

Analisis Sentimen Masyarakat Terhadap Kebijakan IKN Pada Periode Jokowi dan Prabowo Menggunakan Algoritma NBC, SVM, dan K-NN

Nur Shabrina Nasution*, Inggih Permana, Febi Nur Salisah, M Afdal, Megawati Megawati

Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia
Email: ¹*12150321423@students.uin-suska.ac.id, ²ingghipermana@uin-suska.ac.id, ³febinursalisah@uin-suska.ac.id,
⁴m.afdal@uin-suska.ac.id, ⁵megawati@uin-suska.ac.id

Email Penulis Korespondensi: 12150321423@students.uin-suska.ac.id

Submitted: 04/05/2025; Accepted: 31/05/2025; Published: 01/06/2025

Abstrak—Pemindahan Ibu Kota Negara (IKN) dari Jakarta ke Kalimantan Timur menimbulkan beragam respons dari masyarakat Indonesia yang terekam melalui media sosial, khususnya platform X. Penelitian ini bertujuan untuk menganalisis dan membandingkan sentimen masyarakat terhadap kebijakan IKN pada dua periode pemerintahan, yaitu Presiden Joko Widodo dan Presiden Prabowo Subianto, dengan pendekatan machine learning. Tiga algoritma yang digunakan dalam klasifikasi sentimen adalah Naive Bayes Classifier (NBC), Support Vector Machine (SVM), dan K-Nearest Neighbor (K-NN). Proses penelitian meliputi crawling data (masing-masing 600 data per periode), text preprocessing (cleaning, tokenizing, filtering, stemming), pelabelan data menggunakan pendekatan Lexicon-Based dengan kamus InSet, serta pembobotan menggunakan metode TF-IDF. Hasil analisis menunjukkan bahwa pada periode Jokowi, sentimen masyarakat cenderung lebih seimbang, dengan dominasi sentimen negatif (35,9%), diikuti sentimen positif (33,4%) dan netral (30,7%). Sedangkan pada periode Prabowo, sentimen negatif meningkat menjadi 40,3%, sedangkan positif menurun ke 26,3%. Berdasarkan evaluasi akurasi model, pada periode Jokowi algoritma NBC menunjukkan performa terbaik dengan akurasi sebesar 73%, sementara pada periode Prabowo, algoritma SVM unggul dengan akurasi tertinggi mencapai 81%. Temuan ini memberikan gambaran dinamis mengenai persepsi publik terhadap kebijakan IKN di bawah dua pemerintahan yang berbeda.

Kata Kunci: Analisis Sentimen; IKN; Naive Bayes Classifier; Support Vector Machine; K-Nearest Neighbour

Abstract—The relocation of the National Capital City (IKN) from Jakarta to East Kalimantan has generated a variety of responses from the Indonesian people recorded through social media, especially platform X. This study aims to analyze and compare public sentiment towards the IKN policy in two periods of government, namely President Joko Widodo and President Prabowo Subianto. This study aims to analyze and compare public sentiment towards the policy of the National Capital City during two periods of government, namely President Joko Widodo and President Prabowo Subianto, using a machine learning approach. The three algorithms used in sentiment classification are Naive Bayes Classifier (NBC), Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN). The research process includes data crawling (600 data each per period), text preprocessing (cleaning, tokenizing, filtering, stemming), data labeling using Lexicon-Based approach with InSet dictionary, and weighting using TF-IDF method. The results of the analysis show that in the Jokowi period, public sentiment tends to be more balanced, with the dominance of negative sentiment (35.9%), followed by positive sentiment (33.4%) and neutral (30.7%). Whereas in the Prabowo period, negative sentiment increased to 40.3%, while positive decreased to 26.3%. Based on the model accuracy evaluation, in the Jokowi period, the NBC algorithm showed the best performance with an accuracy of 73%, while in the Prabowo period, the SVM algorithm excelled with the highest accuracy reaching 81%. These findings provide a dynamic picture of public perception of IKN policies under two different governments.

Keywords: Sentiment Analysis; IKN; Naive Bayes Classifier; Support Vector Machine; K-Nearest Neighbour

1. PENDAHULUAN

Pemindahan ibu kota dari Jakarta ke Kalimantan Timur diharapkan dapat mengurangi beban Jakarta serta menciptakan pemerataan pembangunan di Indonesia [1]. Namun, kebijakan ini juga menimbulkan berbagai perdebatan, baik dalam aspek ekonomi, sosial, maupun lingkungan [2][3][4]. Masyarakat menyuarakan pendapatnya melalui berbagai platform media sosial, salah satunya adalah X, yang menjadi wadah utama diskusi publik mengenai pemindahan IKN [5]. Oleh karena itu, analisis sentimen terhadap opini masyarakat mengenai kebijakan IKN menjadi penting untuk memahami bagaimana persepsi publik terbentuk dan bagaimana pemerintah dapat meresponsnya secara efektif.

Analisis sentimen merupakan salah satu metode dalam bidang *Natural Language Processing* (NLP) yang bertujuan untuk mengekstrak, memahami, dan mengklasifikasikan opini dalam bentuk teks [6]. Dengan perkembangan teknologi *machine learning*, analisis sentimen kini dapat dilakukan dengan lebih akurat menggunakan berbagai algoritma klasifikasi seperti *Naive Bayes Classifier* (NBC), *Support Vector Machine* (SVM), dan *K-Nearest Neighbour* (K-NN). Ketiga algoritma ini memiliki keunggulan masing-masing dalam pemrosesan data teks. NBC dikenal karena efisiensinya dalam menangani data dalam jumlah besar [7], sementara SVM memiliki performa yang baik dalam mengelola data yang tidak seimbang [8]. Penelitian oleh Widyanto et al. (2023) menunjukkan bahwa SVM mampu meningkatkan akurasi klasifikasi secara signifikan setelah diterapkan teknik penyeimbangan data seperti SMOTE, dari 76,7% menjadi 93,2% [9]. Di sisi lain, K-NN sering digunakan karena kemampuannya dalam menangani data yang mengandung *noise* dengan lebih baik [10]. Dimana pada studi yang dilakukan oleh Yunata et al. (2024) menunjukkan bahwa K-NN tetap mempertahankan akurasi yang tinggi meskipun data mengandung *noise*, terutama saat menggunakan nilai *k* yang optimal dan metrik jarak yang sesuai [11].

Beberapa penelitian sebelumnya telah menganalisis sentimen masyarakat terhadap pemindahan IKN. Diantaranya adalah Pramana et al. (2023) menggunakan algoritma Naive Bayes Classifier dan K-Nearest Neighbors pada 2.404 data komentar Instagram terkait pemindahan Ibu Kota Negara, dengan algoritma KNN menunjukkan akurasi tertinggi sebesar 69,23% dibandingkan dengan NBC yang hanya mencapai 63,09% [12]. Selanjutnya ada Lubis et al. (2024) yang membandingkan Naive Bayes dan K-NN, dengan akurasi tertinggi oleh K-NN sebesar 88,12% [13]. Kemudian ada Muliawan dan Dazki (2023) menemukan Naive Bayes adalah algoritma yang paling akurat (65,26%) dibanding K-NN dan Random Forest [6]. Sementara itu, Nur Hadi (2025) menggunakan SVM, Logistic Regression, dan Naive Bayes, dengan performa terbaik dicapai oleh SVM (80%) [14]. Yang terakhir yaitu Munawaroh et al. menunjukkan Naive Bayes mampu mengklasifikasikan sentimen dengan akurasi 87% [15].

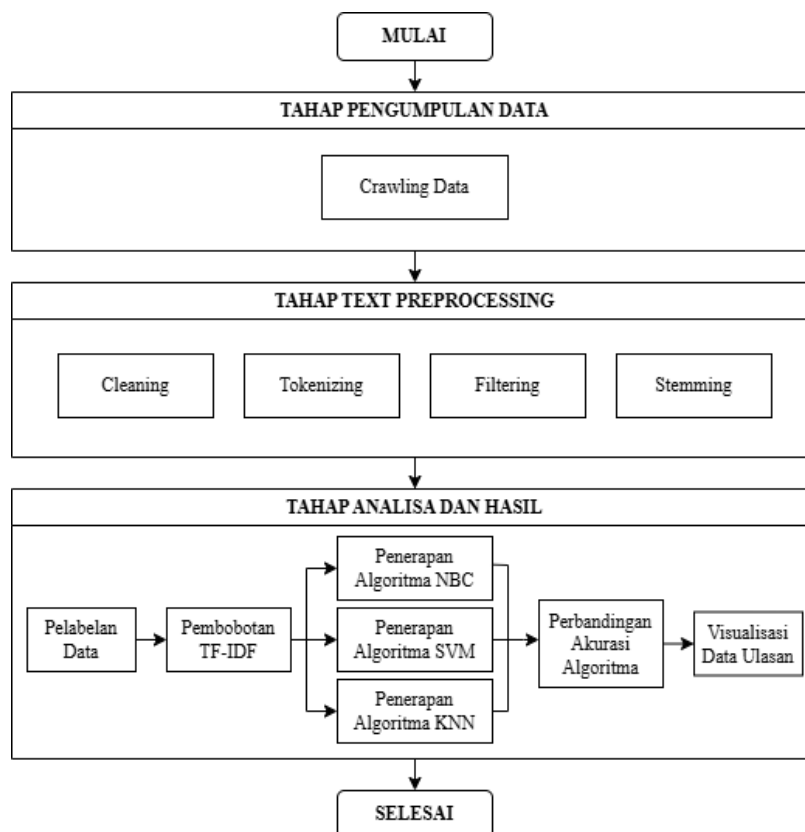
Meski banyak penelitian membahas sentimen terhadap pemindahan IKN, sebagian besar hanya fokus pada masa pemerintahan Presiden Joko Widodo, tanpa mempertimbangkan perubahan opini publik saat kepemimpinan berganti, seperti di era Presiden Prabowo Subianto. Setiawan et al. (2024) menggunakan SVM dan Naive Bayes untuk menganalisis sentimen IKN, namun tidak membandingkan dua periode pemerintahan [1]. Faturrahmi et al. (2023) meneliti sentimen terhadap Prabowo sebagai capres dengan K-NN, tapi tidak membahas isu IKN atau perbandingan antar periode [16]. Segmentasi data berdasarkan periode pemerintahan juga masih jarang, sehingga dinamika opini publik belum tergambar utuh.

Karena itu, penelitian ini bertujuan mengisi kesenjangan tersebut dengan membandingkan sentimen masyarakat terhadap kebijakan pemindahan IKN pada era Presiden Joko Widodo dan Presiden Prabowo Subianto menggunakan pendekatan machine learning. Dimana penelitian ini akan menjadi landasan penting bagi pemangku kebijakan dalam memahami dinamika opini publik terhadap kebijakan pemindahan IKN.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Untuk tahapan-tahapan ataupun proses yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

Berikut merupakan penguraian tahapan-tahapan yang ada pada Gambar 1.

2.1.1 Tahap Pengumpulan Data

Pengumpulan data dilakukan dengan proses Crawling. Crawling merupakan proses otomatis untuk mengumpulkan data dari media sosial seperti Twitter menggunakan API (Application Programming Interface). Dalam proses ini, sistem (crawler) mengirim permintaan ke server Twitter untuk mendapatkan data dalam format JSON, kemudian

menyaring dan menyimpan informasi seperti tanggal, lokasi, bahasa, dan isi tweet secara terstruktur dimana data ini dapat digunakan untuk penelitian dan analisis [17].

2.1.2 Tahap Preprocessing

Preprocessing merupakan langkah awal dalam pengolahan data yang bertujuan untuk mentransformasi data tweet yang tidak terstruktur menjadi format yang terstandar dan siap digunakan dalam proses analisis lebih lanjut, termasuk dalam pengelompokan data ke dalam berbagai cluster [18]. Berikut merupakan uraian dari tahapan-tahapan tersebut.

a. **Cleaning Text**

Pada tahap ini, sebelum dataset digunakan maka akan dilakukan pembersihan terlebih dahulu seperti pembersihan unsur tidak penting, seperti singkatan, symbol, angka, karakter, URL maupun emoticon yang dapat mengganggu proses analisis serta kevalidan hasil penelitian nantinya [19].

b. **Tokenizing**

Tahap ini merupakan tahapan dimana data akan di proses lagi sehingga tanda baca akan dihilangkan sehingga kalimat akan menghasilkan sebuah kata atau frasa individual [20].

c. **Filtering**

Tahap ini merupakan proses ekstraksi kata-kata penting dari hasil tokenisasi melibatkan pemilihan kata-kata yang relevan sambil menghilangkan kata-kata umum yang tidak memiliki makna signifikan, seperti kata hubung dan kata depan [21].

d. **Stemming**

Pada tahap ini, dilakukan penghapusan kata imbuhan, kata ganti, kata depan, dan kata akhir yang sesuai dengan KBBI [22]. Dimana pada tahap ini peneliti akan menggunakan library sastrawi dalam melakukan stemming pada teks Bahasa Indonesia.

2.1.3 Tahap Analisa dan Hasil

Untuk tahapan analisa dan hasil dapat diuraikan sebagai berikut.

a. **Pelabelan Data.**

Tahap ini dilakukan proses pelabelan data Negatif, Positif atau Netral dilakukan menggunakan pendekatan Lexicon Based dengan memanfaatkan Kamus InSet.

b. **Pembobotan Dengan TF-IDF**

Tahap ini memiliki tujuan utama dalam mendapatkan nilai atau bobot dari setiap kata yang telah diekstrak. Teknik ini didasarkan pada frekuensi kemunculan kata (term) dalam dokumen, dengan mengalikan frekuensi term (TF) dan invers frekuensi dokumen (DF) [21]. Semakin sering suatu term muncul, maka semakin besar pula nilai bobot yang diperoleh [23]. Persamaan untuk TF-IDF dapat dilihat pada persamaan 1.

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \tag{1}$$

Persamaan (1) menunjukkan bahwa nilai TF-IDF dari sebuah term t dalam dokumen d diperoleh dengan mengalikan frekuensi kemunculan kata tersebut dalam dokumen (TF) dengan nilai kebalikannya terhadap jumlah dokumen yang mengandung kata tersebut (IDF).

c. **Penerapan Algoritma NBC, SVM dan K-NN.**

Sebelum dilakukan pengklasifikasian algoritma data akan di SMOTE terlebih dahulu untuk menyeimbangkan data, setelah itu data akan dilatih menggunakan algoritma Naïve Baiyes Classifier, Support Vector Machine dan K-Nearest Neighbour.

Naïve Bayes Classifier adalah algoritma klasifikasi yang menggunakan rumus probabilitas (Teorema Bayes) dengan asumsi bahwa setiap fitur bersifat independent dan cocok untuk analisis teks seperti sentiment [24]. Untuk rumus NBC dapat dilihat pada persamaan 2.

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

Persamaan (2) menunjukkan bahwa Naïve Bayes Classifier memprediksi kelas data berdasarkan fitur yang dimiliki. C yang merupakan kelas dan X yang merupakan fitur, dan $P(C | X)$ menunjukkan peluang data termasuk ke dalam kelas C . Perhitungan ini menggunakan $P(X | C)$ (kemungkinan fitur muncul dalam kelas) dan $P(C)$ (peluang awal kelas), sementara $P(X)$ biasanya diabaikan karena sama untuk semua kelas.

Selanjutnya ada algoritma Support Vector Machine (SVM) yang merupakan algoritma klasifikasi yang mencari garis pemisah terbaik (hyperplane) antara dua kelas dengan memaksimalkan jarak (margin) antar data terdekat dari masing-masing kelas [25]. Dalam hal ini, rumus umum untuk SVM linear dapat dituliskan pada persamaan 3.

$$f(x) = sign(w \cdot x + b) \tag{3}$$

Pada Persamaan (3) menunjukkan bahwa di sini, tanda sign memberikan hasil kelas berdasarkan posisi x terhadap hyperplane yang dibentuk oleh w dan b .



Sedangkan Algoritma K-Nearest Neighbors (KNN) adalah algoritma klasifikasi yang menentukan kelas data baru berdasarkan mayoritas kelas dari K data terdekat yang paling mirip, dihitung menggunakan jarak seperti Euclidean distance [26]. Untuk menghitung jarak menggunakan rumus Euclidean distance yang dapat dilihat pada persamaan 4.

$$d(p, q) = \sqrt{\sum_{i=1}^m (p_i - q_i)^2} \tag{4}$$

Berdasarkan persamaan (4) menunjukkan bahwa $d(p,q)$ adalah jarak Euclidean antara titik p dan titik q . Titik p dan q merupakan dua data yang dibandingkan, masing-masing direpresentasikan sebagai vektor fitur. Simbol m menunjukkan jumlah dimensi atau fitur dalam data. Sedangkan p_i dan q_i adalah nilai fitur ke- i dari titik p dan titik q secara berurutan

- d. Perbandingan Performa Algoritma
Perbandingan ini dilakukan untuk mengetahui mana Algoritma yang memiliki performa terbaik, algoritma terbaik ditentukan berdasarkan hasil Akurasi.
- e. Visualisasi dengan Wordcloud
Visualisasi pada Data Komentar Kata-kata akan divisualisasikan dalam bentuk word cloud. Dimana wordcloud merupakan data yang akan divisualisasikan berdasarkan frekuensi kemunculannya di setiap kelas dalam dataset [27].

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pada penelitian ini, data akan dikumpulkan menggunakan Google Collabatory dan Bahasa pemrograman Phyton, yaitu melalui komentar yang diambil dari Platform X dengan menggunakan *keyword* “IKN”. Pada Periode Presiden Jokowi data diambil dengan rentang waktu tahun 2022-2023 sedangkan pada Periode Presiden Prabowo data diambil dari tahun 2024-2025. Dimana data yang didapatkan pada setiap periode berjumlah 600.

3.2 Text Processing

Text preprocessing mencakup Cleaning yang menghapus simbol, angka, URL, dan emotikon kemudian dilanjutkan dengan Tokenizing untuk memecah kalimat menjadi kata atau frasa diikuti oleh Filtering yang menyaring kata-kata penting dan menghilangkan kata umum serta diakhiri dengan Stemming untuk menghapus imbuhan sesuai KBBI.

3.2.1 Periode Jokowi

Hasil Text Processing pada Periode Jokowi dapat dilihat pada Tabel 1.

Tabel 1. Hasil Text Preprocessing Periode Jokowi

Proses	Hasil
Data Awal	@windrh1 @erwinr80 @Kimberley_2024 IKN yg skr baru digarap seluas area monas aja sdh nelan biaya 70 Trilyun. Semua semata demi legacy jokowi. Membangun berdasarkan nafsu. Bukan demi rakyat.
Cleaning	ikn yang sekarang baru dikerjakan seluas area monas saja sudah menelan biaya trilyun semua semata demi warisan jokowi membangun berdasarkan nafsu bukan demi rakyat
Tokenizing	['ikn', 'yang', 'sekarang', 'baru', 'dikerjakan', 'seluas', 'area', 'monas', 'saja', 'sudah', 'menelan', 'biaya', 'trilyun', 'semua', 'semata', 'demi', 'warisan', 'jokowi', 'membangun', 'berdasarkan', 'nafsu', 'bukan', 'demi', 'rakyat']
Filtering	['ikn', 'seluas', 'area', 'monas', 'menelan', 'biaya', 'trilyun', 'warisan', 'jokowi', 'membangun', 'berdasarkan', 'nafsu', 'rakyat']
Stemming	['ikn', 'luas', 'area', 'monas', 'telan', 'biaya', 'trilyun', 'waris', 'jokowi', 'bangun', 'dasar', 'nafsu', 'rakyat']

3.2.1 Periode Prabowo

Hasil Text Processing pada Periode Prabowo dapat dilihat pada Tabel 2.

Tabel 2. Hasil Text Preprocessing Periode Prabowo

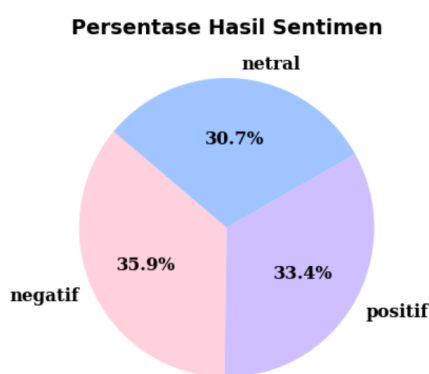
Proses	Hasil
Data Awal	@DS_yantie Abis tekor mau cuci tangan apa mau mundur soal nya kan biasanya gitu tabiat pejabat di indo selalu lari pas udah terjadi masalah contoh IKN yg dibuat si raja Jawa ga ada lanjutan tuh anj
Cleaning	habis rugi mau cuci tangan apa mau mundur soal ya kan biasanya begitu tabiat pejabat di indonesia selalu lari setelah sudah terjadi masalah contoh ikn yang dibuat sih raja jawa tidak ada lanjutan itu anjing

Proses	Hasil
Tokenizing	['habis', 'rugi', 'mau', 'cuci', 'tangan', 'apa', 'mau', 'mundur', 'soal', 'ya', 'kan', 'biasanya', 'begitu', 'tabiat', 'pejabat', 'di', 'indonesia', 'selalu', 'lari', 'setelah', 'sudah', 'terjadi', 'masalah', 'contoh', 'ikn', 'yang', 'dibuat', 'sih', 'raja', 'jawa', 'tidak', 'ada', 'lanjutan', 'itu', 'anjing']
Filtering	['habis', 'rugi', 'cuci', 'tangan', 'mundur', 'ya', 'tabiat', 'pejabat', 'indonesia', 'lari', 'contoh', 'ikn', 'sih', 'raja', 'jawa', 'lanjutan', 'anjing']
Stemming	['habis', 'rugi', 'cuci', 'tangan', 'mundur', 'ya', 'tabiat', 'pejabat', 'indonesia', 'lari', 'contoh', 'ikn', 'sih', 'raja', 'jawa', 'lanjut', 'anjing']

3.3 Pelabelan

3.3.1 Periode Jokowi

Berdasarkan hasil pelabelan dengan Lexicon Based sebanyak 600 data yang didapatkan pada komentar sosial media X memiliki 215 sentimen negatif, 200 sentimen positif, dan 185 sentimen netral. Persentase keberhasilan pelabelan setiap kategori dapat dilihat lebih jelas pada Gambar 2.

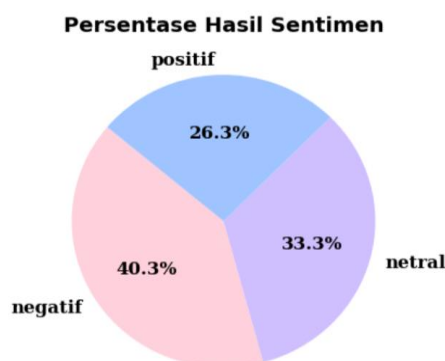


Gambar 2. Persentase Hasil Pelabelan pada Periode Jokowi

Pada Gambar 2 menunjukkan bahwa sentimen tertinggi adalah pada sentimen negatif yaitu sebesar 35,9%, sedangkan sentimen positif sebesar 33,4%, dan sentimen terendah pada netral sebesar 30,7%.

3.3.2 Periode Prabowo

Berdasarkan hasil pelabelan dengan Lexicon Based menggunakan kamus Inset sebanyak 600 data yang didapatkan pada komentar sosial media X memiliki 242 sentimen negatif, 158 sentimen positif, dan 200 sentimen netral. Persentase keberhasilan pelabelan setiap kategori dapat dilihat secara lebih jelasnya pada Gambar 2.



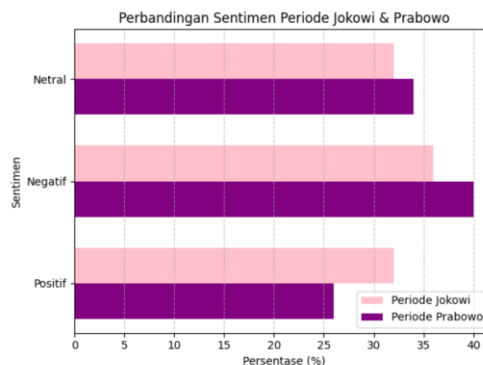
Gambar 3. Persentase Hasil Pelabelan pada Periode Prabowo

Pada Gambar 3 memperlihatkan bahwa sentimen tertinggi adalah pada sentimen negatif yaitu sebesar 40,3%, sedangkan sentimen positif sebesar 26,3%, dan sentimen netral sebesar 33,3%.

3.4 Perbandingan Sentimen IKN pada Periode Jokowi dan Prabowo

Berdasarkan hasil pelabelan menggunakan Lexicon dengan jumlah 600 data pada masing-masing periode, terdapat tiga kategori sentimen yang dianalisis, yaitu sentiment positif, negatif, dan juga netral. Dan dapat diketahui bahwa pada periode Presiden Jokowi, sentimen positif mendominasi sebesar sekitar 33%, diikuti oleh sentimen negatif sekitar 35%, dan netral sekitar 31%. Sementara itu, pada periode Presiden Prabowo, sentimen negatif lebih tinggi dibandingkan yang lain, mencapai lebih dari 40%, diikuti oleh sentimen netral sekitar 34%, dan sentimen positif yang lebih rendah, sekitar 27%. Hal ini menunjukkan adanya pergeseran persepsi publik dari yang sebelumnya relative

seimbang pada masa Jokowi, menjadi lebih cenderung negatif pada masa awal periode Prabowo. Hasil perbandingan dapat dilihat pada Gambar.4



Gambar 4. Perbandingan Sentimen Berdasarkan Hasil Pelabelan Kedua Periode

3.5 Pembobotan dengan TF-IDF

Berikut hasil pembobotan dengan TF-IDF, dapat dilihat pada Tabel 3. dan Tabel 4.

3.5.1 Periode Jokowi

Berikut merupakan beberapa hasil pembobotan atau kata-kata yang sering muncul pada data untuk Periode Jokowi dapat dilihat pada Tabel 3.

Tabel 3. Hasil Pembobotan dengan TF-IDF pada Periode Jokowi

No	Bangun	dukung	ikn	indonesia	Lanjut	...	proyek
1	0.2	0.0	0.2	0.0	0.1	...	0.3
2	0.1	0.3	0.1	0.4	0.2	...	0.2
3	0.4	0.2	0.3	0.1	0.0	...	0.2
...
600	0.0	0.0	0.1	0.0	0.0	...	0.0

3.5.2 Periode Prabowo

Berikut merupakan beberapa hasil pembobotan atau kata-kata yang sering muncul pada data untuk Periode Jokowi dapat dilihat pada Tabel 4.

Tabel 4. Hasil Pembobotan dengan TF-IDF pada Periode Prabowo

No	Bangun	Dana	ikn	investasi	Investor	...	negara
1	0.1	0.1	0.2	0.0	0.0	...	0.0
2	0.3	0.0	0.1	0.0	0.1	...	0.2
3	0.1	0.4	0.3	0.3	0.2	...	0.1
...
600	0	0.0	0.0	0.0	0.0	...	0.1

3.6 Klasifikasi Algoritma NBC, SVM dan K-NN

3.6.1 Pembagian Data

Setelah dilakukan tahap pre-processing data, pelabelan, dan pembobotan dengan TF-IDF, langkah selanjutnya adalah mengimplementasikan klasifikasi dengan menggunakan model Naive Baiyes Classifier, Support Vector Machine dan K-Nearest Neighbour. Sebelum masuk ke proses klasifikasi, dataset terlebih dahulu dibagi menjadi data latih dan data uji dengan metode hold-out validation. Pada penelitian ini, algoritma Naive Bayes dan Support Vector Machine menggunakan pembagian data dengan perbandingan 70:30, di mana 70% data digunakan untuk proses pelatihan (training), dan 30% sisanya digunakan untuk pengujian (testing) performa model. Dan data yang digunakan akan di SMOTE terlebih dahulu, untuk memastikan bahwa data yang digunakan seimbang. Sedangkan untuk algoritma K-Nearest Neighbour metode yang digunakan adalah Cross Validation dengan nilai 10 K-Fold, dengan rentang nilai k mulai dari 3 hingga 21 dengan langkah 2 (3, 5, 7, ..., 21).

3.6.2 Algoritma Naive Baiyes Classifier

Hasil pengujian pada algoritma Naive Baiyes Classifier dapat dilihat pada Tabel.4 Sampai Tabel.7 serta hasil dari Confusion Matrix dapat dilihat pada Gambar 5. dan Gambar 6.

a. Periode Jokowi

Pada Periode Jokowi, akan dilakukan uji coba performa Algoritma NBC pada setiap variasi dengan dua kondisi, yaitu menggunakan metode SMOTE dan tanpa SMOTE. Untuk hasil performa NBC tanpa SMOTE dapat dilihat pada Tabel 4. Sedangkan untuk hasil performa NBC dengan SMOTE dapat dilihat pada Tabel.5

Tabel 4. Hasil Performa Algoritma NBC pada Periode Jokowi tanpa SMOTE

Variasi	Accuracy	Precision	Recall	F-1 Score
GaussianNB	0.62	0.63	0.62	0.62
MultinomialNB	0.63	0.65	0.63	0.63
BernoulliNB	0.49	0.69	0.49	0.48

Tabel 5. Hasil Performa Algoritma NBC pada Periode Jokowi dengan SMOTE

Variasi	Accuracy	Precision	Recall	F-1 Score
GaussianNB	0.73	0.74	0.73	0.73
MultinomialNB	0.65	0.66	0.66	0.65
BernoulliNB	0.63	0.75	0.62	0.61

Berdasarkan Tabel 4 (tanpa SMOTE) dan Tabel 5 (dengan SMOTE), performa algoritma Naive Bayes Classifier (NBC) pada periode Jokowi meningkat setelah penerapan metode SMOTE. Variasi MultinomialNB menunjukkan akurasi sebesar 63% tanpa SMOTE dan tetap menjadi salah satu variasi terbaik. BernoulliNB memiliki presisi tinggi 69% namun recall rendah 49%, menyebabkan F1-Score hanya 48%. Setelah menggunakan SMOTE, GaussianNB menunjukkan peningkatan signifikan dengan akurasi naik dari sebelumnya lebih rendah menjadi 73% (naik sekitar 9–10%), presisi 74%, recall 73%, dan F1-Score 73%. Hal ini menunjukkan SMOTE efektif menyeimbangkan data sehingga model GaussianNB belajar lebih baik dan memberikan prediksi lebih akurat. Secara keseluruhan, SMOTE meningkatkan performa NBC sekitar 9–25%, terutama memperkuat GaussianNB sebagai variasi terbaik pada periode ini.

b. Periode Prabowo

Pada Periode Prabowo, akan dilakukan uji coba performa Algoritma NBC pada setiap variasi dengan dua kondisi, yaitu menggunakan metode SMOTE dan tanpa SMOTE. Untuk hasil performa NBC tanpa SMOTE dapat dilihat pada Tabel 6. Sedangkan untuk hasil performa NBC dengan SMOTE dapat dilihat pada Tabel.

Tabel 6. Hasil Performa Algoritma NBC pada Periode Prabowo tanpa SMOTE

Variasi	Accuracy	Precision	Recall	F-1 Score
GaussianNB	0.62	0.65	0.64	0.64
MultinomialNB	0.52	0.61	0.52	0.48
BernoulliNB	0.46	0.52	0.46	0.40

Tabel 7. Hasil Performa Algoritma NBC pada Periode Prabowo dengan SMOTE

Variasi	Accuracy	Precision	Recall	F-1 Score
GaussianNB	0.67	0.70	0.68	0.67
MultinomialNB	0.69	0.69	0.70	0.69
BernoulliNB	0.59	0.78	0.59	0.55

Berdasarkan Tabel 6 (tanpa SMOTE) dan Tabel 7 (dengan SMOTE), performa algoritma Naive Bayes Classifier (NBC) pada periode Prabowo meningkat setelah penerapan SMOTE. Variasi GaussianNB naik akurasinya dari 62% menjadi 67% (naik 5%), dengan F1-Score dari 64% menjadi 67% (naik 3%). MultinomialNB mengalami peningkatan akurasi dari 52% menjadi 69% (naik 17%) dan F1-Score dari 48% menjadi 69% (naik 21%), menjadikannya variasi terbaik setelah di SMOTE. BernoulliNB juga naik dari akurasi 46% ke 59% (naik 13%) dan F1-Score dari 40% ke 55% (naik 15%). Secara keseluruhan, SMOTE meningkatkan performa model sebesar 3–21% dan memperkuat MultinomialNB sebagai variasi paling optimal dalam klasifikasi sentimen pada periode ini.

3.6.3 Algoritma Support Vector Machine

Hasil pengujian pada algoritma Support Vector Machine dapat dilihat pada Tabel.8 dan Tabel.11 serta hasil dari Confusion Matrix dapat dilihat pada Gambar 7. dan Gambar 8. dibawah ini.

a. Periode Jokowi

Pada Periode Jokowi, akan dilakukan uji coba performa Algoritma SVM pada setiap variasi dengan dua kondisi, yaitu menggunakan metode SMOTE dan tanpa SMOTE. Untuk hasil performa SVM tanpa SMOTE dapat dilihat pada Tabel 8. Sedangkan untuk hasil performa NBC dengan SMOTE dapat dilihat pada Tabel 9.

Tabel 8. Hasil Performa Algoritma SVM pada Periode Jokowi tanpa SMOTE

Kernel	Cost(C)	Gamma	Degree(d)	Accuracy	Precision	Recall	F1-Score
--------	---------	-------	-----------	----------	-----------	--------	----------



Linear	10	0.68	0.71	0.68	0.68
RBF	100	0.1	...	0.68	0.69	0.68	0.68
Polynomial	0.01	10	2	0.59	0.71	0.59	0.59
Sigmoid	100	0.1	...	0.62	0.63	0.62	0.62

Tabel 9. Hasil Performa Algoritma SVM pada Periode Jokowi dengan SMOTE

Kernel	Cost(C)	Gamma	Degree(d)	Accuracy	Precision	Recall	F1-Score
Linear	10	0.70	0.72	0.69	0.69
RBF	100	0.1	...	0.71	0.73	0.71	0.71
Polynomial	0.01	10	2	0.62	0.73	0.63	0.62
Sigmoid	100	0.1	...	0.69	0.71	0.69	0.69

Berdasarkan Tabel 8 (tanpa SMOTE) dan Tabel 9 (dengan SMOTE), dapat disimpulkan bahwa penerapan SMOTE juga meningkatkan performa semua kernel pada algoritma SVM untuk periode Jokowi. Kernel RBF mencatat hasil terbaik setelah menggunakan SMOTE, dengan akurasi, precision, recall, dan F1-score sebesar 71%, meningkat dari sebelumnya 68%. Kernel Linear naik dari 68% ke 70%, Polynomial dari 59% ke 69%, dan Sigmoid dari 63% ke 68%. Secara keseluruhan, SMOTE meningkatkan performa model sebesar 2–10% dan memperkuat kernel RBF sebagai yang paling optimal dalam klasifikasi sentimen pada periode ini.

b. Periode Prabowo

Pada Periode Prabowo, akan dilakukan uji coba performa Algoritma SVM pada setiap variasi dengan dua kondisi, yaitu menggunakan metode SMOTE dan tanpa SMOTE. Untuk hasil performa NBC tanpa SMOTE dapat dilihat pada Tabel 10. Sedangkan untuk hasil performa SVM dengan SMOTE dapat dilihat pada Tabel.11

Tabel 10. Hasil Performa Algoritma SVM pada Periode Prabowo tanpa SMOTE

Kernel	Cost(C)	Gamma	Degree(d)	Accuracy	Precision	Recall	F1-Score
Linear	10	0.62	0.65	0.62	0.61
RBF	100	0.1	...	0.63	0.66	0.63	0.63
Polynomial	0.01	10	2	0.51	0.68	0.51	0.45
Sigmoid	100	0.1	...	0.59	0.61	0.59	0.59

Tabel 11. Hasil Performa Algoritma SVM pada Periode Prabowo dengan SMOTE

Kernel	Cost(C)	Gamma	Degree(d)	Accuracy	Precision	Recall	F1-Score
Linear	1	0.77	0.79	0.77	0.77
RBF	100	0.1	...	0.81	0.83	0.81	0.81
Polynomial	0.01	10	2	0.68	0.80	0.69	0.68
Sigmoid	100	0.1	...	0.79	0.81	0.79	0.79

Berdasarkan Tabel 10 (tanpa SMOTE) dan Tabel 11 (dengan SMOTE), dapat disimpulkan bahwa penerapan SMOTE juga meningkatkan performa semua kernel pada algoritma SVM untuk periode Prabowo. Kernel RBF menunjukkan hasil terbaik setelah menggunakan SMOTE, dengan nilai akurasi, precision, recall, dan F1-score sebesar 81%, meningkat dari sebelumnya 63%. Kernel Linear naik dari 62% ke 77%, Polynomial dari 51% ke 80%, dan Sigmoid dari 59% ke 79%. Secara keseluruhan, SMOTE meningkatkan performa model sebesar 16–29% dan menjadikan RBF sebagai kernel paling optimal untuk klasifikasi sentimen pada periode ini

3.6.4 Algoritma K-Nearest Neighbour

Hasil pengujian pada algoritma K-Nearest Neighbour dapat dilihat pada Tabel.12 sampai Tabel.15

a. Periode Jokowi

Pada Periode Jokowi akan dilakukan uji coba performa Algoritma K-NN pada dua kondisi, yaitu menggunakan metode SMOTE dan tanpa SMOTE. Untuk hasil performa K-NN tanpa SMOTE dapat dilihat pada Tabel 12. Sedangkan untuk hasil performa K-NN dengan SMOTE dapat dilihat pada Tabel.13

Tabel 12. Hasil Performa Algoritma K-NN pada Periode Jokowi tanpa SMOTE

Nilai K	Accuracy	Precision	Recall
3	0.49	0.56	0.49
5	0.50	0.56	0.50
7	0.51	0.55	0.51
9	0.53	0.55	0.53
11	0.52	0.55	0.52
13	0.54	0.57	0.54
15	0.54	0.56	0.54
17	0.55	0.58	0.55



19	0.54	0.59	0.54
21	0.54	0.58	0.54

Tabel 13. Hasil Performa Algoritma K-NN pada Periode Jokowi dengan SMOTE

Nilai K	Accuracy	Precision	Recall
3	0.49	0.56	0.49
5	0.50	0.56	0.50
7	0.51	0.55	0.51
9	0.52	0.55	0.53
11	0.52	0.55	0.52
13	0.54	0.57	0.54
15	0.54	0.56	0.54
17	0.56	0.64	0.56
19	0.55	0.59	0.55
21	0.54	0.58	0.54

Berdasarkan Tabel 12 dan Tabel 13, performa algoritma K-Nearest Neighbor (K-NN) pada periode Jokowi cenderung kurang optimal, baik tanpa maupun dengan penerapan metode SMOTE. Tanpa SMOTE, nilai K terbaik adalah 17 dengan akurasi tertinggi 56%, precision 64%, dan recall 56%. Namun, secara umum akurasi seluruh variasi K tidak melebihi 56%, menunjukkan K-NN kurang efektif pada data yang tidak seimbang. Setelah penerapan SMOTE, performa K-NN tidak mengalami peningkatan signifikan karena data hanya tersedia untuk K = 3 dengan hasil yang sama seperti tanpa SMOTE (akurasi 49%, precision 56%, recall 49%). Dengan demikian, SMOTE tidak terbukti efektif untuk meningkatkan performa K-NN pada dataset ini, dan K-NN bukan pilihan optimal untuk klasifikasi sentimen pada periode Jokowi, baik dengan data seimbang maupun tidak seimbang.

b. Periode Prabowo

Pada Periode Prabowo, akan dilakukan uji coba performa Algoritma K-NN dengan dua kondisi, yaitu menggunakan metode SMOTE dan tanpa SMOTE. Untuk hasil performa K-NN tanpa SMOTE dapat dilihat pada Tabel 14. Sedangkan untuk hasil performa K-NN dengan SMOTE dapat dilihat pada Tabel 15.

Tabel 14. Hasil Performa Algoritma K-NN pada Periode Prabowo tanpa SMOTE

Nilai K	Accuracy	Precision	Recall
3	0.37	0.33	0.37
5	0.36	0.29	0.36
7	0.36	0.22	0.36
9	0.39	0.58	0.39
11	0.42	0.63	0.42
13	0.44	0.62	0.44
15	0.46	0.65	0.46
17	0.45	0.62	0.45
19	0.46	0.63	0.47
21	0.47	0.58	0.47

Tabel 15. Hasil Performa Algoritma K-NN pada Periode Prabowo dengan SMOTE

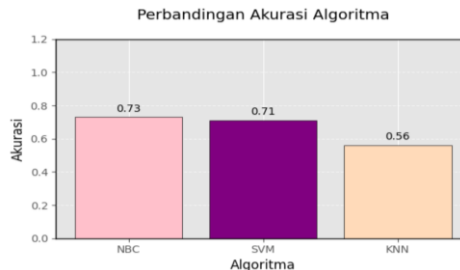
Nilai K	Accuracy	Precision	Recall
3	0.37	0.33	0.37
5	0.36	0.29	0.36
7	0.36	0.22	0.36
9	0.39	0.58	0.39
11	0.42	0.63	0.42
13	0.44	0.62	0.44
15	0.46	0.65	0.46
17	0.45	0.62	0.45
19	0.46	0.63	0.47
21	0.47	0.58	0.47

Berdasarkan Tabel 14 dan Tabel 15, performa algoritma K-Nearest Neighbor (K-NN) pada periode Prabowo menunjukkan hasil yang serupa tanpa dan dengan penerapan metode SMOTE. Akurasi tertinggi dicapai pada nilai K=21 dengan 47%, precision sebesar 58%, dan recall 47%. Secara keseluruhan, nilai akurasi, precision, dan recall di berbagai nilai K relatif rendah dan tidak menunjukkan peningkatan signifikan setelah SMOTE diterapkan. Hal ini mengindikasikan bahwa SMOTE tidak memberikan dampak positif terhadap performa K-NN pada periode Prabowo, dan algoritma ini kurang cocok digunakan untuk klasifikasi sentimen pada data tersebut.

3.6.5 Perbandingan Akurasi Ketiga Algoritma

a. Periode Jokowi

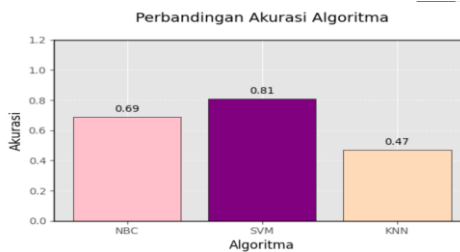
Berdasarkan data terkait IKN pada Periode Pemerintahan Jokowi, dari ketiga algoritma yang digunakan yakni, Naive Bayes Classifier (NBC), Support Vector Machine (SVM), dan K-Nearest Neighbors (KNN), dapat disimpulkan bahwa algoritma Naive Bayes Classifier menunjukkan performa tertinggi dengan akurasi sebesar 73%, seperti ditampilkan pada Gambar 5.



Gambar 5. Perbandingan Akurasi Ketiga Algoritma pada Periode Jokowi

b. Periode Prabowo

Berdasarkan data terkait IKN pada Periode Pemerintahan Prabowo, dari ketiga algoritma yang digunakan yakni, Naive Bayes Classifier (NBC), Support Vector Machine (SVM), dan K-Nearest Neighbors (KNN), dapat disimpulkan bahwa algoritma Support Vector Machine menunjukkan performa tertinggi dengan akurasi sebesar 81%, seperti ditampilkan pada Gambar 6.



Gambar 6. Perbandingan Akurasi Ketiga Algoritma pada Periode Prabowo

3.7 Visualisasi Data

Berikut merupakan hasil visualisasi data komentar terkait topik IKN (Ibu Kota Negara Baru) yang divisualisasikan menggunakan Wordcloud. Visualisasi ini dibagi menjadi tiga kategori, yaitu visualisasi untuk data dengan sentimen positif, sentimen negatif, dan sentimen netral. Visualisasi tersebut dapat dilihat pada Gambar 13 dan Gambar 14.

3.7.1 Periode Jokowi

Berikut merupakan visualisasi data Positif, Negatif dan Netral pada data untuk Periode Presiden Jokowi. Dapat dilihat pada Gambar 7.



Gambar 7. Visualisasi (a) Data Positif, (b) Data Negatif, dan (c) Data Netral pada Periode Jokowi

Berdasarkan Gambar 7 dapat disimpulkan bahwa gambar tersebut menyajikan visualisasi wordcloud untuk masing-masing kategori sentimen positif (a), negatif (b), dan netral (c) dengan topik pemindahan Ibu Kota Negara (IKN) pada periode pemerintahan Presiden Jokowi. Berdasarkan hasil visualisasi, kata "IKN", "dukung", dan "bangun" mendominasi pada sentimen positif, menunjukkan adanya dukungan publik terhadap pembangunan IKN.

Sementara itu, sentimen negatif menampilkan kata seperti "uang", "rakyat", dan "henti", yang mencerminkan kekhawatiran publik terhadap pembiayaan dan dampaknya terhadap masyarakat. Adapun pada sentimen netral, kata-kata seperti "proyek", "simbol", dan "identitas" muncul, yang mengindikasikan adanya pandangan yang bersifat informatif atau deskriptif terhadap isu IKN. Visualisasi ini membantu memberikan gambaran umum persepsi publik di media sosial terhadap pemindahan IKN selama masa pemerintahan Presiden Jokowi.

3.7.2 Periode Prabowo

Berikut merupakan visualisasi data Positif, Negatif dan Netral pada data untuk Periode Presiden Prabowo. Dapat dilihat pada Gambar 8.



Gambar 8. Visualisasi (a) Data Positif, (b) Data Negatif, dan (c) Data Netral pada Periode Prabowo

Berdasarkan Gambar 8 dapat disimpulkan bahwa gambar tersebut menyajikan visualisasi wordcloud untuk data sentimen positif (a), negatif (b), dan netral (c) terkait pemindahan Ibu Kota Negara (IKN) pada periode pemerintahan Prabowo. Pada sentimen positif, kata-kata seperti "IKN", "investor", "bangun", dan "dukung" mendominasi, mengindikasikan optimisme masyarakat terhadap dukungan investasi dan kelanjutan proyek IKN. Sementara itu, word cloud sentimen negatif memperlihatkan kata "uang", "dana", "ikhlas", dan "bengkalai", yang mencerminkan kekhawatiran publik terkait pembiayaan dan potensi kegagalan proyek. Pada sentimen netral, kata-kata seperti "kerja", "undang", dan "investor" muncul sebagai bagian dari wacana yang lebih informatif dan diskursif. Perbandingan visual ini menunjukkan adanya pergeseran narasi dan fokus isu publik dalam menyikapi proyek IKN pada masa pemerintahan Prabowo.

4. KESIMPULAN

Penelitian ini mengungkap dinamika sentimen masyarakat terhadap kebijakan pemindahan Ibu Kota Negara (IKN) pada masa pemerintahan Presiden Joko Widodo dan Presiden Prabowo Subianto. Dengan pendekatan machine learning menggunakan algoritma NBC, SVM, dan K-NN, serta proses pelabelan Lexicon dan pembobotan TF-IDF, ditemukan bahwa sentimen negatif meningkat dari 35,9% pada periode Jokowi menjadi 40,3% pada periode Prabowo. Dari sisi performa, NBC unggul di periode Jokowi dengan akurasi 73%, sedangkan SVM unggul di periode Prabowo dengan akurasi 81%. Evaluasi metrik precision, recall, dan F1-score menunjukkan SVM lebih akurat dalam mengklasifikasikan kelas minoritas. Penggunaan metode SMOTE terbukti meningkatkan performa model secara signifikan, terutama pada algoritma NBC dan SVM. SMOTE efektif dalam menyeimbangkan data tidak seimbang, sehingga menghasilkan model yang lebih adil dan akurat. Penelitian ini menunjukkan pentingnya evaluasi menyeluruh dan teknik penyeimbangan data dalam analisis sentimen media sosial. Studi lanjutan disarankan dengan data lebih besar dan eksplorasi algoritma lain seperti ensemble learning.

REFERENCES

- [1] A. Setiawan and R. R. Suryono, "Analisis Sentimen Ibu Kota Nusantara menggunakan Algoritma Support Vector Machine dan Naïve Bayes," *Edumatic J. Pendidik. Inform.*, vol. 8, no. 1, pp. 183–192, 2024, doi: 10.29408/edumatic.v8i1.25667.
- [2] G. Aji, Z. Arfani, A. M. Sari, R. Septiani, and U. K. H. Abdurrahman Wahid, "Dampak Pemindahan Ibukota Negara Baru terhadap Ekonomi dan Sosial di Provinsi Kalimantan Timur," *J. Ilmu Huk.*, vol. 1, no. 5, pp. 2985–5624, 2023, doi: <https://doi.org/10.572349/kultura.v1i5.474>.
- [3] H. P. Sutanto, "Transformasi Sosial Budaya Penduduk IKN Nusantara," *J. Stud. Kebijak. Publik*, vol. 1, no. 1, pp. 43–56, 2022, doi: <https://doi.org/10.21787/jskp.1.2022.43-56>.
- [4] M. H. Al Habib, A. Dairobby, R. A. Zoya, and R. R. Pramasha, "Dampak Pembangunan Ikn Nusantara: Menjadi Solusi Pemerataan Perekonomian Atau Timbul Permasalahan Lingkungan?," *IJEN Indones. J. Econ. Educ. Econ.*, vol. 02, no. 02, pp. 405–411, 2024, [Online]. Available: <https://jurnal.academiacenter.org/index.php/IJEN/article/view/504>
- [5] S. Budi Setiawan and A. Rahman Isnain, "Sentimen Analisis Masyarakat Terhadap Pembangunan IKN Menggunakan Algoritma Lexicon Based Approach dan Naïve Bayes Samuel," *J. Media Inform. Budidarma*, vol. 8, no. 2, pp. 1019–1030, 2024, doi: 10.30865/mib.v8i2.7605.
- [6] J. Muliawan and E. Dazki, "Sentiment Analysis of Indonesia's Capital City Relocation Using Three Algorithms: Naïve Bayes, Knn, and Random Forest," *J. Tek. Inform.*, vol. 4, no. 5, pp. 1227–1236, 2023, doi: 10.52436/1.jutif.2023.4.5.1436.



- [7] N. Husin, “Komparasi Algoritma Random Forest, Naïve Bayes, dan Bert Untuk Multi-Class Classification Pada Artikel Cable News Network (CNN),” *J. Esensi Infokom J. Esensi Sist. Inf. dan Sist. Komput.*, vol. 7, no. 1, pp. 75–84, 2023, doi: 10.55886/infokom.v7i1.608.
- [8] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, “Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN,” *J. KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [9] A. Widyanto, K. Kusriani, and K. Kusnawi, “Pengaruh Keseimbangan Data terhadap Akurasi Model Support Vector Machine pada Data Set Donor Darah,” *J. Teknol. Terpadu*, vol. 9, no. 2, pp. 79–88, 2023, doi: 10.54914/jtt.v9i2.771.
- [10] A. R. Hakim, W. Gata, A. Z. P. Widodo, O. Kurniawan, and A. R. Syarif, “Analisis Perbandingan Algoritma Machine Learning Terhadap Sentimen Analisis Pemindahan Ibu Kota Negara,” *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 7, no. 2, pp. 179–185, 2023, doi: 10.35870/jtik.v7i2.701.
- [11] A. S. YUNATA, A. HALIM, and H. LUTHFIYAH, “Analisis Pengaruh Noise pada Performa K-Nearest Neighbors Algorithn dengan Variasi Jarak untuk klasifikasi Beban Listrik,” *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 12, no. 3, p. 745, 2024, doi: 10.26760/elkomika.v12i3.745.
- [12] D. Pramana, M. Afdal, M. Mustakim, and I. Permana, “Analisis Sentimen Terhadap Pemindahan Ibu Kota Negara Menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbors,” *J. Media Inform. Budidarma*, vol. 7, no. 3, p. 1306, 2023, doi: 10.30865/mib.v7i3.6523.
- [13] K. A. Lubis, M. T. A. Bangsa, and A. Yudertha, “Analisis Sentimen Opini Masyarakat terhadap Pindahnya Ibu Kota Indonesia dengan menggunakan Klasifikasi Naïve Bayes,” *J. Teknoinfo*, vol. 18, no. 1, pp. 226–238, 2024, [Online]. Available: <https://ejournal.teknokrat.ac.id/index.php/teknoinfo/index>
- [14] N. Hadi and D. Sugiarto, “Analisis Sentimen Pembangunan IKN pada Media Sosial X Menggunakan Algoritma SVM , Logistic Regression dan Naïve Bayes,” *J. Inform. J. Pengemb. IT*, vol. 10, no. 1, pp. 37–49, 2025, doi: 10.30591/jpit.v10i1.7106.
- [15] A. Munawaroh, R. Ridhoi, and R. Rudiman, “Sentiment Analysis Dengan Naïve Bayes Berbasis Orange Terhadap Resiko Pembangunan Ikn,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 587–592, 2024, doi: 10.36040/jati.v8i1.8454.
- [16] Aurumnisva Faturrahmi, Zamahsary Martha, Yenni Kurniawati, and Fadhilah Fitri, “Sentiment Analysis of Prabowo Subianto as 2024 Presidential Candidate on Twitter Using K-Nearest Neighbor Algorithm,” *UNP J. Stat. Data Sci.*, vol. 1, no. 5, pp. 385–391, 2023, doi: 10.24036/ujsds/vol1-iss5/101.
- [17] S. S. Sohail *et al.*, “Crawling Twitter data through API: A technical/legal perspective,” 2021, [Online]. Available: <http://arxiv.org/abs/2105.10724>
- [18] I. Yanti and E. Utami, “Sentiment Analysis Of Indonesia ’ S Capital Relocation Using Word2vec And Long Short-Term Memory Method Analisis Sentimen Pemindahan Ibu Kota Indonesia Menggunakan Word Embedding Word2vec Dan Metode Long Short-Term,” vol. 6, no. 1, pp. 149–157, 2025, doi: <https://doi.org/10.52436/1.jutif.2025.6.1.2712>.
- [19] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.
- [20] D. R. Fathwa Daud, B. Irawan, and A. Bahtiar, “Penerapan Metode Naive Bayes Pada Analisis Sentimen Aplikasi Mcdonalds Di Google Play Store,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 759–766, 2024, doi: 10.36040/jati.v8i1.8784.
- [21] M. Theofany Aulia Anwar, S. Hadi Wijoyo, and W. Hayuhardhika Nugraha Putra, “Implementasi Metode TextRank dan Named Entity Recognition Untuk Ekstraksi Kata Kunci Pada Media Online Berita,” *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 5, no. 1, pp. 34–41, 2024, doi: 10.25126/justsi.v5i1.401.
- [22] Rianto, A. B. Mutiara, E. P. Wibowo, and P. I. Santosa, “Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation,” *J. Big Data*, vol. 8, no. 1, pp. 1–16, 2021, doi: 10.1186/s40537-021-00413-1.
- [23] J. A. Rieuwpassa, S. Sugito, and T. Widiari, “Implementasi Metode Naive Bayes Classifier Untuk Klasifikasi Sentimen Ulasan Pengguna Aplikasi Netflix Pada Google Play,” *J. Gaussian*, vol. 12, no. 3, pp. 362–371, 2024, doi: 10.14710/j.gauss.12.3.362-371.
- [24] D. Darwis, N. Siskawati, and Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [25] F. Abdusyukur, “Penerapan Algoritma Support Vector Machine (Svm) Untuk Klasifikasi Pencemaran Nama Baik Di Media Sosial Twitter,” *Komputa J. Ilm. Komput. dan Inform.*, vol. 12, no. 1, pp. 73–82, 2023, doi: 10.34010/komputa.v12i1.9418.
- [26] A. D. Adhi Putra, “Sentiment Analysis on User Reviews of the Bibit and Bareksa Application with the KNN Algorithm,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021.
- [27] Y. Jin, “Development of Word Cloud Generator Software Based on Python,” *Procedia Eng.*, vol. 174, pp. 788–792, 2017, doi: 10.1016/j.proeng.2017.01.223.