayaran ukt yang belum jelas kapan tglnya dan berapa jumlah minimal uktnya mungkin bisa ditambahkan disiakad jika sudah selesai semester lansung ada jumlah ukt yang jelas dan pembayarannya atau membuat *portal* informasi yang bisa diakses oleh mahasiswa untuk memperoleh informasi tentang pembayaran ukt termasuk batas wktu pembayaran dan metode pembayaran jgan disampaikan ketika sudah ingin pengisian krs agar

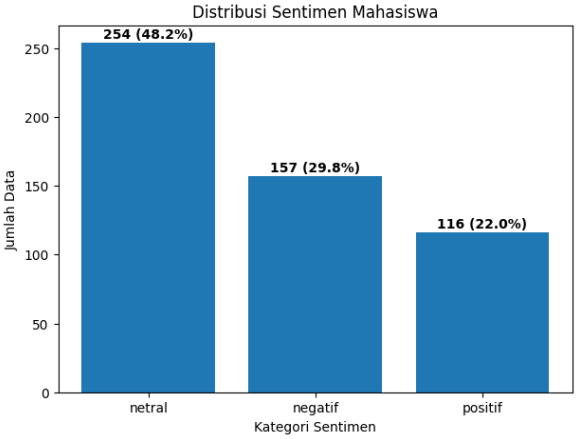
	ingin pengisian krs agar tidak terburu2 dan tetap bisa dikondisikan, dn fasilitas menurut saya mungkin sudah lebih baik dari kemarin2 jadi itu saja yang bisa saya tambahkan:)))	tidak terburu dan tetap bisa dikondisikan dn fasilitas menurut saya mungkin sudah lebih baik dari kemarin jadi itu saja yang bisa saya tambahkan:)))
. . .	. . .	. . .
527	Tingkatkan kedisiplinan dosen dalam mengajar sesuai jadwal.	tingkatkan kedisiplinan dosen dalam mengajar sesuai jadwal.

B. Label Encoding

Kemudian data perlu dilakukan *label encoding* terlebih dahulu sebelum dimodelkan dengan model *Neural Network* IndoBERTweet. Proses *Label Encoding* ini mengubah kata pada kolom label yaitu, positif menjadi angka 1, negatif menjadi 0,dan netra menjadi 2 seperti pada Tabel IV.

TABEL IV
PROSES LABEL ENCODING

NO	Teks	Label	*Label Encoding*
1.	mengenai jadwal kelas yg masih tidak beraturan dengan siakad	Positif	1
2.	Saran untuk pelayanannya lebih objektif	Negatif	0
3.	Libatkan mahasiswa dalam evaluasi pelayanan secara berkala.	Netral	2
4.	Perlu jadwal pengumuman nilai yang lebih tepat waktu agar mahasiswa bisa merencanakan studi dengan lebih baik.	Negatif	0
5.	Perbaiki fasilitas penunjang belajar, seperti jaringan internet/WiFi di kelas dan area umum.	Netral	2
6.	Pengurusan UKT dan KRS berjalan sebagaimana mestinya, namun terkadang masih terjadi kesalahan teknis yang dapat membingungkan mahasiswa baru.	Positif	1
7.	Ada dosen yang cukup konsisten dan disiplin, tetapi masih ada beberapa yang tidak tepat waktu dan kurang terorganisir dalam memberikan materi.	Positif	1
8.	Secara umum layanan akademik di Fakultas Teknik berjalan baik, namun masih bisa ditingkatkan dalam hal kecepatan layanan dan penyebaran informasi.	Negatif	0
9.	Percepat proses pengurusan UKT dan KRS agar tidak membingungkan mahasiswa	Netral	2
10.	Fasilitas kampus cukup layak digunakan, tetapi ada beberapa bagian seperti toilet atau AC yang butuh perhatian lebih.	Positif	1
. . .	. . .	. . .	. . .
527	Tingkatkan kedisiplinan dosen dalam mengajar sesuai jadwal.	Netral	2



Gambar 5. Distribusi sentimen mahasiswa (positif,negatif,netral)

Gambar 5 di atas, menampilkan distribusi kelas sentimen dari total 527 data. Informasi ini penting untuk melihat keseimbangan kelas sebelum proses pelatihan model.

C. Split Data

Pada tahap ini, dataset yang telah melalui proses cleaning dibagi menjadi dua bagian, yaitu data latih (*training set*) dan data uji (*testing set*). Pembagian dilakukan menggunakan fungsi *train_test_split* dari *library scikit-learn* dengan proporsi 80% data latih dan 20% data uji. Dari total 527 data, diperoleh 421 data latih dan 106 data uji. Data latih digunakan untuk membangun dan melatih model klasifikasi, sedangkan data uji berfungsi untuk mengevaluasi performa model yang dihasilkan. Pemisahan data ini penting agar model tidak hanya mengingat pola dari data latih, tetapi juga mampu menggeneralisasi terhadap data baru yang belum pernah dilihat sebelumnya. Dengan demikian, proses *split* data menjadi salah satu langkah krusial dalam memastikan reliabilitas dan validitas hasil analisis sentimen.

D. Embedding Bert

Tabel di atas menampilkan contoh representasi *embedding* hasil pemrosesan teks menggunakan model BERT. Setiap kalimat asli mahasiswa diproyeksikan ke dalam bentuk vektor berdimensi tinggi (768 dimensi) yang merepresentasikan makna semantik kalimat tersebut.

TABEL V
CONTOH REPRESENTASI EMBEDDING BERT(10 DIMENSI LEBIH AWAL)

Kalimat Asli	Dimensi *Embedding* (potong)	
komunikasi antar mahasiswa dan dosen cukup baik tapi perlu ruang dialog yang lebih rutin	[0.82672733 0.3861096 2.7626958 0.82474184 0.03154293	1.5333126 0.106259 -2.1555963 0.39448354 0.56647784] ...
beberapa dosen sudah cukup disiplin meskipun masih ada yang datang terlambat atau jarang masuk	[0.98821026 0.9369668 2.2586918 1.482653 0.57025576	1.182093 -0.0370429 -1.7850853 0.28454986 0.9730654] ...
sistem krs berjalan sesuai jadwal meskipun terkadang ada gangguan teknis di awal masa pengisian	[0.38041988 0.25811455 2.011015 0.5184196	1.8495159 0.48324656 -1.1905895 0.70673835] ...
penggunaan sistem informasi akademik perlu disosialisasikan lebih baik karena masih banyak mahasiswa yang belum memahami prosedur secara menyeluruh	[1.184723 0.7289892 2.4450996 1.1204386 0.28752932	1.4278345 0.48789793 -2.3309548 0.28316712 – 0.66814965] ...
saya merasa puas karena semua pengurusan surat akademik bisa selesai dalam satu hari dengan pelayanan yang ramah	[1.7809695 0.8871504 1.9700454 0.42840832 1.1233228	1.9170977 0.72011286 -1.8535706 -0.09173472 0.9203067] ...

Dalam tabel, hanya ditampilkan 10 dimensi pertama dari *embedding* untuk menjaga keringkasan penyajian, sedangkan dimensi lainnya dilambangkan dengan tanda elipsis (...). *Embedding* yang dihasilkan berupa angka real positif maupun negatif, yang tidak dapat ditafsirkan secara langsung oleh manusia. Namun, nilai- nilai" ini digunakan sebagai input untuk algoritma knn untuk menghitung kedekatan semantik antar kalimat. Misalnya, kalimat dengan makna serupa akan menghasilkan representasi vektor yang relatif berdekatan, sementara kalimat dengan makna yang sangat berbeda akan memiliki jarak vektor yang lebih jauh.

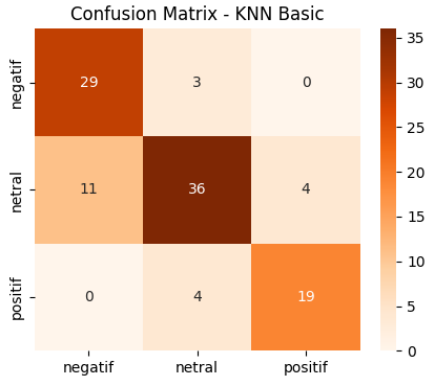
E. Implementasi Model KNN

a. KNN Basic (tanpa Hyperparameter Tuning)

	precision	recall	f1-score	support
negatif	0.72	0.91	0.81	32
netral	0.84	0.71	0.77	51
positif	0.83	0.83	0.83	23
accuracy			0.79	106
macro avg	0.80	0.81	0.80	106
weighted avg	0.80	0.79	0.79	106

Gambar 6. Hasil Evaluasi Model KNN Basic

Gambar 6 tersebut menampilkan hasil evaluasi kinerja model klasifikasi sentimen berdasarkan metrik precision, recall, f1-score, dan jumlah data (support) untuk tiga kategori sentimen, yaitu negatif, netral, dan positif. Pada kelas negatif, model memperoleh precision sebesar 0.72, recall 0.91, dan f1-score 0.81, menandakan kemampuan yang kuat dalam mengenali data negatif. Kelas netral memiliki precision 0.84, recall 0.71, dan f1-score 0.77, menunjukkan bahwa meskipun prediksi netral cukup akurat, model masih melewatkan sebagian data netral. Kelas positif menunjukkan performa stabil dengan precision dan recall yang sama-sama 0.83 dan f1-score 0.83. Secara keseluruhan, gambar tersebut memperlihatkan bahwa model mencapai akurasi 0.79, dengan nilai macro average f1-score 0.80 dan weighted average f1-score 0.79, yang secara umum mengindikasikan performa model yang baik dan relatif seimbang dalam mengklasifikasikan ketiga jenis sentimen.



Gambar 7. *Confusion Matrix* KNN Basic(tanpa Hyperparameter

Gambar 7 *confusion matrix* tersebut menunjukkan kinerja model KNN Basic dalam mengklasifikasikan tiga kategori sentimen, yaitu negatif, netral, dan positif. Model mampu mengidentifikasi kelas negatif dengan cukup baik, ditunjukkan oleh 29 data yang diprediksi benar, meskipun masih terdapat 3 data negatif yang salah diklasifikasikan sebagai netral. Pada kelas netral, model memprediksi dengan benar sebanyak 36 data, namun masih terjadi kesalahan klasifikasi ke kelas negatif sebanyak 11 data dan ke kelas positif sebanyak 4 data. Sementara itu, kelas positif memperoleh 19 prediksi benar, dengan sebagian kecil kesalahan yaitu 4 data positif yang diklasifikasikan sebagai netral. Secara keseluruhan, gambar ini menggambarkan bahwa kinerja model paling optimal pada kelas netral, cukup baik pada kelas negatif dan positif, namun tetap menunjukkan adanya tumpang tindih prediksi antarkelas yang perlu diperbaiki lebih lanjut.

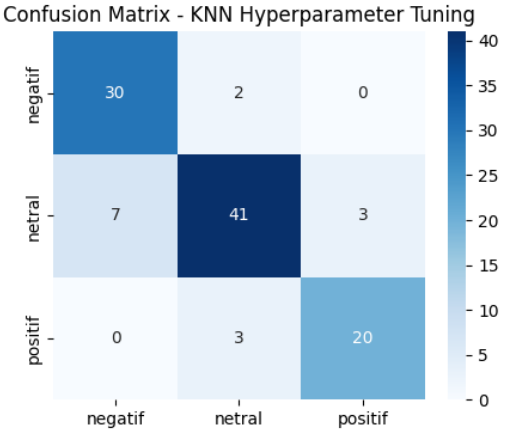
b. KNN + Hyperparameter Tuning

	precision	recall	f1-score	support
negatif	0.81	0.94	0.87	32
netral	0.89	0.80	0.85	51
positif	0.87	0.87	0.87	23
accuracy			0.86	106
macro avg	0.86	0.87	0.86	106
weighted avg	0.86	0.86	0.86	106

Gambar 8. Hasil Evaluasi Model KNN+Hyperparameter Tuning

Gambar 8 tersebut menunjukkan *classification report* yang berisi metrik evaluasi kinerja model dalam mengklasifikasikan sentimen ke dalam tiga kelas: negatif, netral, dan positif. Setiap kelas memiliki nilai *precision*, *recall*, dan *f1-score* yang relatif tinggi, dengan rentang 0,80–0,94. Kelas negatif memperoleh *recall* tertinggi (0,94),

menunjukkan bahwa model sangat baik dalam menemukan data yang benar-benar termasuk kategori negatif. Sementara itu, kelas positif dan netral memiliki kinerja seimbang dengan *f1-score* masing-masing 0,87 dan 0,85. Secara keseluruhan, model memperoleh nilai akurasi sebesar 0,86, serta nilai *macro average* dan *weighted average* yang konsisten di angka 0,86–0,87, menandakan bahwa model mampu melakukan klasifikasi dengan stabil di seluruh kelas pada 106 data uji.

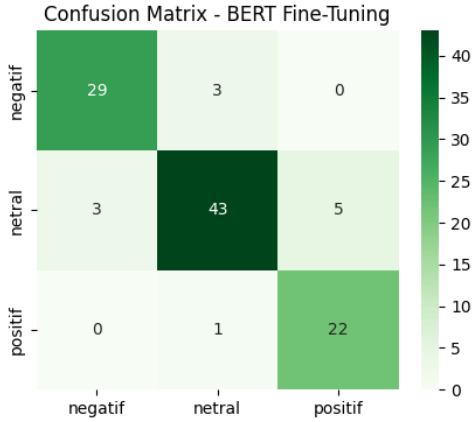


Gambar 9. Confusion Matrix KNN Dengan Hyperparameter Tuning

Gambar 9 tersebut memperlihatkan *confusion matrix* hasil tuning hyperparameter pada algoritma K-Nearest Neighbors (KNN**)** yang digunakan untuk mengklasifikasikan sentimen ke dalam tiga kategori: negatif, netral, dan positif. Matriks ini menunjukkan distribusi prediksi model terhadap data uji. Pada kelas negatif, model berhasil mengklasifikasikan dengan tepat sebanyak 30 data, dengan hanya sedikit kesalahan prediksi ke kelas netral 2 data. Untuk kelas netral, model menghasilkan prediksi benar sebanyak 41 data, namun masih terjadi mis-klasifikasi ke kelas negatif 7 data dan positif 3 data. Sementara itu, pada kelas positif, model mengidentifikasi 20 data secara benar, dengan kesalahan kecil berupa 3 data yang diprediksi sebagai netral. Secara keseluruhan, pola pada confusion matrix ini menggambarkan bahwa model KNN yang telah di-tuning mampu memberikan performa klasifikasi yang cukup baik, dengan tingkat ketepatan prediksi yang dominan berada pada diagonal utama matriks.asiswa dibandingkan pendekatan tradisional seperti KNN.

c. *Hasil Bert*

Gambar 10 tersebut menampilkan *confusion matrix* hasil *fine-tuning* model BERT untuk tugas klasifikasi sentimen dengan tiga kelas: negatif, netral, dan positif. Matriks ini menunjukkan jumlah prediksi benar dan salah yang dilakukan model terhadap data uji. Pada kelas negatif, model berhasil memprediksi dengan benar sebanyak 29 data, dan hanya salah memprediksi 3 data sebagai netral. Untuk kelas netral, performanya sangat baik dengan 43 data terklasifikasi benar, meskipun masih terdapat kesalahan kecil ke kelas negatif 3 data dan positif 5 data. Pada kelas positif, model mengenali 22 data dengan tepat, dan hanya terjadi satu kesalahan prediksi menuju kelas netral.



Gambar 10. Confusion Matrix BERT

Secara keseluruhan, pola *confusion matrix* memperlihatkan bahwa model BERT yang telah di-fine-tune mampu mengidentifikasi kelas dengan akurasi tinggi, ditandai oleh dominasi angka pada diagonal utama. Kinerja yang stabil pada ketiga kelas menunjukkan bahwa BERT memiliki kemampuan representasi teks yang lebih baik dan lebih presisi dibandingkan metode tradisional seperti KNN.

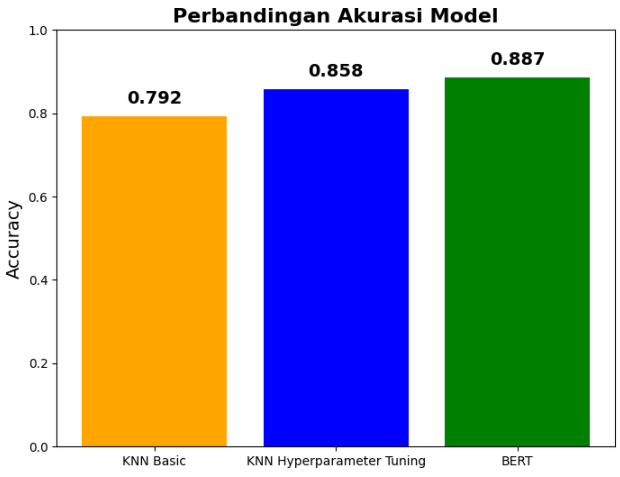
F. Pembahasan

Analisis pada bagian hasil belum hanya bersifat deskriptif, tetapi juga perlu menunjukkan alasan mengapa performa setiap model berbeda. Pada penelitian ini, pembahasan diperkuat dengan melihat pola kesalahan prediksi dan kemampuan masing-masing algoritma dalam menangani karakteristik teks mahasiswa. BERT menunjukkan keunggulan karena mampu memahami konteks kalimat secara mendalam, terutama pada sentimen netral dan kalimat ambigu. Sebaliknya, KNN memiliki keterbatasan pada kalimat panjang dan teks dengan makna implisit karena hanya mengandalkan kedekatan vektor. Pendekatan analitis ini menegaskan bahwa perbedaan performa model tidak hanya ditunjukkan melalui angka akurasi, tetapi juga melalui penjelasan logis mengenai mekanisme kerja masing-masing algoritma.Untuk mengevaluasi kinerja algoritma *K-Nearest Neighbors* (KNN), penelitian ini menggunakan dua pendekatan, yaitu KNN Basic dan KNN dengan *Hyperparameter Tuning*. KNN Basic dijalankan tanpa penyesuaian parameter sehingga hasil yang diperoleh merepresentasikan performa standar dari algoritma tersebut. Sementara itu, pada KNN dengan Hyperparameter Tuning dilakukan optimasi parameter seperti nilai k, distance metric, serta skema pembobotan, dengan tujuan meningkatkan kemampuan model dalam mengenali pola pada data.

Hasil penelitian menunjukkan adanya perbedaan performa yang cukup jelas antara algoritma KNN Basic, KNN setelah dilakukan hyperparameter tuning, dan model BERT dalam mengklasifikasikan sentimen mahasiswa. Pada KNN Basic, akurasi yang diperoleh sebesar 79%, dengan precision 80%, recall 81%, dan f1-score 80%. Hasil ini berada dalam kisaran yang sejalan dengan penelitian Lukman Rizky Darmawan

(2020), yang menunjukkan bahwa KNN mampu mencapai akurasi 86% pada analisis sentimen layanan akademik. Kesamaan ini mengindikasikan bahwa KNN cukup efektif untuk tugas klasifikasi berbasis kemiripan fitur, namun masih kurang optimal ketika dihadapkan pada variasi gaya bahasa mahasiswa yang panjang, tidak baku, dan ambigu. Setelah dilakukan hyperparameter tuning, performa KNN meningkat menjadi 86% dengan nilai precision, recall, dan f1-score yang lebih stabil. Peningkatan ini konsisten dengan temuan Putri Nabila (2023), yang menyatakan bahwa pemilihan parameter KNN secara tepat dapat meningkatkan akurasi dan kestabilan prediksi pada analisis sentimen multibahasa. Meskipun demikian, model ini masih menunjukkan kelemahan terutama pada klasifikasi sentimen netral, di mana beberapa kalimat ambigu cenderung salah prediksi akibat keterbatasan KNN dalam memahami konteks semantik. Berbeda dengan kedua hasil tersebut, model BERT menunjukkan performa terbaik dengan akurasi 88.68%, precision 87.87%, recall 90.19%, dan f1-score 88.79%. Kinerja ini mencerminkan keunggulan BERT dalam memahami konteks kalimat secara mendalam melalui mekanisme attention, sehingga lebih efektif dalam menangani opini mahasiswa yang bersifat campuran atau memiliki makna implisit. Selain itu, BERT lebih stabil dalam membedakan kelas netral yang pada metode tradisional sering menjadi tantangan. keunggulan BERT bukan hanya dibuktikan melalui nilai akurasi, tetapi juga melalui alasan metodologis yang menjelaskan kemampuan model dalam memahami hubungan antar kata dan makna kalimat. Penjelasan ini menegaskan bahwa interpretasi hasil penelitian tidak hanya bersifat deskriptif, tetapi juga didukung oleh logika ilmiah yang kuat. Dengan demikian, penelitian ini memperkuat temuan sebelumnya mengenai efektivitas KNN, sekaligus memberikan kontribusi baru bahwa pendekatan berbasis deep learning seperti BERT/IndoBERT lebih representatif dan konsisten untuk menganalisis opini mahasiswa terhadap layanan kampus

Perbandingan hasil akurasi dari ketiga pendekatan tersebut dapat dilihat pada Gambar 11 berikut.



Gambar 11. Perbandingan Akurasi KNN Basic, KNN HyperparameterTuning dan Bert

Gambar 11 tersebut menampilkan perbandingan akurasi tiga model klasifikasi sentimen, yaitu KNN Basic, KNN dengan *Hyperparameter Tuning*, dan BERT. Grafik batang menunjukkan bahwa setiap model memiliki tingkat akurasi yang berbeda dengan peningkatan signifikan pada metode yang lebih kompleks. Model KNN Basic memperoleh akurasi sebesar 0,792, yang merupakan performa terendah di antara ketiganya. Setelah dilakukan *hyperparameter tuning*, akurasi model KNN meningkat menjadi 0,858, menunjukkan bahwa pengaturan parameter yang optimal mampu memperbaiki kinerja model secara jelas. Sementara itu, model BERT menunjukkan performa terbaik dengan akurasi tertinggi yaitu 0,887, menggambarkan kemampuan representasi bahasa yang lebih unggul dibandingkan metode KNN. Secara keseluruhan, grafik ini mengilustrasikan bahwa kompleksitas dan kemampuan pemodelan yang lebih baik pada BERT menghasilkan akurasi klasifikasi yang lebih tinggi.

Analisis kesalahan menunjukkan bahwa model KNN sering salah mengklasifikasikan kalimat panjang dan deskriptif karena keterbatasannya dalam memahami konteks, misalnya pada kalimat berisi keluhan detail yang justru diprediksi sebagai netral. KNN juga kerap keliru pada kalimat ambigu yang mengandung dua sentimen sekaligus, seperti ucapan apresiasi yang diikuti kritik, sehingga prediksinya cenderung terlalu positif. Sebaliknya, BERT mampu menangani kedua jenis kalimat tersebut dengan lebih baik karena memahami hubungan antar kata secara kontekstual. Namun, BERT masih melakukan kesalahan pada kalimat yang sangat pendek atau terlalu umum akibat kurangnya informasi semantik. Secara keseluruhan, kesalahan terbesar KNN terjadi pada teks panjang dan ambigu, sedangkan BERT cenderung salah pada teks yang minim konteks. Hasil penelitian ini memiliki implikasi praktis yang penting bagi pengelolaan evaluasi layanan kampus. Dengan performa BERT yang lebih unggul, model ini dapat diintegrasikan ke dalam sistem survei kampus untuk melakukan analisis sentimen secara otomatis dan *real-time* sehingga proses evaluasi tidak lagi memerlukan pembacaan manual yang memakan waktu. Model ini juga dapat diterapkan sebagai modul pemantauan kepuasan mahasiswa dalam aplikasi layanan akademik, di mana setiap komentar mahasiswa dapat langsung diklasifikasikan ke dalam kategori positif, negatif, atau netral. Dengan demikian, pengambil kebijakan dapat lebih cepat mengidentifikasi masalah dan melakukan perbaikan layanan secara lebih responsif.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang membandingkan kinerja algoritma *K-Nearest Neighbors* (KNN) dan BERT/IndoBERT dalam mengklasifikasikan sentimen mahasiswa terhadap layanan kampus di Fakultas Teknik UNSULBAR, dapat disimpulkan bahwa model berbasis *deep learning* memberikan performa yang lebih unggul. KNN Basic memperoleh akurasi sebesar 0.79%, meningkat menjadi 0.85% setelah dilakukan *hyperparameter tuning*, namun

performanya masih berada di bawah model BERT. Sementara itu, BERT Fine-Tuning mencapai akurasi tertinggi yaitu 0.88%, serta menunjukkan kemampuan yang lebih baik dalam mengklasifikasikan ketiga kelas sentimen, terutama pada kelas netral yang cenderung sulit diprediksi oleh metode tradisional. Hasil *confusion matrix* juga memperlihatkan bahwa BERT memiliki tingkat kesalahan klasifikasi lebih rendah dibandingkan KNN pada semua kategori sentimen. Temuan ini menegaskan bahwa BERT lebih efektif dalam memahami konteks, struktur, serta variasi bahasa alami mahasiswa yang bersifat campuran, ambigu, dan kontekstual. Oleh karena itu, model BERT direkomendasikan sebagai pendekatan yang lebih akurat dan adaptif untuk implementasi analisis sentimen dalam sistem evaluasi layanan kampus berbasis data.

Saran untuk penelitian selanjutnya adalah memperluas penggunaan model deep learning lain sebagai pembanding, sehingga dapat diketahui sejauh mana performa masing-masing algoritma dalam klasifikasi sentimen. Selain itu, pengujian sebaiknya dilakukan dengan jumlah data yang lebih besar dan lebih bervariasi agar hasil yang diperoleh lebih representatif dan komprehensif. Diperlukan pula pendekatan tambahan seperti aspect-based sentimen analysis atau integrasi metode pembelajaran mesin lain untuk menangkap makna yang lebih mendalam dari opini mahasiswa. Dengan pengembangan lebih lanjut, diharapkan sistem analisis sentimen ini dapat semakin akurat, adaptif, dan memberikan kontribusi nyata bagi peningkatan mutu layanan kampus.

DAFTAR PUSTAKA

[1] S. Masyita, "Kepuasan Mahasiswa Ditinjau dari Kualitas Pelayanan Akademik dan Non Akademik (Studi pada Program Studi Manajemen STIEM Bongaya)," *J. Brand*, vol. 2, no. 1, p. 90, 2020, [Online]. Available: https://forlap.ristekdikti.go.id

[2] D. Suranti, Y. Yupianti, and V. N. Sari, "Implementasi Metode Bayes Dalam Evaluasi Kepuasan Mahasiswa Terhadap Layanan Universitas," *SINTECH (Science Inf. Technol. J.*, vol. 4, no. 1, pp. 12–21, 2021, doi: 10.31598/sintechjournal.v4i1.569.

[3] J. Waruwu, J. Dawolo, D. Laoli, Y. P. Harefa, O. Zebua, and D. C. Lase, "Nias Menggunakan Metode Analisis Regresi," vol. 4, no. 2, pp. 143–148, 2024.

[4] A. Bulu, E. Umar, and P. M. Ate, "Analisis Sentimen Terhadap Sistem Informasi Akademik STIMIKOM Stella Maris Sumba Menggunakan Algoritma Naïve Bayes," *J. Sist. Inf. Dan Inform.*, vol. 2, no. 1, pp. 115–124, 2023, doi: 10.47233/jiska.v1i2.1085.

[5] L. R. Dharmawan, I. Arwani, and D. E. Ratnawati, "Analisis Sentimen pada Sosial Media Twitter Terhadap Layanan Sistem Informasi Akademik Mahasiswa Universitas Brawijaya dengan Metode K-Nearest Neighbor," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, pp. 959–965, 2020, [Online]. Available: http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/7099

[6] P. Nabila, F. Wajidi, and W. Firgiawan, "Analisis Algoritma K-Nearest Neighbor dan Naive Bayes pada Kebijakan Perkuliahan Online dengan Multi Bahasa," *Techno.Com*, vol. 24, no. 2, pp. 524–542, 2025, doi: 10.62411/tc.v24i2.12656.

[7] M. A. Korompot, V. P. Rantung, and ..., "Aplikasi Evaluasi Analisis Sentimen Terhadap Layanan Akademik Menggunakan Algoritma K-Nearest Neighbor Berbasis Web," *... , Educ. ...*, pp. 18–26, 2024, [Online]. Available: https://jibeit.teknikinformatika.org/index.php/jibeit/article/view/60

[8] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," 2019.

[9] I. H. Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 302–307, 2023, doi: 10.30591/jpit.v8i3.5734.

[10] D. A. YULIASARI, "Pengukuran Kepuasan Mahasantri Terhadap Pesantren Mahasiswa Unissula Berbasis Analisis Sentimen Menggunakan Indobert," *Accid. Anal. Prev.*, 2023.

[11] W. Firgiawan and D. Yustianisa, "Penyetelan Hiperparameter untuk Mengoptimalkan Stunting Klasifikasi dengan Algoritma KNN , SVM , dan Naïve Bayes Machine Translated by Google," vol. 19, no. 1, pp. 92–104, 2021.

[12] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *Smatika J.*, vol. 10, no. 02, pp. 71–76, 2020, doi: 10.32664/smatika.v10i02.455.

[13] S. P. Collins *et al.*, "No Title 済無No Title No Title No Title," no. April 2024, pp. 167–186, 2021.

[14] N. Cahyono and Anggista Oktavia Praneswara, "Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes," *Indones. J. Comput. Sci.*, vol. 12, no. 6, pp. 3925–3940, 2023, doi: 10.33022/ijcs.v12i6.3473.

[15] A. Tantoni, "Ekstraksi Label Otomatis untuk Analisis Sentimen dalam Bahasa Indonesia Teks," vol. 15, no. 3, pp. 718–728, 2025.

[16] M. Z. Subarkah, M. Hildha, N. T. Amanda, and E. Zukhronah, "Analisis Sentimen Review Tempat Wisata Pada Data Analisis Sentimen Review Tempat Wisata Pada Data Online Travel Agency Di Yogyakarta Menggunakan Model Neural Network IndoBERTweet Fine Tuning (Analysis of Sentiment Reviews of Tourist Attractions on Online ," *Semin. Nas. Off. Stat.*, vol. 1, no. https://prosiding.stis.ac.id/index.php/semnasoffstat/issue/view/6, pp. 543–552, 2022.

[17] S. Ratnaswari, N. C. Wibowo, D. Satria, and Y. Kartika, "Analisis Sentimen Menggunakan Metode Lexicon-Based Dan Support Vector Machine Pada Presiden Dan Wakil Presiden Indonesia," vol. 13, no. 1, pp. 362–368, 2025.

[18] P. Studi, S. Informasi, and F. Sains, "Prediksi Produksi Batu Andesit Menggunakan Algoritma Support Vector Regression Prediction Of Andesit Stone Production Using," vol. 13, pp. 1893–1901, 2024.

[19] Z. A. Sriyanti, D. S. Y. Kartika, and A. R. E. Najaf, "Implementasi Model Bert Pada Analisis Sentimen Pengguna Twitter Terhadap Aksi Boikot Produk Israel," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, pp. 2335–2342, 2024, doi: 10.23960/jitet.v12i3.4743.

[20] A. C. Surya, N. Aurora, A. Satria, and C. T. Nadeak, "Implementasi sentence transformer berbasis bert untuk edukasi dan informasi tentang hiv/aids," vol. 2, no. 1, pp. 607–614, 2025.

[21] S. Kasus, S. Ransomware, and D. Nasional, "Topic Modelling Berbasis Embedding pada Komentar YouTube," pp. 873–884, 2024.

[22] Joshua Agung Nurcahyo and Theopilus Bayu Sasongko, "Hyperparameter Tuning Algoritma Supervised Learning untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras," *Indones. J. Comput. Sci.*, vol. 12, no. 3, pp. 1351–1365, 2023, doi: 10.33022/ijcs.v12i3.3254.

[23] N. Alvionika, S. Faisal, R. Rahmat, and A. F. N. Masruriyah, "Analisis Sentimen Pada Komentar Instagram Provider By.U Menggunakan Metode K-Nearest Neighbors (KNN)," *J. Algoritm.*, vol. 21, no. 2, pp. 50–63, 2024, doi: 10.33364/algoritma/v.21-2.1672.

[24] P. Studi, T. Informatika, F. Sains, D. A. N. Teknologi, U. Islam, and N. Syarif, "No Title," 2024.

[25] Y. R. Nasution, A. H. Lubis, and T. F. Eliza, "Analisis Sentimen Platform X Mengenai Pro Kontra Rekrutmen Guru Melalui Marketplace Menggunakan Metode Naïve Bayes," *Bull. Comput. Sci. Res.*, vol. 4, no. 6, pp. 406–413, 2024, doi: 10.47065/bulletincsr.v4i6.368.

[26] V. A. Permadi, "Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura," pp. 141–151.

[27] S. M. Natzir, "Perbandingan Kinerja Model Pembelajaran Mesin dalam Prediksi Banjir menggunakan KNN , Naive Bayes , dan Random Forest," vol. 14, no. c, pp. 59–64, 2023.