

Analisis Sentimen Terhadap Pemindahan Ibu Kota Negara (IKN) Pada Platform Twitter Menggunakan Metode Naïve Bayes

Ira Dwi Nugrayani^{#1}, Moch. Hafid^{*2}, Dede Irmayanti^{#3}

Teknik Informatika Sekolah Tinggi Teknologi Wastukencana

Jl. Cikopak no 53 Sadang, Purwakarta, Jawa Barat

¹iradwi98@wastukencana.ac.id

²mhafid@wastukencana.ac.id

³dedeirmayanti@wastukencana.ac.id

Abstract—The relocation of the National Capital (IKN) has been a widely discussed topic on social media, particularly Twitter, with diverse public sentiments. This study analyzes public sentiment toward IKN relocation using the Naïve Bayes algorithm as a classification method. Data was collected through the Twitter API and processed through preprocessing stages, including filtering, tokenization, and transformation. Sentiments were classified into three categories—positive, negative, and neutral—and evaluated using a confusion matrix and the Orange software. The findings indicate that the majority of opinions are negative, with a classification accuracy of 91%. Negative sentiment is primarily driven by concerns regarding the budget, environmental impact, and government policy priorities. Meanwhile, positive sentiment reflects hopes for equitable development and reduced pressure on Jakarta. This study provides insights for the government in understanding public perception and serves as a reference for more effective communication strategies.

Key Words— Sentiment Analysis, Twitter, National Capital, Naïve Bayes, Preprocessing, Confusion Matrix

Abstrak— Pemindahan Ibu Kota Negara (IKN) menjadi topik diskusi luas di media sosial, khususnya Twitter, dengan sentimen publik yang beragam. Penelitian ini menganalisis sentimen masyarakat terhadap pemindahan IKN menggunakan algoritma Naïve Bayes sebagai metode klasifikasi. Data dikumpulkan melalui Twitter API dan diproses melalui preprocessing, meliputi filtering, tokenization, dan transformation. Sentimen diklasifikasikan ke dalam tiga kategori—positif, negatif, dan netral—dan dievaluasi menggunakan confusion matrix serta perangkat lunak Orange. Hasil penelitian menunjukkan mayoritas opini bersifat negatif, dengan akurasi klasifikasi mencapai 91%. Faktor utama sentimen negatif meliputi kekhawatiran terhadap anggaran, dampak lingkungan, dan prioritas kebijakan pemerintah. Sementara itu, sentimen positif didorong oleh harapan pemerataan pembangunan dan pengurangan beban Jakarta. Studi ini memberikan wawasan bagi pemerintah dalam memahami persepsi publik serta sebagai referensi untuk strategi komunikasi yang lebih efektif.

Kata kunci— Analisis Sentimen, Twitter, Ibu Kota Negara, Naïve Bayes, Preprocessing, Confusion Matrix

I. PENDAHULUAN

Perkembangan dunia teknologi informasi dan komunikasi yang pesat tidak lepas dari penyedia layanan web yang menyediakan informasi yang beragam. Informasi yang menyebabkan penambahan data yang kebanyakan berupa data teks dapat dijadikan sumber yang sangat potensial untuk digali lebih dalam. Salah satu contohnya adalah data teks yang diambil dari Twitter [1].

Sebagai media komunikasi masyarakat, jejaring sosial semakin populer digunakan. Salah satu jejaring sosial yang cukup populer yaitu Twitter. Twitter dapat dimanfaatkan sebagai sarana promosi produk, iklan, kampanye politik maupun sebagai sarana menyampaikan pendapat terkait kritik, saran, isu-isu dan opini-opini publik [2].

Dikutip dari Katadata.com (<https://katadata.co.id>). Twitter merupakan salah satu jejaring media sosial yang populer digunakan masyarakat di Indonesia. Menurut laporan Statista, terdapat 18,45 juta pengguna aplikasi yang didirikan oleh Jack Dorsey ini di tanah air per Januari 2022 lalu. Capaian ini menempatkan Indonesia sebagai negara pengguna Twitter terbanyak ke-5 di dunia.

Isu pemindahan IKN menjadi sorotan di media sosial Twitter, khususnya bagi pengguna di Indonesia. Pemindahan ibu kota di tengah pandemi Covid-19 memicu perdebatan, terutama karena banyak pengguna Twitter yang menolak kebijakan tersebut. Sebagian besar diskusi berpusat pada pentingnya pemerintah lebih memprioritaskan penanganan varian baru Covid-19, yang membutuhkan anggaran besar dari Anggaran Pendapatan dan Belanja Negara (APBN) serta Pemberdayaan Ekonomi Nasional (PEN). Isu ini menunjukkan adanya perbedaan pandangan di kalangan masyarakat terkait pemindahan ibu kota. Penulis mengangkat isu ini sebagai fokus penelitian karena relevansi dan perhatian luas dari berbagai media. Diharapkan hasil analisis ini dapat menjadi bahan pertimbangan bagi pemangku kepentingan dalam proses pengambilan keputusan terkait kebijakan pemindahan IKN.

Dalam konteks informatika, analisis sentimen adalah studi yang bertujuan untuk menganalisis opini, sentimen, dan emosi yang terdapat pada dokumen atau data. Tugas dasar dari analisis sentimen adalah untuk mengelompokkan sifat dari teks

yang ada di dalam kalimat maupun pendapat, biasanya terbagi menjadi 3 kelas yaitu negatif, positif dan netral [3].

Terdapat beberapa teknik yang digunakan dalam proses klasifikasi dokumen teks. Pertama, menentukan nilai kemungkinan kemunculan setiap kata dalam klasifikasi dokumen. Kedua, menghitung probabilitas kemunculan setiap kata pada masing-masing kategori dokumen. Ketiga, menetapkan kelas dokumen uji berdasarkan hasil perhitungan dari proses pertama dan kedua [4].

Penelitian ini membahas pemanfaatan opini pengguna Twitter terkait pemindahan ibu kota negara dalam analisis sentimen menggunakan algoritma Naïve Bayes. Opini tersebut kemudian diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Algoritma Naïve Bayes merupakan teknik prediksi probabilitas yang relatif sederhana. Metode ini bekerja dengan memperkirakan kemungkinan terjadinya suatu peristiwa di masa mendatang berdasarkan data historis yang tersedia. Selain itu, algoritma ini mampu mengolah data dalam jumlah besar dengan tingkat akurasi yang tinggi.

II. TINJAUAN PUSTAKA

A. Analisis Sentimen

Analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif atau negatif [4].

B. Pemindahan Ibu Kota Negara

Faktor-faktor yang menjadi alasan pemerintah dalam pemindahan ibu kota ke luar Jawa, yaitu terdapat sekitar 57% penduduk Indonesia terkonsentrasi di Pulau Jawa, kontribusi ekonomi per pulau terhadap Produk Domestik Bruto (PDB) Nasional, krisis ketersediaan air di Pulau Jawa terutama DKI Jakarta dan Jawa Timur. Konversi lahan terbesar terjadi di pulau Jawa, pertumbuhan urbanisasi yang sangat tinggi dengan konsentrasi penduduk terbesar di Jakarta dan Jabodetabekpunjur, dan meningkatnya beban DKI Jakarta [5].

Pemindahan ibu kota memiliki potensi untuk meningkatkan pertumbuhan ekonomi nasional dengan tingkat inflasi yang tepat rendah. Selain itu, pertumbuhan wilayah industri dengan rantai nilai hingga hilir akan meningkatkan pendapatan masyarakat yang pada gilirannya akan menurunkan ketimpangan ekonomi antara pulau Jawa dan pulau lain [6].

C. Twitter

Twitter adalah *microblogging platform* yang sangat populer. Layanan jejaring sosial ini memungkinkan pengguna untuk membagikan 240 karakter melalui suatu *tweet*. Pada pertengahan tahun 2015, jumlah pengguna *Twitter* di Indonesia mencapai lima puluh juta pengguna. Hal ini menjadikan Indonesia sebagai negara dengan pengguna aktif *Twitter* terbanyak [7].

D. Algoritma

Algoritma merupakan rangkaian langkah sistematis dan logis yang digunakan untuk menyelesaikan suatu masalah.

Kata 'logis' menjadi kata kunci dalam algoritma, karena setiap langkah dalam algoritma harus bersifat logis serta dapat ditentukan nilai kebenarannya, yaitu benar atau salah [8].

E. Naïve Bayes

Naïve Bayes merupakan pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas [9]. Rumus teori dari algoritma Naïve Bayes di bawah ini [1].

$$P(X|H) = \frac{P(X|H) \times P(H)}{P(X)}$$

F. Text Mining

Text Mining adalah ilmu yang bertujuan untuk memproses teks agar menjadi informasi, menambang suatu data yang berupa teks yang bersumber dari data tersebut [10]. Data yang biasanya diperoleh dari dokumen dan digunakan untuk mencari kata-kata yang dapat mewakili isi dari dokumen tersebut [11].

G. Confusion Matrix

Confusion matrix merupakan salah satu metode yang digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya, *confusion matrix* berisi informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. Dalam pengukuran kinerja menggunakan *confusion matrix*, terdapat empat istilah utama sebagai hasil proses klasifikasi, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). *True Negative* (TN) menunjukkan jumlah data negatif yang terdeteksi dengan benar, sedangkan *False Positive* (FP) menunjukkan data negatif yang keliru terdeteksi sebagai data positif [12]. Terdapat 3 rumus yang digunakan dalam perhitungan *confusion matrix* di bawah ini [13].

$$(1) \text{ accuracy} = \frac{TP + TN}{\text{Total}}$$

$$(2) \text{ precision} = \frac{TP}{TP + FP}$$

$$(3) \text{ recall} = \frac{TP}{TP + FN}$$

H. Preprocessing

Preprocessing merupakan tahapan sebelum proses klasifikasi yang bertujuan untuk membersihkan, menghilangkan, dan mengubah sumber data, baik berupa karakter non-alfabet maupun kata-kata yang tidak diperlukan, sehingga data yang digunakan lebih optimal dalam proses klasifikasi. Tahapan *preprocessing* meliputi *transformation*, yaitu proses mengubah data input dengan menerapkan transformasi huruf kecil secara default; *tokenization*, yakni metode untuk memecah teks menjadi komponen yang lebih kecil, seperti kata, kalimat, atau bigram; serta *filtering*, yaitu

proses penyaringan untuk menghapus atau memilih kata-kata serta simbol yang tidak diperlukan dalam tahap berikutnya [14] [15].

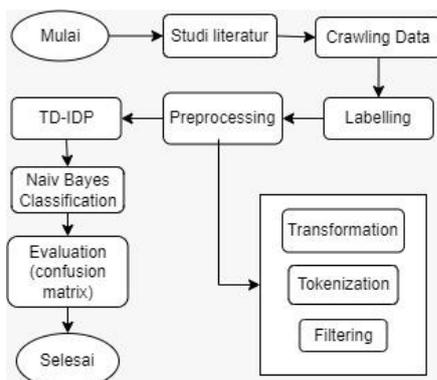
I. Term Frequency Inverse Document Frequency

Term Frequency-Inverse Document Frequency (TF-IDF) adalah sebuah ukuran statistik yang digunakan untuk mengevaluasi tingkat kepentingan suatu kata dalam sebuah dokumen maupun dalam sekumpulan data. Metode TF-IDF berfungsi untuk menghitung bobot setiap kata dan merupakan salah satu teknik yang paling umum digunakan dalam information retrieval [16]. Berikut merupakan rumus TF-IDF:

$$IDF = \text{Log} \frac{tf}{df}$$

III. METODOLOGI PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang dirancang untuk memastikan kelancaran proses dari awal hingga akhir guna mencapai tujuan yang telah ditetapkan. Alur penelitian dapat dilihat pada Gambar 1.



Gambar 1 Alur Penelitian

A. Pengumpulan Data (Crawling)

Dataset yang digunakan dalam penelitian ini diperoleh melalui proses pencarian data. Data yang dikumpulkan berasal dari tweet di media sosial Twitter, sementara proses pengumpulan data dilakukan menggunakan Twitter Application Programming Interface (API). Tweet yang telah diperoleh kemudian disimpan dalam format Excel

B. Pelabelan (Labelling)

Labelling merupakan kondisi ketika seseorang mendapatkan julukan dari orang lain berdasarkan perilakunya. Setelah diberikan, label tersebut menjadi identitas diri yang menggambarkan individu tersebut. Label biasanya diberikan kepada orang-orang yang terlihat berbeda dari yang lain. Dalam proses labelling, data yang telah diproses sebelumnya akan melalui tahap pelabelan. Proses ini dilakukan dalam dua tahap, yaitu pertama menggunakan perangkat lunak Orange, dan kedua menggunakan Microsoft Excel dengan rumus IF (A<0; 'negatif'; IF(A>0; 'positif'; IF(A=0; 'netral'))). Tahap kedua diperlukan karena emosi pada tahap pertama masih dalam bentuk digital dan tidak dapat digunakan sebagai label dalam proses selanjutnya.

C. Preprocessing

Preprocessing teks bertujuan untuk membagi teks menjadi unit-unit yang lebih kecil melalui tahapan seperti transformasi, tokenisasi, dan penyaringan (filtering). Langkah-langkah analisis diterapkan secara berurutan serta dapat diaktifkan atau dinonaktifkan sesuai kebutuhan. Preprocessing data melalui beberapa tahapan, yaitu: (1) Transformation, yaitu proses mengubah data input, meliputi perubahan ke huruf kecil (lowercase), penghapusan aksen (remove accents), parsing, dan penghapusan URL (remove URL); (2) Tokenizing, yakni metode untuk memecah teks menjadi komponen yang lebih kecil dengan membagi teks kata per kata dan mempertahankan simbol tanda baca; serta (3) Filtering, yaitu proses penyaringan untuk menghapus kata-kata yang tidak diperlukan dalam analisis selanjutnya, seperti "dan", "yang", "di", "adalah", serta retweet. Contoh preprocessing dapat dilihat pada Tabel 1.

TABLE 1. ILUSTRASI PREPROCESSING

Preprocessing	Sebelum	Sesudah
Transformation	Naïve	naïve
	<a href.> beberapa teks 	beberapa teks
	sebuah url http://orange.biolab.si/.	sebuah url
Tokenizing	This example	This, example
Filtering	pesan yang dimuat adalah	pesan dimuat

D. Pembobotan Kata (TF-IDF)

Setelah tahap preprocessing, yang mencakup transformasi, tokenisasi, dan penyaringan, proses selanjutnya adalah term weighting, yaitu pemberian bobot pada istilah dalam dokumen sesuai tingkat kepentingannya. Dalam metode TF-IDF, terdapat Term Frequency (TF), yaitu frekuensi kemunculan suatu istilah dalam dokumen tertentu. Selain itu, terdapat Inverse Document Frequency (IDF), yang berfungsi untuk mengurangi dominasi istilah yang sering muncul di berbagai dokumen dengan memperhitungkan kebalikan frekuensi dokumen yang mengandung kata tersebut. TF-IDF merupakan algoritma yang menghitung bobot istilah dengan mengalikan nilai Term Frequency dengan Inverse Document Frequency.

E. Naïve Bayes Classification

Dokumen yang mengandung kategori tertentu, kata-kata dalam dokumen tersebut dapat dilakukan pengklasifikasian, kumpulan dari kata pada dokumen tersebut dapat digunakan untuk menentukan kategori karena memiliki nilai makna tertentu. Metode Naïve Bayes digunakan untuk menentukan sentimen dari dokumen tweet.

F. Evaluasi

Evaluasi pada penelitian ini menggunakan confusion matrix (pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih). Evaluasi dilakukan untuk mengetahui nilai performansi dari sistem yang sudah dibuat berdasarkan hasil dari klasifikasi.

IV. HASIL DAN PEMBAHASAN

A. Menghitung Klasifikasi Naïve Bayes

Berikut adalah contoh data yang akan digunakan, terdapat 9 data tweet yang ditandai dengan 3 positif, 3 negatif, dan 3 netral. Data dapat dilihat pada Tabel 2.

TABLE 2. DATA SAMPLE

Dokumen Tweet	Sentimen
Indonesia maju pembangunan IKN	Positif
IKN pindah kalimantan Indonesia	Positif
proyek IKN nusantara dampak positif Indonesia	Positif
pembangunan IKN nusantara lupakan kultur budaya Kalimantan	Negatif
IKN nusantara kalimantan kota hijau	Negatif
gausah bahas IKN rakyat susah	Negatif
pemindahan IKN pemerata pembangunan	Netral
IKN nusantara kalimantan menunjang ekonomi	Netral
wajah IKN nusantara	Netral

Dari Tabel 2 dilakukan perhitungan dengan menggunakan microsoft excel yang pertama dilakukan adalah menentukan *Term Frequency*, dapat dilihat pada Tabel 3.

TABLE 3. TERM FREQUENCY

Dokumen	Term	Term / Kata							
1	Indonesia	1	1	1	1	0	0	0	0
2	maju	1	0	0	1	1	1	0	0
3	pembangunan	1	0	0	1	0	0	1	1
4	IKN	0	0	1	1	0	0	0	1
5	pindah	0	0	0	1	0	1	0	1
6	kalimantan	0	0	0	1	0	0	0	0
7	proyek	0	0	1	1	0	0	0	0
8	nusantara	0	0	0	1	0	1	0	1
9	jawa	0	0	0	1	0	0	0	1

Dari Tabel 3. Selanjutnya dilakukan perhitungan jumlah kata unik yang muncul dan jumlah kata pada setiap kategori. Dapat dilihat pada Tabel 4.

TABLE 4. KATEGORISASI

Kategori	Jumlah dokumen per Kategori	Jumlah dokumen	Probabilitas (Kategori)	Jumlah kata per Kategori
Positif	3	9	0,33	14
Negatif	3	9	0,33	16
Netral	3	9	0,33	12
Jumlah Kata Unik		24		

Kata unik merupakan kata yang sering muncul di setiap dokumen. Setelah mendapatkan jumlah kata unik dan jumlah kata pada setiap kategori, yang dilakukan selanjutnya adalah menghitung probabilitas pada setiap kategori. Berikut rumus menghitung probabilitas:

$$P(kata) = \frac{\text{Jumlah satu kata} + 1}{\text{jml kata kategori} + \text{jml keseluruhan}}$$

Dari tabel TF di atas didapatkan hasil jumlah kata, dalam kategori positif 14 kata, negatif 16 kata dan netral 12 kata dengan jumlah kata unik sebanyak 24 kata. Cara menghitung probabilitas penulis akan mencontohkan untuk kata IKN dimana 1 tambah 3 lalu dibagi 14 yang didapat dari jumlah setiap kategori positif lalu di tambah 24 yang didapat dari jumlah kata unik.

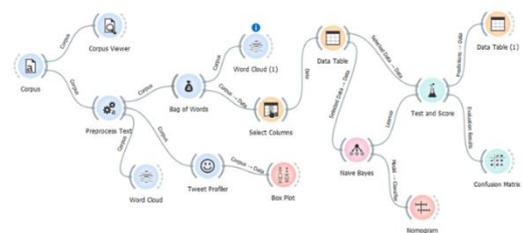
$$(ikn | positif) = \frac{3 + 1}{14 + 24} = \frac{4}{38} = 0,105$$

$$(ikn | negatif) = \frac{3 + 1}{16 + 24} = \frac{4}{60} = 0,067$$

$$(ikn | netral) = \frac{3 + 1}{12 + 24} = \frac{4}{36} = 0,111$$

B. Pengujian Naïve Bayes pada *Tool Orange*

Tool Orange digunakan untuk memodelkan analisis sentimen. Dimulai dari penyiapan data, corpus, *preprocessing*, *bag of words*, algoritma *Naïve Bayes*, dan validasi *test score*. Skema pengujian dapat dilihat pada Gambar 2.



Gambar 2. Skema Analisis Sentimen

Hasil klasifikasi Algoritma *Naïve Bayes*, perhitungan *accuracy*, *precision*, *recall*, dan *f1-score* digunakan untuk membuktikan hasil dari performa algoritma *Naïve Bayes*. Klasifikasi rata-rata untuk seluruh Kelas Sentimen masing-masing adalah *accuracy* 91%, *precision* 91%, *recall* 91%, dan *f1-score* 91%. Detail tiap kelas dapat dilihat pada Tabel 4.

TABLE 4. DETAIL HASIL KLASIFIKASI

Kategori	Accuracy	Precision	Recall	f1-Score
Positif	0.93	0.93	0.89	0.91
Negatif	0.95	0.85	0.91	0.88
Netral	0.93	0.92	0.92	0.92

C. Evaluasi

Evaluasi dilakukan untuk memverifikasi keakuratan hasil klasifikasi. Pada tahap ini, digunakan *confusion matrix* untuk menghitung nilai akurasi, presisi, dan *recall*. Hasil *confusion matrix* dari pengujian terhadap keseluruhan data *tweet* dapat dilihat pada Tabel 5.

TABLE 5. CONFUSION MATRIX

	Negatif	Netral	Positif
Negatif	140	10	3
Netral	7	272	16
Positif	17	12	256

Berikut adalah perhitungan manual akurasi:

$$Accuracy = \frac{TP + TNetral + TN}{\text{Jumlah Data}} = \frac{256 + 272 + 140}{733} = \frac{668}{733} = 0,91132333$$

Berikut adalah perhitungan manual *precision*:

TABLE 6. CONFUSION MATRIX

	Negatif	Netral	Positif
TP	140	272	256
FP	10 + 3	7 + 16	17 + 12
Precision	140/(140+13)	272/(272+23)	256/(256+29)
	0,915	0,922	0,898

$$All\ Precision = \frac{Precision\ A + B + C}{Jumlah\ Kelas}$$

$$All\ Precision = \frac{0,915 + 0,922 + 0,898}{3} = \frac{2,735}{3} = 0,911$$

Berikut adalah perhitungan manual *recall*:

TABLE 7. CONFUSION MATRIX

	Negatif	Netral	Positif
TP	140	272	256
FP	7 + 17	10 + 12	3 + 16
Recall	140/(140+24)	272/(272+22)	256/(256+19)
	0,854	0,918	0,931

$$Recall = \frac{Recall\ A + B + C}{Jumlah\ Kelas}$$

$$Recall = \frac{0,854 + 0,918 + 0,931}{3} = \frac{2,703}{3} = 0,901$$

D. Visualisasi

Visualisasi adalah suatu rekayasa dalam pembuatan gambar, diagram atau animasi untuk penampilan suatu informasi. Visualisasi digambarkan dengan *word cloud* pada sentimen keseluruhan dokumen negatif positif dan netral tentang Pindahan IKN. Hasil dari Visualisasi ditampilkan pada Gambar 3 dan Gambar 4.

Weight	Word
767	ikn
264	indonesia
208	sambut
208	g20
166	ayoindonesiapastipulih
165	pembangunan
150	nusantara
102	kota
68	yg
56	pemindahan
56	dukung
54	negara
51	pemerintah
47	hijau
46	kalimantan

Gambar 3. Frekwensi kemunculan kata Pindahan IKN



Gambar 4. Word cloud Pindahan IKN

Pada Gambar 3 menjelaskan yang mana kata yang berukuran paling besar dan berada di tengah adalah kata yang memiliki frekuensi tertinggi word cloud seperti IKN dan Indonesia.

V. KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen terhadap pemindahan IKN di platform *Twitter*, sebanyak 733 data telah melewati tahap *preprocessing* yang mencakup *filtering*, *tokenization*, dan *transformation*. Selanjutnya, data diklasifikasikan menggunakan algoritma *Naive Bayes* dan dievaluasi dengan *confusion matrix* menggunakan perangkat lunak *Orange*. Hasil penelitian menunjukkan bahwa opini masyarakat mengenai pemindahan IKN cenderung negatif, dengan akurasi sebesar 95%, presisi 85%, dan tingkat keberhasilan (*recall*) 91%. Nilai ini mengindikasikan bahwa algoritma *Naive Bayes* memiliki kinerja yang cukup baik dalam memproses dokumen *Twitter*, mengingat persentase akurasi melebihi 70%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa sentimen masyarakat terhadap pemindahan IKN di platform *Twitter* tergolong negatif.

REFERENSI

- [1] F. Ratnawati, "Implementasi algoritma naive bayes terhadap analisis sentimen opini film pada twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018.
- [2] N. Hardi, Y. Alkahfi, P. Handayani, W. Gata, and M. R. Firdaus, "Analisis sentimen physical distancing pada twitter menggunakan text mining dengan algoritma naive bayes classifier," *Sistemasi*, vol. 10, no. 1, p. 131, 2021.
- [3] S. Afrizal, H. N. Irmanda, N. Falih, and I. N. Isnainiyah, "Implementasi metode Naive Bayes untuk analisis sentimen warga Jakarta terhadap mass rapid transit," *Inform. J. Ilmu Komput.*, vol. 15, no. 3, p. 157, 2020.
- [4] H. Hermanto and A. Noviriandini, "Analisa sentimen terhadap belajar online pada masa Covid-19 menggunakan algoritma support vector machine berbasis particle swarm optimization," *J. Inform. Kaputama*, vol. 5, no. 1, pp. 129–136, 2021.
- [5] N. S. Wardani, A. Prahutama, and P. Kartikasari, "Analisis sentimen pemindahan ibu kota negara dengan klasifikasi naive bayes untuk model Bernoulli dan multinomial," *J. Gaussian*, vol. 9, no. 3, pp. 237–246, 2020.
- [6] S. A. F. Silalahi, "Dampak ekonomi dan risiko pemindahan ibu kota negara," *Info Singk. Bid. Ekon. Dan Kebijak. Publik*, vol. 9, no. 16, pp. 19–24, 2019.
- [7] E. Y. Hidayat, R. W. Hardiansyah, and A. Affandy, "Analisis sentimen twitter untuk menilai opini terhadap perusahaan publik menggunakan algoritma deep neural network," *J. Nas. Teknol. dan Sist. Inf.*, vol. 7, no. 2, pp. 108–118, 2021.
- [8] D. Numaningsih and A. A. Permana, "Rancangan aplikasi pengamanan data dengan algoritma advanced encryption standard (AES)," *J. Tek. Inform.*, vol. 11, no. 2, pp. 177–186, 2018.
- [9] M. F. Rifai, H. Jatnika, and B. Valentino, "Penerapan algoritma Naive Bayes pada sistem prediksi tingkat kelulusan peserta sertifikasi Microsoft Office Specialist (MOS)," *Petir*, vol. 12, no. 2,

- pp. 131–144, 2019.
- [10] W. Athira Luqyana, I. Cholissodin, and R. S. Perdana, “Analisis sentimen cyberbullying pada komentar Instagram dengan metode klasifikasi support vector machine,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4704–4713, 2018.
- [11] R. K. Mulia, “Text mining analysis dan sentiment analysis dengan menggunakan metode naïve bayes classifier,” 2021.
- [12] E. Prasetyo, “Mengukur kinerja algoritma klasifikasi dengan confusion matrix,” *Achmatim.net*, 2017. .
- [13] D. Normawati and S. A. Prayogi, “Implementasi naïve bayes classifier dan confusion matrix pada analisis sentimen berbasis teks pada twitter,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [14] F. A. Muttaqin and A. M. Bachtiar, “Implementasi teks mining pada aplikasi pengawasan penggunaan internet anak ‘Dodo Kids browser,’” *J. Ilm. Komput. dan Inform.*, pp. 1–8, 2016.
- [15] R. A. raffaidy Wiguna and A. I. Rifai, “Analisis text clustering masyarakat di twitter mengenai Omnibus Law menggunakan Orange data mining,” *J. Inf. Syst. Informatics*, vol. 3, no. 1, pp. 1–12, 2021.
- [16] V. Amrizal, “Penerapan metode term frequency inverse document frequency (TF-IDF) dan cosine similarity pada sistem temu kembali informasi untuk mengetahui syarah hadits berbasis web (Studi kasus: Hadits Shahih Bukhari-Muslim),” *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, 2018.