

ANALISIS SENTIMEN KOMENTAR NETIZEN TERHADAP VIDEO PEMINDAHAN DAN PEMBANGUNAN IBU KOTA NUSANTARA DI TIKTOK MENGGUNAKAN METODE *NAÏVE BAYES*

Adrian Sagostian¹, Laurensius Yuda², Yandi³

¹²³Program Studi Teknik Informatika, Universitas Widya Dharma Pontianak
 Jl. HOS Cokroaminoto, Pontianak – Kalimantan Barat, Indonesia

Email: ¹adriansagostian05@gmail.com, ²laurensiusyuda11@gmail.com, ³yandi23421721@gmail.com

ABSTRAK

Pemindahan Ibu Kota Negara (IKN) ke Kalimantan Timur, yang dikenal sebagai Ibu Kota Nusantara, telah menjadi topik diskusi yang luas di berbagai platform media sosial. TikTok, sebagai salah satu *platform* berbagai video yang populer di Indonesia, menjadi medium bagi Masyarakat untuk mengespresikan opini mereka terkait program Pembangunan IKN. Penelitian ini dilakukan menggunakan Klasifikasi Sentimen dengan menerapkan algoritma Naive Bayes pada komentar TikTok. *Dataset* yang digunakan terdiri dari 1.810 komentar berbahasa Indonesia yang dikumpulkan melalui Teknik *web scraping* dari video TikTok yang membahas pemindahan IKN. Setiap komentar dikategorikan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Proses analisis mencakup tahapan pra-pemrosesan teks, seperti labeling, pembersihan data, tokenisasi, dan stemming, sebelum dilakukan pelatihan model klasifikasi menggunakan algoritma Naïve Bayes. Hasil penelitian ini menunjukkan bahwa mayoritas komentar bersentimen negatif. Model Naïve Bayes yang digunakan menghasilkan tingkat akurasi sebesar 65%. Penelitian ini diharapkan menjadi referensi dalam memahami opini publik secara *real-time* serta dalam pengembangan model klasifikasi sentimen yang lebih akurat pada *platform* media sosial.

Kata kunci: Analisis Sentimen, IKN, komentar netizen, Naïve Bayes, TikTok

ABSTRACT

The relocation of the National Capital City (IKN) to East Kalimantan, known as the Capital City of the archipelago, has become a topic of widespread discussion on various social media platforms. TikTok, as one of the popular video platforms in Indonesia, is a medium for the public to express their opinions regarding the IKN Development program. This research was conducted using Sentiment Classification by applying the Naive Bayes algorithm to TikTok comments. The dataset used consisted of 1,810 comments in Indonesian collected through web scraping techniques from TikTok videos discussing the transfer of the IKN. Each comment is categorized into three sentiment classes, which are positive, negative, and neutral. The analysis process includes pre-processing stages of text, such as labeling, data cleaning, tokenization, and stemming, before training the classification model using the Naïve Bayes algorithm. The results of this study show that the majority of comments have negative sentiments. The Naïve Bayes model used produces an accuracy rate of 65%. This research is expected to be a reference in understanding public opinion in real-time as well as in developing a more accurate sentiment classification model on social media platforms.

Keywords: Sentiment Analysis, IKN, netizens comments, Naïve Bayes, TikTok

1. PENDAHULUAN

Latar Belakang

Pemindahan Ibu Kota Negara (IKN) Nusantara sudah melalui perjalanan yang sangat panjang serta banyak sekali dinamika-dinamika dalam pemindahan IKN Nusantara tersebut. Hal ini juga merupakan berita/isu nasional yang sangat kontroversial. Pada tanggal 29 Agustus 2019 Presiden Joko Widodo mengumumkan untuk pemindahan Ibu Kota Baru ke Nusantara dengan wilayah administratif yang mencakup Kabupaten Penajam Paser Utara dan Kabupaten Kutai Kartanegara di Provinsi Kalimantan Timur. Pemindahan IKN ke Nusantara ini mempunyai alasan tersendiri, yaitu padatnya penduduk di Pulau Jawa. Padatnya penduduk di Pulau Jawa inilah yang menjadi alasan mengapa Presiden Joko Widodo memindahkan ibu kota negara ke Nusantara tepatnya di Kalimantan Timur [1]. Pembangunan IKN dimulai pada tahun 2022 dengan target pemindahan yaitu pada tahun 2024 hingga 2045. Pemindahan Ibu Kota Negara Indonesia ini menuai opini pro dan kontra dari masyarakat [2]. Hal ini banyak disampaikan oleh masyarakat pada salah satu *platform* media sosial yang sangat populer yaitu tiktok. Dalam video pemindahan ibu kota yang diunggah oleh akun Takasi.info, komentar yang disampaikan oleh



para pengguna TikTok pada kolom komentarnya berjumlah 7804. Banyaknya komentar ini menjadi penting untuk diteliti karena mencerminkan persepsi netizen secara langsung dan *real time*, sekaligus menjadi sumber data yang relevan untuk memahami dukungan maupun penolakan terhadap pembangunan IKN Nusantara tersebut. Sehingga sering digunakan untuk melakukan penilaian yang merujuk pada suatu topik tertentu yang dikenal dengan analisis sentimen.

Analisis sentimen merupakan salah satu cabang dari *Text mining*, yang dimana pada proses ini menganalisa dokumen atau sumber teks untuk mengelompokkan suatu opini atau tulisan ke dalam opini positif, netral maupun negatif secara otomatis [3]. *Text mining* dapat didefinisikan sebagai suatu proses untuk mengekstrak informasi yang berguna dari suatu sumber data melalui identifikasi dan eksplorasi pola tertentu [4]. Cara kerja analisis sentimen dalam mengambil data dibagi menjadi tiga langkah, yaitu klasifikasi, evaluasi, dan visualisasi hasil. Dengan cara ini, kita bisa mengetahui apakah masyarakat mendukung, menolak, atau netral terhadap kebijakan pemindahan ibu kota tersebut [5].

Salah satu metode yang digunakan untuk analisis sentimen komentar tentang pemindahan Ibu Kota Negara (IKN) Nusantara adalah *Naïve Bayes* yang merupakan salah satu metode klasifikasi. Klasifikasi tersebut akan mengetahui apakah komentar video pemindahan dan pembabangan IKN Nusantara tersebut bersifat positif, netral, atau negatif [6]. Metode *Naïve Bayes* dipilih karena merupakan metode klasifikasi probabilistik yang terkenal [7]. Metode ini juga merupakan metode yang sederhana, namun efisien dengan beragam aplikasi dunia nyata, khususnya klasifikasi teks [8].

Pada penelitian sebelumnya terkait analisis sentimen dengan menggunakan metode *Naïve Bayes* mendapatkan nilai akurasi sebesar 87% dengan dominan komentar positif. Penelitian ini juga menghasilkan nilai *Precision* sebesar 91%, *Recall* 97% dan nilai F-Measure 93% [3]. Penelitian lainnya mendapatkan proses analisis Sentimen dengan menggunakan metode *Naïve Bayes* dan menghasilkan nilai akurasi sebesar 98.80% dari 585 data uji [9].

Penelitian ini bertujuan untuk menganalisis komentar-komentar netizen di *platform* TikTok terkait video pemindahan dan pembangunan IKN, apakah tanggapan para netizen positif, negatif, atau netral. Selain itu juga menguji efektivitas algoritma *Naïve Bayes* dalam menganalisis sentimen data teks komentar terkait isu nasional. Dengan demikian dapat diketahui seberapa besar dukungan atau penolakan masyarakat terhadap pemabangunan IKN Nusantar di Kalimantan Timur. Selain itu, analisis serupa lebih banyak dilakukan di Twitter, Instagram, atau Youtube, sehingga penelitian pada kali memberikan gambaran terkait keakuratan *Naïve Bayes* dalam menganalisis sentimen di TikTok. Penggunaan algoritma *Naïve Bayes* juga dapat membantu dalam merspon dinamika opini secara *real-time*.

2. METODE PENELITIAN

Data

Data ini dimulai dengan pengumpulan data komentar penonton diperoleh menggunakan aplikasi TikTok sebanyak 1.810 baris komentar dari 1 buah video. Judul video dan dan nama *channel* TikTok yang dipaparkan melalui Tabel 1 data komentar penonton.

Tabel 1. Daftar video sumber data

No	Channel TikTok	Judul Video	Tanggal Unggah	Jumlah Komentar
1	Takasi.info	Habiskan Rp 2T, patung Garuda IKN Berubah Warna Otomatis Dari Mencekam Jadi Elegan	28 April 2025	7804

Tahap Analisis Data

Tahapan Analisis Data Dalam pembuatan model klasifikasi, metode yang digunakan yakni *Naïve Bayes*. Pada data, dilakukan *pre-processing* (*drop duplicates*, *case folding*, *cleaning*, *convert slangword*, *tokenizing*, *stemming*, *stopword removal*, dan *drop NA*) [10], dan pelabelan data yang dilakukan secara manual.

Pre-Processing

Proses *pre-processing* merupakan tahapan memperbaiki data teks yang tidak terstruktur menjadi data yang dapat dianalisis. Tahapan *pre-processing* pada data adalah sebagai berikut.

- Drop Duplicates* untuk menghilangkan duplikat komentar yang sama dari *user* yang sama.
- Case Folding* untuk mengubah semua teks yang mengandung huruf kapital menjadi huruf kecil.
- Cleaning-Remove Unnecessary Character* untuk menghilangkan *username* yang dimulai dengan *mention*, karakter simbol dan URL.
- Cleaning-Remove Emoji* untuk menghilangkan emoji.
- Cleaning-Remove Non-Alpha Numeric* untuk menghapus tanda baca.
- Convert Slangwords* untuk mengubah kata slang menjadi kata baku menurut Kamus Besar Bahasa Indonesia.
- Tokenizing* untuk memecah teks menjadi kata per kata.
- Stemming* untuk menghilangkan kata imbuhan sehingga diperoleh kata dasar.

- i. *Stopwords Removal* untuk menghilangkan kata-kata yang tidak penting dalam teks dengan menghapus kata pada komentar yang terdapat dalam daftar *stopwords*.
- j. *Drop NA* untuk menghilangkan baris komentar kosong.

Pelabelan Data

Setelah melalui tahap *preprocessing*, kemudian dilakukan pelabelan data. Pelabelan data teks terbagi menjadi 3 kelas sentimen yaitu Positif, Negatif, Netral. Pelabelan ini dilakukan secara manual.

Pembobotan kata TF-IDF

Dataset yang sudah di *preprocessing* sebelumnya kemudian diproses kembali ke dalam bentuk format bilangan biner sehingga dapat dikenal oleh sistem.

Klasifikasi Naïve Bayes

Pada tahapan ini akan dilakukannya proses klasifikasi dengan algoritma yang digunakan. Adapun algoritma yang digunakan adalah algoritma Naïve Bayes. Pada tahap ini data yang sudah melewati proses *crawling* dan *labeling* akan diproses menggunakan algoritma yang dipilih yaitu Naïve Bayes.

Naïve Bayes memiliki bentuk umum yaitu:

$$P(C|X) = \frac{P(X|C) \cdot P(c)}{P(x)}$$

Keterangan:

- X : Data dengan *class* yang belum diketahui
- C : Hipotesis data yang merupakan suatu *class* spesifik
- P(C|X) : Probabilitas hipotesis C berdasarkan kondisi (*posterior probability*)
- P(c) : Probabilitas hipotesis C (*prior probability*)
- P(x|c) : Probabilitas berdasarkan kondisi pada hipotesis
- P(x) : Probabilitas c

3. HASIL DAN PEMBAHASAN

Pre-processing Data

Data komentar TikTok tentang pemindahan ibu kota negara yang berjumlah sebesar 7804 komentar memasuki tahap *pre-processing* terlebih dahulu untuk menyeragamkan isi dari data yang diperoleh.

Tabel 1. Hasil *pre-processing*

Tahap	Komentar
Kalimat awal	@HAFIZ PAKAI Z:sekelas nyoman nuarta di kritik karyanya sama orang yang cuman bisa gambar gunung 😊
<i>Case Folding</i>	@hafiz pakai z:sekelas nyoman nuarta di kritik karyanya sama orang yang cuman bisa gambar gunung 😊
<i>Cleaning</i>	hafiz pakai z sekelas nyoman nuarta di kritik karyanya sama orang yang cuman bisa gambar gunung
<i>Convert Slangwords</i>	hafiz pakai z sekelas nyoman nuarta dikritik karyanya oleh orang yang hanya bisa menggambar gunung
<i>Tokenizing</i>	['hafiz', 'pakai', 'z', 'sekelas', 'nyoman', 'nuarta', 'dikritik', 'karyanya', 'oleh', 'orang', 'yang', 'hanya', 'bisa', 'menggambar', 'gunung']
<i>Ste Netrall mming</i>	['hafiz', 'pakai', 'z', 'kelas', 'nyoman', 'nuarta', 'kritik', 'karya', 'oleh', 'orang', 'yang', 'hanya', 'bisa', 'gambar', 'gunung']
<i>Removal Stopwords</i>	['hafiz', 'pakai', 'z', 'kelas', 'nyoman', 'nuarta', 'kritik', 'karya', 'orang', 'bisa', 'gambar', 'gunung']

Pelabelan Data

Setelah melalui tahap *pre-processing*, kemudian dilakukan pelabelan data. Pelabelan data teks terbagi menjadi 3 kelas sentimen yaitu Positif, Negatif, Netral. Pelabelan ini dilakukan secara manual dipilih agar besar dan terang, memudahkan proses pembacaan. Ikuti dan sesuaikan pada model tabel berikut:

Tabel 2. Distribusi Sentimen

No	Sentimen	Jumlah Komentar	Persentase
1	Positif	426	23.54%
2	Netral	354	19.56%
3	Negatif	1030	56.91%

Berdasarkan Tabel 2 diketahui bahwa hasil dari *pre-processing* dan pelabelan kelas sentimen kelas negatif terbanyak yaitu sebesar 1030 komentar, sedangkan kelas sentimen positif sebanyak 426 komentar dan sentimen netral sebanyak 354 komentar dari total komentar sebanyak 1810 komentar.

Pembobotan kata TF-IDF

Sebelum proses klasifikasi dilakukan, *Term Frequency-Inverse Document Frequency* (TF-IDF) digunakan untuk pembobotan kata. Metode TF-IDF menghitung seberapa penting suatu kata dalam sebuah dokumen relatif terhadap keseluruhan koleksi dokumen.

$$\text{IDF}(t) = \log\left(\frac{N}{\text{DF}(t)}\right)$$

Tabel 3. Hasil pembobotan

No	Kata	DF	IDF	TF	TF-IDF
1	Proyek	1	0.954	0.1	0.0954
2	PHK	1	0.954	0.2	0.1908
3	Garuda	1	0.954	0.33	0.3178
4	Keren	1	0.954	0.5	0.477
5	Bangun	1	0.954	0.1	0.0954

Pengujian hasil

Setelah melakukan proses pembobotan kata, maka dilakukan tahap pengujian untuk melihat *precision*, *recall*, *F1-score*, dan *support*-nya.

Tabel 4. Detail laporan klasifikasi

No	Label	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
1	Negatif	0.62	1.00	0.77	209
2	Netral	0.33	0.02	0.03	66
3	Netral	0.00	0.00	0.00	1
4	Positif	0.96	0.30	0.46	87

Pada Tabel 4 *recall*-nya sangat rendah untuk netral dan positif karena banyak komentar netral dan positif yang salah diklasifikasi sebagai negatif. *Precision* positif tinggi ketika model sebuah komentar positif, kemungkinan besar benar, tetapi *recall* rendah artinya banyak yang tidak terdeteksi.

Hasil akurasi menggunakan visualisasi matrix

Visualisasi *confusion matrix* berikut menunjukkan hasil klasifikasi sentimen komentar menggunakan algoritma Naïve Bayes. Terlihat bahwa model sangat baik dalam mengenali komentar negatif, namun masih kurang akurat dalam mendeteksi komentar netral dan positif. Untuk mengetahui akurasinya dengan metode Confusion matrix dapat menggunakan rumus berikut:

$$\text{Accuracy} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Data}} \times 100\%$$

Tabel 5. Hasil Confusion Matrix

No	Aktual\ Prediksi	Negatif	Netral	Positif
1	Negatif	209	0	0
2	Netral	64	1	1
4	Positif	55	6	26

Dari perhitungan data testing 363, maka dapat dihitung nilai akurasinya sebagai berikut

$$\text{Accuracy} = \frac{209+1+26}{209+66+87} \times 100$$

$$\text{Accuracy} = \frac{236}{363} \times 100$$

$$\text{Accuracy} = \frac{236}{363} \approx 0.65 \times 100$$

$$\text{Accuracy} = 65\%$$

Model Naïve Bayes memiliki akurasi sebesar 65%, yang berarti sebanyak 236 dari 363 komentar berhasil diklasifikasikan dengan benar ke dalam kelas sentimen yang sesuai. Nilai ini menunjukkan bahwa performa model

Naïve Bayes cukup baik, terutama dalam mengenali komentar negatif, meskipun masih perlu peningkatan pada kelas netral dan positif.

4. SIMPULAN

Penelitian ini berhasil melakukan analisis sentimen terhadap 1.810 komentar netizen di *platform* TikTok mengenai pemindahan dan pembangunan Ibu Kota Nusantara menggunakan algoritma Naïve Bayes. Proses analisis meliputi tahapan pengumpulan data, *pre-processing* teks (*case folding, cleaning, convert slangword, tokenizing, stemming, stopword removal*), pelabelan manual, pembobotan kata dengan metode TF-IDF, serta klasifikasi sentimen. Hasil klasifikasi menunjukkan bahwa sentimen negatif mendominasi komentar netizen sebesar 56,91%, sedangkan sentimen positif sebesar 23,54% dan sentimen netral sebesar 19,56%. Model Naïve Bayes yang digunakan menghasilkan akurasi sebesar 65%, dengan recall tinggi pada kelas sentimen negatif, namun recall pada sentimen netral dan positif relatif rendah sehingga masih terdapat keterbatasan dalam mendeteksi kedua sentimen tersebut secara optimal.

DAFTAR PUSTAKA

- [1] N. S. Wardani, A. Prahutama, and P. Kartikasari, "Analisis Sentimen Pemindahan Ibu Kota Negara dengan Klasifikasi Naïve Bayes Untuk Model Bernoulli dan Multinomial," *Jurnal Gaussian*, vol. 9, no. 3, pp. 237-246, Aug. 2020. doi:10.14710/j.gauss.9.3.237-246.
- [2] S. F. Huwaida, R. Kusumawati, dan B. Isnaini, "Analisis Sentimen Komentar YouTube terhadap Pemindahan Ibu Kota Negara Menggunakan Metode Naïve Bayes," *Jambura Journal of Informatics*, vol. 6, no. 1, hlm. 26–39, Apr 2024, doi: 10.37905/jji.v6i1.24718.
- [3] H. A. R. Harpizon, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, and F. Syafria., "Analisis Sentimen Komentar di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, pp. 131-140, Feb. 2023. doi: 10.19184/isj.v8i1.33937.
- [4] A. Fitriadin dan A. Sidiq Purnomo, "Analisis Sentimen Masyarakat Terhadap Pandemi Covid-19 pada Sosial Media Menggunakan Naïve Bayes Classifier," *INFORMAL Informatics Journal*, vol. 8, no. 1, pp. 51–57, Apr 2024, doi: 10.37905/jji.v6i1.24718.
- [5] S. F. Huwaida, R. Kusumawati, and B. Isnaini, "Analisis Sentimen Komentar YouTube terhadap Pemindahan Ibu Kota Negara Menggunakan Metode Naïve Bayes," *Jambura Journal of Informatics*, vol. 6, no. 1, pp. 26–39, Apr 2024, doi: 10.37905/jji.v6i1.24718.
- [6] A. Safira, dan F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *Jurnal Sistem Informasi* vol. 5, no. 1, pp. 59–70, Jan. 2023, doi: 10.31849/zn.v5i1.12856.
- [7] Z. Li, R. Li, dan G. Jin, "Sentiment analysis of Danmaku Videos Based on Naïve Bayes and Sentiment Dictionary," *IEEE Access*, vol. 8, pp. 75073–75084, 2020, doi: 10.1109/ACCESS.2020.2986582.
- [8] K. Redaksi, A. Af Supianto, B. Riset dan Inovasi Nasional BRIN, E. Rosana Widasari, dan I. Wahyuni, "UB Oicial BITS Webmail UB News Dewan Editor", doi: 10.25126/jtiik.2023105.
- [9] F. M. Sarimole and W. Septian, "Analisis Sentimen Masyarakat Terhadap Isu Penundaan Pemilu 2024 pada Twitter dengan Metode Naive Bayes dan Support Vector Machine," *Jurnal Sains dan Teknologi*, vol. 5, no. 3, pp. 890-899, Feb. 2024, doi: 10.55338/saintek.v5i1.2789.
- [10] L. Hermawan, M. B. Ismiati, "Pembelajaran Text Preprocessing berbasis Simulator Untuk Mata Kuliah Information Retrieval," *Jurnal Transformatika*, vol. 17, no. 2, pp. 188–199, 2025, doi: 10.26623/transformatika.v17i2.1705.