

Comparison of K-Nearest Neighbors and Naive Bayes Classifier Algorithms in Sentiment Analysis of 2024 Election in Twitter (X)

Lola Enjelia ^{1*}, Yana Cahyana ^{2*}, Rahmat ^{3*}, Deden Wahiddin ^{4*}

* Teknik Informatika, Universitas Buana Perjuangan Karawang

if21.lolaenjelia@mhs.ubpkarawang.ac.id ¹, yana.cahyana@ubpkarawang.ac.id ², rahmat@ubpkarawang.ac.id ³,
deden.wahiddin@ubpkarawang.ac.id ⁴

Article Info

Article history:

Received 2025-05-02

Revised 2025-06-09

Accepted 2025-06-19

Keyword:

*Sentiment analysis,
K-Nearest Neighbor,
Naive Bayes,
2024 Election,
Twitter (X).*

ABSTRACT

This study compares the performance of the K-Nearest Neighbors (K-NN) and Naive Bayes Classifier (NBC) algorithms in sentiment analysis of the 2024 Regional Election (Pilkada) based on Indonesian local data sourced from platform X. A total of 1,187 tweets were collected through crawling, followed by extensive preprocessing and manual sentiment labeling by a professional linguist to ensure data validity and reliability. The study highlights NBC's superior accuracy (81.05%) compared to K-NN (75.26%), largely due to the characteristics of short-text social media data that align with NBC's independence assumptions. Key terms identified through TF-IDF analysis include "pilkada", "2024", and "damai" in positive sentiment, while "mahkamah konstitusi" and "kalah" dominated negative sentiment. The results imply that although public discourse largely supports the election process, critical sentiments toward election dispute issues persist. These findings offer practical implications for election authorities, policymakers, and digital campaign strategists, particularly in optimizing public communication strategies, early detection of potential conflicts, and designing public opinion monitoring systems based on real-time sentiment analysis. By leveraging high-quality labeled local data, this study makes a significant contribution to modeling public opinion dynamics in Indonesia during political events.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Dalam sistem demokrasi Indonesia, Pemilihan Kepala Daerah (PILKADA) menjadi momen penting untuk memilih gubernur, bupati, dan wali kota sebagai pemimpin daerah. Proses ini memiliki peran vital dalam menentukan arah kebijakan dan pemerintahan daerah dalam periode tertentu. Namun, pasca pemilihan, sering kali muncul dinamika politik lanjutan. Perselisihan terkait hasil pemilu, aksi protes dari pendukung calon yang tidak terpilih, serta konflik internal dalam pemerintahan dapat berdampak pada stabilitas sosial dan politik. Ketidakpuasan terhadap hasil pemilu kerap berujung pada gugatan ke Mahkamah Konstitusi atau lembaga pengawas pemilu, yang berpotensi menimbulkan ketegangan, terutama bila isu-isu tersebut meluas melalui media sosial. Dalam konteks ini, analisis sentimen menjadi sangat penting untuk menangkap pandangan publik, terutama di tengah tantangan seperti penyebaran disinformasi,

polarisasi opini, dan kampanye negatif yang marak di platform digital.

Opini publik memainkan peran yang sangat penting dalam menentukan hasil akhir pemilihan. Media sosial, khususnya platform X (sebelumnya Twitter), X merupakan situs jejaring sosial yang terkenal untuk menyebarkan informasi dan pandangan. Pengguna memanfaatkannya untuk mengungkapkan aktivitas, pemikiran, atau perasaan mereka [1]. Platform X telah menjadi ruang bagi publik untuk secara terbuka menyuarakan opini, kritik, maupun dukungan terhadap calon kepala daerah. Situasi ini menimbulkan tantangan dalam menafsirkan makna dan tujuan dari setiap unggahan yang dibagikan. Beragam pandangan, baik yang mendukung maupun yang menentang, mencerminkan aspirasi serta ekspektasi masyarakat terhadap pelaksanaan PILKADA, termasuk terkait sosok calon, afiliasi partai, dan isu-isu yang menyertainya. Oleh sebab itu, dibutuhkan upaya

analisis sentimen guna mengekstrak informasi-informasi penting dari konten tersebut.

Analisis sentimen, cabang dari *Natural Language Processing* (NLP), adalah metode untuk mengekstrak, memproses, serta melakukan identifikasi otomatis terhadap opini dalam data teks yang tidak terstruktur dalam kalimat [2]. Proses ini bertujuan untuk mengeksplorasi dan mengolah data guna menemukan informasi sentimen, yang kemudian diklasifikasikan menjadi klasifikasi positif, negatif, atau netral [3]. Analisis sentimen memungkinkan kita untuk memisahkan teks menjadi fakta dan opini, di mana opini mencerminkan perasaan dan sentimen subjektif [4]. Seiring perkembangan teknologi, analisis sentimen telah memanfaatkan algoritma pembelajaran mesin dan teknik NLP untuk otomatisasi dan skala, sehingga menjadi efisien, objektif, dalam mengelola teks yang besar [5].

Studi ini memanfaatkan pendekatan *K-Nearest Neighbor* (KNN) dan *Naive Bayes Classifier* (NBC) sebagai metode untuk melakukan *analisis sentimen* terhadap data yang diperoleh. KNN adalah teknik pembelajaran mesin yang mengklasifikasikan data berdasarkan kategori mayoritas dari k tetangga terdekat dalam dataset pelatihan [6]. Dalam konteks analisis sentimen PILKADA, KNN mengelompokkan *tweet* atau pendapat publik ke dalam kategori positif, negatif, atau netral, sesuai pandangan mereka terhadap calon atau isu politik. NBC adalah strategi pengelompokan berdasar pada *Teorema Bayes* melalui anggapan mengenai ketidakbergantungan antar fitur [7]. NBC efisien untuk mengolah data besar dan sering digunakan dalam pembelajaran mesin [8]. Dalam konteks PILKADA, NBC dapat menganalisis sentimen publik dengan mengklasifikasikan *tweet* atau opini menjadi kategori positif, negatif, atau netral berdasarkan data yang sudah dianalisis [9].

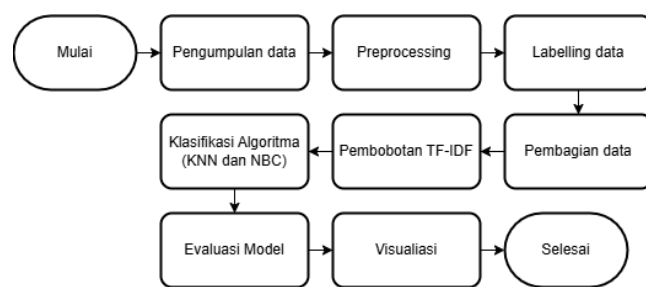
Pemilihan kedua algoritma ini didasarkan pada keunggulannya dalam mengolah data teks serta kemampuannya dalam melakukan *klasifikasi* dengan tingkat *akurasi* yang relatif tinggi, sebagaimana dibuktikan oleh sejumlah penelitian terdahulu. Studi oleh Furqan et al. (2022) mencatat *akurasi K-Nearest Neighbor (KNN)* sebesar 67,2% dalam *analisis sentimen PILKADA DKI 2017* melalui platform X. Dalam konteks *e-commerce*, Afdal & Elita (2022) melaporkan *akurasi KNN* sebesar 72,85% dan *Naive Bayes* sebesar 66,66% dalam *analisis sentimen toko online Lazada*. Adapun Kurnianengsing et al. (2024) menghasilkan *akurasi naive bayes* sebanyak 60% dalam analisis sentimen terhadap pemilihan kepala daerah serentak 2024.

Studi ini bertujuan untuk mengelompokkan sentimen masyarakat terkait *PILKADA 2024* pada tiga klasifikasi, yakni positif, negatif, dan netral, dan melakukan perbandingan performa algoritma KNN dan NBC menerapkan metrik evaluasi berupa *akurasi*, *presisi*, dan *recall*. Selain itu, penelitian ini juga berupaya mengeksplorasi serta mengidentifikasi isu-isu yang mencuat selama penyelenggaraan *PILKADA 2024*, termasuk opini publik mengenai calon kepala daerah, afiliasi partai politik, dan dinamika sosial-politik yang memengaruhi persepsi

masyarakat. Lebih dari sekadar perbandingan efektivitas kedua algoritma, penelitian ini juga bertujuan untuk menggali pandangan publik terkait isu-isu spesifik selama masa kampanye, seperti pandangan terhadap isu lingkungan dan ekonomi. Diharapkan temuan pada studi ini mampu menyajikan pandangan menyeluruh pada opini publik. di media sosial serta membantu dalam menentukan metode *analisis sentimen* yang paling tepat untuk konteks Pilkada.

II. METODE

Proses penelitian dimulai dengan pengumpulan data melalui crawling, kemudian dilakukan *Preprocessing* untuk membersihkan data. Selanjutnya data diberi label dalam tahap *labelling* terbagi keda data pelatihan dan pengujian. Data pelatihan untuk melatih algoritma seperti KNN dan NBC, sedangkan pengujian digunakan untuk mengukur akurasi model. Selanjutnya, dilakukan evaluasi dan visualisasi data guna mempermudah pemahaman hasil analisis. Langkah-langkah penelitian yang dilaksanakan ditampilkan dalam gambar 1.



Gambar 1. Tahapan penelitian

A. Pengumpulan data

Studi ini menerapkan metode crawling untuk mengumpulkan informasi dari platform X. Dalam proses crawling, peneliti menggunakan pustaka *tweet-hearvest*. *tweet-hearvest* adalah konsep Atau sarana yang digunakan untuk mengambil data dari X. Biasanya, alat ini dirancang untuk mengekstrak dan menyimpan *tweet* berdasarkan kata kunci, tagar, pengguna tertentu, atau kriteria lainnya.

Untuk memperoleh *tweet* yang berhubungan dengan *PILKADA 2024*, pencarian dilakukan melalui tagar *Pilkada2024*. Penelitian ini berhasil mengumpulkan total 1.203 data untuk dijadikan dataset, dataset ini disimpan dalam bentuk CSV. Rentang waktu data yang berhasil dikumpulkan adalah dari tanggal 08 Desember 2024 hingga 30 Desember 2024. Dataset ini mencakup komentar yang memiliki sentimen positif, negatif, dan netral

B. Preprocessing

reprocessing adalah tahap pertama sebelum teks dianalisis, Yang berfungsi untuk meningkatkan keteraturan dan kebersihan data teks guna mendukung keakuratan dalam proses analisis lebih lanjut. Memainkan peran penting dalam mengatasi kemungkinan kesalahan saat Mengekstrak fitur atau atribut dari teks *tweet* dapat memberikan dampak positif

yang signifikan terhadap peningkatan kinerja dalam *analisis sentimen*.

Prapemrosesan teks adalah langkah yang krusial, karena proses ini berkontribusi langsung terhadap efektivitas algoritma yang diterapkan. Data yang lebih bersih akan mendukung performa algoritma secara maksimal [10]. *Preprocessing* memiliki beberapa tahapan, diantaranya :

- 1) *Data Cleaning* : Langkah membersihkan data untuk menghilangkan beberapa komponen yang umum ditemukan dalam teks namun tidak mengandung arti substansial. [11]. Proses penghapusan seperti url, penyebutan, emotikon, dan tagar [12].
- 2) *Case Folding*: Proses case folding merujuk pada tahap untuk merubah karakter dalam teks ke format huruf kecil. menghilangkan perbedaan yang timbul dari variasi penggunaan huruf kapital. Langkah ini menjaga keutuhan teks, yang sangat penting untuk analisis selanjutnya. Case folding dilakukan dengan menggunakan fungsi `.lower()` dari pustaka *pandas* di Python, sehingga setiap karakter dalam teks berubah menjadi kecil untuk proses analisis [13].
- 3) *Normalization* : Mengubah, menghapus, atau memodifikasi kata atau istilah yang disingkat [14].
- 4) *Tokenizing* : tahapan pembagian teks dipisahkan ke dalam komponen-komponen yang lebih kecil. Memanfaatkan koma atau spasi untuk memisahkan elemen dalam teks [12]
- 5) *Stopwords Removal*: *Stopwords* merupakan kosakata yang umum ditemukan dalam teks, tetapi perannya dalam mendukung analisis sangat terbatas. seperti kata "dan", "atau", dan "adalah". Menghilangkan *stopwords* dapat meningkatkan kinerja model dengan mengurangi jumlah kata yang perlu diproses [15].
- 6) *Stemming* Tahapan *stemming* Langkah yang digunakan untuk mengubah kata terinfleksi ke dasarnya [16]. Prosedur *Stemming* ini memanfaatkan *library sastrawi* [17].

C. Pelabelan data

Dalam penelitian ini, Proses pemberian label pada data dilakukan secara manual dengan melibatkan ahli bahasa yakni Adelya Daniyah, M.Pd. hal ini guna memastikan tingkat akurasi yang tinggi, konsistensi, dan validitas anotasi sentimen. Setiap data teks hasil preprocessing dianalisis secara mendalam dengan mempertimbangkan struktur linguistik, makna kontekstual, serta nuansa emosional yang terkandung di dalamnya.

Ahli bahasa menerapkan pedoman klasifikasi yang telah distandardisasi, mencakup definisi operasional untuk kategori sentimen positif, negatif, dan netral, guna mengurangi subjektivitas individu selama proses anotasi. Selain memberikan label polaritas, pelabelan ini juga bertujuan mengidentifikasi ambiguitas makna dan konteks yang dapat mempengaruhi hasil klasifikasi algoritma di tahap selanjutnya, sehingga menghasilkan dataset yang berkualitas tinggi untuk kebutuhan pelatihan dan pengujian model.

Proses klasifikasi komentar dilakukan dengan menganalisis setiap teks berdasarkan aspek semantik, sintaksis, dan konteks wacana, guna mengidentifikasi kecenderungan sentimen yang tersembunyi. Komentar-komentar yang telah melewati proses *preprocessing* selanjutnya diklasifikasikan secara terstruktur menjadi tiga kelas utama, yaitu positif, negatif, dan netral. dengan berpedoman terhadap kriteria linguistik yang telah dirumuskan sebelumnya. Penekanan diberikan pada analisis makna implisit serta ekspresi emosional, sehingga setiap kategori sentimen dapat diidentifikasi secara komprehensif dan akurat. Pendekatan berbasis ahli bahasa ini diharapkan mampu menghasilkan dataset berlabel yang valid, meningkatkan kinerja model, serta memperkuat keandalan analisis sentimen berbasis machine learning dalam penelitian ini.

D. Pembagian data

Dalam proses pembagian dataset memisahkan data berlabel jadi dua bagian, mencakup data pelatihan dan data pengujian. Pemisahan ini dilakukan melalui pendekatan berbasis persentase, di mana sekitar 80% dari seluruh dataset diperuntukkan bagi data pelatihan, sedangkan sekitar 20% lainnya digunakan untuk pengujian. Dengan cara ini, mayoritas data dapat dimanfaatkan untuk mengoptimalkan proses pembelajaran model, sementara sebagian kecil lainnya digunakan untuk menguji keakuratan model yang telah dibangun.

E. Pembobotan TF-IDF

Proses yang dikenal sebagai penilaian *TF-IDF* (*Term Frequency-Inverse Document Frequency*) Berperan dalam mendefinisikan tingkat bobot bagi tiap kata dengan tujuan menilai tingkat kepentingan suatu kata dalam sebuah dokumen. digunakan ukuran yang disebut *TF-IDF*. Tingkat kemunculan suatu kata pada tiap dokumen. yang menggambarkan seberapa dominan kata tersebut, disebut *TF* (*Term Frequency*). Sementara itu, *DF* (*Document Frequency*) mengukur frekuensi kemunculan kata tersebut di sejumlah dokumen, menunjukkan betapa umum atau spesifiknya kata tersebut. Kata yang banyak ditemukan dalam dokumen diasumsikan mempunyai tingkat relevansi dominan, sedangkan istilah yang jarang digunakan dianggap kurang signifikan, jika kata itu muncul berulang kali di berbagai dokumen, maka relevansinya akan berkurang [3].

F. Implementasi Algoritma K-Nearest Neighbors

Algoritma *K-Nearest Neighbors* (*KNN*) ialah teknik dasar yang sering dimanfaatkan sebagai langkah pertama dalam mempelajari *machine learning*. Kelebihan dari *KNN* terletak pada kesederhanaannya serta kemampuannya untuk menangani data yang tidak linier atau yang belum terklasifikasi dengan jelas. Namun, kelemahan dari metode ini adalah penurunan kinerja pada dataset besar, karena algoritma perlu menghitung jarak antara setiap pasangan data pelatihan. [18]. Untuk menentukan label kategori atau nilai

dari suatu sampel, *K-Nearest Neighbors* (KNN) mencari K Memanfaatkan data dengan jarak terdekat untuk menentukan label atau nilai yang sesuai. Pendekatan ini mengelompokkan objek berdasarkan data yang paling dekat. KNN juga menghitung jarak kedekatan data untuk memperkirakan nilai bagi *query* baru. Metode *K-Nearest Neighbors* dapat diterapkan untuk menentukan kedekatan antar data dengan langkah-langkah berikut:

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (1)$$

Penjelasan:

d = nilai jarak

a = data uji

b = data sampel

i = indeks variable

n = banyaknya variabel

G. Implementasi Algoritma Naïve Bayes Classifier

Naive Bayes Classifier mengaplikasikan *Teorema Bayes* serta mengandalkan prinsip probabilitas dasar, dengan anggapan bahwa tiap fitur tidak saling bergantung. Algoritma ini diterapkan dalam berbagai masalah, seperti analisis sentimen, penyaringan *email spam*, pengelompokan *email* otomatis, pengurutan *email* berdasarkan prioritas, serta pengkategorian dokumen.

Metode *Naive Bayes* yang paling sesuai untuk pengklasifikasian teks atau dokumen adalah *multinomial Naive Bayes*, yang mempertimbangkan tidak hanya kemunculan kata, tetapi juga frekuensinya. Teknik ini diperkenalkan oleh ilmuwan asal Inggris, *Thomas Bayes*. Dibawah menunjukkan rumus klasifikasi data menggunakan *Naive Bayes* :

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (2)$$

Keterangan:

X : data yang tidak terlabeli.

C : asumsi mengenai kategori kelas yang mungkin untuk data X.

$P(C|X)$: *posterior probability*, yaitu peluang hipotesis C benar setelah mempertimbangkan data X.

$P(C)$: *prior probability* dari hipotesis C, yakni peluang awal tanpa memperhitungkan data X.

$P(X|C)$: *likelihood*, yang merujuk pada kemunculan data X dengan asumsi bahwa hipotesis C adalah benar.

$P(X)$: *evidence*, yang mengacu pada total data X tanpa memperhitungkan kelas yang terkait.

H. Evaluasi Model

Model evaluasi adalah langkah untuk menilai performa dari sebuah model *machine learning* setelah proses pelatihan dengan data yang tersedia. Pentingnya evaluasi model terletak pada kemampuannya dalam menunjukkan seberapa baik model dapat melakukan prediksi, sekaligus membantu dalam memilih model yang paling sesuai untuk suatu permasalahan [19].

Beberapa metrik yang sering digunakan dalam *klasifikasi* meliputi *akurasi*, Metode evaluasi ini mencakup *accuracy*,

Mengukur sejauh mana model menghasilkan prediksi yang benar dengan frekuensi tertentu. *precision* , yang menentukan tingkat ketepatan model dalam mengklasifikasikan data ke pada Kelas positif; *recall*, yang menilai efektivitas model dalam mengidentifikasi seluruh kasus positif; serta *F1-score*, yang mengelompokkan *precision* serta *recall* dalam menilai secara adil kemampuan model.

Rumus yang diterapkan :

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

Keterangan:

TP (*True Positive*): jumlah kasus positif yang berhasil diklasifikasikan secara akurat oleh model sebagai positif.

TN (*True Negative*): jumlah kasus negatif yang benar dikategorikan oleh model sebagai negatif.

FP (*False Positive*): jumlah kasus negatif yang keliru diklasifikasikan sebagai positif oleh model.

FN (*False Negative*): jumlah kasus positif yang salah diprediksi sebagai negatif oleh model.

I. Visualisasi

Setelah kedua algoritma diterapkan, dilakukan visualisasi data dengan memanfaatkan *wordcloud* dan grafik pada setiap kategori sentimen (positif, netral, dan negatif) yang diperoleh dari data latih yang telah dilatih [12].

WorldCloud adalah salah satu metode untuk memvisualisasikan yang menunjukkan keterkaitan antara frekuensi kata, dengan meningkatkan ukuran kata yang lebih sering digunakan secara cepat [20].

III. HASIL DAN PEMBAHASAN

A. Crawling Data

Studi ini mengumpulkan data dari X untuk memperoleh *tweet* yang berkaitan dengan Pilkada 2024, dengan total 1.203 *tweet*. Data hasil pengumpulan tersebut disimpan pad file CSV.

B. Pembersihan Data

Setelah data terkumpul, proses selanjutnya yaitu pengolahan data. untuk menghilangkan informasi tidak relevan agar kualitas data dalam analisis tetap terjaga. Tahap ini sangat penting, terutama karena data yang digunakan berupa teks yang sering kali mengandung noise, seperti karakter khusus, spasi berlebih, data duplikat, serta kata-kata yang tidak bermakna dalam analisis sentimen.

Proses pembersihan dilakukan dengan berbagai metode, seperti menghapus tanda baca, tautan, *mention*, dan karakter

non-alfabet yang tidak mendukung penelitian. Setelah melalui tahap ini, jumlah data yang semula sebanyak 1.203 *tweet*, terdapat 16 *tweet* duplikat, mengalami penyusutan menjadi 1.187 *tweet*. Data yang telah dibersihkan ini siap untuk tahap berikutnya, yaitu *Preprocessing*, guna memperbaiki akurasi dalam analisis sentimen.

C. Preprocessing

Preprocessing data merupakan tahap penting pada analisis data dirancang guna mengatur, menghapus, dan memodifikasi dataset agar lebih mudah dimengerti dan dianalisis dan dapat digunakan oleh algoritma pengolahan data atau model yang akan diterapkan pada tahap analisis selanjutnya.

Pada tabel 1 memuat hasil dari preprosesing yang terdiri dari tahapan *Cleaning*, *case folding*, *Normalization*, *stopword removal*, dan *Stemming*.

TABEL I
DATA HASIL PREPROCESSING

Contoh data	Menang 9 dari 13 Kontestasi Partai Golkar Rajai Pilkada 2024 Se-Sulteng https://t.co/L1k3PbkkQX
<i>Cleaning data</i>	Menang dari Kontestasi Partai Golkar Rajai Pilkada Se Sulteng
<i>Case Folding</i>	menang dari kontestasi partai golkar rajai pilkada se sulteng
<i>Normalization</i>	menang dari kontestasi partai golkar rajai pilkada se sulawesi tengah
<i>Tokenisasi</i>	'menang', 'dari', 'kontestasi', 'partai', 'golkar', 'rajai', 'pilkada', 'se', 'sulawesi', 'tengah'
<i>Stopword Removal</i>	'menang', 'kontestasi', 'partai', 'golkar', 'rajai', 'pilkada', 'sulawesi', 'tengah'
<i>Stemming</i>	'menang', 'kontes', 'partai', 'golkar', 'raja', 'pilkada', 'sulawesi', 'tengah'

D. Hasil Labelling

Pelabelan data dilakukan secara manual oleh ahli bahasa terhadap 1.187 data, terdiri atas 745 data berlabel positif, 260 data berlabel netral, dan 182 data berlabel negatif. Proses ini bertujuan untuk memastikan tingkat akurasi dan konsistensi yang tinggi pada klasifikasi. Untuk data dianalisis secara mendalam dengan mempertimbangkan konteks linguistik dan makna tersirat dalam teks, kemudian diberi label sentimen sesuai dengan pedoman anotasi yang telah dirumuskan sebelumnya.

TABEL II
HASIL LABELLING

No	Text	Label
1	komisi 1 dprd provinsi gorontalo kpud intensif awas siar pilkada 2024	Positive
2	pilkada serentak 2024 sejarah baru indonesia	Positive
3	berkat lancar pilkada 2024 sukses gelar patut apresiasi seluruh pihak libat dalam masuk diri sendiri milik suara jeju air caca tengker sandra dewi liverpool	Positive
4	paripurna dewan wakil rakyat singgung jalan pilkada serentak 2024 pilkada damai	Negative

Keterlibatan ahli bahasa dalam tahap pelabelan ini diharapkan dapat meningkatkan validitas data, sehingga dapat digunakan secara optimal dalam proses pelatihan model analisis sentimen. Hasil pelabelan data latih oleh ahli bahasa dan distribusi kelas data disajikan dalam tabel 2 dan 3 dibawah ini.

TSBEL III
DISTRIBUSI KELAS DATA

kelas	jumlah data	presentase
positif	745	62,7%
netral	260	21,9%
negatif	182	15,3%
total	1.187	100%

Distribusi ini menunjukkan adanya ketidakseimbangan kelas (*class imbalance*), di mana jumlah data berlabel positif jauh lebih banyak dibandingkan dengan data berlabel netral maupun negatif. Ketidakseimbangan ini berpotensi memengaruhi kinerja model klasifikasi, khususnya dalam mengklasifikasikan data dari kelas minoritas. Oleh karena itu, perlu dipertimbangkan penggunaan metode penanganan ketidakseimbangan kelas, seperti penerapan metrik evaluasi tambahan (*Precision*, *Recall*, dan *F1-score*) untuk memperoleh gambaran performa model yang lebih akurat.

E. Hasil Pembagian Data

Dataset terbagi dalam dua komponen utama, yaitu data pelatihan dan data pengujian, yang masing-masing memiliki peran dalam proses pelatihan algoritma dan pengujian performanya. Dari total 1.187 data, sebanyak 949 data (80,00%) digunakan untuk pelatihan, sedangkan 238 data (20,00%) dialokasikan untuk pengujian. Metode ini diterapkan guna menjamin proses evaluasi model dilakukan secara objektif terhadap data tidak digunakan selama pelatihan.

F. Hasil TF-IDF

Perhitungan *Term Frequency-Inverse Document Frequency* (TF-IDF) dilakukan untuk mengidentifikasi kata-kata yang paling berkontribusi dalam dokumen berdasarkan frekuensi kemunculan dan tingkat kepentingannya di seluruh korpus data. Lima kata dengan nilai TF-IDF tertinggi disajikan pada Tabel 4 berikut.

TABEL IV
PERHITUNGAN TF-IDF

Kata	TF-IDF	Frekuensi
pilkada	20,417871	277
2024	17,101545	183
damai	12,259926	58
pilih	11,380211	101
umum	10,417801	82

Berdasarkan hasil perhitungan TF-IDF, kata "*pilkada*" memiliki nilai tertinggi sebesar 20,417871 dengan frekuensi 277 kali, diikuti oleh kata "*2024*", "*damai*", "*pilih*", dan "*umum*". Kata-kata ini menunjukkan fokus utama data pada

peristiwa Pilkada 2024 serta nilai-nilai yang diharapkan, seperti pemilihan umum yang damai dan partisipatif. Dominasi kata-kata tersebut sejalan dengan tema data yang berkaitan dengan konteks waktu dan harapan terhadap pelaksanaan pilkada. Selain itu, hasil ini mendukung proses ekstraksi fitur dalam klasifikasi sentimen, di mana kata-kata dengan bobot TF-IDF tinggi menjadi indikator penting dalam mengklasifikasikan sentimen menjadi positif, netral, dan negatif.

G. Klasifikasi Model Algoritma Naive Bayes dan K-Nearest Neighbors

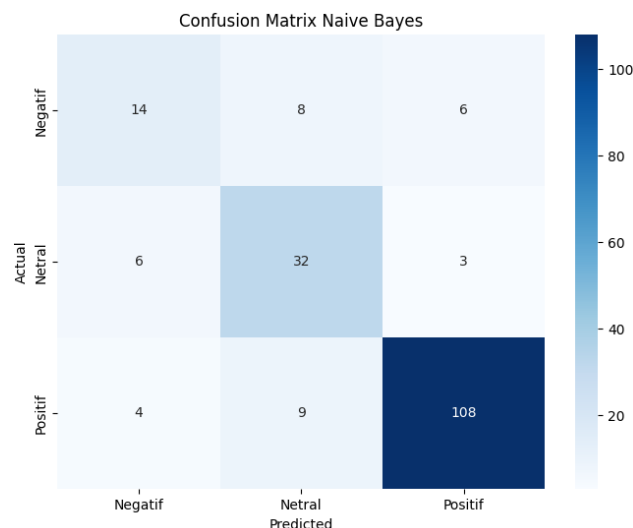
Studi ini menerapkan dua algoritma yaitu *Naive Bayes* dan *K-Nearest Neighbor* (KNN) telah diterapkan dalam menganalisis sentimen Pilkada yang terbagi ke dalam tiga kelas: negatif, netral, dan positif.

Klasifikasi yang dilakukan menggunakan algoritma *Naive Bayes* berhasil mencapai akurasi sebesar 81,05%. Performa terbaik dicapai pada kelas Positif, dengan nilai *precision* sebesar 0,92, *recall* sebesar 0,89, dan *f1-score* sebesar 0,91 dari total 121 data positif. Pada kelas Netral, model juga menunjukkan performa yang cukup baik dengan *precision* sebesar 0,65, *recall* sebesar 0,78, dan *f1-score* sebesar 0,71 dari 41 data. Sementara itu, untuk kelas Negatif, model menghasilkan *precision* sebesar 0,58, *recall* sebesar 0,50, dan *f1-score* sebesar 0,54 dari 28 data yang tersedia. Secara keseluruhan, nilai rata-rata makro (*macro average*) untuk *precision*, *recall*, dan *f1-score* masing-masing adalah 0,72. Hasil klasifikasi terlihat pada gambar berikut.

	precision	recall	f1-score	support
Negative	0.58	0.50	0.54	28
Netral	0.65	0.78	0.71	41
Positive	0.92	0.89	0.91	121
accuracy			0.81	190
macro avg	0.72	0.72	0.72	190
weighted avg	0.81	0.81	0.81	190

Gambar 1. Hasil Klasifikasi Algoritma *Naive Bayes*

Dalam *confusion matrix naive bayes*, model berhasil mengklasifikasikan 108 data positif dengan benar, 32 data netral dengan benar, dan 14 data negatif dengan benar. Namun, terdapat beberapa kesalahan klasifikasi, yaitu 6 data negatif diklasifikasikan sebagai positif dan 8 sebagai netral, 6 data netral diklasifikasikan sebagai negatif dan 3 sebagai positif, serta 4 data positif diklasifikasikan sebagai negatif dan 9 sebagai netral. Secara menyeluruh, model *Naive Bayes* menghasilkan performa yang baik, terutama dalam mengklasifikasikan data positif dan netral, meskipun masih perlu peningkatan dalam membedakan kelas negatif secara lebih akurat.



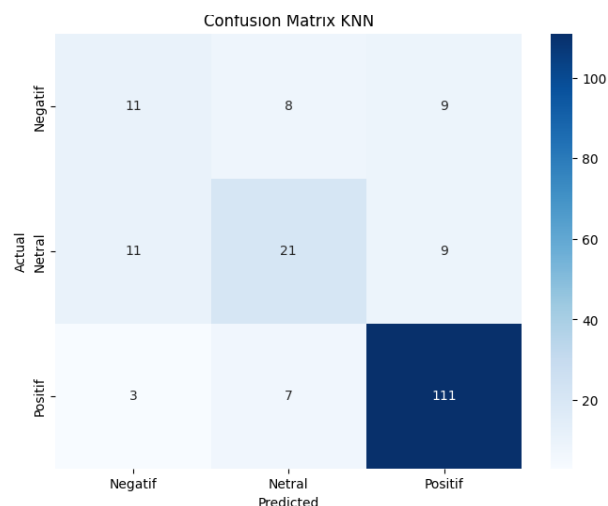
Gambar 2. Hasil *confusion matrik Naive Bayes*

Pada sisi lain, klasifikasi yang dilakukan menggunakan algoritma *K-Nearest Neighbor* (KNN) berhasil mencapai akurasi sebesar 75,26%. Performa terbaik kembali dicapai pada kelas Positif, dengan nilai *precision* sebesar 0,86, *recall* sebesar 0,92, dan *f1-score* sebesar 0,89 dari total 121 data positif. Untuk kelas Netral, model mencatat *precision* sebesar 0,58, *recall* sebesar 0,51, dan *f1-score* sebesar 0,55 dari 41 data yang tersedia. Pada kelas Negatif, model memiliki performa yang lebih rendah, dengan *precision* sebesar 0,44, *recall* sebesar 0,39, dan *f1-score* sebesar 0,42 dari 28 data. Berdasarkan perhitungan, nilaimakro (*macro average*) untuk *precision*, *recall*, dan *f1-score* secara berurutan adalah 0,63, 0,61, dan 0,62.

	precision	recall	f1-score	support
Negative	0.44	0.39	0.42	28
Netral	0.58	0.51	0.55	41
Positive	0.86	0.92	0.89	121
accuracy			0.75	190
macro avg	0.63	0.61	0.62	190
weighted avg	0.74	0.75	0.74	190

Gambar 3. Hasil klasifikasi algoritma *K-Nearest Neighbor*

Berdasarkan *confusion matrix* untuk algoritma *K-Nearest Neighbor* (KNN), model berhasil mengklasifikasikan 111 data positif dengan benar, 21 data netral dengan benar, dan 11 data negatif dengan benar. Namun, terdapat beberapa kesalahan klasifikasi, yaitu 9 data negatif diklasifikasikan sebagai positif dan 8 sebagai netral, 11 data netral diklasifikasikan sebagai negatif dan 9 sebagai positif, serta 3 data positif diklasifikasikan sebagai negatif dan 7 sebagai netral. Secara umum, model KNN memiliki performa yang cukup baik dalam mengklasifikasikan data positif, tetapi masih perlu perbaikan dalam membedakan antara kelas negatif dan netral. *Confusion matrix* untuk KNN ditampilkan pada gambar 4.

Gambar 4. Hasil *confusion matrik K-Nearest Neighbor*

H. Evaluasi model Algoritma Naive Bayes dan K-Nearest Neighbors

Berdasarkan hasil evaluasi performa model, Evaluasi model dilakukan dengan metode hold-out split, Pembagian data dilakukan dengan alokasi 80% untuk data latih dan 20% untuk data uji. Selain itu, untuk keperluan optimasi hyperparameter seperti pemilihan nilai *alpha* pada Naive Bayes dan jumlah tetangga (*n_neighbors*) pada KNN, digunakan teknik k-fold cross-validation sebanyak 5 lipatan (5-fold)

Berdasarkan hasil optimasi hyperparameter menggunakan GridSearchCV, pada model Multinomial Naive Bayes diperoleh nilai *alpha* terbaik sebesar 0.1. Nilai ini memungkinkan smoothing dilakukan secara minimal, sehingga distribusi frekuensi kata penting dalam analisis sentimen tetap terjaga.

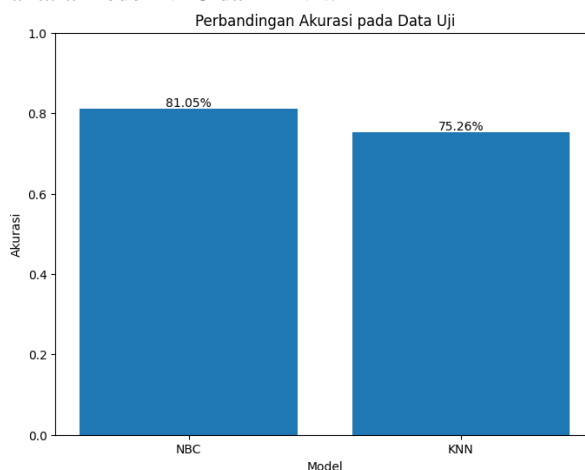
Sedangkan pada model K-NN, diperoleh nilai *n_neighbors* optimal sebesar 5, yang mampu menyeimbangkan sensitivitas terhadap pola lokal dan ketahanan terhadap data *outlier*.

Evaluasi model menunjukkan bahwa *Multinomial Naive Bayes* unggul dibandingkan KNN dalam hal akurasi dan konsistensi prediksi, dengan akurasi akhir mencapai 81,05% untuk *Multinomial Naive Bayes* dan 75,26% untuk KNN. Perbandingan hasil antara model dapat dianalisis lebih lanjut melalui tabel yang terlampir di bawah ini.

TABEL V
HASIL PERBANDINGAN MODEL ALGORITMA

Model	Kelas Sentimen	Precision	Recall	F1-Score
Naive Bayes	Positif	0,92	0,89	0,91
Naive Bayes	Netral	0,65	0,78	0,71
Naive Bayes	Negatif	0,58	0,50	0,54
KNN	Positif	0,86	0,92	0,89
KNN	Netral	0,58	0,51	0,55
KNN	Negatif	0,44	0,39	0,42

Gambar 5, disajikan perbandingan akurasi dalam diagram antara model NBC dan K-NN.



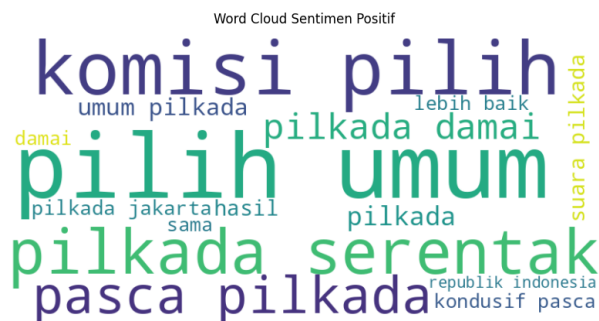
Gambar 5. Perbandingan akurasi model algoritma

Multinomial Naive Bayes lebih unggul karena kemampuannya dalam mengelola data teks dengan fitur tinggi secara efisien, menjaga distribusi kata penting, memberikan probabilitas pada prediksi untuk meningkatkan konsistensi, serta ketahanan terhadap data outlier, sedangkan K-NN lebih sensitif terhadap fluktuasi data dan kesulitan menangani data dengan banyak fitur. Perbandingan hasil antara model dapat dianalisis lebih lanjut melalui tabel yang terlampir di bawah ini.

I. Visualisasi

Tahap terakhir dari proses ini adalah visualisasi data. Visualisasi *word cloud* pada Gambar X menunjukkan distribusi kata-kata yang paling sering muncul dalam masing-masing kategori sentimen, yaitu positif, negatif, dan netral, yang berkaitan dengan topik Pilkada.

Pada kategori sentimen positif, kata-kata seperti “komisi pilih”, “pilih umum”, dan “pilkada serentak” tampak mendominasi. Kemunculan kata-kata tersebut mencerminkan adanya narasi yang mendukung pelaksanaan Pilkada secara demokratis dan terorganisir, serta persepsi publik yang cenderung optimis terhadap proses pemilu lokal. Gambar dibawah ini menampilkan hasil *word cloud* sentimen positif.

Gambar 6. *word cloud* sentimen positif

- Analisis Sentimen Pengguna Twitter Terhadap Aplikasi Sirekap,” *Fountain Of Informatics Journal*, Vol. 9, No. 2, Pp. 65–71, 2024, Doi: <https://dx.doi.org/10.21111/Fij.V9i2.12717>.
- [9] R. Noviana And I. Rasal, “Penerapan Algoritma *Naive Bayes* Dan Svm Untuk Analisis Sentimen Boy Band Bts Pada Media Sosial Twitter,” *Jurnal Teknik Dan Science*, Vol. 2, No. 2, Pp. 51–60, 2023, Doi: <https://doi.org/10.56127/Jts.V2i2.791>.
- [10] S. Wahyu, “Perbandingan Model Algoritma Klasifikasi Pada Analisis Sentimen Opini Masyarakat Terhadap Layanan Kereta Cepat Jakarta Bandung (The Whoosh),” *Proceeding Konik (Konferensi Nasional Ilmu Komputer)*, Vol. 6, Pp. 218–225, 2023.
- [11] A. Agustian And F. Nurapriani, “Analisis Sentimen, Text Mining Penerapan Analisis Sentimen Dan *Naive Bayes* Terhadap Opini Penggunaan Kendaraan Listrik Di Twitter,” *Jurnal Tika*, Vol. 7, No. 3, Pp. 243–249, 2022, Doi: <https://doi.org/10.51179/Tika.V7i3.1550>.
- [12] Y. Khoiruddin, A. Fauzi, And A. M. Siregar, “Analisis Sentimen Gojek Indonesia Pada Twitter Menggunakan Algoritme *Naive Bayes* Dan Support Vector Machine,” *Progresif: Jurnal Ilmiah Komputer*, Vol. 19, No. 1, Pp. 391–400, 2023.
- [13] N. A. Nevrada And M. A. Syaputra, “Sentiment Analysis Of Telegram App Reviews On Google Play Store Using The Support Vector Machine (Svm) Algorithm,” *Journal Of Applied Informatics And Computing*, Vol. 9, No. 1, Pp. 96–105, 2025.
- [14] Y. R. Nasution, S. Suhardi, And I. H. Satrio, “Penerapan Algoritma Klasifikasi *Naive Bayes* Untuk Analisis Sentimen Tentang Pemilu 2024,” *Elkom: Jurnal Elektronika Dan Komputer*, Vol. 17, No. 2, Pp. 495–502, 2024, Doi: <https://doi.org/10.51903/Elkom.V17i2.2053>.
- [15] A. R. Gunawan And R. F. A. Aziza, “Sentiment Analysis Using Lstm Algorithm Regarding Grab Application Services In Indonesia,” *Journal Of Applied Informatics And Computing*, Vol. 9, No. 2, Pp. 322–332, 2025, Doi: <https://doi.org/10.30871/Jaic.V9i2.8696>.
- [16] J. Setyanto And T. B. Sasongko, “Sentiment Analysis Of Sirekap Application Users Using The Support Vector Machine Algorithm,” *Journal Of Applied Informatics And Computing*, Vol. 8, No. 1, Pp. 71–76, 2024, Doi: <https://doi.org/10.30871/Jaic.V8i1.7772>.
- [17] A. A. Muttaqin, S. Alam, And M. A. Komara, “Analisis Sentimen Isu Kecurangan Pemilu 2024 Berdasarkan Opini Pada Media Sosial Twitter Menggunakan Metode Crisp-Dm Dengan Algoritma *Naive Bayes Classifier*,” *Jati (Jurnal Mahasiswa Teknik Informatika)*, Vol. 8, No. 5, Pp. 8764–8772, 2024, Doi: <https://doi.org/10.30871/Jaic.V8i1.7772>.
- [18] M. Fadil, “Perbandingan Metode *K-Nearest Neighbors* Dan *Naive Bayes Classification* Dalam Klasifikasi Sentimen Masyarakat Terhadap Pemilu Presiden Tahun 2024,” Nov. 2024, Doi: <http://dx.doi.org/10.35889/Jutisi.V13i3.2068>.
- [19] A. K. Subagyo, H. S. H. Syahputra, Z. F. Khoiri, And S. Sukmadiningtyas, “Analisis Sentimen Masyarakat Terhadap Pelantikan Kabinet Merah Putih Pada Media Sosial,” In *Proceedings Of The National Conference On Electrical Engineering, Informatics, Industrial Technology, And Creative Media*, 2024, Pp. 1080–1089.
- [20] T. P. R. Sanjaya, A. Fauzi, And A. F. N. Masruriyah, “Analisis Sentimen Ulasan Pada E-Commerce Shopee Menggunakan Algoritma *Naive Bayes* Dan Support Vector Machine,” *Infotech: Jurnal Informatika & Teknologi*, Vol. 4, No. 1, Pp. 16–26, 2023, Doi: <https://doi.org/10.37373/Infotech.V4i1.422>.