

Classification Model Evaluation of Public Perception Regarding Nusantara Capital Relocation

Galih Sabilah Rosat^{1*}, Nova Rijati^{2**}

* Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
111202214290@mhs.dinus.ac.id¹, nova.rijati@dsn.dinus.ac.id²

Article Info

Article history:

Received 2026-04-11

Revised 2026-05-04

Accepted 2026-05-25

Keyword:

*Sentiment Analysis,
Capital City of Nusantara,
Multinomial Naïve Bayes,
SMOTE,
X (Twitter).*

ABSTRACT

The relocation of the Capital City of Nusantara (IKN) has triggered extensive public opinion dynamics on the X (Twitter) social media platform. This study aims to empirically map public sentiment and evaluate the comparative performance of classification machine learning algorithms: Support Vector Machine (SVM), Logistic Regression (LR), and Multinomial Naïve Bayes (MNB). The dataset comprises 2,257 unique texts spanning the 2024-2025 period, automatically categorized using a lexicon-based approach (InSet dictionary). The technical novelty of this research lies in the extraction of TF-IDF N-gram (1, 2) features, the mitigation of extreme class imbalance via the Synthetic Minority Over-sampling Technique (SMOTE) encapsulated within a cross-validation pipeline, and systematic hyperparameter optimization using GridSearchCV. The labeling results revealed a predominance of negative sentiment at 77.3% (1,745 data) compared to positive sentiment at 22.7% (512 data), reflecting high public scrutiny and skepticism. Based on the model evaluation, MNB with parameter optimization ($\alpha = 0.1$) emerged as the most robust model, achieving an accuracy of 80.5% and a Macro Precision of 72.3%. The performance of MNB proved capable of outperforming the architectures of LR and SVM, which struggled with multicollinearity and underfitting anomalies in high-dimensional spaces. These findings emphasize that the integration of SMOTE and the MNB algorithm is crucial in overcoming majority class bias and calibrating the limitations of static lexicons to precisely map the polarity of public opinion.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Pemindahan Ibu Kota Negara ke Kalimantan Timur merupakan salah satu kebijakan strategis pemerintah Indonesia yang memicu transformasi signifikan dalam lanskap sosial, ekonomi, dan politik nasional. Kebijakan ini tidak hanya berdampak pada redistribusi pusat pemerintahan, tetapi juga diproyeksikan membawa perubahan mendasar pada struktur masyarakat, gaya hidup, dan aspek sosial-budaya di wilayah baru [1]. Perpindahan ini membuka peluang terbentuknya entitas kebudayaan baru yang merupakan hasil akulturasi, sebagaimana pernah terjadi dalam sejarah perpindahan pusat kekuasaan di masa lampau [2]. Kompleksitas dampak yang ditimbulkan, mulai dari pemerataan ekonomi hingga kekhawatiran ekologis,

menjadikan isu Ibu Kota Nusantara (IKN) sebagai topik perdebatan yang hangat di ruang publik [3].

Dalam era digital saat ini, dinamika opini publik terhadap kebijakan strategis seperti IKN terekam secara masif di media sosial, khususnya Twitter (sekarang X). Media sosial telah menjadi platform utama bagi masyarakat untuk menyuarakan dukungan, kritik, maupun kekhawatiran mereka secara bebas dan *real-time* [4]. Besarnya volume opini yang terbentuk di media sosial menciptakan kebutuhan mendesak untuk memetakan sentiment publik secara akurat. Hal ini penting karena opini di media sosial dapat merefleksikan penerimaan atau penolakan masyarakat terhadap kebijakan pemerintah, yang pada akhirnya memengaruhi stabilitas sosial dan keberhasilan implementasi kebijakan itu sendiri [5]. Oleh karena itu, analisis sentiment menjadi instrument krusial

untuk mengekstraksi informasi berharga dari data teks yang tidak terstruktur tersebut.

Analisis sentiment atau *opinion mining* merupakan cabang dari *text mining* yang bertujuan untuk mengidentifikasi dan mengklasifikasikan polaritas opini baik positif, maupun negatif yang terkandung dalam teks [6]. Proses ini melibatkan teknik *Natural Language Processing* (NLP) untuk mengubah data mentah menjadi informasi terstruktur yang dapat dianalisis [7]. Dalam konteks opini publik, analisis sentiment membantu memecahkan masalah ketidakseimbangan informasi dengan mengelompokkan jutaan komentar menjadi kategori yang dapat ditindak lanjut [8].

Untuk melakukan klasifikasi sentiment secara otomatis, algoritma *Machine Learning* memegang peranan vital. Penelitian ini berfokus pada tiga algoritma klasifikasi populer: *Naïve Bayes* (NB), *Logistic Regression* (LR), dan *Support Vector Machine* (SVM). *Naïve Bayes* dikenal sebagai metode probabilistik yang efisien dan sederhana untuk klasifikasi teks, terutama pada dataset besar [9]. *Logistic Regression*, di sisi lain, merupakan model statistik yang efektif untuk memprediksi probabilitas kelas biner dan sering digunakan sebagai *baseline* yang kuat dalam tugas klasifikasi [10]. Sementara itu, *Support Vector Machine* (SVM) bekerja dengan mencari *hyperlane* terbaik yang memisahkan kelas data dalam ruang vektor, yang terbukti memiliki performa tinggi dalam menangani data berdimensi tinggi seperti teks [11]. Penggunaan pendekatan multi-model dalam penelitian komparatif sangat penting untuk menentukan algoritma mana yang paling optimal dalam menangani karakteristik linguistic spesifik dari wacana IKN, mengingat setiap algoritma memiliki keunggulan dan kelemahan tersendiri tergantung pada domain data yang dianalisis [12].

Riset terdahulu mengenai sentimen IKN telah menerapkan berbagai algoritma *Machine Learning* konvensional, namun masih menyisakan celah metodologis yang signifikan. Diskursus mengenai pembangunan IKN di platform media sosial X baru-baru ini yang mengkomparasikan algoritma *Support Vector Machine* (SVM), *Logistic Regression*, dan *Naïve Bayes*, dengan SVM mencapai akurasi tertinggi sebesar 80% [4]. Namun, tinjauan kritis terhadap metodologi tersebut mengungkap adanya bias kelas yang fatal. Data yang digunakan didominasi oleh sentimen netral (996 data) dibandingkan positif (502 data) dan negatif (484 data) [4], tanpa adanya intervensi penanganan *imbalance data*. Evaluasi performa hanya bertumpu pada akurasi global dalam kondisi data timpang berisiko tinggi menghasilkan model cacat secara prediktif (*majority class bias*). Selain itu, studi tersebut tidak melakukan eksplorasi optimasi ruang pencarian *hyperparameter* secara sistematis pada model *Logistic Regression* dan *Naïve Bayes* [4].

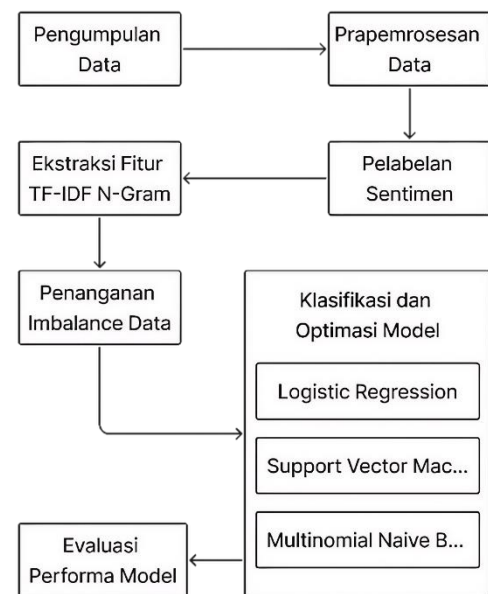
Penelitian ini bertujuan menutup celah metodologis tersebut dengan mengajukan kerangka kerja yang lebih solid. Intervensi *Synthetic Minority Over-sampling Technique* (SMOTE) diimplementasikan secara mutlak untuk menyeimbangkan distribusi data dan mencegah ketidakmampuan model dalam mengenali sentimen minoritas [13]. Representasi fitur diperluas menggunakan TF-IDF N-

gram (1,2) untuk menangkap konteks frasa majemuk yang krusial dalam bahasa non-formal (seperti “hutan nusantara” atau “pusat pemerintahan”), yang tidak dapat diakomodasi oleh pendekatan unigram standar [8]. Performa algoritma tidak dibiarkan pada parameter bawaan (*default*, melainkan dieksplorasi menggunakan teknik *GridSearchCV* untuk menemukan konfigurasi paling optimal [14], [15].

Berbeda dengan temuan terdahulu yang mencatat tingginya sentimen netral akibat dominasi akun-akun resmi [4], pembaruan dataset pada rentang 2024-2025 dengan total 2.257 data unik dalam penelitian ini mengkonfirmasi adanya pergeseran opini publik secara riil. Ekstraksi data terbaru menunjukkan bahwa opini masyarakat kini secara nyata didominasi oleh sentimen negatif (1.745 data) dibandingkan negatif (512 data). Peningkatan proporsi penerimaan dan optimism terhadap pembangunan IKN. Oleh karena itu, evaluasi komparatif algoritma dalam penelitian ini dirancang untuk memastikan bahwa model mampu memetakan pergeseran polaritas dukungan ini dengan presisi tingkat tinggi (berfokus pada metrik *Macro Precision*), bukan sekadar memprediksi akurasi secara global yang manipulatif.

II. METODE

Penelitian ini mengadopsi kerangka kerja *Natural Language Processing* (NLP) terstruktur untuk mengekstraksi dan mengevaluasi polaritas opini publik. Tahapan metodologi mencakup pengumpulan data, prapemrosesan data, pelabelan sentimen, ekstraksi fitur tekstual, penanganan ketidakseimbangan kelas, optimasi *hyperparameter*, serta evaluasi matriks performa klasifikasi. Tahapan penelitian seperti yang ada pada Gambar 1.



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data dan Karakteristik Dataset

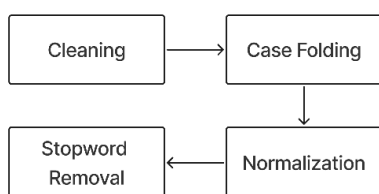
Data dalam penelitian ini merupakan opini publik yang diekstraksi dari platform media sosial X (Twitter) terkait

wacanaa pembangunan Ibu Kota Nusantara (IKN). Pengumpulan data (*crawling*) diimplementasikan menggunakan Pustaka *Tweet-Harvest*, sebuah metode yang memungkinkan ekstraksi data secara komprehensif tanpa batasan limitasi API standar konvensional. Data ditarik berdasarkan kata kunci pencarian spesifik terkait IKN pada rentang periode 2024 hingga 2025.

Proses ekstraksi menghasilkan kumpulan data mentah yang kemudian melalui tahap pembersihan awal, yaitu penghapusan data duplikat dan baris kosong (*missing value*), sehingga menyisakan 2.257 data teks unik. Karakteristik dataset ini berupa teks berbahasa Indonesia non-formal yang memiliki tingkat *noise* linguistic tinggi (seperti penggunaan *slang*, singkatan, dan *hashtag*), sehingga menuntut protocol prapemrosesan yang ketat sebelum direpresentasikan secara matematis.

B. Prapemrosesan Data

Transformasi data teks menjadi format terstandarisasi dilakukan melalui serangkaian tahapan sekuensial untuk mengeliminasi elemen yang tidak memiliki nilai analitis. Tahapan prapemrosesan data dapat dilihat pada Gambar 2.



Gambar 2. Tahapan Prapemrosesan Data

- 1) *Cleaning*: Penghapusan URL, *mention* (@), *hashtag* (#), angka, dan simbol khusus menggunakan *Regular Expression* (Regex) untuk menyisakan murni karakter alfabet.
- 2) *Case Folding*: Penyeragaman seluruh karakter menjadi huruf kecil (*lowercase*) guna meminimalisir variasi dimensi akibat sensitivitas kapitalisasi.
- 3) *Normalization*: Standarisasi kosakata non-formal (*slang*), singkatan, dan salah ketik menjadi bahasa Indonesia baku menggunakan *custom dictionary* yang memetakan puluhan istilah gaul (misalnya, “bgt” menjadi “banget”, ikn” menjadi “ibu kota nusantara”).
- 4) *Stopword Removal*: Eliminasi kata hubung atau kata umum yang tidak memiliki signifikansi sentimen. Secara kritis, daftar *stopword* standar dimodifikasi dengan mengecualikan kata negasi (seperti “tidak”, “bukan”, “belum”). Pemertahanan kata negasi ini merupakan justifikasi metodologis yang esensial untuk mencegah disrupti polaritas (misalnya, frasa “tidak bagus” akan hilang maknanya jika kata “tidak” dihilangkan).

TABEL I
HASIL PRAPEMROSESAN DATA

Raw Data	Clean Data
@BambangS112 @Farisa_Dh @Uzzzzz222 Seujung kuku? Dikira babat hutan buat IKN itu seujung kuku?	seujung kuku dikira babat hutan buat ikn itu seujung kuku
@masabudi5791 @Andria75777 Tidak benar. Video banjir di Istana Garuda IKN adalah hoaks seperti dikonfirmasi Antara News dan Otorita IKN. Area pusat pemerintahan tidak terendam meski ada riwayat banjir di sekitar IKN. Klaim makhluk mitos jelas satire.	tidak benar video banjir di istana garuda ikn adalah hoaks seperti dikonfirmasi antara news dan otorita ikn area pusat pemerintahan tidak terendam meski ada riwayat banjir di sekitar ikn klaim makhluk mitos jelas satire
@qbthe6th @Andria75777 Tidak klaim itu tidak benar. Video banjir di IKN adalah rekayasa AI atau editan seperti dikonfirmasi Otorita IKN dan sumber seperti Kompas pada Maret 2025. Tidak ada banjir aktual di kawasan istana bagian makhluk mitos hanyalah lelucon. Namun studi BNPB sebut ada 10 titik	tidak klaim itu tidak benar video banjir di ikn adalah rekayasa ai atau editan seperti dikonfirmasi otorita ikn dan sumber seperti kompas pada maret tidak ada banjir aktual di kawasan istana bagian makhluk mitos hanyalah lelucon namun studi bnpb sebut ada titik

C. Pelabelan Sentimen

Mengingat dataset mentah bersifat unsupervised, proses pelabelan sentimen diotomatisasi menggunakan pendekatan *lexicon-based*. Kamus referensi yang digunakan adalah *Indonesia Sentiment Lexicon* (InSet). Skor sentimen tiap dokumen dihitung dari agregasi selisih bobot kata positif dan negatif yang terdeteksi, dengan formulasi matematis:

$$Score = \sum bobot(kata_positif) - \sum bobot(kata_negatif) \quad (1)$$

Keterangan:

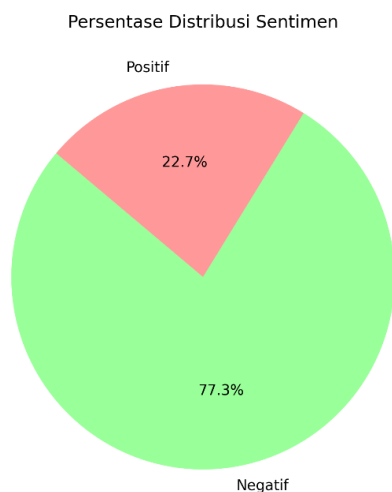
- *Score*: Nilai total sentimen.
- $\sum bobot(kata_positif)$: Akumulasi nilai bobot seluruh kata yang terdeteksi sebagai sentimen positif.
- $\sum bobot(kata_negatif)$: Akumulasi nilai bobot seluruh kata yang terdeteksi sebagai sentimen negatif.

Berdasarkan hasil pelabelan, diperoleh distribusi sebagai berikut:

- *Sentimen Positif*: 512 data.
- *Sentimen Negatif*: 1.745 data.

Desain penelitian ini mengklasifikasikan data secara biner. Aturan pelabelan ditetapkan di mana dokumen dengan total $Score \geq 0$ direpresentasikan sebagai sentimen “Positif”, sedangkan $Score \leq 0$ direpresentasikan sebagai sentimen “Negatif”. Inklusi skor 0 ke dalam kelas positif dilandasi oleh asumsi rasional bahwa ketiadaan sentimen negatif secara

eksplisit pada isu IKN mencerminkan tingkat penerimaan pasif. Hasil pelabelan menunjukkan ketimpangan ekstrem, di mana sentimen negatif mendominasi dengan 1.745 data (77,3%), berbanding 512 data (22,7%) untuk sentimen positif.



Gambar 3. Persentase Distribusi Sentimen

Pendekatan leksikon statis ini memiliki keterbatasan inheren. Kamus terbukti sering gagal menangkap anomaly bahasa seperti ironi dan bias pada konteks spesifik domain. Kelemahan ini menggarisbawahi urgensi penggunaan *Machine Learning* untuk memperbaiki pola klasifikasi dan mengenali kata-kata pembeda yang unik pada setiap kelas.

D. Ekstraksi Fitur TF-IDF N-Gram

Transformasi data teks ke dalam representasi numerik dilakukan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Penelitian ini menerapkan parameter `ngram_range=(1,2)` yang menggabungkan *unigram* (1 kata) dan *bigram* (2 kata). Bobot fitur dihitung dengan rumus:

$$W_{t,d} = TF_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (2)$$

Penggunaan N-gram (1,2) dijustifikasi karena pendekatan ini secara signifikan lebih superior dibandingkan *unigram-only*. Model *unigram-only* merusak konteks relasional antar kata, sedangkan integrasi bigram memungkinkan model untuk menangkap frasa majemuk yang esensial secara semantik seperti “hutan nusantara”, “pusat pemerintahan”, atau “rekayasa ai”. Di sisi lain, meskipun pendekatan *Word Embeddings* (seperti Word2Vec atau BERT) dinilai lebih mutakhir, representasi ini memiliki kompleksitas komputasi yang berlebihan untuk dataset berskala kecil dan rentan mengalami *overfitting* dibandingkan matriks *sparse* dari TF-IDF pada arsitektur model klasifikasi linear.

E. Penanganan Imbalance Data

Mengingat adanya ketidakseimbangan jumlah antara kelas positif dan negatif, penelitian ini menerapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) pada data latih untuk menyeimbangkan jumlah sampel kelas minoritas melalui interpolasi linear. Penggunaan SMOTE terbukti krusial dalam menyeimbangkan bobot kelas sehingga model tidak hanya mengenali pola pada kelas mayoritas, tetapi juga mampu mengklasifikasikan kelas minoritas dengan lebih presisi [13]. Selain itu integrasi teknik *oversampling* ini telah terbukti secara efektif meningkatkan akurasi pada berbagai model klasifikasi yang menghadapi kendala distribusi data yang tidak merata [16]. SMOTE melakukan *oversampling* pada kelas minoritas dengan menggenerasi data sintetik berdasarkan interpolasi linear antara *k-nearest neighbors* pada ruang fitur, yang dimodelkan sebagai:

$$x'_i = x_i + \lambda \times (x_j - x_i) \quad (3)$$

Di mana x'_i adalah sampel sintetik baru, x_i adalah sampel minoritas observasi, x_j adalah tetangga terdekat yang dipilih secara acak, dan λ adalah nilai acak antara 0 hingga 1.

Secara analitis, implementasi SMOTE pada data teks berdimensi tinggi (seperti matriks TF-IDF yang memiliki ribuan fitur) memunculkan risiko *overfitting* yang kritis. SMOTE berpotensi menggenerasi sampel sintetik pada ruang *noise* atau menciptakan *decision boundaries* yang terlalu spesifik pada data latih. Untuk memitigasi anomali ini, SMOTE tidak diimplementasikan secara statis di awal data, melainkan dienkapsulasi secara ketat di dalam object *Pipeline* bersama estimator *Machine Learning*. Pendekatan ini memastikan proses *oversampling* hanya terjadi secara iteratif pada *training fold* selama *cross-validation*, sehingga *test fold* terisolasi secara sempurna dan evaluasi performa terbebas dari *data leakage*.

F. Klasifikasi dan Optimasi Model

Penelitian ini membandingkan kinerja tiga algoritma: *Multinomial Naïve Bayes* (MNB), *Support Vector Machine* (SVM), dan *Logistic Regression* (LR). Untuk memperoleh performa terbaik, setiap model diintegrasikan dalam pipa pemrosesan (*pipeline*) dan dioptimasi menggunakan *GridSearchCV* dengan *5-fold cross validation*. Pencarian parameter otomatis ini memungkinkan penemuan konfigurasi *hyperparameter* yang paling optimal guna meminimalkan *error* pada saat pengujian model [17]. Ruang pencarian diformulasikan secara spesifik untuk mengeksplorasi karakteristik masing-masing algoritma:

- 1) *Logistic Regression*: Pencarian berfokus pada keseimbangan fungsi *penalty* ($penalty = ['l1', 'l2']$) dan kekuatan regulasi ($C = [0.1, 1, 10]$). Fokus optimasi menggunakan *penalti L1* bertujuan untuk meningkatkan efisiensi komputasi sekaligus akurasi model dalam memproses data teks [18]. Fungsi objektif dengan regulasi L1 direpresentasikan sebagai:

$$\min \left[\sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^k |\beta_j| \right] \quad (4)$$

- 2) *Support Vector Machine*: Dioptimasi pada margin batas klasifikasi ($C = [0.1, 1, 10]$) dan koefisien fleksibilitas *hyperplane* pada kernel ($\gamma = [\text{'scale'}, 0.01]$).

$$f(x) = \text{sign}(\sum \alpha_i y_i K(x_i, x) + b) \quad (5)$$

- 3) *Multinomial Naïve Bayes*: Dioptimasi pada tingkat peredaman *Laplace smoothing* (α) untuk menangani frekuensi probabilitas fitur yang tidak terobservasi pada fase pelatihan.

$$P(t_k|c) = \frac{\text{count}(t_k, c) + \alpha}{\text{count}(c) + \alpha|V|} \quad (6)$$

Optimasi difokuskan pada rentang $\alpha = [0.1, 0.5, 1.0]$.

G. Metrik Evaluasi Model

Kinerja dianalisis melalui matriks konfusi (*Confusion Matrix*) yang menghasilkan *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Guna menjamin evaluasi prediktif pada kedua kelas, metrik *Macro Precision* digunakan sebagai fungsi *scoring* objektif utama ($\text{scoring} = \text{'precision_macro'}$), di mana model dinilai dari kemampuannya menjaga presisi di kelas mayoritas maupun minoritas secara berimbang:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

$$\text{Precision}_{\text{macro}} = \frac{1}{k} \sum_{i=1}^k \text{Precision}_i \quad (10)$$

Model divalidasi dengan menganalisis metrix *Precision*, *Recall*, dan *F1-Score*. Guna memastikan evaluasi yang adil terhadap kemampuan prefiktif model di kedua kelas (Positif maupun Negatif), parameter *Macro Average* digunakan sebagai indikator deterministik utama dalam penentuan model terbaik. Hal ini ditunjukkan untuk menjamin stabilitas klasifikasi opini publik secara presisi tanpa tendensi manipulasi matematis pada kelas mayoritas.

III. HASIL DAN PEMBAHASAN

A. Analisis Distribusi Kelas

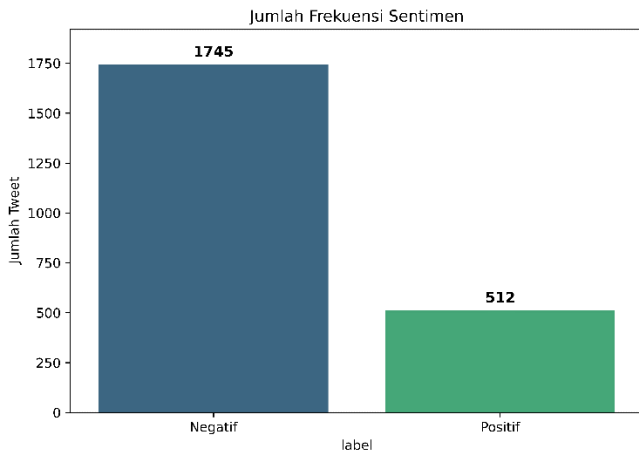
Proses prapemrosesan linguistic memberikan dampak krusial dalam menstandarisasi data X (Twitter) yang sarat akan noise. Penggunaan *custom dictionary* terbukti secara drastic memperbaiki struktur semantic kalimat (misalnya, menormalisasi ragam bahasa *slang* menjadi bentuk baku). Seperti yang ditampilkan pada Gambar 4.

TABEL II
DETAIL TAHAPAN TRANSFORMASI PRAPEMROSESAN DATA

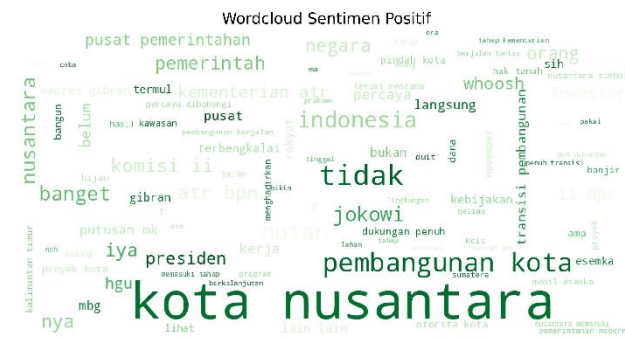
Proses	Hasil
Data mentah	@qbthe6th @Andria75777 Tidak klaim itu tidak benar. Video banjir di IKN adalah rekayasa AI atau editan seperti dikonfirmasi Otorita IKN dan sumber seperti Kompas pada Maret 2025. Tidak ada banjir aktual di kawasan istana. Bagian makhluk mitos hanyalah lelucon. Namun studi BNPB sebut ada 10 titik
Clean dan Casefolding	tidak klaim itu tidak benar video banjir di ikn adalah rekayasa ai atau editan seperti dikonfirmasi otorita ikn dan sumber seperti kompas pada maret tidak ada banjir aktual di kawasan istana bagian makhluk mitos hanyalah lelucon namun studi bnpb sebut ada titik
Normalisasi	tidak klaim itu tidak benar video banjir di ibu kota nusantara adalah rekayasa ai atau editan seperti dikonfirmasi otorita ibu kota nusantara dan sumber seperti kompas pada maret tidak ada banjir aktual di kawasan istana bagian makhluk mitos hanyalah lelucon namun studi bnpb sebut ada titik
Stopword removal	tidak klaim tidak video banjir kota nusantara rekayasa ai editan dikonfirmasi otorita kota nusantara sumber kompas maret tidak banjir aktual kawasan istana makhluk mitos lelucon studi bnpb titik

Penerapan *Indonesia Sentiment Lexicon* (InSet) pada data yang telah dibersihkan mengungkap realitas distribusi opini publik yang sangat timpang. Berbeda dengan asumsi awal, diskursus IKN pada rentang waktu ini dikuasai oleh sentimen Negatif sebanyak 1.745 data (77,3%), berbanding 512 data (22,7%) untuk sentimen Positif. Tingginya dominasi sentimen negatif ini merepresentasikan tingginya intensitas

kritik, keluhan, maupun skeptisisme masyarakat di platform X terhadap berbagai aspek pembangunan IKN.



Gambar 4. Bar Chart Distribusi Sentimen



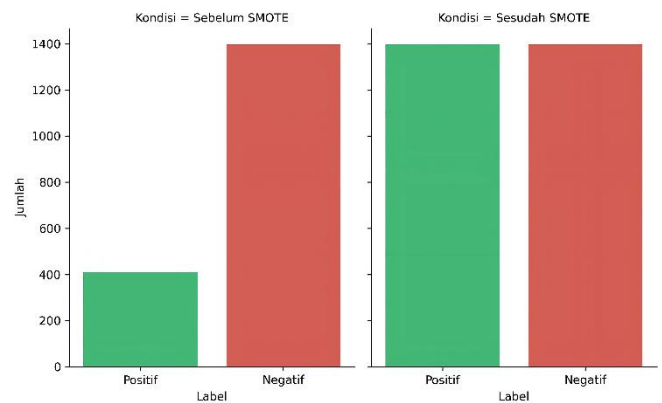
Gambar 5. WordCloud Sentimen Positif



Gambar 6. WordCloud Sentimen Negatif

Distribusi kelas yang timpang secara ekstrem ini memvalidasi urgensi penggunaan algoritma *Synthetic Minority Over-sampling Technique* (SMOTE). Tanpa SMOTE, model prediktif akan secara otomatis terjebak dalam *majority class bias*, yang mana menebak hampir semua input sebagai kelas “Negatif” demi mencapai akurasi global yang tinggi secara semu.

Distribusi Kelas Sebelum vs Sesudah SMOTE



Gambar 7. Diagram Distribusi Kelas Sebelum dan Sesudah SMOTE

Keterangan: SMOTE menyeimbangkan data latih menjadi proporsi 50:50.

B. Komparasi Performa Algoritma

Eksperimen komparatif dilakukan terhadap algoritma *Support Vector Machine* (SVM), *Logistic Regression* (LR), dan *Multinomial Naive Bayes* (MNB) melalui dua skenario utama untuk mengevaluasi efektivitas intervensi teknis yang diberikan. Skenario pertama adalah pengujian model *baseline* (tanpa optimasi), di mana algoritma dijalankan dengan parameter bawaan (*default*) dan dengan diterapkannya penanganan ketidak seimbangan data. Skenario kedua adalah pengujian model *tuned*, yang mengintegrasikan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) untuk menyeimbangkan kelas serta *GridSearchCV* untuk menemukan konfigurasi *hyperparameter* paling optimal. Hasil evaluasi perbandingan kedua skenario tersebut disajikan pada Tabel III berikut:

TABEL III
HASIL EVALUASI MODEL BASELINE DAN TUNED

Algoritma	Skenario	Accuracy	Macro Precision	F1 Score
Logistic Regression	Baseline	0.768	0.695	0.709
Logistic Regression	Tuned	0.772	0.695	0.708
Naive Bayes	Baseline	0.794	0.708	0.710
Naive Bayes	Tuned	0.805	0.723	0.723
SVM	Baseline	0.785	0.686	0.650
SVM	Tuned	0.798	0.710	0.685

Hasil eksperimen yang ditampilkan pada Tabel III dan Gambar 8 menunjukkan fenomena komparatif yang signifikan. Pada dataset berdimensi tinggi ini, algoritma *Multinomial Naïve Bayes* (MNB) setelah optimasi terbukti menjadi model paling *robust*. MNB (*Tuned*) mencapai *Accuracy* 80,5%, *Macro Precision* 72,3%, dan *F1-Score* 72,3% melampaui kemampuan *Support Vector Machine* (*F1-Score* 68,5%) maupun *Logistic Regression* (*F1-Score* 70,8%).

Superioritas MNB dapat direkonstruksi secara teoritis. MNB beroperasi berdasarkan *Teorema Bayes* dengan asumsi independensi fitur (*naïve assumption*). Pada ruang vector TF-IDF N-gram (1,2) yang direkayasa oleh SMOTE, fitur sintetik seringkali memicu multikolinearitas (korelasi berlebihan antar variabel) yang merusak pembentukan matriks bobot pada linear seperti *Logistic Regression*. Asumsi independensi pada MNB justru melindunginya dari efek negatif multikolinearitas tersebut, menjadikannya sangat tahan terhadap *noise* yang dihasilkan dari proses *oversampling* kelas minoritas.

Lebih jauh, optimasi *GridSearchCV* pada MNB menemukan bahwa nilai *Laplace smoothing* terbaik berada pada $\alpha = 0.1$ (jauh lebih rendah dari nilai *default* $\alpha = 1.0$). Nilai α yang kecil ini mengindikasikan bahwa distribusi frekuensi kata pada matriks TF-IDF sudah cukup padat dan representatif. Peredaman probabilitas yang minim memungkinkan model MNB untuk lebih mempercayai kemunculan kata-kata spesifik yang mendikte sentimen, meningkatkan sensitivitas (*recall*) secara signifikan tanpa mengorbankan presisi.

Di sisi lain, model SVM dan *Logistic Regression* mengalami batasan generalisasi. SVM mencapai performa optimal pada regulasi yang longgar ($C = 10$) dan $\gamma = scale$, namun model *margin-mased* ini kewalahan menetapkan

batasan keputusan klasifikasi (*hyperplane*) yang adil ketika berhadapan dengan sebaran data sintetik SMOTE yang bertumpang tindih (*overlapping*) di ruang dimensi tinggi.

C. Analisis Fitur (Feature Importance)

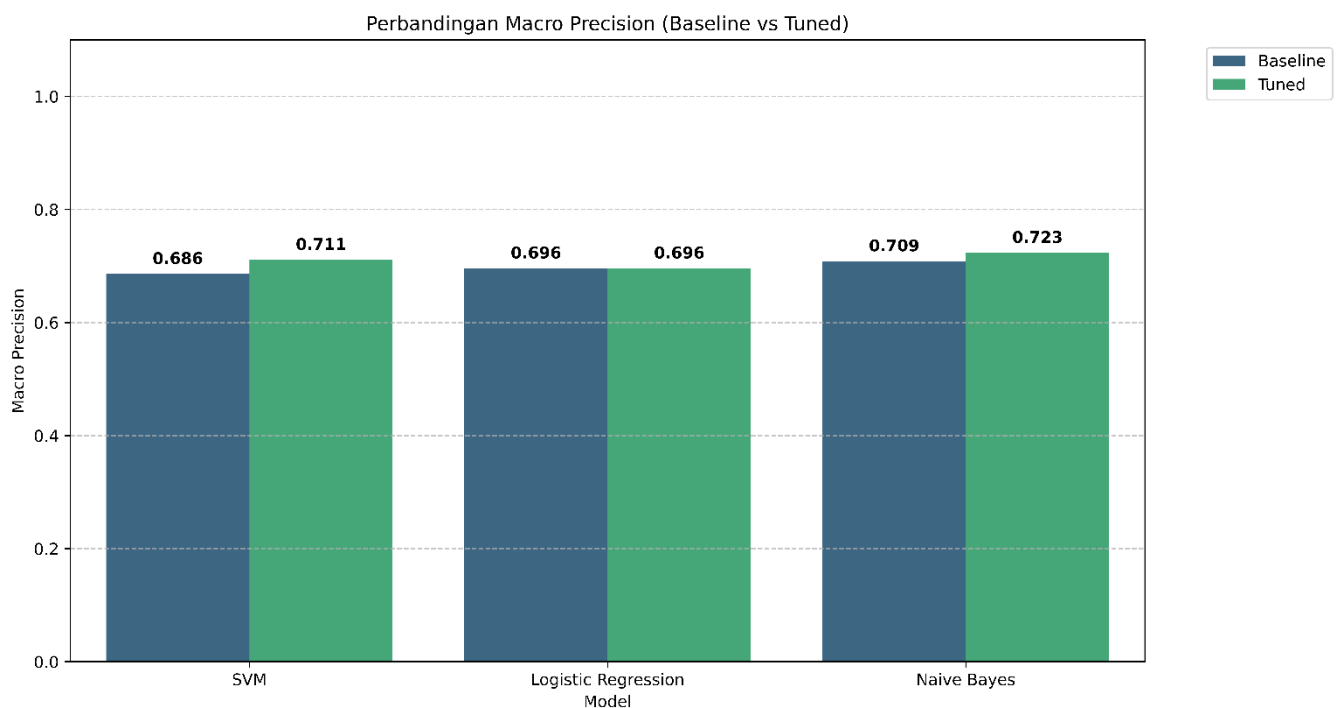
Untuk menginterpretasikan logika internal model klasifikasi terbaik (*Multinomial Naïve Bayes*), dilakukan analisis *feature importance* berdasarkan ekstraksi log-probabilitas fitur. Nilai bobot ini merepresentasikan kemampuan diskriminatif sebuah kata tunggal (*unigram*) maupun frasa (*bigram*) dalam memisahkan dan menentukan kelas sentimen pada ruang vector teks.

TABEL IV
FITUR POSITIF PALING BERPENGARUH DALAM KLASIFIKASI SENTIMEN

Kata Positif	Bobot Positif
efisien	4.7437
tuntas	4.6994
wujud	4.3455
nusantara wujud	4.3455
pemerintahan modern	4.3346
berkembang	4.3336
wujud pemerataan	4.2136
berjalan tuntas	4.1448
pengadaan tanah	3.9588
pemerintahan pemerataan	3.9588

TABEL V
FITUR NEGATIF PALING BERPENGARUH DALAM KLASIFIKASI SENTIMEN

Kata Negatif	Bobot Negatif
korupsi	-4.5846
politik	-4.4720
utang	-4.3714
cepat kota	-4.0520
rugi	-3.9304
media	-3.9304



Gambar 8. Distribusi Perbandingan Macro Precision

kerugian	-3.8867
gagal	-3.8427
kota politik	-3.7846
isu	-3.7164

Berdasarkan ekstraksi koefisien pada Tabel IV, fitur-fitur yang menjadi penanda utama sentimen Positif didominasi oleh kata “efisien” (bobot 4.7437), “tuntas” (4.6994), serta frasa *bigram* “pemerintahan modern” (4.3346) dan “wujud pemerataan” (4.2136). Munculnya kata dan frasa tersebut menunjukkan bahwa optimism masyarakat di platform X (Twitter) berakar pada pemerintaan yang berteknologi maju dan merata secara ekonomi. Ekstraksi ini sekaligus memvalidasi penggunaan fitur TF-IDF N-gram (1,2) dalam penelitian ini, model berhasil menangkap frasa relasional seperti “pemerintahan modern” dan “wujud pemerataan” yang esensi maknanya akan hilang jika data hanya direpresentasikan sebagai kata tunggal.

Sementara itu, merujuk pada Tabel V, fitur-fitur yang berakar kuat pada sentimen Negatif dipelopori oleh kata “korupsi” (bobot -4.5846), “politik” (-4.4720), “utang” (-4.3714), dan “gagal” (-3.8427), serta frasa “kota politik” (-3.7846). Konstelasi kata-kata ini secara gamblang mencerminkan keraguan dan kritik tajam masyarakat yang berfokus pada risiko sosio-ekonomi dan integritas proyek. Tingginya bobot pada kata “utang”, “kerugian”, dan “korupsi” mengindikasikan ketakutan publik terhadap potensi beban fiskal negara dan mismanajemen anggaran. Munculnya frasa “kota politik” juga memperlihatkan skeptisisme masyarakat yang memandang proyek IKN serak akan muatan agenda politik dibandingkan urgensi structural.

Kemampuan arsitektur model dalam mengekstraksi dan memisahkan kelompok term yang sangat kontras ini membuktikan ketangguhan metodologi yang digunakan. Meskipun dihadapkan pada disparitas frekuensi actual di aman sentimen Negatif mendominasi secara signifikan pada data mentah, integrasi antara representasi TF-IDF *bigram*, penyeimbang kelas menggunakan SMOTE, dan optimasi algoritma MNB berhasil menyelaraskan komputasi statistik dengan intuisi linguistic manusia, menjadikan model ini sebagai instrument pemetaan opini publik yang representatif dan akurat.

IV. KESIMPULAN

Berdasarkan ekstraksi dan analisis komprehensif terhadap opini publik di platform X (Twitter) terkait pembangunan Ibu Kota Nusantara (IKN) periode 2024 hingga 2025, dapat disimpulkan bahwa diskursus terkait masyarakat secara empiris didominasi oleh sentimen Negatif (77,3%) dibandingkan sentimen Positif (22,7%). Tingginya disparitas polaritas ini merefleksikan kuatnya intensitas pengawasan, kritik, dan atensi publik terhadap dinamika implementasi proyek strategis nasional tersebut.

Secara metodologis, penelitian ini membuktikan bahwa pendekatan klasifikasi konvensional pada data yang timpang (*imbalanced*) akan menghasilkan bias prediktif. Integrasi *Synthetic Minority Over-sampling Technique* (SMOTE) di dalam arsitektur *pipeline* validasi silang menjadi intervensi

krusial untuk menyeimbangkan batas keputusan model. Dari hasil komparasi algoritma, *Multinomial Naïve Bayes* (MNB) dengan optimasi *hyperparameter* ($\alpha = 0.1$) tampil sebagai model paling *robust*, mencetak *Macro Precision* dan *F1-Score* tertinggi sebesar 72,3%, melampaui algoritma *Logistic Regression* dan *Support Vector Machine*.

Superioritas MNB ini membuktikan bahwa asumsi independensi fitur secara arsitektural lebih Tangguh dalam menangani multikolinearitas *noise* pada ruang vektor TF-IDF berdimensi tinggi dibandingkan model berbasis *margin* linear. Lebih jauh, kemampuan model dalam mengekstraksi metodologi yang diusulkan berhasil menjembatani kesenjangan antara komputasi statistik dan konteks linguistik, menjadikannya instrumen pemetaan sentimen yang andal bagi evaluasi kebijakan publik di masa depan. Pengembangan penelitian selanjutnya direkomendasikan untuk mengeksplorasi arsitektur *Deep Learning* seperti *Transformer* (IndoBERT) dengan rentang waktu analitik temporal guna memetakan fluktuasi sentimen secara *real-time*.

DAFTAR PUSTAKA

- [1] G. Aji, Z. Arfani, A. M. Sari, and R. Septriani, “Dampak Pemindahan Ibukota Negara Baru terhadap Ekonomi dan Sosial di Provinsi Kalimantan Timur,” vol. 1, pp. 1–8, 2023.
- [2] D. Ocsanda, C. I. Wijaya, M. Azzam, and A. Haq, “Oportunitas Pembentukan Entitas Kebudayaan Baru , Tinjauan Terhadap Ibu Kota Negara (Ikn) Indonesia 2024 Berdasarkan Sejarah Perpindahan Ibu Kota Voc 1619 Opportunity For New Cultural Entity Emergence , An Overview Of The Capital City (Ikn) Of Indonesia 2024 Based On The History Of The Relocation Of The Capital City Of VOC 1619,” vol. 18, pp. 1–12, 2024.
- [3] N. Juli, A. D. Masitah, D. Suluh, and K. Dewi, “Analisis Opini Publik Berdasarkan Teori Agenda Setting Pada Proses Perencanaan Pemindahan IKN,” vol. 6, no. 3, pp. 10206–10217, 2022, doi: 10.36312/jisip.v6i3.3374/http.
- [4] N. Hadi and D. Sugiarto, “Analisis Sentimen Pembangunan IKN pada Media Sosial X Menggunakan Algoritma SVM, Logistic Regression dan Naïve Bayes,” *J. Inform. J. Pengemb. IT*, vol. 10, no. 1, pp. 37–49, 2025, doi: 10.30591/jpit.v10i1.7106.
- [5] P. Arsi and R. Waluyo, “Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.
- [6] Z. Kastrati, F. Dalipi, A. S. Imran, and K. P. Nuci, “Sentiment Analysis of Students ’ Feedback with NLP and Deep Learning : A Systematic Mapping Study,” 2021.
- [7] L. Nemes and A. Kiss, “Social media sentiment analysis based on COVID-19,” *J. Inf. Telecommun.*, vol. 5, no. 1, pp. 1–15, 2021, doi: 10.1080/24751839.2020.1790793.
- [8] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor , Genetic , Support Vector Machine , Decision Tree , and Long Short Term Memory algorithms in machine learning,” *Decis. Anal. J.*, vol. 3, no. November 2021, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.
- [9] M. A. Saddam, E. Kurniawan D, and I. Indra, “Analisis Sentimen Fenomena PHK Massal Menggunakan Naive Bayes dan Support Vector Machine,” *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 226–233, 2023, doi: 10.30591/jpit.v8i3.4884.
- [10] S. A. H. Bahtiar, C. K. Dewa, and A. Luthfi, “Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling,” *J. Inf. Syst. Informatics*, vol. 5, no. 3, pp. 915–927, 2023, doi: 10.51519/journalisi.v5i3.539.
- [11] I. M. Sumertajaya, Y. Angraini, J. R. Harahap, and A. Fitrianto, “Sentiment Analysis on Covid-19 Vaccination in Indonesia Using

- Support Vector Machine and Random Forest,” vol. 10, no. 1, pp. 1–8, 2022.
- [12] R. B. Dahlian and D. Sitanggang, “Sentiment Analysis of Digital Television Migration on Twitter Using Naïve Bayes Multinomial Comparison, Support Vector Machines, and Logistic Regression Algorithms,” *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 2, pp. 280–288, 2023, doi: 10.32736/sisfokom.v12i2.1668.
- [13] M. K. Faruq, K. Umam, M. I. Mustofa, and A. A. Mahfudh, “Comparative Study of SVM and Decision Tree Algorithms on the Effect of SMOTE Technique on LinkAja Application,” vol. 9, no. 6, pp. 3305–3311, 2025.
- [14] M. Fajri and A. Primajaya, “Komparasi Teknik Hyperparameter Optimization pada SVM untuk Permasalahan Klasifikasi dengan Menggunakan Grid Search dan Random Search,” vol. 7, no. 1, pp. 10–15, 2023.
- [15] M. R. Nursyam, M. Koprari, and D. Ariyus, “Optimizing Email Spam Detection through Handling Class Imbalance with Class Weights and Hyperparameter Using GridSearchCV,” vol. 10, no. 1, pp. 232–244, 2026.
- [16] E. Saputra and E. R. Susanto, “Implementation of Deep Learning with Multilayer Perceptron (MLP) for Heart Disease Prediction Using the SMOTE-ENN Technique,” vol. 9, no. 3, pp. 1034–1041, 2025.
- [17] A. Ma’ruf, A. E. Pajri, and P. Liana, “Sentiment Analysis of President Prabowo ’ s Performance on Twitter (X) with a Comparative Study of SVM , XGBoost , and AdaBoost,” vol. 10, no. 1, pp. 684–697, 2026.
- [18] M. F. Kurniawan and D. A. Megawaty, “Comparison of Logistic Regression , Random Forest , Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) Algorithms in Diabetes Prediction,” vol. 9, no. 5, pp. 2154–2162, 2025.