

Analisis Sentimen Pemindahan Ibu Kota Indonesia Menggunakan K-Nearest Neighbor

Asri Khayatul Iman^{*1}, Erik Iman Heri Ujjianto²

¹Informatika, Fakultas Sains & Teknologi, Universitas Teknologi Yogyakarta, Indonesia

²Magister Teknologi Informasi, Universitas Teknologi Yogyakarta, Indonesia

Email: ¹khayatulasri10@gmail.com, ²erik.iman@uty.ac.id

Abstrak

Pemindahan Ibu Kota Indonesia dari Jakarta ke Kalimantan Timur telah memunculkan beragam opini di masyarakat. Penelitian ini bertujuan untuk menganalisis sentimen publik terkait rencana tersebut dengan menggunakan algoritma K-Nearest Neighbor (KNN). Data diperoleh dari media sosial Twitter menggunakan teknik crawling, kemudian melalui tahapan pengumpulan data, pemrosesan data, pembagian data menjadi data latih dan data uji, serta pelatihan model. Algoritma KNN digunakan untuk mengklasifikasikan sentimen menjadi dua kategori, yaitu positif dan negatif. Berdasarkan 4241 data tweet yang di analisis, sekitar 71,5% menunjukkan sentimen positif, sementara 28,5% lainnya bersifat negatif. Tweet positif umumnya mengandung kata seperti "kota modern" dan "pembangunan merata", sementara tweet negatif mengkhawatirkan masalah seperti "biaya" dan "dampak lingkungan". Penelitian ini memperlihatkan bagaimana analisis sentimen dapat membantu dalam memahami persepsi masyarakat terhadap perpindahan ibu kota Indonesia. Analisis sentimen dapat memahami pandangan masyarakat terhadap kebijakan besar. Hasil ini dapat membantu pemerintah dalam merumuskan kebijakan yang lebih responsif terhadap kekhawatiran publik. Analisis sentimen ini memberikan cara yang efektif bagi pemerintah untuk mengevaluasi dan merespons persepsi publik terhadap perpindahan ibu kota untuk mengambil keputusan yang lebih baik.

Kata kunci: Analisis sentimen, K-Nearest Neighbor, Opini publik, Pemindahan ibu kota, Twitter

Sentiment Analysis of the Relocation of Indonesia's Capital Using K-Nearest Neighbor

Abstract

The relocation of Indonesia's capital from Jakarta to East Kalimantan has sparked a variety of opinions among the public. This study aims to analyze public sentiment regarding this plan using the K-Nearest Neighbor (KNN) algorithm. Data was collected from Twitter using a crawling technique, followed by stages of data collection, data processing, splitting the data into training and testing sets, and model training. The KNN algorithm was used to classify sentiment into two categories: positive and negative. Based on 4,241 tweet data analyzed, approximately 71.5% showed positive sentiment, while the remaining 28.5% was negative. Positive tweets commonly contained words such as "modern city" and "equal development," while negative tweets expressed concerns such as "cost" and "environmental impact." This study demonstrates how sentiment analysis can help in understanding public perception of the capital relocation. Sentiment analysis provides insights into public views on major policies. These results can help the government formulate more responsive policies that address public concerns. This sentiment analysis offers an effective way for the government to evaluate and respond to public perceptions of the capital relocation, enabling better decision-making.

Keywords: Capital relocation, K-Nearest Neighbor, Public opinion, Twitter, Sentiment analysis

1. PENDAHULUAN

Presiden Joko Widodo resmi mendeklarasikan pemindahan ibu kota ke Kalimantan Timur pada Agustus 2019. Namun, keputusan ini memicu berbagai respon dari masyarakat yang terbagi menjadi dua, antara yang mendukung dan yang menolak pemindahan ibu kota. Saat ini, masyarakat dapat menyampaikan pendapat atau memberikan penilaian melalui media sosial. Salah satu platform yang sering digunakan untuk menyampaikan pendapat adalah Twitter. Twitter merupakan platform komunitas informal dengan basis pengguna dan informasi yang luas dan terus berkembang. Banyak pengguna menggunakan Twitter untuk mengekspresikan pandangan mereka karena platform ini memungkinkan untuk berbagi pesan singkat yang dikenal sebagai tweet. Ada potensi untuk melakukan analisis lebih lanjut dengan menguraikan opini dan sentimen pengguna secara komputasi[1].

Pemindahan Ibu Kota Indonesia dari Jakarta ke Kalimantan Timur menjadi isu yang menuai beragam respon di masyarakat. Kebijakan ini diambil untuk mengurangi beban Jakarta sebagai pusat pemerintahan dan perekonomian, serta memberikan peluang bagi daerah lain untuk berkembang. Namun, kebijakan ini juga menimbulkan kekhawatiran, terutama terkait tingginya biaya, potensi kerusakan lingkungan, dan dampak sosial di ibu kota baru. Menurut beberapa studi, pemindahan ibu kota juga dapat berdampak pada redistribusi sumber daya dan merangsang pertumbuhan ekonomi di daerah yang sebelumnya terabaikan[2]. Namun, kebijakan ini menghadapi tantangan besar dalam hal pelaksanaan, biaya, serta dampaknya terhadap masyarakat sekitar[3].

Analisis sentimen adalah alat yang berguna untuk memahami perspektif individu terhadap kebijakan ini. Analisis sentimen adalah komponen pemrosesan bahasa alami yang berupaya mengkategorikan opini publik berdasarkan sentimen positif dan negatif. Penelitian sebelumnya menunjukkan bahwa analisis sentimen melalui media sosial, khususnya Twitter, dapat memberikan pemahaman yang lebih baik mengenai opini publik dalam konteks kebijakan pemerintah[4]. Twitter menawarkan kemampuan berbagi menarik yang dapat dimanfaatkan untuk beragam aplikasi, termasuk pengambilan informasi. Fitur unik dari platform ini membuat pengambilan informasi menjadi mudah. Twitter adalah platform utama dalam penelitian kami karena karakteristik real-time dan kapasitasnya untuk mewakili sentimen publik yang luas[5]. Pendekatan K-Nearest Neighbor digunakan untuk kategorisasi dalam penelitian ini karena kesederhanaannya. Prosedur ini bergantung pada metodologi pembobotan yang sederhana dan mudah untuk dijalankan, dimodifikasi, dan dipahami. Selain itu, pendekatan ini menunjukkan tingkat akurasi yang tinggi.

Penelitian tentang analisis sentimen semakin maju dalam domain penambangan teks, seperti yang terlihat dari banyak publikasi jurnal yang menggunakan pengklasifikasi Naïve Bayes dan KNN, yang dilakukan oleh berbagai peneliti sebelumnya. Berbagai metode lain seperti Naïve Bayes atau SVM telah digunakan, tetapi penelitian ini menunjukkan potensi KNN dalam memberikan hasil lebih akurat dengan dataset besar. Penelitian ini menggunakan metode K-Nearest Neighbor (KNN) untuk mengkaji sentimen masyarakat terhadap pemindahan ibu kota. Metode KNN dipilih karena kemanjurannya dalam klasifikasi data, khususnya untuk data tekstual seperti opini media sosial. Sejumlah penelitian sebelumnya telah menunjukkan bahwa teknik analisis sentimen yang memanfaatkan media sosial menawarkan pemahaman yang lebih menyeluruh tentang perspektif individu mengenai berbagai situasi.

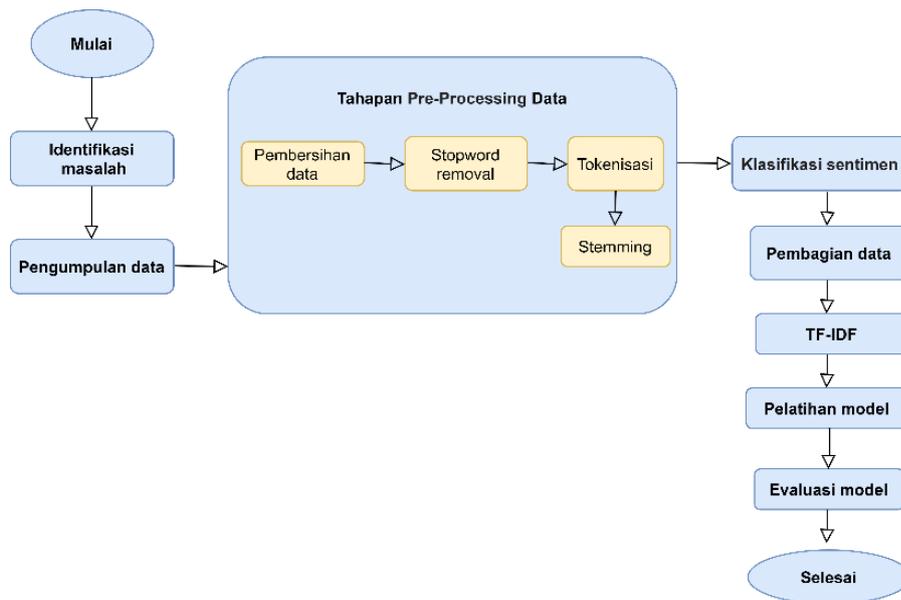
Penelitian terdahulu dengan topik yang terkait di antaranya analisis sentimen relokasi ibu kota Nusantara menggunakan algoritma Naïve Bayes dan KNN dengan menggunakan 800 data, sehingga memperoleh akurasi metode Naïve Bayes 82,27% dan metode KNN 88,12%[6]. Penelitian tambahan analisis sentimen terkait pemindahan ibu kota Indonesia di Twitter menggunakan algoritma K-Nearest Neighbor (K-NN) dengan dataset sebanyak 1.000 entri. Kajian tersebut meliputi penentuan nilai K optimal pada algoritma K-NN untuk mengidentifikasi parameter terbaik. Pada pengumpulan data awal, algoritma K-NN mencapai akurasi sebesar 62% dengan K=9, sedangkan pada pengumpulan data selanjutnya memperoleh akurasi sebesar 64% dengan K=5[7]. Penelitian sebelumnya, yaitu dengan judul perbandingan metode KNN vs Naïve Bayes untuk analisis sentimen pada pengguna e-wallet aplikasi Dana menggunakan fitur ekstraksi TF-IDF, menunjukkan bahwa hasil KNN dan Naïve Bayes memiliki akurasi yang berbeda berdasarkan sumber label data. Pada data yang diberi label model lexicon, akurasi KNN mencapai 78% dan Naïve Bayes 74%. Namun pada data yang diberi label expert, akurasi kedua metode meningkat menjadi 96%[8]. Penelitian sebelumnya melakukan analisis sentimen terhadap persepsi masyarakat terhadap elektabilitas Pranowo pada pemilu 2024 di Twitter dengan memanfaatkan algoritma KNN dan Naïve Bayes. Temuan menunjukkan bahwa K-Nearest Neighbor mencapai akurasi lebih tinggi yaitu 99%, melampaui Naïve Bayes yang mencapai akurasi 96%[9]. Penelitian sebelumnya yang dilakukan oleh Muhidin, D, membandingkan kemanjuran metode Support Vector Machine dan K-Nearest Neighbor dalam analisis sentimen terhadap kebijakan new normal. Hasil pengujian menunjukkan bahwa metode K-NN menunjukkan akurasi yang unggul sebesar 72,96%, berbeda dengan SVM yang sebesar 71,19%[10]. Penelitian Pertiwi M, W menunjukkan bahwa pendekatan KNN memiliki akurasi yang lebih unggul dibandingkan dengan SVM dan Naïve Bayes, dengan KNN mencapai akurasi 90,76% dan AUC 0,939, SVM mencapai akurasi 89,03% dan AUC 0,5, dan Naïve Bayes mencatat akurasi sebesar 78,16% dan AUC sebesar 0,567[11].

Penelitian ini mengusulkan pendekatan K-Nearest Neighbor (KNN) untuk melakukan analisis sentimen terhadap opini masyarakat atas pemindahan ibu kota Indonesia. Studi ini menggunakan data yang bersumber dari Twitter, saluran utama untuk mengumpulkan opini publik mengenai topik sosial dan kebijakan pemerintah. Sentimen opini publik akan dikategorikan menjadi dua klasifikasi: emosi positif dan negatif. Analisis ini bertujuan untuk mengidentifikasi berbagai persoalan serta kelebihan dan kekurangan yang terkait dengan usulan pemindahan ibu kota. Hasil yang diharapkan dari analisis sentimen ini bertujuan untuk memberikan pemahaman yang lebih mendalam tentang perspektif masyarakat, sehingga dapat berfungsi sebagai dasar untuk meninjau dan mempertimbangkan pengambilan keputusan dan perumusan kebijakan di masa depan.

Penelitian ini bertujuan untuk mengklasifikasikan sentimen publik terkait pemindahan ibu kota Indonesia menggunakan KNN dan memberikan rekomendasi berdasarkan hasil analisis.

2. METODE PENELITIAN

Penelitian ini menggunakan analisis sentimen publik terkait pemindahan Ibu Kota Indonesia. Pendekatan ini dipilih karena dinilai efektif untuk memahami pola dan karakteristik sentimen publik yang diungkapkan melalui media sosial, khususnya Twitter. Algoritma ini memungkinkan pengelompokan opini publik ke dalam kategori sentimen positif dan negatif menggunakan teknik machine learning, yaitu K-Nearest Neighbor (KNN). Algoritma KNN dipilih karena keandalannya dalam menangani klasifikasi data teks, terutama pada data yang berukuran besar dan tidak terstruktur [12]. Berikut ini adalah tahapan penelitian yang digunakan.



Gambar 1. Tahapan Penelitian

1. Identifikasi Masalah

Pemindahan ibu kota dari Jakarta ke Kalimantan Timur memunculkan beragam respons publik, mulai dari dukungan terhadap tujuan pemerataan hingga kekhawatiran tentang biaya, dampak lingkungan, dan perubahan sosial. Untuk memahami sentimen masyarakat terhadap kebijakan ini, diperlukan analisis sentimen melalui data Twitter yang merefleksikan opini publik secara real-time. Dalam hal ini, metode K-Nearest Neighbor (KNN) adalah salah satu teknik klasifikasi yang telah lama diterapkan dan dikenal karena kesederhanaannya. Metode ini juga dikenal sebagai Instance-Based Learning, KNN mengelompokkan objek dengan mengukur jarak antara objek tersebut dan objek lainnya di dalam data[13]. Penelitian ini bertujuan untuk menilai akurasi KNN dalam analisis sentimen serta memberikan gambaran umum persepsi publik sebagai pertimbangan dalam pengambilan kebijakan.

2. Pengumpulan Data

Metode pengumpulan data yang digunakan adalah web crawling, dengan sumber data berasal dari media sosial Twitter. Dalam teknik web crawling, pengumpulan data dilakukan dengan menggunakan kata kunci pencarian ‘IKN’ dan token akses ‘6046705ad539e1102d901f49193288647b08f8a9’. Token ini digunakan untuk mengakses API Twitter sehingga peneliti dapat mendapatkan data sesuai kata kunci yang telah ditentukan. Data yang diperoleh berupa data mentah yang akan melalui tahap pemrosesan lebih lanjut untuk memastikan kualitas dan kelayakannya dalam analisis mendalam. Dalam proses pengumpulan ini, peneliti melakukan crawling pada rentang tanggal 29 September hingga 5 Oktober 2024. Dari proses ini, berhasil dikumpulkan 4.241 data [14]. Alur pengambilan twets dapat dilihat pada Gambar 2.



Gambar 2. Alur Pengambilan Data

3. **Pre-Processing**

Dataset yang telah diperoleh kemudian melalui proses preprocessing data. Tahap pertama adalah:

- a. *Pembersihan data*, bertujuan untuk menghapus elemen-elemen teks yang tidak relevan, seperti simbol, angka, tautan URL, tag HTML, atau karakter khusus lainnya. Proses ini penting untuk memastikan bahwa data yang digunakan dalam analisis tidak mengandung elemen yang dapat mengganggu atau mendistorsi hasil klasifikasi sentimen.
- b. *Stopword removal* digunakan untuk membantu mengurangi kompleksitas data dan fokus pada kata-kata penting yang menggambarkan sentimen. Dengan menghapus kata-kata umum seperti "dan", "atau", "yang", yang tidak memberikan kontribusi signifikan terhadap analisis sentimen, kita dapat meningkatkan efisiensi dan akurasi model.
- c. *Tokenization* adalah proses memecah teks menjadi bagian-bagian lebih kecil yang disebut token, seperti ('ikn', 'adalah', 'kota', 'modern', dan seterusnya). Tokenisasi ini memudahkan model untuk memahami dan memproses setiap kata atau frasa yang memiliki makna dalam konteks analisis sentimen.
- d. *Stemming* yang bertujuan untuk mereduksi kata menjadi bentuk dasar (stem). Sebagai contoh, kata “membangun”, “membangun”, “pengembangan” semuanya diubah menjadi bentuk dasar “bangun” [15]. Proses ini membantu menyederhanakan variasi kata yang memiliki makna serupa, sehingga memperkecil dimensi data dan memudahkan model dalam pengklasifikasian.

Tabel 1. Contoh Pre-Processing

Tweet mentah	Pre-Processing
jujur mirip ikn bang #NKRI	jujur mirip ikn bang nkri
@macantutulijo Iyaa jadi burung garuda di IKN	macantutulijo iyaa jadi burung garuda di ikn
@dialsdrank IKN mau robohin gedung	dialsdrank ikn mau robohin gedung

4. **Klasifikasi Sentimen**

Setelah tahap pre-processing data, langkah selanjutnya adalah melakukan klasifikasi untuk menentukan apakah tweet tersebut mengandung sentimen positif atau negatif. Pada proses ini, peneliti menggunakan kode pemrograman di Google Colab untuk mempermudah klasifikasi sentimen secara otomatis. Data yang telah diproses diurutkan berdasarkan sentimen yang terkandung dalam tweet tersebut, sehingga menghasilkan opini masyarakat terhadap perpindahan Ibu Kota Indonesia. Tabel berikut menunjukkan hasil klasifikasi sentimen dari tweet terkait perpindahan ibu kota, yang menggambarkan sentimen positif dan negatif berdasarkan pendapat pengguna Twitter[16].

Tabel 2. Pengklasifikasian Sentimen

Creaned_Text	Sentiment
jujur ikn bang nkri	Negatif
macantutulijo iyaa burung garuda ikn	Positif
agbudisan akbarfaizal anakkolong anaktangsiad	Negatif
dialsdrank ikn robohin gedung	Positif

5. **Pembagian Data**

Dalam analisis sentimen, pembagian data merupakan langkah penting untuk memastikan kemampuan generalisasi model yang dibangun. Umumnya, data dibagi menjadi dua bagian, yaitu 80% untuk pelatihan dan 20% untuk pengujian. Data pelatihan digunakan untuk melatih model KNN agar dapat mengklasifikasikan komentar menjadi positif atau negatif, sementara data pengujian digunakan untuk mengukur kinerja model pada data yang belum pernah dihadapi sebelumnya, sehingga mengurangi risiko overfitting. Pembagian ini bertujuan untuk memastikan model tidak hanya efektif pada data pelatihan, tetapi juga mampu memproses data baru dengan akurat[17].

6. **TF-IDF**

Metode Term Frequency-Inverse Document Frequency (TF-IDF) merupakan teknik pembobotan yang diterapkan pada dataset yang telah melalui proses praproses. Dalam metode ini, digunakan pendekatan N-gram yang mencakup Unigram dan Bigram untuk membantu merepresentasikan teks dalam bentuk numerik. Melalui penelitian sebelumnya [18], diketahui bahwa pendekatan TF-IDF yang digabungkan dengan algoritma K-Nearest Neighbor (KNN) dapat meningkatkan performa klasifikasi teks. Selain efektif, metode TF-IDF juga menawarkan skema pembobotan yang sederhana namun mampu menghasilkan hasil klasifikasi yang cukup baik[19].

7. Pelatihan Model

Pelatihan model K-Nearest Neighbors (KNN) dalam penelitian ini mencakup proses pembelajaran dari data pelatihan guna mengenali pola-pola sentimen yang terdapat dalam data. Model ini disesuaikan dan dioptimalkan untuk meminimalkan kesalahan prediksi, sehingga dapat mengklasifikasikan sentimen dengan lebih akurat. Pemilihan nilai K merupakan salah satu parameter penting dalam algoritma KNN. Dalam penelitian ini, nilai K=3 dipilih. Pemilihan K yang lebih kecil akan membuat model lebih sensitif terhadap data pelatihan dan berisiko terjebak pada overfitting, sementara K yang lebih besar dapat menyebabkan model menjadi terlalu umum (underfitting). Untuk memastikan pemilihan K yang optimal, dilakukan proses validasi silang (cross-validation) untuk mengevaluasi kinerja model pada berbagai nilai K. Dalam hal ini, K=3 dipilih karena memberikan keseimbangan terbaik antara akurasi dan generalisasi, seperti yang diperoleh dari eksperimen validasi pada dataset yang digunakan [20].

8. Evaluasi Model

Evaluasi model dalam penelitian ini bertujuan untuk menilai sejauh mana model dapat memprediksi sentimen dengan akurat. Metrik yang digunakan untuk mengukur kinerja model meliputi akurasi, presisi, recall, dan F1-score[21]. Akurasi menghitung proporsi prediksi yang benar dengan rumus:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

TP merujuk pada True Positive, TN pada True Negative, FP pada False Positive, dan FN pada False Negative. Presisi mengukur tingkat ketepatan dalam memprediksi kelas positif, dengan rumus sebagai berikut:

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Sementara itu, recall mengukur kemampuan model untuk mendeteksi semua kasus positif dengan menggunakan rumus

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

Terakhir, skor F1 memberikan keseimbangan antara presisi dan recall dengan menggunakan rumus:

$$F1 - Score = 2x \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Akurasi dalam penelitian ini dihitung secara keseluruhan, sedangkan presisi, recall, dan f1-score dievaluasi secara spesifik untuk setiap kategori tweet. Dengan demikian nilai TP, TN, FP, FN berubah untuk setiap kelas tweet.

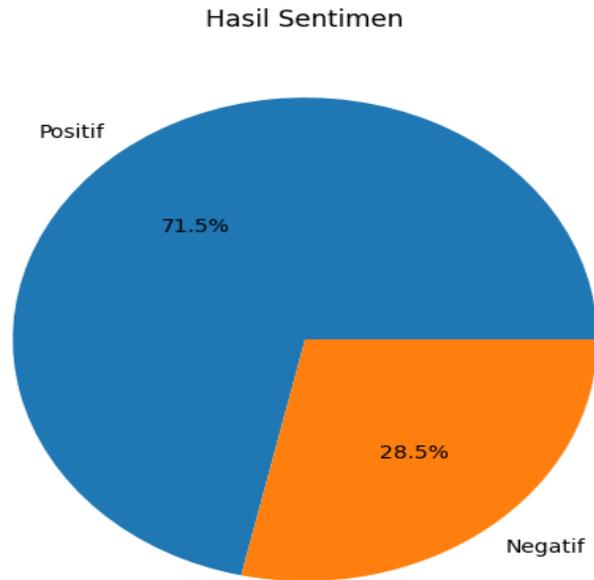
3. HASIL DAN PEMBAHASAN

Penelitian ini menganalisis sentimen dari 4.241 tweet yang membahas pemindahan Ibu Kota Indonesia dari Jakarta ke Kalimantan Timur. Data tersebut diklasifikasikan ke dalam dua kategori utama, yaitu sentimen positif dan negatif. Tabel 3 menunjukkan distribusi sentimen yang diperoleh dari analisis tersebut.

Tabel 3. Distribusi Klasifikasi Sentimen

Klasifikasi Sentimen	Jumlah tweet	Presentase (%)
Positif	3.032	71.5 %
Negatif	1.209	28.5 %
Jumlah	4241	100

Dari hasil analisis pada tabel 3, dapat dilihat bahwa mayoritas tweet (71,5%) memiliki sentimen positif terkait rencana pemindahan ibu kota. Hal ini menunjukkan bahwa masyarakat cenderung mendukung kebijakan ini, kemungkinan besar karena adanya harapan pemerataan pembangunan dan mengurangi beban yang selama ini ditanggung Jakarta. Sebaliknya, hanya 28,5% dari total tweet yang mengungkapkan sentimen negatif, yang mencerminkan kekhawatiran akan biaya dan dampak lingkungan dari kebijakan tersebut.



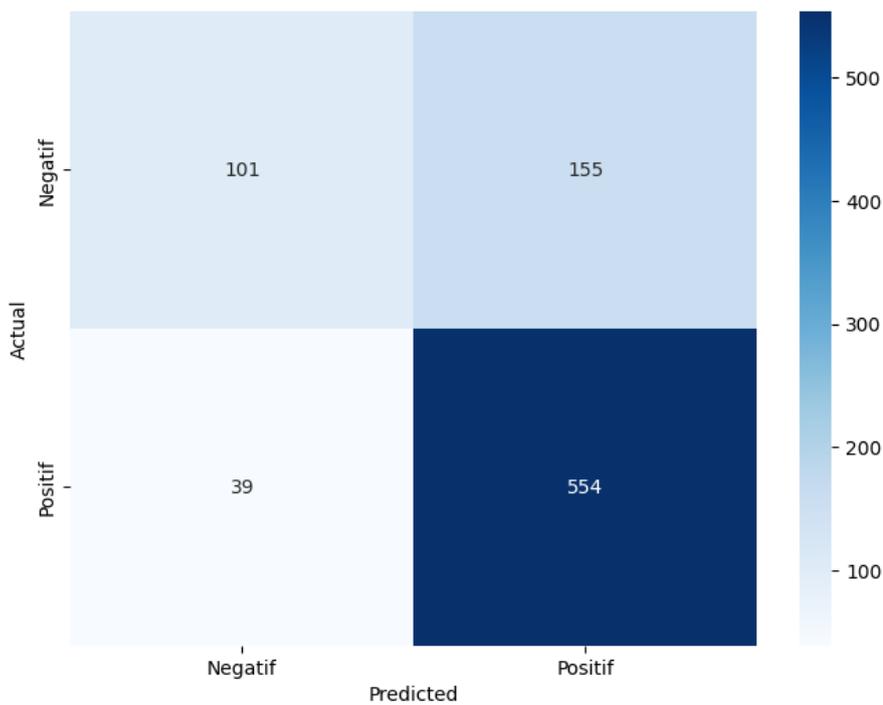
Gambar 3. Grafik Distribusi Sentimen

Gambar 3 menunjukkan hasil analisis sentimen dengan metode KNN, di mana 71,5% opini bersifat positif dan 28,5% negatif. Mayoritas opini cenderung positif terhadap isu yang dianalisis.

Tabel 4 Hasil Klasifikasi

Tweet_class	Precision	Recall	F1-Score	support
Positif	78%	93%	85%	593
Negatif	72%	39%	51%	256

Proses KNN menghasilkan akurasi sebesar 77% yang diperoleh dari confusion matrix. Adapun hasil precision, recall, f1-score dan support dapat dilihat pada tabel 4 :



Gambar 4. Confusion Matrix

Hasil analisis sentimen terhadap pemindahan ibu kota Indonesia menggunakan algoritma K-Nearest Neighbor (KNN) menunjukkan akurasi yang relatif baik, namun terdapat perbedaan signifikan antara klasifikasi sentimen positif dan negatif. Hasil evaluasi menunjukkan bahwa KNN berhasil mengklasifikasikan sentimen positif dengan lebih baik, sementara sentimen negatif memiliki performa yang lebih rendah dengan skor $F1=0.51$. Skor $F1$ yang rendah pada kategori sentimen negatif ini menjadi titik perhatian penting untuk melakukan analisis lebih mendalam terhadap penyebabnya.

Salah satu faktor utama yang dapat menjelaskan performa rendah pada sentimen negatif adalah kompleksitas bahasa yang digunakan oleh pengguna Twitter. Tweet yang mengandung sentimen negatif sering kali memiliki konstruksi kalimat yang lebih kompleks, seperti penggunaan ironi, sarkasme, atau pernyataan yang ambigu. Bahasa sehari-hari di Twitter sering kali tidak terstruktur dengan baik dan mengandung kata-kata yang dapat memiliki makna ganda tergantung pada konteksnya. Hal ini menyulitkan model KNN untuk secara konsisten mengidentifikasi sentimen yang sebenarnya dari teks tersebut. Misalnya, pengguna mungkin menggunakan kalimat yang terdengar positif, tetapi dengan konteks yang negatif, seperti "Pindah ibu kota? Biaya pasti melambung tinggi!".

Selain itu, ketidakseimbangan data antara sentimen positif dan negatif juga berperan dalam menurunnya akurasi pada kategori negatif. Dalam dataset yang digunakan, terdapat lebih banyak tweet yang mengungkapkan pendapat positif tentang pemindahan ibu kota, seperti harapan akan pembangunan yang merata dan kemajuan kota baru. Sebaliknya, tweet yang mengungkapkan sentimen negatif, seperti kekhawatiran terhadap biaya dan dampak lingkungan, jauh lebih sedikit. Model KNN cenderung lebih "terlatih" untuk mengenali pola dalam kategori yang lebih sering muncul, sehingga sentimen negatif yang lebih jarang dapat terlewat atau terkategori dengan kurang tepat.

Perbandingan dengan penelitian lain yang menggunakan algoritma berbeda, seperti Naive Bayes dan *Support Vector Machine* (SVM), memberikan wawasan lebih lanjut tentang potensi algoritma KNN dalam menangani masalah ini. Penelitian sebelumnya yang menggunakan Naive Bayes dan SVM juga menunjukkan hasil yang lebih baik dalam klasifikasi sentimen positif, tetapi performa mereka pada sentimen negatif tidak selalu lebih baik daripada KNN. Sebagai contoh, sebuah studi yang menggunakan SVM melaporkan skor $F1$ yang lebih rendah untuk sentimen negatif (0.45) dibandingkan dengan KNN yang mencapai 0.51 [1][3]. Namun, Naive Bayes cenderung menunjukkan kinerja yang sedikit lebih baik dalam beberapa kasus, dengan skor $F1$ untuk sentimen negatif mencapai 0.56 [22]. Perbandingan ini menunjukkan bahwa meskipun KNN memiliki beberapa keterbatasan, seperti kesulitan dalam menangani data yang tidak seimbang atau kompleksitas bahasa, KNN tetap menjadi pilihan yang solid untuk analisis sentimen jika data dipersiapkan dengan baik.

Untuk meningkatkan performa pada sentimen negatif, beberapa langkah perlu dipertimbangkan. Salah satunya adalah peningkatan jumlah data negatif. Dengan meningkatkan jumlah data yang mewakili sentimen negatif secara lebih proporsional, model KNN dapat belajar mengenali pola-pola yang lebih jelas dalam kategori ini. Selain itu, peningkatan teknik pra-pemrosesan, seperti menggunakan pendekatan lemmatization yang lebih canggih atau analisis semantik, dapat membantu model dalam menangani kompleksitas bahasa dan mengurangi kesalahan klasifikasi. Penggunaan model yang lebih kompleks seperti SVM atau bahkan model berbasis deep learning, seperti LSTM (*Long Short-Term Memory*), juga bisa menjadi alternatif untuk menangani masalah klasifikasi sentimen negatif.

Penurunan performa pada sentimen negatif dalam penelitian ini dapat dianggap sebagai tantangan yang umum dalam analisis sentimen, terutama ketika menggunakan dataset yang terbatas atau tidak seimbang. Namun, dengan perbaikan pada kualitas data dan penggunaan teknik optimasi yang lebih lanjut, seperti penyesuaian parameter KNN atau eksplorasi algoritma lain, model ini dapat lebih efektif dalam mengidentifikasi dan mengklasifikasikan sentimen negatif. Analisis lebih lanjut mengenai ketidakseimbangan data dan teknik peningkatan model dapat memberikan arah yang lebih jelas untuk penelitian selanjutnya dalam topik ini.

Hasil penelitian ini memberikan wawasan yang penting bagi pemerintah terkait bagaimana publik merespons rencana pemindahan ibu kota ke Kalimantan Timur. Mayoritas sentimen positif yang ditemukan dalam analisis ini dapat memberikan legitimasi bagi pemerintah untuk melanjutkan proyek pemindahan ibu kota, karena menunjukkan bahwa sebagian besar masyarakat mendukung kebijakan tersebut. Hal ini juga dapat menjadi dasar bagi pemerintah untuk memperkuat komunikasi publik dan lebih menekankan manfaat positif dari pemindahan ibu kota, seperti pemerataan pembangunan dan pengurangan beban Jakarta. Dengan pemahaman yang lebih mendalam tentang sentimen masyarakat, pemerintah dapat merancang strategi sosialisasi yang lebih efektif untuk menjawab kekhawatiran yang ada, seperti masalah biaya dan dampak lingkungan.

Di sisi lain, temuan mengenai sentimen negatif yang cukup signifikan menunjukkan bahwa ada sejumlah kekhawatiran yang perlu diatasi oleh pemerintah. Implikasi praktisnya adalah penting bagi pemerintah untuk mengidentifikasi dan merespon secara langsung isu-isu yang mengemuka, seperti biaya tinggi dan dampak lingkungan, melalui kebijakan atau program yang lebih transparan. Dengan memahami kekhawatiran tersebut, pemerintah dapat merancang kebijakan yang lebih responsif terhadap isu-isu masyarakat, serta melakukan

perbaikan dalam hal komunikasi dan pengelolaan proyek. Analisis sentimen ini memberi pemerintah cara yang efektif untuk mengevaluasi persepsi publik dan merumuskan keputusan kebijakan yang lebih tepat sasaran dan diterima oleh masyarakat.

4. KESIMPULAN

Berdasarkan hasil analisis sentimen terhadap tweet mengenai pemindahan ibu kota Indonesia, mayoritas opini publik di Twitter cenderung mendukung rencana tersebut. Sebanyak 71,5% tweet menunjukkan sentimen positif, mencerminkan harapan masyarakat terhadap pemerataan pembangunan dan pengurangan beban Jakarta. Di sisi lain, sekitar 28,5% tweet mengandung sentimen negatif, yang mencerminkan kekhawatiran terkait biaya tinggi, dampak lingkungan, dan potensi risiko sosial dari pemindahan ibu kota. Hasil ini menunjukkan pentingnya analisis sentimen dalam memahami respons publik terhadap kebijakan besar seperti pemindahan ibu kota. Penelitian ini juga memberikan wawasan yang berharga bagi pemerintah untuk merespons kekhawatiran masyarakat dan merumuskan kebijakan yang lebih responsif. Hasil yang diperoleh dengan menggunakan algoritma KNN menunjukkan akurasi yang cukup baik, yakni sebesar 77%.

DAFTAR PUSTAKA

- [1] P. Painem and A. I. E. Prayogo, "Penerapan Algoritma Naïve Bayes untuk Analisis Pandangan Masyarakat terhadap Pemindahan Ibu Kota Indonesia," *Informatik: Jurnal Ilmu Komputer*, vol. 18, no. 3, pp. 249-258, 2022.
- [2] G. A. Saputri and D. Alita, "Analisis Sentimen Twitter Terhadap Pemindahan Ibu Kota Negara Menggunakan Support Vector Machine," *Jurnal Informatika*, vol. 9, no. 3, 2024. DOI: <https://doi.org/10.30591/jpit.v9i3.6612>
- [3] S. D. Prasetyo, S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, vol. 10, no. 1, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [4] A. Siregar, "Analisis Sentimen Pindah Ibu Kota Negara (IKN) Baru pada Twitter Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," *Faktor Exacta*, vol. 16, no. 3, Oct. 2023, doi: 10.30998/faktorexacta.v16i3.16703.
- [5] H. D. Al Assyam and F. N. Hasan, "Analisis sentimen Twitter terhadap perpindahan ibu kota negara ke IKN nusantara menggunakan orange data mining," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, pp. 341-349, 2023. doi: 10.30865/KLIK.V4I1.957
- [6] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, pp. 1-7, 2023. DOI: 10.35134/komtekinfo.v10i1.330
- [7] M. I. Ramadhon, "Analisis sentimen pemindahan Ibu Kota Indonesia di media sosial twitter menggunakan metode algoritma K-Nearest Neighbor (K-NN)," Skripsi Sarjana, Fakultas Sains dan Teknologi, Universitas Islam Negeri Syarif Hidayatullah, Jakarta, 2020.
- [8] M. R. E. Rayhan, R. Rudiman, and F. Y. Fendy, "Perbandingan Metode K-Nearest Neighbor (KNN) dan Naïve Bayes terhadap Analisis Sentimen pada Pengguna E-Wallet Aplikasi Dana menggunakan Fitur Ekstraksi TF-IDF," *Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, vol. 18, no. 2, pp. 139-159, 2024.
- [9] D. Sandi, E. Utami, and K. Kusnawi, "Analisis Sentimen Publik Terhadap Elektabilitas Ganjar Pranowo di Tahun Politik 2024 di Twitter dengan Algoritma KNN dan Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 7, no. 3, pp. 1097-1108, 2023. DOI: <https://doi.org/10.30865/mib.v7i3.6298>
- [10] D. Muhidin and A. Wibowo, "Perbandingan kinerja algoritma Support Vector Machine dan K-Nearest Neighbor terhadap analisis sentimen kebijakan new normal," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 5, no. 2, pp. 153-159, 2020. DOI: <http://dx.doi.org/10.30998/string.v5i2.6715>
- [11] M. W. Pertiwi, "Analisis sentimen opini publik mengenai sarana dan transportasi mudik tahun 2019 pada twitter menggunakan algoritma naïve bayes, neural network, KNN dan SVM," *Inti Nusa Mandiri*, vol. 14, no. 1, pp. 27-32, 2019.
- [12] M. H. Asnawi, I. Firmansyah, R. Novian, and R. S. Pontoh, "Perbandingan algoritma naïve bayes, k-NN, dan SVM dalam pengklasifikasian sentimen media sosial," in *E-Prosiding Seminar Nasional Statistika, Departemen Statistika FMIPA Universitas Padjadjaran*, vol. 10, pp. 20-20, Dec. 2021. DOI: 10.1234/pns.v10i.85

-
- [13] S. Lestari, M. Mupaat, and A. Erfina, "Analisis Sentimen Masyarakat Indonesia terhadap Pemindahan Ibu Kota Negara Indonesia pada Twitter," *JUSIFO (Jurnal Sistem Informasi)*, vol. 8, no. 1, pp. 13-22, 2022. DOI: <https://doi.org/10.19109/jusifo.v8i1.12116>
- [14] A. Y. Rahman and F. Marisa, "Analysis Sentiment Pada Review Hotel Labuan Bajo Menggunakan Metode Bidirectional Long Short Term Memory (BI-LSTM)," in *National Conference on Electrical, Informatics and Industrial Technology (NEIIT)*, vol. 1, no. 1, Jul. 2024.
- [15] S. Lestari and S. Saepudin, "Analisis sentimen vaksin sinovac pada twitter menggunakan algoritma Naive Bayes," in *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*, vol. 1, pp. 163-170, Sep. 2021. DOI: <https://doi.org/10.35870/jtik.v6i4.530>
- [16] S. Rabbani, D. Safitri, N. Rahmadhani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153-160, 2023. DOI: <https://doi.org/10.57152/malcom.v3i2.897>
- [17] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based framework for text categorization," *Procedia Engineering*, vol. 69, pp. 1356-1364, 2014. DOI: 10.1016/j.proeng.2014.03.129
- [18] R. S. Amardita, A. Adiwijaya, and M. D. Purbolaksono, "Analisis Sentimen terhadap Ulasan Paris Van Java Resort Lifestyle Place di Kota Bandung Menggunakan Algoritma KNN," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 1, pp. 62-68, 2022. DOI: <https://doi.org/10.30865/mib.v7i4.6648>
- [19] A. J. Arifin and A. Nugroho, "Uji Akurasi Penggunaan Metode KNN dalam Analisis Sentimen Kenaikan Harga BBM pada Media Twitter," *Progresif: Jurnal Ilmiah Komputer*, vol. 19, no. 2, pp. 700-709, 2023.
- [20] D. Pramudita, Y. Akbar, and T. Wahyudi, "Analisis Sentimen Terhadap Program Kartu Indonesia Pintar Kuliah pada Media Sosial X Menggunakan Algoritma Naive Bayes," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1420-1430, 2024. DOI: <https://doi.org/10.57152/malcom.v4i4.1565>
- [21] M. N. Muttaqin and I. Kharisudin, "Analisis sentimen aplikasi gojek menggunakan support vector machine dan k nearest neighbor," *UNNES Journal of Mathematics*, pp. 22-27, 2021. <https://doi.org/10.21456/vol14iss3pp247-255>
- [22] P. Arsi, B. A. Kusuma, and A. Nurhakim, "Analisis Sentimen Pindah Ibu Kota Berbasis Naive Bayes Classifier," *Jurnal Informatika UPGRIS*, vol. 7, no. 1, 2021. DOI:10.26877/jiu.v7i1.7636.