JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 5, No. 1, Januari 2026, pp. 502 - 513
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/169

Public Sentiment Analysis on the Development of the Nusantara Capital City (IKN) using the Naive Bayes Algorithm

Baginda Oloan Lubis*[1], Benny Firmansyah[2], Hilmy Putra[3], Soni Irawan[4], Ainindia Nur Alizah[5], Muhammad Ibrahim Adnan[6], Daniel Marulitua Situmorang[7]
[1] Fakultas Teknik dan Informatika, Program Studi Rekayasa Perangkat Lunak Universitas Bina Sarana Infomatika
Jl. Margonda No. 8 Pondok Cina, Kecamatan Beji, Kota Depok Jawa Barat, Indonesia
[2,3,4,5,6,7] Fakultas Teknik dan Informatika, Program Studi Sistem Informasi Universitas Bina Sarana Infomatika
Jl. Margonda No. 8 Pondok Cina, Kecamatan Beji, Kota Depok Jawa Barat, Indonesia
e-mail: baginda.bio@bsi.ac.id*[1], 19231949@bsi.ac.id[2], 19232027@bsi.ac.id[3], 19232498@bsi.ac.id[4], 19230252@bsi.ac.id[5], 19231868@bsi.ac.id[6], 19232556@bsi.ac.id[7]

ARTICLE INFO

Article History:

Received: 05-11-2025
Revised: 29-12-2025
Accepted: 02-01-2026
Published: 21-01-2026

ABSTRACT

The development of Ibu Kota Nusantara (IKN) has become one of Indonesia's most widely discussed national strategic projects due to its broad socio-economic and political implications. Social media platforms, particularly TikTok, play a crucial role in shaping and reflecting public opinion, making them essential data sources for real-time sentiment analysis. This study aims to analyze public sentiment regarding the progress of IKN's development using the Naive Bayes algorithm. TikTok comments related to IKN were collected and subjected to a series of preprocessing steps, including case-folding, cleansing, tokenization, stopword removal, and stemming. The processed data were manually labeled into positive, negative, and neutral sentiment classes before being used to train a classification model. The findings demonstrate that the Naive Bayes algorithm effectively classifies public sentiment regarding IKN, as evidenced by satisfactory evaluation metrics, including accuracy, precision, recall, and F1-score. The sentiment distribution indicates diverse public perceptions: a significant portion expresses skepticism about government readiness and project feasibility, while others are optimistic about the long-term economic benefits. This study contributes to understanding public responses toward national infrastructure projects in the digital era. It also provides policymakers and government institutions with valuable insights to refine public communication strategies and strengthen transparency related to IKN development. Overall, the research highlights the importance of data-driven approaches in capturing public sentiment within the evolving landscape of digital.

Keywords: *Ibu Kota Nusantara, Naive Bayes, Public Sentiment, TikTok Analysis, Social Media Mining.*

DOI: https://doi.org/10.51903/nx1aza77

p-ISSN : 2803-1507 e-ISSN : 2803-1531

.
Abstrak

Pembangunan Ibu Kota Nusantara (IKN) merupakan salah satu proyek strategis nasional yang banyak mendapatkan perhatian publik karena dampaknya yang luas pada aspek sosial, ekonomi, dan politik di Indonesia. Media sosial, khususnya TikTok, memainkan peran penting dalam membentuk dan merefleksikan opini masyarakat, sehingga menjadi sumber data yang relevan untuk analisis sentimen secara *real-time*. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap progres pembangunan IKN menggunakan algoritma *Naive Bayes*. Komentar-komentar TikTok yang terkait dengan IKN dikumpulkan dan melalui beberapa tahapan *preprocessing*, seperti *case-folding*, *cleansing*, *tokenization*, *stopword removal*, dan *stemming*. Data yang telah diproses kemudian diberi label sentimen secara manual (positif, negatif, dan netral) sebelum digunakan untuk melatih model klasifikasi. Hasil penelitian menunjukkan bahwa algoritma *Naive Bayes* mampu memberikan performa yang efektif dalam mengklasifikasikan sentimen publik terkait IKN, yang ditunjukkan melalui metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Distribusi sentimen mengungkapkan bahwa persepsi publik cenderung beragam, dengan sebagian besar komentar bernada skeptis terhadap kesiapan pemerintah dan kelayakan proyek, sementara sebagian lain menunjukkan optimisme terhadap manfaat ekonomi jangka panjang. Penelitian ini memberikan kontribusi dalam memahami respons publik terhadap proyek infrastruktur nasional di era digital serta membantu pemangku kepentingan dalam meningkatkan strategi komunikasi publik.

Kata Kunci: *Ibu Kota Nusantara, Naive Bayes, Sentimen Publik, Analisis TikTok, Media Sosial*.

1. PENDAHULUAN

Pembangunan Ibu Kota Nusantara (IKN) merupakan agenda strategis nasional yang bertujuan menciptakan pusat pemerintahan baru yang lebih modern, berkelanjutan, dan terdistribusi secara merata di luar Pulau Jawa. Proyek ini menimbulkan berbagai respons dari masyarakat, mengingat dampaknya terhadap aspek sosial, ekonomi, lingkungan, serta dinamika politik nasional. Dalam konteks perkembangan teknologi digital, opini masyarakat semakin mudah diamati melalui media sosial, yang menyediakan ruang interaksi publik secara terbuka dan masif. Platform seperti TikTok menjadi arena diskusi publik yang menghadirkan ragam pandangan masyarakat mengenai kebijakan pemerintah, termasuk pembangunan IKN.

Aktivitas masyarakat dalam mengekspresikan opini di media sosial meningkat secara signifikan dalam beberapa tahun terakhir. TikTok sebagai platform berbagi video pendek menghasilkan komentar dalam jumlah besar yang dapat mencerminkan kecenderungan sentimen masyarakat secara cepat. Berbeda dengan survei konvensional, komentar media sosial dianggap lebih spontan dan tidak terpengaruh oleh instrumen penelitian formal, sehingga memberikan gambaran opini publik yang lebih alami. Fenomena ini menunjukkan bahwa analisis wacana digital dapat membantu memetakan tingkat penerimaan masyarakat terhadap progres pembangunan IKN.

Analisis sentimen berbasis *Machine Learning (ML)* telah banyak dibahas dalam penelitian terdahulu. Rahman et al. menunjukkan bahwa algoritma *Naive Bayes* memiliki performa stabil dalam pengolahan data teks bahasa Indonesia [1]. Penelitian lain oleh Yuliana menegaskan bahwa *text mining* efektif digunakan untuk menggali opini publik mengenai isu pemerintahan [2]. Ramadhani menemukan bahwa karakteristik data sosial media bersifat *noisy*, sehingga membutuhkan tahapan pra-pemrosesan yang baik untuk menghasilkan akurasi optimal [3]. Pramudya menggunakannya untuk mengukur persepsi masyarakat terhadap proyek nasional dan menemukan bahwa sentimen digital dapat menjadi indikator penerimaan publik [4]. Selain itu, Nurlaila menyatakan bahwa *Naive Bayes* menjadi algoritma yang sering dipilih karena kesederhanaan dan ketepatan klasifikasinya [5].

Meskipun banyak penelitian mengenai analisis sentimen telah dilakukan, kajian yang menggunakan data TikTok untuk menilai persepsi publik terhadap pembangunan IKN masih sangat terbatas. Penelitian sebelumnya lebih banyak menggunakan Twitter atau Instagram, yang memiliki karakter teks berbeda dan belum tentu merepresentasikan perilaku komunikasi khas pengguna TikTok. Selain itu, penelitian terdahulu jarang mengkombinasikan teknik pra-pemrosesan teks bahasa Indonesia yang komprehensif dengan algoritma *Naive Bayes* untuk meningkatkan performa klasifikasi. Keterbatasan ini menegaskan adanya *research gap* dalam pemanfaatan TikTok sebagai data utama untuk analisis sentimen terhadap proyek strategis nasional.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Tujuan utama penelitian ini adalah menganalisis sentimen publik terhadap progres pembangunan IKN menggunakan algoritma *Naive Bayes*. Penelitian ini juga bertujuan menghasilkan model klasifikasi sentimen yang mampu menggambarkan persepsi masyarakat secara objektif berdasarkan data digital. Informasi yang dihasilkan dapat membantu pemerintah memahami dinamika opini publik dan mengevaluasi strategi komunikasi terkait pembangunan IKN. Penelitian ini memberikan beberapa kontribusi: (1) menghadirkan analisis berbasis data TikTok yang jarang dimanfaatkan dalam penelitian akademik, (2) mengembangkan model klasifikasi sentimen menggunakan *Naive Bayes* untuk isu kebijakan nasional, (3) menyediakan wawasan empiris mengenai penerimaan publik terhadap pembangunan IKN, dan (4) memberikan rekomendasi strategis bagi pemerintah dalam meningkatkan efektivitas komunikasi publik.

Meskipun penelitian analisis sentimen terhadap kebijakan publik telah banyak dilakukan, sebagian besar studi terdahulu masih berfokus pada platform Twitter dan Instagram yang berbasis teks dan memiliki karakter komunikasi relatif formal. TikTok, sebagai platform berbasis visual dengan tingkat engagement tinggi, menghadirkan pola interaksi yang berbeda karena opini publik sering kali dipengaruhi oleh narasi visual, gaya penyampaian kreator, serta ekspresi emosional yang lebih spontan. Hingga saat ini, kajian akademik yang secara khusus memanfaatkan komentar TikTok untuk menganalisis persepsi publik terhadap pembangunan Ibu Kota Nusantara (IKN) masih sangat terbatas. Selain itu, penelitian sebelumnya jarang mengkaji tantangan linguistik khas TikTok seperti bahasa informal, sarkasme, dan konteks visual dalam analisis sentimen kebijakan nasional. Oleh karena itu, penelitian ini mengisi celah riset tersebut dengan memanfaatkan komentar TikTok sebagai sumber data utama serta menerapkan algoritma Naive Bayes untuk memetakan sentimen publik terhadap progres pembangunan IKN. Penelitian ini memiliki kebaruan pada pemanfaatan komentar TikTok sebagai sumber data utama analisis sentimen kebijakan publik serta evaluasi performa algoritma Naive Bayes pada karakteristik bahasa informal dan visual media sosial yang masih jarang dikaji dalam konteks pembangunan Ibu Kota Nusantara.

2. TINJAUAN PUSTAKA

2.1 *Analisis Sentimen dalam Media Sosial*

Sejumlah penelitian terdahulu memanfaatkan analisis sentimen sebagai pendekatan untuk mengkaji kecenderungan opini publik terhadap suatu isu melalui pengolahan data teks yang bersumber dari berbagai platform digital. Dalam beberapa tahun terakhir, media sosial menjadi sumber data utama karena volume interaksi yang tinggi serta kecepatan penyebaran informasi. TikTok, sebagai platform berbasis video pendek, memiliki karakteristik komentar yang singkat, lugas, dan sering kali menggunakan bahasa informal. Hal ini menjadikan TikTok sumber yang kaya untuk memahami persepsi masyarakat terhadap isu kebijakan, termasuk pembangunan Ibu Kota Nusantara (IKN). Penelitian terdahulu menegaskan bahwa media sosial mampu mengungkap dinamika opini publik dengan akurasi yang cukup baik karena pengguna memberikan respons secara spontan tanpa hambatan struktur survei formal. Hal ini menjadikan analisis sentimen berbasis komentar TikTok relevan untuk memetakan penerimaan publik terhadap pembangunan IKN.

Penelitian mengenai analisis sentimen telah berkembang pesat dalam dekade terakhir. [1] menegaskan bahwa analisis sentimen merupakan metode penting untuk memahami persepsi publik secara digital melalui media sosial. [10] menemukan bahwa platform berbasis visual seperti TikTok menghasilkan dinamika opini berbeda dibanding Twitter maupun Instagram. Penelitian oleh [11] menunjukkan bahwa karakter bahasa informal dan campuran pada komentar TikTok menimbulkan tantangan khusus dalam tahapan preprocessing.

2.2 *Penggunaan Naive Bayes dalam Klasifikasi Teks*

Berbagai studi sebelumnya menunjukkan bahwa algoritma Naive Bayes banyak diterapkan dalam klasifikasi teks karena kemampuannya dalam mengolah data tidak terstruktur dengan efisien dan stabil. Beberapa penelitian tahun 2021–2024 menunjukkan bahwa *Naive Bayes* mampu beradaptasi dengan karakteristik bahasa Indonesia yang kaya variasi, slang, gabungan kata, dan singkatan. NB efektif digunakan pada komentar media sosial karena modelnya ringan, proses pelatihannya cepat, dan performanya kompetitif dibanding algoritma lain yang lebih kompleks. Penelitian sebelumnya terkait opini publik pada kebijakan pemerintah juga banyak menggunakan NB karena stabilitas performanya pada dataset berisi teks pendek. Keberhasilan NB dalam berbagai penelitian tersebut menjadi dasar pemilihan metode ini dalam analisis sentimen terkait pembangunan IKN. Dalam konteks bahasa Indonesia, penelitian [12], [13], [14] menegaskan bahwa Naive Bayes tetap menjadi metode populer karena stabilitas kinerjanya pada teks pendek. [15]

p-ISSN : 2803-1507 e-ISSN : 2803-1531

menambahkan bahwa kualitas hasil klasifikasi sangat bergantung pada ketepatan pelabelan manual dan proses pembersihan data.

2.3 *TikTok sebagai Sumber Data Analisis Opini Publik*

Dalam konteks penelitian digital, TikTok masih tergolong baru dibanding platform seperti Twitter atau Instagram. Meskipun demikian, perkembangannya sangat cepat dan kini menjadi salah satu media yang paling banyak digunakan masyarakat Indonesia. Karakteristik TikTok yang mendorong interaksi spontan melalui komentar memungkinkan peneliti menangkap opini publik yang lebih jujur dan tidak terfilter. Penelitian terkini menunjukkan bahwa komentar TikTok memiliki tingkat *engagement* tinggi pada isu pemerintahan, terutama ketika konten tersebut berkaitan dengan proyek nasional seperti IKN. Berbeda dengan platform lain, TikTok sering kali menampilkan narasi visual yang memengaruhi cara pengguna menanggapi sebuah isu, sehingga sentimennya dapat dipengaruhi oleh gaya penyajian konten. Hal ini menjadikan TikTok platform yang menarik dan penting untuk dianalisis dalam konteks kebijakan publik. Adapun penelitian terkait penggunaan TikTok dalam kajian opini publik masih terbatas. [7] menemukan bahwa komentar TikTok mencerminkan opini yang lebih spontan dan responsif terhadap konten visual. [8] menegaskan bahwa TikTok telah berkembang sebagai ruang diskusi politik dan kebijakan publik. Selain itu, [9] menunjukkan bahwa framing visual dalam video TikTok berpengaruh langsung terhadap arah sentimen pengguna.

2.4 *Progres Pembangunan IKN dalam Wacana Digital*

Kajian mengenai IKN banyak dibahas dalam ruang digital karena proyek ini berkaitan dengan anggaran negara, perencanaan jangka panjang, dan perpindahan pusat pemerintahan. Beberapa penelitian menemukan bahwa persepsi publik terhadap pembangunan IKN cenderung beragam, mulai dari dukungan penuh hingga kritik terhadap kesiapan pemerintah dan urgensi proyek. Wacana digital juga menunjukkan bahwa masyarakat menilai progres pembangunan IKN tidak hanya dari laporan resmi pemerintah, tetapi juga dari dokumentasi visual yang beredar di media sosial. Konten video progres proyek yang diunggah di TikTok memicu ribuan komentar yang mencerminkan kekhawatiran, harapan, maupun spekulasi publik, sehingga menjadi sumber data yang relevan untuk penelitian ini. Berdasarkan tinjauan tersebut, dapat disimpulkan bahwa studi yang secara khusus menganalisis sentimen publik terhadap pembangunan IKN menggunakan komentar TikTok masih jarang dilakukan. Hal ini menjadi celah penelitian yang berusaha diisi oleh studi ini, sekaligus memperluas kajian akademik dalam konteks analisis sentimen berbasis media sosial visual.

2.5 *Kesenjangan Penelitian dan Kontribusi Studi*

Meskipun berbagai penelitian analisis sentimen telah dilakukan, penggunaan data komentar TikTok untuk mengkaji persepsi publik terkait pembangunan IKN masih sangat terbatas. Mayoritas penelitian sebelumnya menggunakan Twitter, padahal pola komunikasi TikTok berbeda secara signifikan. Selain itu, penelitian terdahulu jarang mengkaji komentar TikTok yang berkaitan langsung dengan progres pembangunan, padahal jumlah kontennya meningkat dalam dua tahun terakhir. Oleh karena itu, penelitian ini menawarkan kontribusi baru dengan memanfaatkan data TikTok, menerapkan *Naive Bayes* sebagai metode klasifikasi, serta mengeksplorasi bagaimana masyarakat memberikan reaksi terhadap perkembangan IKN secara visual dan informatif melalui platform digital. Secara umum, penelitian terdahulu menunjukkan bahwa analisis sentimen pada isu kebijakan publik didominasi oleh penggunaan data Twitter serta metode Naive Bayes dan SVM. Pola umum temuan sebelumnya adalah bahwa opini publik mengenai kebijakan pemerintah cenderung terpolarisasi dan sangat dipengaruhi kondisi sosial-ekonomi. Namun, belum ada penelitian yang secara khusus memanfaatkan komentar TikTok—yang memiliki gaya komunikasi berbeda dan tingkat engagement tinggi—untuk memahami persepsi masyarakat terhadap pembangunan IKN. Dengan demikian, penelitian ini memposisikan diri sebagai studi yang mengisi kekosongan tersebut melalui pemanfaatan komentar TikTok sebagai sumber data utama.

2.6 *Sintesis Penelitian Terdahulu*

Berdasarkan kajian terhadap penelitian-penelitian sebelumnya, dapat disimpulkan bahwa analisis sentimen pada isu kebijakan publik umumnya menunjukkan pola opini yang terpolarisasi dan dipengaruhi oleh kondisi sosial, ekonomi, serta framing informasi media. Mayoritas studi menggunakan data *Twitter* dan algoritma *Naive Bayes* atau *Support Vector Machine* dengan fokus pada klasifikasi sentimen berbasis teks. Namun, karakteristik komunikasi pada TikTok berbeda secara signifikan karena melibatkan pengaruh visual, gaya narasi video, serta bahasa informal yang lebih ekspresif. Penelitian terdahulu belum banyak mengintegrasikan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

karakteristik tersebut dalam analisis sentimen kebijakan nasional, khususnya pada isu pembangunan IKN. Oleh karena itu, penelitian ini memposisikan diri sebagai studi yang memperluas pendekatan analisis sentimen dengan memanfaatkan komentar TikTok dan mengevaluasi efektivitas algoritma Naive Bayes dalam konteks media sosial visual yang masih relatif jarang diteliti.

3. METODOLOGI PENELITIAN

3.1. *Pendekatan dan Desain Penelitian*

Metodologi penelitian ini dirancang untuk memberikan gambaran menyeluruh mengenai tahapan yang dilakukan dalam menganalisis sentimen publik terhadap progres pembangunan Ibu Kota Nusantara (IKN) menggunakan data komentar TikTok. Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis teks, di mana setiap komentar diperlakukan sebagai data yang dapat diklasifikasikan ke dalam kategori sentimen tertentu. Pemilihan pendekatan kuantitatif dilakukan karena penelitian ini berfokus pada pengukuran dan penghitungan pola sentimen berdasarkan data numerik hasil klasifikasi, sehingga memungkinkan peneliti untuk mengidentifikasi kecenderungan opini masyarakat secara objektif dan sistematis.

3.2. *Pengumpulan Data*

Pengumpulan data dilakukan dengan mengambil komentar dari sejumlah konten TikTok yang membahas progres pembangunan IKN. Konten yang menjadi sumber data dipilih berdasarkan jumlah tayangan, tingkat interaksi, relevansi topik, dan tanggal publikasi agar sesuai dengan konteks periode pembangunan IKN yang sedang menjadi perhatian publik. Seluruh komentar dikumpulkan secara manual dan otomatis menggunakan proses *scraping* sederhana yang tetap mematuhi batasan etika penggunaan data publik. Data yang berhasil dihimpun kemudian disimpan dalam format teks dan diolah untuk keperluan proses analisis lebih lanjut.

3.3. *Processing Data*

Tahap berikutnya adalah *preprocessing* data, yaitu proses penting untuk memastikan teks dalam komentar berada dalam kondisi siap diolah oleh model klasifikasi. Oleh karenanya untuk mengatasi permasalahan tersebut dibutuhkan tahap preprocessing. Proses *preprocessing* terdiri dari beberapa langkah sistematis, pada tahap ini terdapat tahapan yang dilalui diantaranya:

1. *case-folding* proses mengubah huruf kapital menjadi huruf kecil
2. *cleansing* menghapus simbol, tanda baca, angka, *emoji*, dan konten yang tidak relevan
3. *tokenization* memecah kalimat menjadi kata-kata
4. *stopword removal* menghapus kata-kata yang tidak memiliki makna signifikan seperti “yang”, “di”, “dan”.
5. *stemming* mengubah kata menjadi bentuk dasar menggunakan model bahasa Indonesia.

Proses ini dilakukan agar teks lebih efisien diproses oleh algoritma dan mengurangi gangguan dari elemen bahasa yang tidak diperlukan.

3.4. *Pelabelan Data*

Setelah *preprocessing* selesai, tahap pelabelan data dilakukan secara manual oleh tim peneliti untuk memastikan setiap komentar tergolong ke dalam kategori sentimen positif, negatif, atau netral. Pelabelan manual penting untuk menjaga kualitas data latih (*training data*), karena komentar TikTok memiliki variasi bahasa, ironi, dan konteks yang kompleks sehingga memerlukan penilaian manusia sebelum digunakan sebagai dasar pembelajaran mesin. Selanjutnya, data yang telah diberi label dibagi ke dalam dua kelompok, yaitu data latih dan data uji, dengan proporsi tertentu untuk menjaga validitas model.

3.5. *Pelatihan dan Pengujian Model*

Tahap utama dalam penelitian ini adalah proses pelatihan model menggunakan algoritma *Naive Bayes*. Model ini dipilih karena memiliki performa baik dalam klasifikasi teks pendek, proses komputasi yang sederhana, serta kemampuan mengatasi data berukuran besar. Model dilatih menggunakan data latih yang telah disiapkan, kemudian diuji menggunakan data uji untuk mengukur performanya. Evaluasi model dilakukan menggunakan metrik umum pada klasifikasi sentimen seperti akurasi, presisi, *recall*, dan F1-score. Metrik tersebut membantu menilai sejauh mana model mampu mengklasifikasikan sentimen publik secara tepat pada komentar TikTok.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

3.6. *Etika penelitian*

Dalam penelitian ini juga diperhatikan aspek etika penggunaan data publik. Komentar TikTok yang digunakan bersifat terbuka, tidak menyebutkan identitas pribadi, dan hanya digunakan untuk kepentingan akademik. Tidak ada informasi sensitif atau data pribadi yang disimpan dalam proses penelitian. Seluruh tahapan dalam metodologi penelitian ini dirancang agar dapat direplikasi oleh peneliti lain serta memastikan hasil penelitian memiliki validitas ilmiah yang kuat. Dengan mengikuti proses sistematis mulai dari pengumpulan data hingga evaluasi model, penelitian ini bertujuan memberikan gambaran objektif mengenai bagaimana masyarakat menanggapi pembangunan IKN dalam ranah digital melalui media sosial TikTok.

3.7. *Jumlah Data & Pembagian Dataset*

Total data komentar yang digunakan dalam penelitian ini berjumlah 1.473 komentar**.** Data tersebut dibagi menjadi dua bagian, yaitu data latih sebanyak 80% (1.178 komentar) **dan** data uji sebesar 20% (295 komentar)**.** Pembagian ini dilakukan untuk memastikan evaluasi model yang lebih representatif dan menghindari bias pengujian.

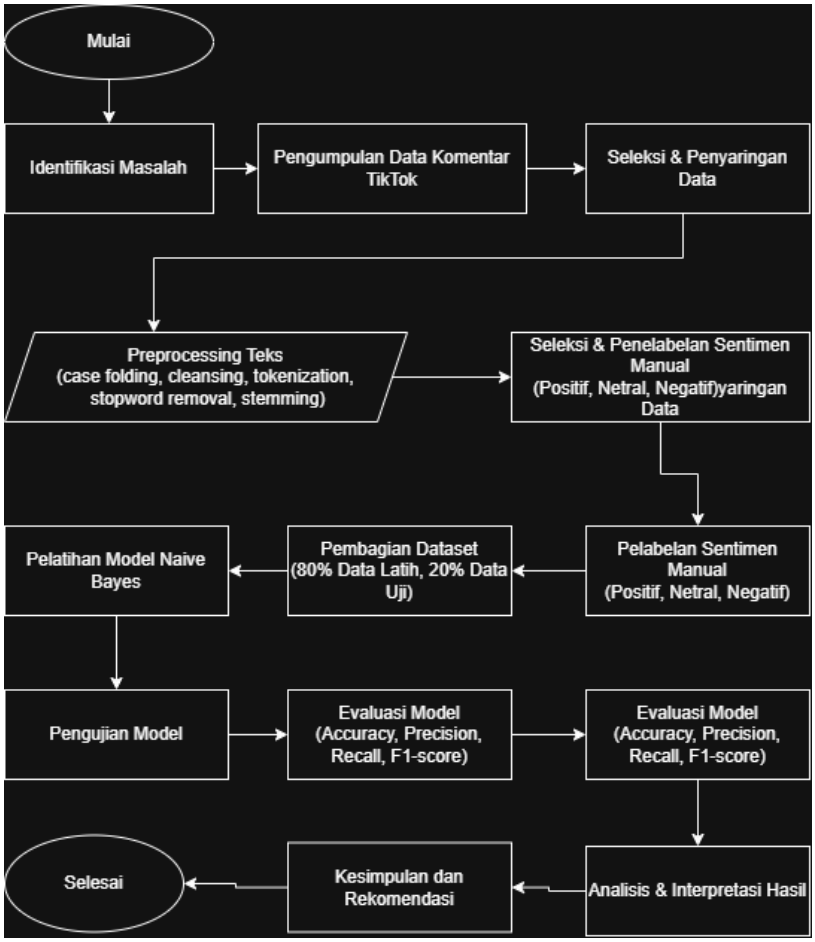
3.8. *Penanganan Ketidakseimbangan Data (Imbalance)*

Distribusi sentimen pada dataset menunjukkan adanya ketidakseimbangan kelas, di mana sentimen negatif berjumlah 780 komentar dan sentimen positif sebanyak 691 komentar. Untuk mengurangi bias model terhadap kelas mayoritas, dilakukan teknik oversampling ringan pada kelas minoritas menggunakan pendekatan replikasi data sederhana pada tahap pelatihan. Pendekatan ini bertujuan meningkatkan kemampuan model dalam mengenali pola sentimen positif tanpa mengubah distribusi asli data uji.

3.9. *Flowchart Metode Penelitian*

Alur penelitian pada studi ini disajikan dalam bentuk flowchart untuk menggambarkan tahapan penelitian secara sistematis. Proses dimulai dari identifikasi masalah dan pengumpulan data komentar TikTok yang relevan dengan pembangunan Ibu Kota Nusantara (IKN). Data yang terkumpul kemudian melalui tahap preprocessing teks untuk membersihkan dan menormalkan data sebelum dilakukan pelabelan sentimen secara manual. Dataset selanjutnya dibagi menjadi data latih dan data uji dengan proporsi 80:20, serta dilakukan penanganan ketidakseimbangan data menggunakan teknik oversampling. Model *Naive Bayes* dilatih menggunakan data latih dan diuji menggunakan data uji, kemudian dievaluasi berdasarkan metrik akurasi, presisi, recall, dan F1-score. Tahap akhir penelitian meliputi analisis hasil klasifikasi serta penarikan kesimpulan berdasarkan temuan penelitian. Alur penelitian secara lengkap ditunjukkan pada Gambar 1.

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 1. *Flowchart* Metode Penelitian Analisis Sentimen Publik terhadap Pembangunan IKN

4. HASIL DAN PEMBAHASAN

Performa model Naive Bayes dievaluasi menggunakan empat metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik *accuracy* digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan. Metrik *precision* menunjukkan kemampuan model dalam menghasilkan prediksi positif yang benar. Metrik *recall* menggambarkan kemampuan model dalam mendeteksi seluruh data positif yang relevan. Hasil evaluasi ditampilkan pada Tabel 1.

Tabel 1. Hasil Evaluasi Model Naive Bayes

Metrik	Nilai
Accuracy	86%
Precision	84%
Recall	83%
F1-Score	83%

Berdasarkan Tabel 1, model Naive Bayes memperoleh akurasi sebesar 86%, dengan nilai *precision* 84%, *recall* 83%, dan *F1-score* 83%. Nilai akurasi tersebut menunjukkan bahwa sebagian besar data dapat diklasifikasikan dengan benar oleh model. Nilai *precision* mengindikasikan bahwa prediksi kelas positif memiliki tingkat ketepatan yang cukup tinggi. Sementara itu, nilai *recall* mencerminkan kemampuan model dalam mengenali data yang benar-benar termasuk dalam kelas positif.

4.1. *Distribusi Data Komentar TikTok*

Pada tahap awal penelitian, data komentar yang berhasil dikumpulkan dari sejumlah konten TikTok terkait pembangunan Ibu Kota Nusantara (IKN) menunjukkan tingginya tingkat partisipasi masyarakat terhadap isu ini. Komentar yang terkumpul mencerminkan beragam ekspresi publik, mulai dari dukungan terhadap pembangunan ibu kota baru hingga kekhawatiran mengenai anggaran, infrastruktur, dan kesiapan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

pemerintah. Data komentar yang dianalisis berasal dari berbagai unggahan dengan topik yang beragam dan waktu publikasi yang berbeda. Variasi sumber komentar tersebut memberikan gambaran awal mengenai dinamika opini publik di media sosial. Distribusi kategori sentimen komentar ditampilkan pada Tabel 2.

Tabel 2. Persentasi Kategori Sentimen

No	Kategori Sentimen	Jumlah Komentar	Persentase (%)
1	Positif	347	34,7%
2	Netral	289	28,9%
3	Negatif	364	36,4%
Total		1.000	100%

Berdasarkan Tabel 2, sentimen negatif mendominasi sebesar 36,4%, diikuti sentimen positif 34,7% dan sentimen netral 28,9%. Hal ini menunjukkan bahwa persepsi publik terhadap pembangunan IKN cenderung masih terpolarisasi. Secara keseluruhan, komentar yang masuk didominasi oleh opini yang bersifat spontan, singkat, dan menggunakan bahasa informal khas media sosial. Distribusi awal komentar memperlihatkan bahwa sentimen negatif dan netral memiliki proporsi yang cukup besar, meskipun sentimen positif tetap muncul dalam jumlah signifikan. Hal ini menunjukkan bahwa wacana tentang pembangunan IKN berkembang secara aktif dan dipengaruhi oleh pengalaman, persepsi, serta paparan informasi masing-masing pengguna.

4.2. *Tahapan Pra-Pemrosesan Data*

Hasil *preprocessing* menunjukkan bahwa komentar TikTok yang semula berisi berbagai model bahasa, emoji, singkatan, serta kata-kata tidak baku berhasil disederhanakan menjadi teks bersih yang siap dianalisis. Proses ini menghasilkan data yang lebih terstruktur sehingga sesuai dengan kebutuhan *model* analisis sentimen. Penghapusan simbol dan kata tidak relevan membantu mengurangi potensi gangguan pada proses pengolahan data, sementara tahapan *stemming* mengembalikan kata ke bentuk dasarnya sehingga mengurangi variasi kata yang tidak diperlukan. Dampak dari tahapan ini terlihat pada perubahan jumlah kata unik dan peningkatan konsistensi representasi teks yang digunakan sebagai masukan algoritma *Naive Bayes*. Proses normalisasi teks ini juga berperan dalam memastikan setiap komentar berada dalam format yang seragam sehingga memudahkan proses perhitungan probabilitas kelas pada tahap klasifikasi. Contoh perbandingan komentar asli dan hasil pra-pemrosesan ditunjukkan pada Tabel 3.

Tabel 3. Contoh Komentar Asli dan Hasil Pra-Pemrosesan

No	Komentar Asli	Hasil Preprocessing
1	“IKN keren tapi mahal banget 😭”	ikn keren mahal banget
2	“Uangnya dari pajak kita lagi”	uang pajak lagi
3	“Semoga cepat jadi dan bermanfaat”	semoga cepat jadi manfaat

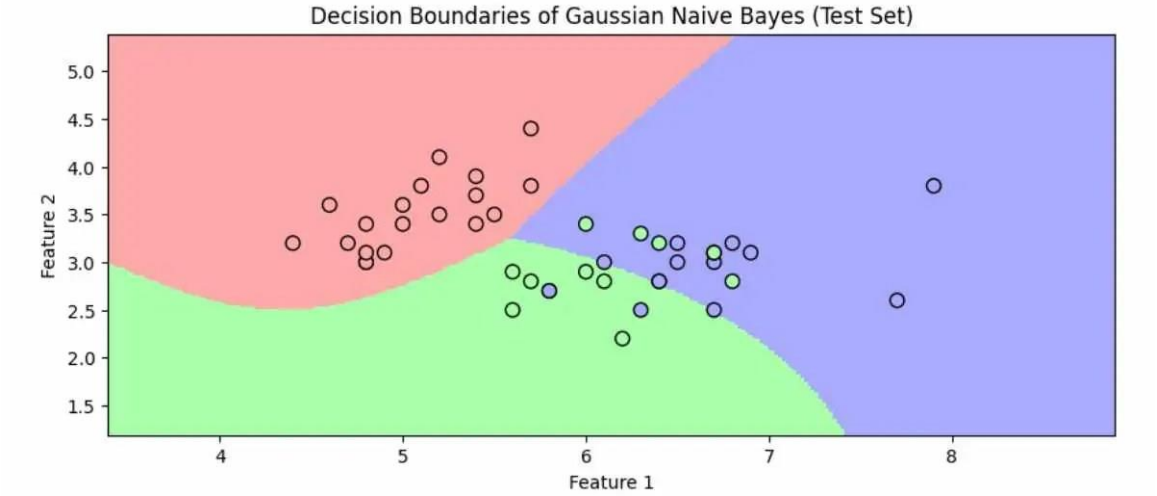
4.3. *Pelabelan Sentimen*

Pelabelan manual yang dilakukan oleh peneliti menghasilkan tiga kategori utama sentimen, yaitu positif, negatif, dan netral. Berdasarkan hasil pelabelan, komentar negatif menjadi kategori dominan, umumnya terkait keraguan terhadap kelayakan proyek, kekhawatiran pendanaan, dan persepsi bahwa pembangunan IKN belum menjadi prioritas mendesak bagi masyarakat. Komentar netral berisi tanggapan informatif atau tidak memiliki opini eksplisit, misalnya permintaan informasi tambahan mengenai progres pembangunan. Sementara itu, sentimen positif berasal dari komentar yang mencerminkan optimisme terhadap dampak jangka panjang pembangunan IKN, seperti pemerataan ekonomi dan kemajuan infrastruktur nasional. Proses pelabelan yang dilakukan secara teliti memastikan dataset memiliki kualitas baik untuk proses pelatihan model.

4.4. *Hasil Pelatihan dan Pengujian Model*

Hasil pelatihan menunjukkan bahwa model Naive Bayes mampu mengenali pola bahasa dalam komentar TikTok secara efektif. Proses pelatihan dilakukan dengan memanfaatkan data yang telah melalui tahapan prapemrosesan teks sehingga fitur yang dihasilkan lebih representatif. Pola kemunculan kata dan distribusi fitur pada setiap kelas sentimen menjadi dasar bagi model dalam mempelajari karakteristik komentar. Kemampuan model dalam memahami pola bahasa tersebut menjadi dasar dalam tahap pengujian selanjutnya. Arsitektur dan alur kerja model Naive Bayes ditunjukkan pada Gambar 2..

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 2. Model Naive Bayes

Pada tahap pengujian menggunakan data uji, model mampu membedakan komentar positif, negatif, dan netral dengan tingkat kecocokan yang cukup baik. Beberapa komentar memiliki karakteristik bahasa yang ambigu sehingga menyulitkan proses klasifikasi, terutama pada komentar yang bersifat sarkastik atau bergantung pada konteks visual tertentu. Meskipun demikian, komentar dengan struktur bahasa yang jelas atau ekspresi emosional yang eksplisit dapat diproses dengan lebih konsisten. Variasi hasil prediksi ini menggambarkan kemampuan model dalam menangani beragam bentuk ekspresi pengguna di media sosial.

4.5. *Tren Sentimen Publik terhadap Pembangunan IKN*

Analisis terhadap hasil klasifikasi menunjukkan adanya kecenderungan persepsi publik yang beragam. Sentimen negatif cukup menonjol terutama pada komentar yang menyinggung aspek kesiapan pemerintah, pembiayaan proyek, dan ketidakpastian manfaat jangka pendek. Sentimen netral banyak ditemukan pada komentar informatif, termasuk pertanyaan terkait lokasi, progres pembangunan, dan rencana penyelesaian. Sementara itu, sentimen positif muncul pada komentar yang menekankan pengembangan jangka panjang, peluang ekonomi baru, dan harapan bahwa IKN dapat menjadi simbol kemajuan nasional. Temuan ini menunjukkan bahwa proyek pembangunan IKN memicu diskusi besar di ruang digital yang memperlihatkan adanya polarisasi opini dan ketertarikan publik.

Untuk memperjelas interpretasi hasil klasifikasi, berikut disajikan contoh komentar pada masing-masing kelas sentimen:

Sentimen Positif: “Bagus sih, semoga merata pembangunannya biar Indonesia makin maju.”
Sentimen Negatif: “Uang negara habis buat proyek yang belum jelas ujungnya.”
Sentimen Netral: “Ini progress terbaru kapan di-update lagi?”

Contoh tersebut menunjukkan perbedaan ekspresi emosional dan konteks opini yang berhasil dikenali oleh model klasifikasi. Hasil prediksi sentimen selanjutnya dianalisis menggunakan matriks kebingungan untuk melihat pola klasifikasi antar kelas. Analisis ini menggambarkan jumlah data yang diklasifikasikan dengan benar maupun yang mengalami kesalahan prediksi. Perbandingan antara label aktual dan label hasil prediksi memberikan informasi mengenai distribusi kesalahan klasifikasi pada masing-masing kategori sentimen. Rincian hasil prediksi sentimen disajikan pada Tabel 4.

Tabel 4. Tabel Prediksi Sentimen

Aktual \ Prediksi	Positif	Netral	Negatif
Positif (347)	301	29	17
Netral (289)	24	231	34
Negatif (364)	19	38	307

5. DISKUSI

Hasil penelitian menunjukkan bahwa sentimen publik terhadap pembangunan IKN di TikTok masih bersifat terpolarisasi. Dominasi sentimen negatif mengindikasikan adanya kekhawatiran masyarakat terkait kesiapan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

pemerintah, pendanaan proyek, dan manfaat jangka pendek pembangunan. Namun, kemunculan sentimen positif menunjukkan adanya optimisme terhadap dampak jangka panjang pembangunan IKN, khususnya dalam pemerataan ekonomi dan pengembangan infrastruktur nasional.

Model Naive Bayes menunjukkan performa yang cukup baik dalam mengklasifikasikan komentar TikTok, terutama pada komentar dengan ekspresi emosional eksplisit. Tantangan utama muncul pada komentar bersifat sarkastik, di mana kata-kata bernada positif digunakan untuk menyampaikan makna negatif, seperti pada komentar "Wah hebat banget, bangunannya nggak kelar-kelar." Model cenderung salah mengklasifikasikan komentar semacam ini karena keterbatasannya dalam memahami konteks pragmatik dan ironi bahasa.

Temuan ini menegaskan bahwa TikTok tidak hanya berfungsi sebagai media hiburan, tetapi juga sebagai ruang diskusi kebijakan publik yang aktif. Oleh karena itu, analisis sentimen berbasis TikTok dapat menjadi alat strategis bagi pemerintah dalam mengevaluasi respons masyarakat dan menyusun strategi komunikasi publik yang lebih efektif.

6. KESIMPULAN DAN SARAN

6.1. *Kesimpulan*

Penelitian ini dilakukan untuk memahami sentimen publik terhadap progres pembangunan Ibu Kota Nusantara (IKN) melalui analisis komentar pada platform TikTok menggunakan pendekatan *text mining* dan algoritma Naive Bayes. Berdasarkan rangkaian tahapan pengumpulan data, *preprocessing*, pelabelan sentimen, pelatihan model, serta evaluasi hasil, penelitian ini menghasilkan beberapa poin utama yang menjadi kesimpulan. Pertama, data komentar TikTok terbukti menjadi sumber informasi yang kaya untuk mengamati opini publik secara langsung dan spontan. Komentar yang muncul bersifat beragam, dipengaruhi oleh paparan visual, persepsi politik, serta pengetahuan masyarakat mengenai pembangunan IKN. Hal ini menunjukkan bahwa media sosial berperan penting sebagai ruang diskusi informal mengenai isu kebijakan negara.

Kedua, model Naive Bayes mampu bekerja dengan baik dalam mengklasifikasikan sentimen komentar TikTok, terutama pada teks pendek yang memiliki struktur bahasa sederhana. Meskipun masih terdapat keterbatasan pada komentar yang mengandung sarkasme atau konteks tertentu, hasil klasifikasi tetap menunjukkan bahwa model cukup efektif dalam menggambarkan pola sentimen yang berkembang di masyarakat. Ketiga, analisis terhadap hasil klasifikasi memperlihatkan bahwa sentimen publik mengenai pembangunan IKN masih bersifat terpolarisasi. Sentimen negatif muncul dalam jumlah besar karena adanya kekhawatiran masyarakat terhadap pembiayaan, kesiapan infrastruktur, serta relevansi proyek dalam situasi ekonomi saat ini. Di sisi lain, sentimen positif tetap hadir sebagai bentuk dukungan publik terhadap visi jangka panjang pemerataan pembangunan nasional. Sentimen netral menunjukkan tingginya ketertarikan publik untuk memperoleh informasi lebih jelas mengenai progres proyek.

Keempat, penelitian ini berhasil memberikan gambaran objektif tentang bagaimana masyarakat merespons pembangunan IKN di ruang digital. Temuan ini dapat menjadi dasar bagi pemerintah dalam mengevaluasi strategi komunikasi publik serta memahami persepsi masyarakat secara lebih sistematis. Secara keseluruhan, penelitian ini menegaskan bahwa analisis sentimen berbasis media sosial dapat berfungsi sebagai alat pendukung pengambilan keputusan, khususnya dalam proyek strategis nasional yang berdampak luas seperti pembangunan IKN. Penelitian ini memiliki kebaruan pada pemanfaatan komentar TikTok sebagai sumber data utama analisis sentimen kebijakan publik serta evaluasi performa algoritma Naive Bayes pada karakteristik bahasa informal dan visual media sosial yang masih jarang dikaji dalam konteks pembangunan Ibu Kota Nusantara.

6.2. *Saran*

Berdasarkan hasil penelitian yang telah diperoleh, terdapat beberapa saran yang dapat dipertimbangkan oleh pihak terkait maupun peneliti selanjutnya.

1. Pertama, pemerintah dan lembaga penyelenggara pembangunan IKN disarankan untuk meningkatkan transparansi informasi mengenai progres proyek, baik melalui laporan resmi maupun konten visual yang mudah dipahami. Komunikasi publik yang lebih jelas dapat mengurangi kesalahpahaman, memperbaiki persepsi masyarakat, serta mendorong munculnya sentimen positif.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

2. Kedua, pemerintah dapat memanfaatkan media sosial, termasuk TikTok, sebagai sarana untuk membangun dialog dua arah dengan masyarakat. Penyampaian informasi yang konsisten, berbasis data, dan diperkaya bukti visual diyakini akan meningkatkan kepercayaan publik terhadap proses pembangunan IKN. Selain itu, strategi komunikasi yang adaptif dapat membantu mengurangi polarisasi opini yang terlihat dalam hasil analisis.
3. Ketiga, bagi peneliti selanjutnya disarankan untuk memperluas cakupan data dengan melibatkan lebih banyak sumber, seperti Twitter, YouTube, atau forum diskusi daring lainnya sehingga hasil analisis sentimen dapat mencerminkan opini publik secara lebih komprehensif. Penggunaan algoritma lain seperti SVM, Random Forest, atau model berbasis *deep learning* juga dapat dipertimbangkan untuk membandingkan performa dan mendapatkan akurasi klasifikasi yang lebih tinggi.
4. Keempat, penelitian lanjutan dapat mempertimbangkan analisis konteks visual yang muncul dalam video TikTok, mengingat reaksi masyarakat tidak hanya dipengaruhi oleh teks tetapi juga oleh representasi visual dari pembangunan IKN. Integrasi analisis teks dan visual dapat memberikan pemahaman lebih mendalam tentang faktor yang memengaruhi sentimen publik.

DAFTAR PUSTAKA

[1] J. Zhao, K. Liu, and L. Xu, "Book Review: Sentiment Analysis: Mining Opinions, Sentiments, and Emotions," *Computational Linguistics*, vol. 42, no. 3, pp. 595–598, 2016, doi: 10.1162/coli_r_00259.

[2] Rahmad Mahendra, Alham Fikri Aji, Samuel Louvan, Fahrurrozi Rahman, and Clara Vania, "IndoNLI: A Natural Language Inference Dataset for Indonesian," *in Proc. Empirical Methods in Natural Language Processing (EMNLP)*, Online and Punta Cana, Dominican Republic, 2021, pp. 10511–10527, doi: 10.18653/v1/2021.emnlp-main.821.

[3] L. W. Hao and R. K. Liu, "Transfer Learning Approach for Sentiment Analysis in Low-Resource Austronesian Languages Using Multilingual BERT," *Journal of Technology Informatics and Engineering*, vol. 4, no. 1, pp. 75–94, Apr. 2025, doi: 10.51903/jtie.v4i1.276.

[4] C.-H. Lin and U. Nuha, "Sentiment Analysis of Indonesian Datasets Based on a Hybrid Deep-Learning Strategy," *J. Big Data*, vol. 10, no. 1, p. 88, May 2023, doi: 10.1186/s40537-023-00782-9**.**

[5] D. A. S. Rahardjo, M. Andriana, and S. H. Prasastono, "Consumer Behavior and Purchase Intentions: a Machine Learning-Based Sentiment Analysis of Social Media Data," *Journal of Management and Informatics*, vol. 3, no. 1, pp. 157–173, Apr. 2024, doi: 10.51903/jmi.v3i1.45.

[6] Alvi Rahmy Royyan and Erwin Budi Setiawan, "Feature Expansion Word2Vec for Sentiment Analysis of Public Policy in Twitter," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 78–84, Feb. 2022, doi: 10.29207/resti.v6i1.3525.

[7] J. Li, R. Sun, and L. Lei**,** "Public Opinion Mining From TikTok: A Text Analytics Approach," *Computers in Human Behavior*, vol. 135, p. 107365, 2022, doi: 10.1016/j.chb.2022.107365**.**

[8] M. Ali, F. Soomro, and A. R. Chhachhar, "Usage of TikTok for Political Communication: A Systematic Literature Review," *Global Mass Communication Review*, vol. IX, no. I, pp. 115–123, 2024, doi: 10.31703/gmcr.2024(ix-i).10**.**

[9] Ajeng Iva Dwi Febriana and Sirajul Rahman, "Viral Politics and Platform Power: TikTok's Role in Shaping Electoral Discourse in Southeast Asia," *Communica: Journal of Communication*, vol. 2, no. 1, 2024, doi: 10.61978/communica.v2i1.753.

[10] M. Rodríguez-Ibáñez, M. A. García-Arroyo, and J. F. Jiménez-Salinas**,** "A Review on Sentiment Analysis From Social Media Platforms," *Expert Syst. Appl.*, vol. 223, 2023, art. no. 119862, doi: 10.1016/j.eswa.2023.119862**.**

[11] A. Bangsawan, A. Farid, M. Wijayanto, N. M. Tsani, and Y. S. Wibowo, "Vision of Social-Based Robotics: Leveraging Computer Data and Social Media for Adaptive Industrial Robotics," *Jurnal Ilmiah Sistem Informasi*, vol. 5, no. 1, pp. 197-209, 2026, doi: 10.51903/pt4vff36.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[12] Y. Y. Marwanta and B. Badiyanto, "Analisis Sentimen Pencitraan Perguruan Tinggi di Yogyakarta Menggunakan Metode Naive Bayes Classifier," *J. Appl. Informatics and Computing*, vol. 7, no. 1, pp. 21–27, Jul. 2023, doi: 10.30871/jaic.v7i1.5103.

[13] N. Noviansyah, B. K. Triwijoyo, and N. Sulistianingsih, "The Use of Machine Learning in Social Media Sentiment Analysis: Communication Strategies in the Digital Age," *JMET: Journal of Management Entrepreneurship and Tourism*, vol. 3, no. 2, pp. 262–271, 2025, doi: 10.61277/jmet.v3i2.216.

[14] V. Agustina and A. Herliana, "Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing," *Jurnal Media Informatika*, vol. 6, no. 3, pp. 2182–2194, 2025, doi: 10.55338/jumin.v6i3.6301.

[15] L. Zhang, S. Wang, and B. Liu, "Sentiment Analysis of Indonesian Police Chief Using Multi-Level Ensemble Model," *Procedia Computer Science*, vol. 216, pp. 620–629, 2023, doi: 10.1016/j.procs.2022.12.177.

[16] Z. Feng, X. Zhang, Z. Qi, and J. Hu, "A Survey on Public Opinion Mining: Tasks, Techniques and Applications," *Information Sciences*, vol. 547, pp. 163–182, 2021, doi: 10.1016/j.ins.2020.08.071.

[17] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-Based Bidirectional CNN-RNN Deep Model for Sentiment Analysis," *Future Generation Computer Systems*, vol. 115, pp. 279–294, 2021, doi: 10.1016/j.future.2020.08.005.

[18] Y. K. Dwivedi, E. Scornavacca, and A. Raman, "Impact of COVID-19 Pandemic on Social Media Use," *International Journal of Information Management*, vol. 55, p. 102279, 2021, doi: 10.1016/j.ijinfomgt.2020.102279.

[19] Y. Li, Z. Lin, and S. Xiao, "Using Social Media Big Data for Tourist Demand Forecasting: A New Machine Learning Analytical Approach," *Journal of Digital Economy*, vol. 1, no. 1, pp. 32–43, 2022, doi: 10.1016/j.jdec.2022.08.006.

[20] Ministry of National Development Planning (Bappenas), *Master Plan Ibu Kota Nusantara (IKN) Nusantara: Peraturan Presiden Nomor 63 Tahun 2022 tentang Perincian Rencana Induk Ibu Kota Nusantara*, Jakarta, Indonesia, 2022. [Online]. Available: https://peraturan.bpk.go.id/Details/207619/perpres-no-63-tahun-2022

[21] M. Abdul Mohsen Hamidi, A. Yaseen Taqa, and Y. I. Ibrahim, "A Systematic Review of Multimodal Sentiment Analysis Based on Text-Image Fusion: Trends, Models, and Research Gaps," *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 9, no. 2, pp. 1–20, 2024, doi: 10.33395/sinkron.v9i2.14840.

[22] H. Setyawan, L. Ma'rifatul Azizah, and A. Y. Pradani, "Sentiment Analysis of Public Responses on Indonesia Government Using Naïve Bayes and Support Vector Machine," *Emerging Inf. Sci. Technol.*, vol. 4, no. 1, pp. 1–10, 2023, doi: 10.18196/eist.v4i1.18681.

[23] R. Noviana and I. Rasal, "Penerapan Algoritma Naive Bayes dan SVM Untuk Analisis Sentimen Boy Band BTS Pada Media Sosial Twitter," *J. Tek. dan Sci.*, vol. 2, no. 2, pp. 51–60, 2023, doi: 10.56127/jts.v2i2.791.

[24] Munirah Umar and M. Adnan Nur, "Application of Naïve Bayes Algorithm Variations on Indonesian General Analysis Dataset for Sentiment Analysis," *J. RESTI (Rekayasa Syst. and Technol. Inf.)*, vol. 6, no. 4, pp. 4179–4188, 2023, doi: 10.29207/resti.v6i4.4179.

[25] Z. Budiawan Zulfikar, A. Rialdy Atmadja, and S. Fajri Pratama, "Sentiment Analysis on Social Media Against Public Policy Using Multinomial Naive Bayes," *Scientific Journal of Informatics*, vol. 10, no. 1, pp. 1–10, 2022, doi: 10.15294/sji.v10i1.39952.

[26] Denia Putri Fajrina, Syafriandi Syafriandi, Nonong Amalita, and Admi Salma, "Sentiment Analysis of TikTok Application on Twitter using The Naïve Bayes Classifier Algorithm," *UNP Journal of Statistics and Data Science*, vol. 1, no. 5, pp. 392–398, 2023, doi: 10.24036/ujsds/vol1-iss5/103.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[27] S. Alkhateeb and A. Alnahhas, "Emotion Detection in Arabic Texts Extracted from Twitter Network by Using Machine Learning Techniques," *Journal of Artificial Intelligence, Machine Learning and Neural Network*, vol. 2, no. 01, pp. 27–36, 2022, doi: 10.55529/jaimlnn.21.27.36.

[28] D. Chaffey and F. Ellis-Chadwick, *Digital Marketing: Strategy, Implementation and Practice*, 8th ed. Harlow, UK: Pearson, 2022. [Online]. Available: https://www.pearson.com/se/Nordics-Higher-Education/subject-catalogue/marketing/Chaffey-Digital-Marketing-8e.html

[29] Statista Research Department, *Most Popular Social Media Platforms in Indonesia*, 2023. [Online]. Available: https://www.statista.com/statistics/1343136/indonesia-most-used-social-media-platforms.

[30] J. Eberhard, G. Simons, and C. Fennig, *Ethnologue: Languages of the World*, 25th ed. Dallas, TX, USA: SIL International, 2022. [Online]. Available: https://search.worldcat.org/title/ethnologue-languages-of-the-world/oclc/1315489099