

## Analisis Sentimen Tanggapan Masyarakat Tentang Garuda IKN Menggunakan Metode Naive Bayes

Nyoman Alvia Wirayani<sup>1\*</sup>, Ni Made Ary Esta Dewi Wirastuti<sup>1</sup>, Ida Bagus Gede Manuaba<sup>1</sup>

<sup>1</sup>Program Studi Teknik Elektro, Universitas Udayana, Indonesia.

---

### Artikel Info

#### Kata Kunci:

Analisis Sentimen;  
Garuda IKN;  
Naive Bayes.

#### Keywords:

*Sentiment Analysis;*  
*Garuda IKN;*  
*Naive Bayes.*

---

#### Riwayat Artikel:

Submitted: 24 November 2024

Accepted: 4 Maret 2025

Published: 8 Maret 2025

**Abstrak:** Ibu Kota Nusantara (IKN) adalah proyek besar yang bertujuan memindahkan ibu kota negara dari Jakarta ke Kalimantan Timur untuk mengurangi beban Jakarta dan mendorong pertumbuhan ekonomi di Indonesia. Salah satu ikon utama dari IKN adalah Patung Garuda, simbol kekuatan dan kebanggaan nasional. Namun, desain Patung Garuda ini mendapat berbagai kritik dari masyarakat, mulai dari aspek estetika hingga penggunaan anggaran yang dianggap berlebihan. Kritik ini banyak disampaikan melalui media sosial, terutama Twitter, yang menjadi platform utama untuk mengungkapkan pandangan secara terbuka. Penelitian ini bertujuan menganalisis sentimen publik terhadap desain Patung Garuda IKN melalui data dari Twitter dengan menggunakan metode Naive Bayes. Hasilnya menunjukkan bahwa metode Naive Bayes memberikan akurasi sebesar 0,82 dalam klasifikasi sentimen. Penelitian ini juga membuktikan bahwa Naive Bayes lebih akurat dibandingkan pustaka TextBlob. Melalui analisis ini, diharapkan dapat diperoleh pemahaman lebih mendalam mengenai pandangan publik dan isu-isu utama yang menjadi fokus kritik terhadap proyek ini. Hasil penelitian diharapkan dapat menjadi masukan berharga untuk perencanaan dan pengambilan keputusan proyek serupa di masa depan.

**Abstract:** *The Nusantara Capital City (IKN) is a significant project aimed at relocating the nation's capital from Jakarta to East Kalimantan to alleviate Jakarta's challenges while fostering national economic growth. One of IKN's key icons is the Garuda Statue, symbolizing strength and national pride. However, the statue's design has sparked various public criticisms, ranging from aesthetic concerns to perceptions of excessive budget allocation. Twitter, in particular, has become a prominent platform for openly sharing these opinions. This study analyzes public sentiment toward the Garuda Statue design in IKN using Twitter data and the Naive Bayes method. The findings indicate that the Naive Bayes method achieves 82% accuracy in sentiment classification, outperforming the TextBlob library. This analysis aims to provide a deeper understanding of public opinions and the primary issues driving criticism of this project. The results are expected to offer valuable insights for planning and decision-making in future large-scale projects.*

---

#### Corresponding Author:

Nyoman Alvia Wirayani

Email: [alviawirayani@gmail.com](mailto:alviawirayani@gmail.com)

---

## PENDAHULUAN

Ibu Kota Nusantara (IKN) adalah proyek besar yang bertujuan untuk memindahkan ibu kota negara dari Jakarta ke lokasi baru di Kalimantan Timur (Puspitasari, 2023). Meskipun pemindahan tersebut baru akan terwujud pada tahun 2045, perencanaan sarana pendukungnya sudah dimulai sejak tahun 2022 (Jiuhardi et al., 2023) menjadi salah satu topik yang banyak dibahas dan mendapatkan perhatian luas dari masyarakat. Langkah ini diambil untuk mengurangi beban Jakarta, yang saat ini mengalami berbagai permasalahan seperti kemacetan, kepadatan penduduk, dan penurunan kualitas lingkungan. Pemerintah juga memandang pemindahan ini sebagai bagian dari strategi Pembangunan yang lebih merata dan mendorong pertumbuhan ekonomi di wilayah Indonesia. Salah satu ikon yang direncanakan untuk menjadi pusat perhatian di IKN adalah Patung Garuda yang merepresentasikan kekuatan, kedaulatan, dan kebanggaan nasional. Namun sejak patung tersebut di publikasikan muncul berbagai kritik terhadap desain Garuda IKN mencakup berbagai aspek mulai dari estetika yang dianggap tidak mencerminkan budaya lokal hingga penggunaan anggaran yang dinilai berlebihan. Kritik ini banyak disampaikan melalui media sosial, khususnya Twitter, yang telah menjadi platform utama bagi masyarakat untuk menyuarakan opini mereka secara terbuka. Dengan tingginya intensitas diskusi publik tentang proyek ini, penting untuk memahami pola sentimen masyarakat terhadap Patung Garuda IKN. Pernyataan Masalah Kontroversi yang muncul terkait Patung Garuda IKN mencerminkan adanya kesenjangan antara tujuan pemerintah dengan persepsi masyarakat. Jika kritik dan pandangan masyarakat ini diabaikan, terdapat risiko berkurangnya kepercayaan publik terhadap proyek IKN, yang dapat memengaruhi keberhasilan implementasinya secara keseluruhan. Pemahaman yang lebih dalam terhadap sentimen publik menjadi sangat penting untuk menjembatani kesenjangan ini, sekaligus memberikan masukan bagi pengambilan keputusan strategis yang lebih inklusif.

Saat ini situs jejaring sosial menjadi tempat yang paling banyak menghasilkan atau mengambil informasi baik menggunakan perangkat, ponsel, laptop, tab, dan gadget untuk mendapatkan data apa pun kapan pun dan dari mana pun (Saini et al., 2019). Twitter adalah platform jejaring sosial di mana setiap hari diposting banyak pesan instan (tweet) yang menarik minat para peneliti untuk menganalisis data Twitter untuk berbagai informasi yang sedang banyak diperbincangkan (Adwan et al., 2020). Layanan ini menarik karena keterbukaannya, batasan 140 karakter per pesan, dan penggunaan tagar sehingga memudahkan dalam mengumpulkan informasi (Bouazizi & Ohtsuki, 2019). Twitter menjadi sarana utama bagi masyarakat untuk menyuarakan opini mereka terkait Istana Garuda IKN. Di platform ini, berbagai pandangan, baik yang mendukung maupun yang menentang disampaikan secara terbuka, sehingga menjadi sumber data yang kaya untuk dianalisis. Dengan adanya beragam opini yang tersebar di media sosial, metode analisis sentimen dapat digunakan untuk mengidentifikasi pola-pola perasaan masyarakat terhadap proyek ini.

Berbagai penelitian telah mengeksplorasi tentang analisis sentimen menggunakan metode *Naïve Bayes* yang menunjukkan keberhasilannya dalam menganalisis platform media sosial, misalnya (Darwis et al., 2021) menggunakan algoritma *Naïve Bayes* untuk menganalisis sentiment data Twitter BMKG Nasional dengan hasil uji akurasi pada metode *Naïve Bayes* sebesar 69,97%. Penelitian (Alfandi Safira & Hasan, 2023) menggunakan algoritma *Naïve Bayes* untuk menganalisis sentiment Masyarakat terhadap Paylater dengan akurasi sebesar 91%, Penelitian (Sari & Wibowo, 2019) menggunakan *Naïve Bayes* untuk menganalisis sentiment Pelanggan Toko Online JD.ID dengan hasil uji akurasi sebesar 96,44%, Penelitian (Astari et al., 2020) menggunakan Metode *Naïve Bayes* untuk menganalisis sentiment mengenai Dampak Virus Corona dengan hasil akurasi sebesar 67%, dan penelitian (Fais Sya' bani et al., 2022) menggunakan Algoritma *Naïve Bayes* untuk menganalisis calon presiden 2024 dengan hasil akurasi sebesar 73,68%.

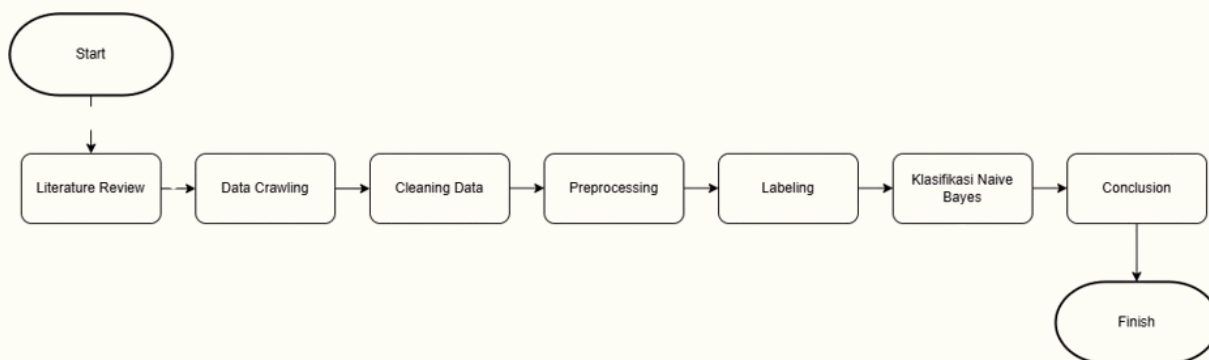
Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap cuitan-cuitan di Twitter yang berkaitan dengan Istana Garuda IKN dengan menggunakan metode *Naïve Bayes*. Metode ini akan membantu mengelompokkan sentimen masyarakat, baik yang positif, negatif, maupun netral, serta mengidentifikasi isu-isu utama yang menjadi fokus kritik. *Naïve Bayes* dipilih karena memiliki keunggulan dalam mengolah data teks pendek seperti tweet, algoritma ini juga dikenal sederhana namun efisien, sehingga cocok untuk menangani data dalam jumlah besar dengan akurasi yang

kompetitif. Melalui analisis ini, diharapkan dapat diperoleh pemahaman yang lebih mendalam tentang pandangan publik terhadap desain Garuda IKN, yang pada akhirnya dapat menjadi masukan berharga dalam perencanaan dan pengambilan keputusan proyek serupa di masa depan.

## METODE

Penelitian ini menggunakan Teknik *text mining* yang digunakan dalam mengklasifikasi analisis sentimen publik terhadap *tweet* yang bersifat positif, negatif, atau netral. Data diperoleh dari hasil *crawling* data *tweet* pada aplikasi Twitter kemudian diolah dan dimasukkan ke dalam *database*. Hasil *crawling* pada Twitter berdasarkan kata kunci yang dimasukkan dengan memanfaatkan API (Application Programming Interface) pencarian yang disediakan Twitter. Twitter API digunakan untuk mengumpulkan data teks dengan menggunakan akses kunci token sehingga mengakses data untuk pekerjaan proyek (Mandloi & Patel, 2020).

Alur penelitian yang dapat dilihat pada diagram alir sebagaimana yang disajikan pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

- a. Bagian Mencari literatur yang relevan terakit analisis sentimen
- b. Crawling data: Mengambil data mentah dari twitter dengan kata kunci #GarudaIKN
- c. Cleaning Data: Menghilangkan simbol atau karakter khusus agar lebih siap untuk di analisis
- d. Preprocessing: Mengubah data mentah menjadi format yang dipahami meliputi: Tokenisasi, Normalisasi, Stopword, dan Stemming
- e. Labeling: Pemberian label pada dataset untuk mengkategorikan data, seperti sentimen positif, negatif, atau netral dengan menggunakan Textblob
- f. Klasifikasi Naïve Bayes: Menentukan label menggunakan metode Naïve Bayes

### A. Dataset

Dataset adalah kumpulan *tweet* dari Twitter yang dikumpulkan dengan kata kunci terkait. Dataset ini digunakan untuk menganalisis sentimen dengan mengklasifikasikan tweet ke dalam tiga kategori sentimen utama: positif, negatif, atau netral (Guo et al., 2023). Dataset mencakup informasi seperti ID tweet, teks, metadata (seperti waktu posting), dan terkadang data lokasi pengguna (E. Chen et al., 2022).

### B. Twitter

Twitter sebagai platform media sosial dengan lebih dari 500 juta pengguna secara global. Platform ini berfungsi sebagai alat penting untuk berbagi berita, mendiskusikan ide, dan menyampaikan pandangan tentang berbagai peristiwa dunia. Selain itu, Twitter juga menjadi sumber informasi yang signifikan, dengan banyak pengguna, termasuk masyarakat umum dan sumber resmi yang berbagi berita, opini, serta informasi[13]

Twitter adalah layanan *microblogging* populer yang memungkinkan pengguna untuk berbagi, menyampaikan, dan menginterpretasikan secara *real-time*, singkat, dan sederhana. Layanan

*microblogging* telah menjadi platform yang paling terkenal dan paling umum digunakan. (Elbagir & Yang, 2020)

### C. Hashtag

*Hashtag* adalah tanda yang diawali dengan simbol "#" yang digunakan pada platform Twitter untuk mengelompokkan topik tertentu agar lebih mudah ditemukan oleh pengguna lain yang tertarik pada topik tersebut. Penggunaan hashtag sebagai alat untuk mengumpulkan data terkait dengan cara mengidentifikasi kata kuncinya (Abilov et al., 2021). *Hashtag* merupakan bagian penting dari analisis sentimen karena mereka dapat digunakan untuk memahami topik yang sedang dibahas dan bagaimana informasi tersebar di platform (Lopez & Gallemore, 2021).

### D. Google Collab Research

*Google collab* digunakan untuk mengimplementasikan algoritma klasifikasi dalam *Python*. Pustaka yang diimpor dalam pekerjaan ini adalah *numpy*, *matplotlib*, dan *pandas* (Balaraman, 2020). *Google Colab* digunakan sebagai platform untuk menjalankan berbagai eksperimen pemrograman *Python*, khususnya untuk pengumpulan data, prapemrosesan, dan analisis sentimen.

Platform ini mendukung pemrograman berbasis *Python* secara online. *Google Colab* sangat berguna bagi peneliti dan pengembang untuk mengerjakan proyek berbasis data seperti analisis tweet dan pustaka *Python* seperti *Tweepy* dan *TextBlob*. Salah satu kelebihanannya adalah akses gratis ke GPU, yang sangat membantu dalam aplikasi pembelajaran mendalam mengenai deep learning (Mattia C, 1862).

### E. Python

*Python* adalah bahasa pemrograman yang sangat populer dan mudah dibaca serta dipahami karena sintaksisnya yang sederhana (Chromiński et al., 2021). *Python* menjadi pilihan utama dalam ilmu data, pembelajaran mesin (machine learning), dan pembelajaran mendalam (deep learning) karena menyediakan menyediakan berbagai Pustaka(alat) dalam mendukung analisis data sehingga menjadikannya sangat cocok untuk proyek penelitian dan industri yang kompleks (Blank & Deb, 2020).

*Python* digunakan untuk mengumpulkan, membersihkan, dan menganalisis data dari Twitter, termasuk penggunaan pustaka seperti *nlTK* untuk membersihkan data dan *TextBlob* untuk analisis sentiment (Kaur & Sharma, 2020).

### F. Sentimen

Analisis sentimen merupakan metode pemrosesan bahasa alami yang digunakan untuk mengidentifikasi dan memahami emosi dalam teks. Proses ini melibatkan pengukuran dua aspek utama: polaritas, yang mengindikasikan apakah suatu teks memiliki nada positif, negatif, atau netral dengan rentang dari -1 (sangat negatif) hingga +1 (sangat positif), sedangkan subjektivitas mengukur seberapa banyak teks tersebut merupakan opini pribadi atau fakta (Garcia & Berton, 2021). Analisis sentimen adalah cabang ilmu komputer yang merupakan bagian dari data mining, sering dimanfaatkan di dunia bisnis untuk memahami pendapat pelanggan secara langsung terhadap suatu produk (Asno Azzawagama Firdaus et al., 2023).

Menurut (Wang et al., 2022) Sentimen dianalisis dengan melihat apakah kata-kata dalam teks menunjukkan perasaan positif, netral, atau negatif. Analisis ini membantu kita memahami emosi dalam teks dan dapat digunakan untuk berbagai keperluan, seperti mempelajari pasar, memprediksi tren, dan melihat bagaimana orang bereaksi terhadap suatu peristiwa.

### G. Naïve Bayes

*Naïve Bayes* adalah salah satu algoritma *data mining* yang sering digunakan untuk klasifikasi. Algoritma ini didasarkan pada asumsi bahwa semua atribut bersifat independen satu sama lain jika diberikan kelas tertentu (S. Chen et al., 2020). Algoritma Naïve Bayes sering digunakan dalam analisis sentimen untuk mengklasifikasikan teks, seperti menentukan apakah suatu ulasan bersifat positif,

negatif, atau netral, dengan memanfaatkan fitur berupa kata-kata yang terdapat dalam teks tersebut (Yunus et al., 2023).

Algoritma yang sederhana dan sangat efisien secara komputasi karena menggunakan asumsi bahwa semua atribut saling independen. Hal ini membuat perhitungan probabilitas menjadi lebih mudah dan cepat (S. Chen et al., 2020). Pengklasifikasi *Naive Bayes* mengasumsikan bahwa keberadaan fitur tertentu dalam sebuah kelas tidak bergantung pada fitur lainnya. Sebagai contoh, buah yang berwarna merah, berbentuk bulat, dan memiliki diameter sekitar tiga inci mungkin dianggap sebagai apel (Alsaeedi & Khan, 2019).

## HASIL DAN PEMBAHASAN

### A. Crawling Data

*Crawling* data adalah teknik untuk mengambil data mentah dari Twitter. Metode ini menggunakan skrip untuk mengakses halaman Twitter secara langsung dalam pengambilan data dalam jumlah besar (Antonakaki et al., 2021). Pada penelitian ini, pengumpulan data tweet dari twitter menggunakan *Collab Research Google* dengan kata kunci yang digunakan yaitu "GarudaIKN" dan disimpan dalam format CSV dengan nama file GarudaIKN.csv. Setelah proses crawling sudah selesai, dataset masih berupa data mentah dan mengandung kolom atau fitur yang tidak akan digunakan dalam klasifikasi. Detail dari data yang diperoleh melalui crawling dapat dilihat pada Tabel berikut:

Tabel 1. Atribut Dataset Awal

No	Atribut
1.	conversation_id_str
2.	created_at
3.	favorite_count
4.	full_text
5.	id_str
6.	image_url
7.	in_reply_to_screen_name
8.	lang
9.	location
10.	quote_count
11.	reply_count
12.	retweet_count
13.	tweet_url
14.	user_id_str
15.	

Dari dataset di atas terdapat 15 atribut hasil *crawling* data. Namun, dalam proses klasifikasi tidak semua atribut tersebut digunakan sehingga perlu dilakukan seleksi atribut. Atribut yang digunakan *full\_text*, *username*, *created\_at*

Tabel 2. Atribut Dataset Terpilih

No	Full_text	Created_at	Username
1.	Patung Istana Garuda IKN: Habis Gelap Terbitlah Hijau	Tue Sep 10 23:05:56 +0000 2024	detik_properti
2.	Patung Garuda IKN Berubah Warna Jadi Biru Toska	Tue Sep 10 13:15:05 +0000 2024	GazanaPublika

Pada tabel di atas terdapat 3 Atribut yang dipilih untuk melakukan analisis sentiment, yaitu *Full\_text* berisikan deskripsi lengkap dari tweet yang mencerminkan opini atau sentimen pengguna.

Atribut ini penting untuk analisis sentimen karena merupakan sumber utama data teks yang akan dianalisis, *created\_at* berisikan catatan waktu dan tanggal saat tweet dibuat. Informasi ini berguna untuk analisis temporal, seperti mengidentifikasi tren sentimen dalam rentang waktu tertentu, dan *username* menyediakan informasi tentang akun pengguna. Meskipun tidak digunakan langsung dalam analisis sentimen, atribut ini dapat membantu dalam melacak sumber data atau melakukan analisis lanjutan berdasarkan pengguna tertentu.

## B. Cleaning Data

Pada proses *cleaning* data simbol-simbol khusus seperti tanda seru, tanda baca, angka, dan karakter khusus lainnya dihapus, sementara semua kata diubah menjadi huruf kecil (Stephen & Prabu, 2019). Proses ini bertujuan untuk menghilangkan elemen-elemen yang tidak relevan atau mengganggu dari data yang tidak terstruktur dari Twitter agar data lebih siap untuk dianalisis (Pradha et al., 2019).

Tabel 3. Dataset Setelah Cleaning Data

No	Full_text	Created_at	Username
1.	patung istana garuda ikn habis gelap terbitlah hijau	Tue Sep 10 23:05:56 +0000 2024	detik_properti
2.	patung garuda ikn berubah warna jadi biru toska	Tue Sep 10 13:15:05 +0000 2024	GazanaPublika

Berdasarkan tabel 3, hasil dari *cleaning* data terlihat bahwa simbol “.” semula berada pada kalimat “Patung Istana Garuda ikn:” telah dibersihkan dan seluruh kata diubah menjadi huruf kecil.

## C. Preprocessing

*Preprocessing* adalah proses mengubah data mentah menjadi format yang dipahami oleh mesin, proses ini melakukan pembersihan dan transformasi data menjadi bentuk yang lebih terstruktur dan bersih, seperti menghapus *URL*, simbol *Hashtag*, karakter khusus karena kata-kata ini tidak relevan untuk dianalisis lebih lanjut (Adwan et al., 2020). Pada langkah ini, akan dilakukan pembersihan dataset mulai dari *URL*, *retweet*, *mention*, dan beberapa karakter khusus seperti “^%#-@” karena kata-kata ini tidak relevan untuk dianalisis lebih lanjut (Nezhad & Deihimi, 2022). Setiap proses *preprocessing* sebagai berikut:

### 1. Tokenisasi

*Tokenisasi* digunakan sebagai langkah awal dalam analisis sentimen sebelum memasuki langkah-langkah pemrosesan lebih lanjut (Yasen & Tedmori, 2019). Proses *tokenisasi* adalah tahapan memecah paragraf besar menjadi potongan teks atau kata-kata (Ahuja et al., 2019). Berdasarkan tabel 4 hasil dari proses *tokenisasi* terlihat bahwa semula berbentuk kalimat telah diubah dengan memisahkannya menjadi perkata.

Tabel 4. Dataset Setelah Proses Tokenisasi

No	Full_text	Created_at	Username
1.	patung, istana, garuda, ikn, habis, gelap, terbitlah, hijau	Tue Sep 10 23:05:56 +0000 2024	detik_properti
2.	patung, garuda, ikn, berubah, warna, jadi, biru, toska	Tue Sep 10 13:15:05 +0000 2024	GazanaPublika

### 2. Normalisasi

Proses normalisasi bertujuan untuk merapikan kata-kata yang tidak baku, seperti bahasa gaul, mengganti simbol dan angka, contohnya angka ‘3’ menjadi ‘tiga’ (Najiyah, 2023). Proses normalisasi sangat penting untuk menangani teks yang berantakan (*noisy text*) dan mengubahnya menjadi bentuk standar (Kumar, 2024).

Tabel 5. Dataset Setelah Proses Normalisasi

No	Full_text	Created_at	Username
1.	patung istana garuda ikn habis gelap terbitlah hijau	Tue Sep 10 23:05:56 +0000 2024	detik_properti
2.	patung garuda ikn berubah warna jadi biru toska	Tue Sep 10 13:15:05 +0000 2024	GazanaPublika

Pada penelitian ini kalimat yang mengandung kata tidak baku dan Bahasa gaul telah diubah. Seperti " yg "=" yang ", " sm " = " sama ", ' sptnya ' = ' sepertinya ', " ogah "=" males ", 'gue'='saya', 'w'='saya', 'kpala' = 'kepala'

### 3. Stopword

*Stopword* adalah proses menghilangkan kata kata yang tidak memiliki arti, kata yang tidak dapat digunakan untuk memprediksi sesuatu. Contoh *Stopword* termasuk “dan”, “yang”, “di”, “dari”, dan “adalah”(Ladani & Desai, 2020). *Stopword removal* dilakukan untuk menghapus kata-kata yang tidak membawa informasi signifikan untuk mendeteksi sentiment (Palomino, 2022).

Tabel 6. Dataset Setelah Proses Stopword

No	Full_text	Created_at	Username
1.	patung istana garuda ikn habis gelap terbitlah hijau	Tue Sep 10 23:05:56 +0000 2024	detik_properti
2.	patung garuda ikn berubah warna jadi biru toska	Tue Sep 10 13:15:05 +0000 2024	GazanaPublika

Berdasarkan tabel 6 hasil dari proses *stopword* terhadap kalimat di atas tidak terdapat perubahan, karena tidak ada kata yang mengandung “dan”, “yang”, “di”, “dari”, dan “adalah”

#### 4. Stemming

*Stemming* atau *lemmatisasi* adalah cara untuk menemukan bentuk dasar dari kata-kata yang memiliki variasi atau bentuk turunan (Pandey & Deorankar, 2019). Menurut (Shehu et al., 2021) Algoritma stemming umumnya menggunakan aturan-aturan tertentu dan sering kali melibatkan penghapusan akhiran dari kata. Misalnya, algoritma stemming akan mengubah kata-kata seperti "fishing," "fished," dan "fisher" menjadi bentuk dasarnya, yaitu "fish". Berdasarkan tabel 7, hasil dari proses *stemming* terlihat bahwa semula terdapat kata "terbitlah" menjadi kata "terbit".

Tabel 7. Dataset Setelah Proses Stemming

No	Full_text	Username
1.	patung istana garuda ikn habis gelap terbit hijau	detik_properti
2.	patung garuda ikn ubah warna jadi biru toska	GazanaPublika

#### D. Labeling

*Labeling* adalah proses penentuan dan pemberian label pada dataset untuk mengkategorikan data, seperti sentimen positif, negatif, atau netral (Tama et al., 2019). Sebelum proses labelling *full\_text* di translate terlebih dahulu ke Bahasa Inggris untuk membantu memastikan akurasi dalam klasifikasi sentimen. Selain itu, dengan menerjemahkan teks. Dengan menerjemahkan teks, model dapat lebih mudah mengenali pola bahasa, sehingga menghasilkan analisis sentimen yang lebih konsisten.

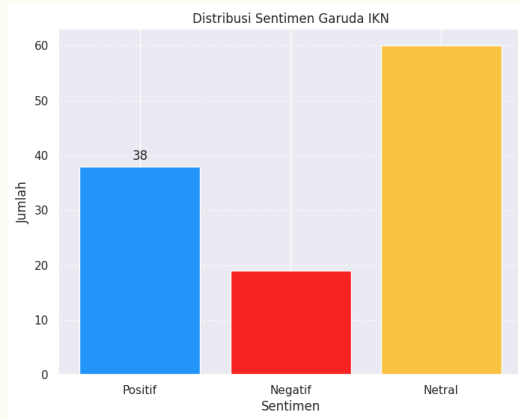
Tabel 8. Labeling Dataset

No	Full_text	Tweet_english	Polaritas	Klasifikasi
1.	patung istana garuda ikn habis gelap terbit hijau	garuda castle statue ikn run out dark rose green	0.08333333333333331	Positif
2.	patung garuda ikn ubah warna jadi biru toska	garuda ikn statue changes color to toska blue	0.0	Netral



Pada tahapan labelling untuk mengetahui klasifikasi termasuk kategori positif, negatif, atau Netral dapat dilihat pada nilai polaritasnya, Jika:

1. Nilai polaritas lebih dari 0.0, maka teks dikategorikan sebagai "positif".
2. Nilai polaritas sama dengan 0.0, teks masuk dalam kategori "netral"
3. Nilai polaritas kurang dari 0.0, teks dikategorikan sebagai "negatif"



Gambar 3. Grafik Distribusi Sentimen Garuda IKN

Berdasarkan grafik distribusi sentimen di atas, jumlah klasifikasi yang diperoleh adalah 38 untuk sentimen positif, 19 untuk sentimen negatif, dan 60 untuk sentimen netral.

### 5. Visualisasi

Proses visualisasi yaitu menampilkan informasi secara visual untuk membantu pengguna dalam mengeksplorasi data yang digunakan untuk memudahkan pengguna dalam memahami data sentimen, seperti positif, negatif, dan netral (Almjawel, 2019). Dengan menggunakan teknik ini, kita bisa melihat kata kunci yang paling sering digunakan (Saini et al., 2019).



Gambar 4. Visualiasasi

### 6. Klasifikasi Naïve Bayes

*TextBlob* merupakan pustaka *Python* yang melakukan klasifikasi sentimen menjadi positif, negatif, atau netral. Namun, *TextBlob* hanya mendukung bahasa Inggris. Jadi, untuk mengklasifikasikan kalimat dalam bahasa Indonesia, perlu diterjemahkan terlebih dahulu ke bahasa Inggris. Setelah diterjemahkan, *TextBlob* mengklasifikasikan sentimen berdasarkan nilai polaritas: positif jika nilai lebih dari 0, netral jika 0, dan negatif jika kurang dari 0. (Mas Diyasa et al., 2021) Pada penelitian ini Akurasi test menggunakan metode *Naïve Bayes* sebesar 0.82. Hasil Klasifikasi menggunakan *Naïve bayes* sebagai berikut.

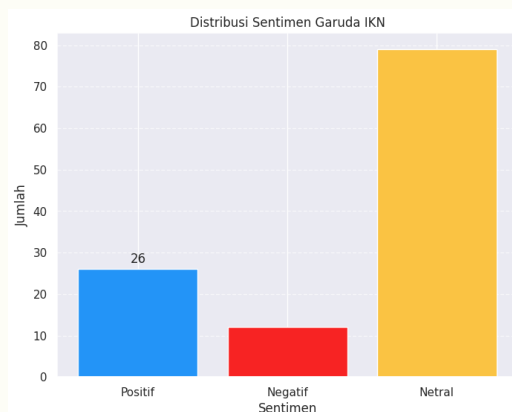
Tabel 9. Dataset Klasifikasi Naïve Bayes

No	Full_text	Tweet_english	Klasifikasi	Klasifikasi Naïve Bayes
1.	patung istana garuda ikn habis gelap terbit hijau	garuda castle statue ikn run out dark rose green	Positif	Positif
2.	patung garuda ikn ubah warna jadi biru toska	garuda ikn statue changes color to toska blue	Netral	Positif

Berdasarkan tabel di atas terdapat perbedaan jenis kategori antara klasifikasi menggunakan *textblob* (Pustaka Python untuk menganalisis) dengan *Naïve Bayes*. Perbedaan ini disebabkan oleh perbedaan mendasar dalam cara kedua metode ini menganalisis teks dan menangkap konteks serta pola bahasa. *TextBlob* lebih cepat dan sederhana karena hanya menggunakan daftar kata, tetapi kurang peka terhadap konteks. Sementara itu, *Naïve Bayes* bisa memberikan hasil yang lebih akurat karena mampu memahami bagaimana kata-kata digunakan dalam konteks yang berbeda. Itulah mengapa hasil dari kedua metode ini bisa berbeda, terutama jika teks yang dianalisis memiliki makna yang lebih kompleks.

Sentimen netral dalam klasifikasi menggunakan *TextBlob* dan *Naïve Bayes* dipengaruhi oleh beberapa faktor penting yang berkaitan dengan konteks dan pola bahasa. Sentimen netral biasanya muncul dari teks yang tidak mengandung emosi atau opini yang jelas, seperti informasi atau fakta yang disampaikan secara deskriptif tanpa penilaian subjektif. Dalam kasus ini, teks sering kali hanya memberikan laporan atau deskripsi tanpa menunjukkan perasaan positif atau negatif. Metode *TextBlob*, yang bergantung pada daftar kata statis untuk menganalisis sentimen, cenderung menetapkan sentimen netral pada teks yang tidak mengandung kata-kata emosional yang terdapat dalam pustaka standarnya. Namun, pendekatan ini kurang mampu menangkap konteks kalimat, terutama jika teks memiliki makna yang lebih kompleks atau menggunakan pola bahasa yang ambigu.

Di sisi lain, metode *Naïve Bayes*, yang menggunakan probabilitas kata berdasarkan data pelatihan, dapat menangkap konteks penggunaan kata dalam kalimat. Namun, sentimen netral tetap dapat muncul jika model tidak menemukan cukup bukti probabilitas yang kuat untuk mengklasifikasikan teks sebagai positif atau negatif. Hal ini sering terjadi pada teks yang singkat atau memiliki pola bahasa yang ambigu, seperti penggunaan kata-kata umum tanpa muatan emosional yang signifikan. Oleh karena itu sentimen netral dipengaruhi oleh minimnya emosi yang tersurat dalam teks, konteks yang tidak ditangkap. Perbedaan hasil antara *TextBlob* dan *Naïve Bayes* dalam kategori netral mencerminkan keterbatasan masing-masing metode dalam memahami kompleksitas bahasa dan menangkap nuansa emosi dalam teks yang dianalisis. Berdasarkan grafik distribusi sentimen IKN dengan metode *Naïve Bayes*, jumlah klasifikasi yang diperoleh adalah 26 untuk sentimen positif, 12 untuk sentimen negatif, dan 79 untuk sentimen netral. Hasil ini disajikan pada Gambar 5:



Gambar 5. Grafik Distribusi Sentimen Garuda IKN Naive Bayes

## KESIMPULAN

Penelitian ini menganalisis sentimen masyarakat Indonesia terhadap desain Patung Garuda di Ibu Kota Nusantara (IKN) melalui platform Twitter dengan menggunakan metode Naive Bayes. Sentimen diklasifikasikan menjadi tiga kategori utama, yaitu positif, negatif, dan netral, berdasarkan data tweet yang dikumpulkan menggunakan kata kunci #GarudaIKN melalui Google Colab Research. Metode Naive Bayes menghasilkan tingkat akurasi sebesar 0,82, yang menunjukkan keandalannya dalam mengklasifikasikan sentimen secara lebih akurat dibandingkan pustaka TextBlob. Hasil penelitian ini memberikan kontribusi yang signifikan dalam memahami persepsi publik terhadap proyek pembangunan nasional. Secara praktis, penelitian ini dapat membantu pemangku kepentingan untuk lebih memahami opini publik, sehingga dapat dijadikan dasar untuk meningkatkan kualitas komunikasi dan interaksi dengan masyarakat. Wawasan yang diperoleh memungkinkan pengambil kebijakan untuk menyusun strategi komunikasi yang lebih efektif, terutama dalam merespons sentimen negatif yang dapat menimbulkan konflik atau menghambat keberhasilan proyek. Selain itu, hasil ini juga memberikan landasan bagi perencanaan yang lebih baik dalam proyek-proyek serupa di masa depan. Dengan memahami pola sentimen publik, pemerintah atau pengembang proyek dapat mengambil keputusan yang lebih tepat sasaran, baik dalam konteks penyampaian informasi maupun dalam mitigasi isu-isu sensitif. Penelitian ini menegaskan pentingnya pemanfaatan teknologi analitik data dalam pengambilan keputusan strategis untuk proyek infrastruktur berskala besar, sehingga dapat memastikan keberhasilan proyek sekaligus membangun kepercayaan masyarakat.

## DAFTAR PUSTAKA

- Abilov, A., Hua, Y., Matatov, H., Amir, O., & Naaman, M. (2021). VoterFraud2020: a Multi-modal Dataset of Election Fraud Claims on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 15, 901–912. <https://doi.org/10.1609/icwsm.v15i1.18113>
- Adwan, O. Y., Al-Tawil, M., Huneiti, A. M., Shahin, R. A., Abu Zayed, A. A., & Al-Dibsi, R. H. (2020). Twitter sentiment analysis approaches: A survey. *International Journal of Emerging Technologies in Learning*, 15(15), 79–93. <https://doi.org/10.3991/ijet.v15i15.14467>
- Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The impact of features extraction on the sentiment analysis. *Procedia Computer Science*, 152, 341–348. <https://doi.org/10.1016/j.procs.2019.05.008>
- Alfandi Safira, & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier. *ZONAsi: Jurnal Sistem Informasi*, 5(1), 59–70. <https://doi.org/10.31849/zn.v5i1.12856>
- Almjawel, A. (2019). Books ' Reviews. *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, 1–6.

- Alsaeedi, A., & Khan, M. Z. (2019). A study on sentiment analysis techniques of Twitter data. *International Journal of Advanced Computer Science and Applications*, 10(2), 361–374. <https://doi.org/10.14569/ijacsa.2019.0100248>
- Antonakaki, D., Fragopoulou, P., & Ioannidis, S. (2021). A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. *Expert Systems with Applications*, 164(February 2020), 114006. <https://doi.org/10.1016/j.eswa.2020.114006>
- Astari, N. M. A. J., Dewa Gede Hendra Divayana, & Gede Indrawan. (2020). Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Dan Informatika (JSI)*, 15(1), 27–29. <https://doi.org/10.30864/jsi.v15i1.332>
- Balaraman, S. (2020). Comparison of Classification Models for Breast Cancer Identification using Google Colab. *Preprints.*, May, 1–11. <https://doi.org/10.20944/preprints202005.0328.v1>
- Blank, J., & Deb, K. (2020). Pymoo: Multi-Objective Optimization in Python. *IEEE Access*, 8, 89497–89509. <https://doi.org/10.1109/ACCESS.2020.2990567>
- Bouazizi, M., & Ohtsuki, T. (2019). Multi-class sentiment analysis on twitter: Classification performance and challenges. *Big Data Mining and Analytics*, 2(3), 181–194. <https://doi.org/10.26599/BDMA.2019.9020002>
- Chen, E., Deb, A., & Ferrara, E. (2022). #Election2020: the first public Twitter dataset on the 2020 US Presidential election. *Journal of Computational Social Science*, 5(1), 1–18. <https://doi.org/10.1007/s42001-021-00117-9>
- Chen, S., Webb, G. I., Liu, L., & Ma, X. (2020). A novel selective naïve Bayes algorithm. *Knowledge-Based Systems*, 192(xxxx), 105361. <https://doi.org/10.1016/j.knosys.2019.105361>
- Chromiński, K., Benko, L., Hernández-Figueroa, Z. J., González-Domínguez, J. D., & Rodríguez-del-Pino, J. C. (2021). Python Fundamentals. *Python Fundamentals*, c. <https://doi.org/10.17846/fpvai-2021-14>
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Elbagir, S., & Yang, J. (2020). *Sentiment Analysis on Twitter with Python's Natural Language Toolkit and VADER Sentiment Analyzer*. 0958, 63–80. [https://doi.org/10.1142/9789811215094\\_0005](https://doi.org/10.1142/9789811215094_0005)
- Fais Sya' bani, M. R., Enri, U., & Padilah, T. N. (2022). Analisis Sentimen Terhadap Bakal Calon Presiden 2024 Dengan Algoritme Naïve Bayes. *JURIKOM (Jurnal Riset Komputer)*, 9(2), 265. <https://doi.org/10.30865/jurikom.v9i2.3989>
- Firdaus, A. A., Yudhana, A., & Riadi, I. (2023). Analisis Sentimen Pada Proyeksi Pemilihan Presiden 2024 Menggunakan Metode Support Vector Machine. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 236-245. <https://doi.org/10.51454/decode.v3i2.172>
- Garcia, K., & Berton, L. (2021). Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Applied Soft Computing*, 101, 107057. <https://doi.org/10.1016/j.asoc.2020.107057>
- Guo, Y., Das, S., Lakamana, S., & Sarker, A. (2023). An aspect-level sentiment analysis dataset for therapies on Twitter. *Data in Brief*, 50, 109618. <https://doi.org/10.1016/j.dib.2023.109618>
- Jiuhardi, J., Rosyadi, R., Wijayanti, T. C., & Fitriadi, F. (2023). What Is the Perspective of Millennial Tourists on the Revitalization of Tourism Development in New Ikn? *Planning Malaysia*, 21(1), 232–246. <https://doi.org/10.21837/PM.V21I25.1236>

- Kaur, C., & Sharma, A. (2020). Social issues sentiment analysis using python. *Proceedings of the 2020 International Conference on Computing, Communication and Security, ICCCS 2020*, 1–6. <https://doi.org/10.1109/ICCCS49678.2020.9277251>
- Kumar, S. (2024). Sentiment Analysis. *Python for Accounting and Finance*, 265–300. [https://doi.org/10.1007/978-3-031-54680-8\\_17](https://doi.org/10.1007/978-3-031-54680-8_17)
- Ladani, D. J., & Desai, N. P. (2020). Stopword Identification and Removal Techniques on TC and IR applications: A Survey. *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020*, 466–472. <https://doi.org/10.1109/ICACCS48705.2020.9074166>
- Lopez, C. E., & Gallemore, C. (2021). An augmented multilingual Twitter dataset for studying the COVID-19 infodemic. *Social Network Analysis and Mining*, 11(1), 1–14. <https://doi.org/10.1007/s13278-021-00825-0>
- Mandloi, L., & Patel, R. (2020). Twitter sentiments analysis using machine learning methods. *2020 International Conference for Emerging Technology, INCET 2020*, 1–5. <https://doi.org/10.1109/INCET49848.2020.9154183>
- Mas Diyasa, I. G. S., Marini Mandenni, N. M. I., Fachrurrozi, M. I., Pradika, S. I., Nur Manab, K. R., & Sasmita, N. R. (2021). Twitter Sentiment Analysis as an Evaluation and Service Base On Python Textblob. *IOP Conference Series: Materials Science and Engineering*, 1125(1), 012034. <https://doi.org/10.1088/1757-899x/1125/1/012034>
- Mattia C. (1862). Preprint not peer reviewed. *The Lancet Pschch*, 1–35.
- Najiyah, I. (2023). Analisis Sentimen Tanggapan Masyarakat Indonesia Tentang Kenaikan Bbm Menggunakan Metode Artificial Neural Network. *Jurnal Responsif: Riset Sains Dan Informatika*, 5(1), 92–100. <https://doi.org/10.51977/jti.v5i1.1061>
- Nezhad, Z. B., & Deihimi, M. A. (2022). Diabetes & Metabolic Syndrome : Clinical Research & Reviews Twitter sentiment analysis from Iran about COVID 19 vaccine. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 16(1), 102367. <https://doi.org/10.1016/j.dsx.2021.102367>
- Palomino, M. A. (2022). *Evaluating-the-Effectiveness-of-Text-PreProcessing-in-Sentiment-AnalysisApplied-Sciences-Switzerland.pdf*.
- Pandey, S. V., & Deorankar, A. V. (2019). A Study of Sentiment Analysis Task and It's Challenges. *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*, 1–5. <https://doi.org/10.1109/ICECCT.2019.8869160>
- Pradha, S., Halgamuge, M. N., & Tran Quoc Vinh, N. (2019). Effective text data preprocessing technique for sentiment analysis in social media data. *Proceedings of 2019 11th International Conference on Knowledge and Systems Engineering, KSE 2019*, 1–8. <https://doi.org/10.1109/KSE.2019.8919368>
- Puspitasari, D. A. (2023). Tracing Trends and Patterns of IKN Words in Media and Twitter: A Linguistic Corpus Study (Menelusuri Tren dan Pola Kata IKN di Media dan Twitter: Kajian Korpus Linguisti). *Gramatika STKIP PGRI Sumatera Barat*, 9(2), 239–257. <https://doi.org/10.22202/jg.2023.v9i2.6947>
- Saini, S., Punhani, R., Bathla, R., & Shukla, V. K. (2019). Sentiment Analysis on Twitter Data using R. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 68–72. <https://doi.org/10.1109/ICACTM.2019.8776685>
- Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, 10(2), 681–686.
- Shehu, H. A., Sharif, M. H., Sharif, M. H. U., Datta, R., Tokat, S., Uyaver, S., Kusetogullari, H., & Ramadan, R. A. (2021). Deep Sentiment Analysis: A Case Study on Stemmed Turkish Twitter

- Data. *IEEE Access*, 9, 56836–56854. <https://doi.org/10.1109/ACCESS.2021.3071393>
- Stephen, J. J., & Prabu, P. (2019). Detecting the magnitude of depression in Twitter users using sentiment analysis. *International Journal of Electrical and Computer Engineering*, 9(4), 3247–3255. <https://doi.org/10.11591/ijece.v9i4.pp3247-3255>
- Tama, V. O., Sibaroni, Y., & Adiwijaya. (2019). Labeling Analysis in the Classification of Product Review Sentiments by using Multinomial Naive Bayes Algorithm. *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012036>
- Wang, Y., Guo, J., Yuan, C., & Li, B. (2022). Sentiment Analysis of Twitter Data. *Applied Sciences (Switzerland)*, 12(22), 1–14. <https://doi.org/10.3390/app122211775>
- Yasen, M., & Tedmori, S. (2019). Movies reviews sentiment analysis and classification. *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2019 - Proceedings*, 860–865. <https://doi.org/10.1109/JEEIT.2019.8717422>
- Yunus, M., Biddinika, M. K., & Fadlil, A. (2023). Optimasi Algoritma Naïve Bayes Menggunakan Fitur Seleksi Backward Elimination untuk Klasifikasi Prevalensi Stunting. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 278-285. <https://doi.org/10.51454/decode.v3i2.188>