

Analisis Sentimen Twitter Terhadap Pemindahan Ibu Kota Negara Menggunakan Support Vector Machine

Gita Aprinda Saputri¹, Debby Alita¹

¹Program Studi Sistem Informasi, Universitas Teknokrat Indonesia,,

¹Jl. ZA. Pagar Alam No.9-11, Kota Bandar Lampung, Kode pos 35132, Indonesia

Info Artikel

Riwayat Artikel:

Received 2024-03-06

Revised 2024-11-28

Accepted 2024-12-01

Corresponding Author:

Gita Aprinda Saputri

Email: g.aprindasp20@gmail.com



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

Abstract – The Indonesian government announced plans to move the capital from Jakarta to East Kalimantan due to the high population burden and economic contribution on the island of Java. Statistical data shows that the island of Java has a large population, reaching 151.59 million people or around 56.10% of the total population of Indonesia, and will provide a large participation in national GDP in 2021. Moving the capital city is seen as a step. . for the sake of equal distribution of population and economy throughout Indonesia. Rapid urbanization on the island of Java, especially in the buffer areas of the capital city of Jakarta, is one of the main reasons behind this decision. This research uses data from the social media platform Twitter to analyze sentiment using 2 categories, namely positive and negative sentiment regarding the relocation of the National Capital, analyzed using the Support Vector Machine method. In this study, the SVM kernel type was used, namely a linear kernel with an accuracy of 92.70%, then improved with Stratified k-Fold Cross Validation, getting 100% accuracy in iterations 1 and 5. The classification results using the Support Vector Machine method are stated that the linear kernel has better accuracy. This sentiment analysis provides insight into the public's views on the proposed measure. This research can be used as material for consideration of future government policy regarding relocating the capital city.

Keywords: sentiment analysis, East Kalimantan, kernel, moving capital, support vector machine.

Abstrak – Pemerintah Indonesia mengumumkan rencana Pemindahan ibu kota dari Jakarta ke Kalimantan Timur karena tingginya beban penduduk dan kontribusi perekonomian di Pulau Jawa. Data statistik menunjukkan Pulau Jawa mempunyai total penduduk yang besar, mencapai 151,59 juta jiwa atau sekitar 56,10% dari total penduduk Indonesia, dan memberi partisipasi yang besar pada PDB nasional pada tahun 2021. Pemindahan ibu kota dinilai sebagai sebuah langkah. untuk pemerataan persebaran penduduk dan perekonomian di seluruh Indonesia. Urbanisasi yang pesat di Pulau Jawa, khususnya di daerah penyangga Ibu Kota Jakarta, jadi salah satu alasan utama di balik keputusan ini. Penelitian ini menggunakan data dari platform media sosial *Twitter* untuk analisis sentimen menggunakan 2 kategori yaitu sentimen positif dan negatif terkait Pemindahan Ibu Kota Negara dianalisis memakai metode *Support Vector Machine*. Pada penelitian ini dilakukan jenis kernel SVM yaitu kernel *linier* dengan akurasi sebesar 92,70% lalu diperbaiki dengan *Stratified k-Fold Cross Validation* mendapatkan akurasi 100% pada iterasi 1 dan 5 Hasil klasifikasi memakai metode *Support Vector Machine* menyatakan bahwa kernel *linier* memiliki akurasi lebih baik. Analisis sentimen ini memberi wawasan mengenai pandangan masyarakat pada rencana pemindahan tersebut. Penelitian ini dapat menjadi bahan pertimbangan kebijakan pemerintah ke depannya terkait Pemindahan ibu kota.

Kata Kunci: analisis sentimen, Kalimantan Timur, kernel, pemindahan ibu kota, support vector machine.

I. PENDAHULUAN

Pemerintah sudah mengumumkan rencana Pemindahan Ibukota dari Jakarta ke Kalimantan Timur, dengan alasan *berdasarkan* data yang menunjukkan beban kependudukan dan kontribusi ekonomi yang terlalu tinggi di Pulau Jawa. Menurut BPS tahun 2020, populasi penduduk Pulau Jawa menggapai 151,59 juta atau sekitar 56,10% dari total populasi Indonesia, sementara Pulau Jawa menyumbang 57,89% dari PDB nasional pada tahun 2021, dengan PDB per kapita DKI Jakarta melebihi 4x lipat PDB rata-rata nasional. Oleh karena itu, Pemindahan ini dianggap sebagai upaya untuk meratakan distribusi penduduk dan ekonomi di seluruh Indonesia. Urbanisasi yang pesat di Pulau Jawa, terutama di daerah penyangga Ibukota Jakarta seperti Bogor, Depok, Tangerang, Bekasi, Puncak, dan Cianjur, jadi salah satu alasan utama untuk Pemindahan Ibukota. Sekitar 32 juta penduduk tinggal di wilayah yang dikenal sebagai Jabodetabek punjur. Alasan terakhir terkait dengan masalah lingkungan di Jakarta, seperti banjir, longsor, abrasi air laut, dan polusi, juga

jadi pertimbangan. Namun, rencana ini mendapat kritik karena dinilai kurang matang oleh beberapa pihak. Diringrat menyoroti kelemahan dalam Pemindahan Ibukota ke luar Jakarta, dengan mengatakan bahwa hanya Pemindahan ke luar Jawa saja tidak cukup untuk memajukan pertumbuhan ekonomi di wilayah timur Indonesia [1].

Sejumlah penelitian telah dilakukan untuk mengevaluasi kesiapan dan dampak Pemindahan Ibu Kota Negara Indonesia. Riset termasuk kesiapan Pemerintah Provinsi Kalimantan Tengah di Palangka Raya, analisis kesiapan masyarakat di Kecamatan Sepaku dan Samboja, serta ancaman terhadap bahasa lokal dan strategi konservasi kebudayaan. Namun, penelitian mengenai aspek kebudayaan yang fokus pada catatan historis dan potensi dampak budaya masih sedikit. Sejarah perpindahan ibu kota sebelumnya, seperti Pemindahan Ibu Kota VOC dari Ambon ke Batavia, menjadi acuan penting dalam memahami proses munculnya entitas budaya baru. Diperlukan penelitian lanjutan, terutama dengan pendekatan arkeologi lanskap yang didukung oleh data sejarah dan antropologi, untuk memahami kemungkinan munculnya entitas budaya baru di IKN Indonesia pada tahun 2024 [2].

Pemindahan Ibu Kota Negara memiliki dampak positif dan negatif, terutama dalam konteks ekonomi Indonesia yang belum stabil akibat pandemi COVID-19. Studi menunjukkan bahwa stimulus ekonomi untuk Pemindahan IKN belum memberi pengaruh yang tinggi pada perkembangan ekonomi nasional, meskipun provinsi yang bersangkutan mengalami kenaikan pertumbuhan ekonomi. Selain aspek ekonomi, Pemindahan ibu kota juga harus dipertimbangkan dari segi sosial, lingkungan, geopolitik, dan pertahanan, dengan perencanaan yang serius untuk menghindari dampak sosial dan budaya yang mungkin terjadi. Perlu kesiapan dan antisipasi dari semua elemen terlibat, termasuk pemerintah, pakar, media, dan aparat keamanan, dengan strategi yang mencakup pertahanan maritim untuk menjaga keamanan nasional dan kedaulatan negara [3].

Pada era digital sekarang ini media sosial khususnya *Twitter* telah jadi sarana utama bagi masyarakat Indonesia untuk menyuarakan pandangan mereka terhadap isu-isu terkini. Menurut survei APJII tahun 2020, sekitar 10% dari pengguna media sosial di Indonesia aktif menggunakan *Twitter* [4]. Salah satu isu yang kontroversial adalah rencana Pemindahan IKN ke Kalimantan Timur, yang menimbulkan beragam pendapat di masyarakat. Meskipun bertujuan untuk meningkatkan persaingan global, keputusan ini memicu debat karena alasan-alasan seperti kepadatan penduduk di Pulau Jawa, kontribusi ekonomi dari pulau tersebut, dan masalah lingkungan. Keputusan ini, yang diumumkan oleh Presiden Joko Widodo pada tahun 2019 dan disahkan dalam RPJMN Tahun Anggaran 2020-2024, mencerminkan perbedaan sudut pandang di masyarakat, tercermin dalam diskusi di media sosial seperti *Twitter* [5].

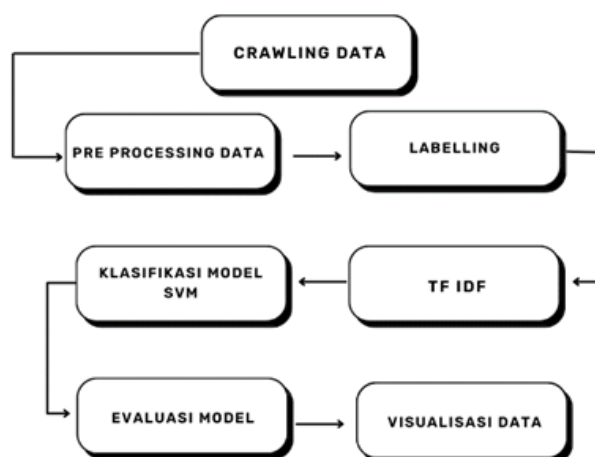
Twitter memungkinkan pengguna untuk berkomunikasi dalam waktu nyata dengan pesan berisi pendapat, ekspresi, atau reaksi terhadap isu-isu terkini, termasuk rencana Pemindahan IKN ke Kalimantan yang menjadi topik hangat pada 2019 [6]. Bahkan, saat Presiden Joko Widodo mengumumkan logo resmi IKN pada Mei 2023, debat seputar Pemindahan IKN kembali mencuat dengan beragam ekspresi [7]. Selain Pemindahan fisik, pembahasan tentang nama baru "Nusantara" juga menciptakan diskusi yang kritis di masyarakat, mengingat makna geografis dan simbolisnya. Keputusan ini didasarkan pada landasan hukum yang kuat, seperti yang diatur dalam UU No 3 Tahun 2022, dengan visi menjadikan Ibu Kota Nusantara sebagai kota dunia yang berkelanjutan, ekonomi Indonesia yang maju, serta simbol identitas nasional yang menggambarkan keberagaman bangsa Indonesia [8].

Dari sekumpulan opini yang berkaitan dengan topik di atas dapat dinyatakan bahwa terdapat opini positif, negatif dan beberapa di antaranya bersifat netral [9], [10]. Berbagai analisis terdahulu yang sudah menyatakan keberhasilan metode SVM dalam melakukan analisa sentimen terhadap berbagai topik yang berkaitan dengan kebijakan pemerintah dan layanan publik di media sosial *Twitter*. Berdasarkan penelitian yang dilakukan oleh Herwiansyah dan Arita Witanti menemukan bahwa SVM efektif dalam mengukur respons masyarakat terhadap kebijakan vaksinasi COVID-19 memperoleh akurasi 98%, *F1 Score* 93%, *Precision* 88% dan *Recall Score* sebesar 99 % belum adanya library yang mendukung untuk menghitung popularitas dengan data teks yang berbahasa indonesia [11] Selain itu, Penelitian dilakukan oleh [12] menunjukkan bahwa SVM dapat digunakan dengan baik dalam mengklasifikasi sentimen terkait kebijakan *lockdown* pemerintah Jakarta dengan menggunakan ekstraksi fitur *tf-idf* mendapat tingkat akurasi mencapai 74%, *Precision* 75%, *Recall* sebesar 92%, dan nilai *F1-Score* sebesar 83%. Cukup baik tetapi perlu adanya optimasi agar mendapat akurasi yang lebih tinggi. Pada penelitian selanjutnya [13] menemukan bahwa SVM pada multiclass mampu secara optimal menganalisis sentimen opini publik berbahasa Indonesia di twitter. Hasil yang didapat akurasi 92%, *Precision* sebesar 86,93%, *Recall* sebesar 92%, *F1-Score* sebesar 89,00%

dapat membuat penerjemah Bahasa agar bersifat continue. Selain itu [14] menggunakan SVM untuk menganalisis sentimen terhadap layanan Indihome di *Twitter* dengan *Confusion Matrix* mendapat hasil akurasi senilai 87%, *Precision* 86%, *Recall* 95% dan nilai *F1-Score* senilai 90%. Dari hasil yang didapat bahwa tingkat kepuasan pengguna layanan indihome cukup rendah. Penelitian oleh [15] mengenai penggunaan aplikasi shopee menggunakan metode SVM menunjukkan kinerja yang baik dari SVM dalam menghasilkan akurasi sebesar 98% dan *F1-Score* 98% dengan menggunakan parameter C. Penelitian ini ditujukan untuk menganalisa sentimen *Twitter* pada Pemindahan IKN dengan menggunakan metode SVM. SVM digunakan untuk mengukur tingkat keakuratan dalam menganalisis opini masyarakat yang diperoleh dari platform media sosial *Twitter*. Berdasarkan temuan-temuan penelitian sebelumnya, penelitian ini akan menganalisis sentimen masyarakat terhadap Pemindahan Ibu kota Negara di Media Sosial *Twitter* menggunakan metode *Support Vector Machine* berfokus pada pandangan Masyarakat melalui proses data kelas yang tidak seimbang dengan memanfaatkan *Stratified k-Fold* dalam mendapatkan akurasi yang tinggi.

II. METODE

Untuk mencapai hasil yang optimal dari penelitian ini, beberapa langkah penting harus dilakukan. Proses dimulai dengan pengumpulan data melalui *crawling*, yang kemudian diikuti dengan tahap *pre-processing*. Data yang telah melalui tahap *pre-processing* akan diberi label dalam dua kategori, kategori positif dan negatif. Selanjutnya, data akan menjalani proses pembobotan kata menggunakan *TF-IDF*. Setelah itu, data akan diklasifikasikan menggunakan metode *Support Vector Machine* (SVM). Setelah diklasifikasikan, evaluasi model akan dilakukan menggunakan *Stratified k-Fold Cross Validation*. Bisa ditinjau pada Gambar. 1 dibawah ini menunjukkan langkah-langkah penelitian yang dilakukan.



Gambar. 1 Tahapan Penelitian

A. Crawling Data

Dalam penelitian ini, pengumpulan data dilakukan dengan menggunakan mesin *python* untuk mengakses data *tweet* pengguna *Twitter* sesuai dengan topik mengenai Pemindahan Ibu Kota Negara. Proses pengumpulan data dilakukan dengan menggunakan *keywords* yaitu “IKN”, “Ibukota”, “Pindah Ibukota” sesuai dengan yang sedang diperbincangkan oleh di kalangan masyarakat. Proses pengumpulan data ini, dilakukan dalam rentang tertentu, dari tanggal 15-19 Desember 2023. Rentang waktu ini dipilih karena topik ini sedang banyak dibicarakan pada saat itu di *twitter*. Secara keseluruhan *tweet* terkumpul sebanyak 4193 *tweet*. Setelah pengumpulan data selesai, maka dilakukan proses deteksi duplikat untuk menghilangkan *tweet* yang mengandung informasi yang sama, sehingga total yang didapat sebanyak 3839 *tweet*. Selanjutnya, dilakukan tahap *preprocessing* data bertujuan untuk membersihkan dan mengubah format agar data dapat diolah ketahapan analisis selanjutnya.

B. Preprocessing Data

Preprocessing ditujukan untuk merubah kumpulan data yang belum diproses jadi kumpulan data yang siap digunakan.

- Dalam tahap cleansing, tahapan ini untuk membersihkan teks tweet dari komponen-komponen yang tidak relevan dalam analisis sentimen dapat dianggap sebagai noise dalam data contohnya seperti Username, Karakter HTML, Emoji, URL dan Hastag.
- Selanjutnya, tahapan case folding semua huruf dalam teks tweet diubah menjadi huruf kecil.
- Setelah itu, masuk ke proses tahapan tokenizing, memecah teks tweet menjadi kata-kata yang berdiri sendiri dengan memperhatikan batasan tanda baca dan spasi.
- Pada tahapan stopword removal ini dilakukan penghapusan kata-kata yang dianggap tidak terlalu bermakna seperti kata sambung dan kata umum lainnya.
- Proses terakhir, tahap stemming proses mengurai kata-kata menjadi bentuk kata dasar. Langkah ini dilakukan untuk membuat proses analisis data lebih mudah.

C. Labeling Data

Pada tahap ini, peneliti menggunakan proses label data untuk mengkategorikan *tweet* yang mengandung sentimen positif atau negatif terkait dengan Pemindahan Ibu Kota Negara. Proses label ini dilakukan dengan memanfaatkan Vader Lexicon. Dalam penggunaan Vader Lexicon tweet akan ber label positif jika mengungkapkan kepuasan atau kegembiraan sedangkan berlabel negatif diberikan jika mengungkapkan ketidakpuasan atau kekecewaan.

D. TF – IDF

Bobot kata dalam teks sebanding dengan frekuensi istilahnya, yang didefinisikan sebagai berapa kali sebuah kata muncul dalam dokumen. kata tersebut direpresentasikan kedalam bentuk vektor dan TF-IDF dan dibuat tabel informasi dokumen yang berisi term frequency (TF), dokumen frequency (DF), dan IDF dari masing- masing term [16].

$$W_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (1)$$

N = jumlah dokumen

df_i = jumlah kemunculan term terhadap i.

tf_{ij} = jumlah kemunculan i di j

E. Klasifikasi Model SVM

SVM merupakan teknik yang dipakai untuk mengklasifikasi teks ke dalam berbagai kategori sentimen. Pada awalnya, *SVM* digunakan pada dua kelas data untuk klasifikasi. Kemajuan *Support Vector Machine* (SVM) dapat mencakup penggabungan klasifikasi multi-kelas [17]. *SVM* dapat menangani data yang tidak berbentuk *linear* dengan memanfaatkan fungsi kernelnya, yang membantu dalam memetakan data ke dalam ruang dimensi fitur yang lebih tinggi untuk memungkinkan pemisahan yang baik antara kategori sentimen. Keunggulan *SVM* meliputi kemampuan generalisasi yang kuat dan kemampuan menghasilkan klasifikasi yang akurat. Algoritma ini berhasil menangani masalah yang berkaitan dengan pola non-linear karena mampu memanfaatkan konsep kernel untuk memproyeksikan data ke dalam ruang dimensi yang lebih besar [18]. Dapat dilihat pada Tabel I dibawah ini.

TABEL I RUMUS KERNEL

Rumus Kernel	Persamaan
Linear	$K(x_i, x) = x^T x \quad (2)$
Polynomial	$K(x_i, x) = (y \cdot x^T x + r)^p, y > 0 \quad (3)$
RBF	$K(x_i, x) = \exp(-\gamma x_i - x ^2), \gamma > 0 \quad (4)$
Sigmoid	$K(x_i, x) = \tanh(\gamma x^T x + r) \quad (5)$

F. Evaluasi Model

Pada penelitian ini, untuk mengevaluasi model dilakukan dengan *confusion matrix*. *confusion matrix* banyak digunakan dalam penelitian untuk mengevaluasi hasil dan mengukur kinerja suatu metode klasifikasi dengan parameter yang diukur adalah akurasi, presisi, recall, dan f-measure [19] Sedangkan *Stratified K-Fold Cross Validation* dilakukan untuk mengetahui performa algoritma dengan percobaan klasifikasi

sebanyak 5 kali dengan parameter yang sama sehingga mendapatkan hasil akurasi yang maksimal [20] Untuk mengevaluasi jumlah sampel dari semua jenis label data yang tidak seimbang sedangkan *confusion matrix* ini menyatakan total prediksi yang benar (*true positive* dan *true negative*) dan jumlah prediksi yang salah (*false positive* dan *false negative*). Dapat dilihat pada Tabel II.

TABEL II CONFUSION MATRIX

Nilai Aktual	Nilai Prediksi	
	Negatif	Positif
Negatif	TN	FP
Positif	FN	TP

G. Visualisasi Data

Visualisasi *data* ialah proses mendeskripsikan informasi data secara visual. Dengan menggunakan *wordcloud* atau *histogram*, visualisasi data dapat membantu menyoroti kata-kata kunci dalam setiap kategori sentimen. *Wordcloud* dipakai untuk melihat kata-kata yang sering muncul, *wordcloud* positif dapat menunjukkan pola positif, sedangkan kata-kata yang sering muncul dalam *wordcloud* negatif dapat menunjukkan ketidakpuasan.

III. HASIL DAN PEMBAHASAN

A. Crawling Data

Penelitian ini memakai Bahasa pemrograman *Python* untuk *crawling data* teks dari media sosial *Twitter* dengan proses pencarian data dilakukan dengan keyword yang digunakan "IKN", "Ibukota", dan "Pindah Ibukota". Data yang dikumpulkan berisi opini dari masyarakat mengenai kebijakan pemerintah terhadap Pemindahan ibu kota negara. Pengumpulan data dilakukan berhasil mengumpulkan total 4193 data dengan waktu pengumpulan dari 15-19 Desember 2023. Selanjutnya, data akan disimpan dalam format *CSV* untuk analisis lebih lanjut. Dapat dilihat pada Gambar. 2 Hasil *Crawling Data*.

No	Full text	Username
1	@putrasetyawan @kalotwdiem @barubikinlol seiring waktu tu brp lma bang? realita paling jls dah, pulau jawa itu, masih byk jg yg jalan & fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg? puluhan tahun smp jalan & fasilitasnya terbangun jg? knp ga lgsg aja bangun fasilitas di kalimantan?	tosee0810
2	Kalau pembangunannya inklusif, nggak perlu nunggu pindah ibukota dulu baru ada akses transportasi umum, pendidikan, kesehatan, hingga sanitasi yang merata.	permatadnda
3	@Alessandra_gy @mimih6mei @jokowi @bengkeldodo @alextham878 @Minietweets_ @yosephrosario_ @digifeeder @Raden_Wijaya008 @rojeoz @th_sitie @gtobing2903 lihat desain IKN ini emang uda layak ibukota pindah karena lebih strategis dan nyaman	ArabellaDe_
4	@xqfparaboyy6 Ganti2 mulu narasinya, kemarin bilang gak akan mindahin ibukota, trus ganti akan dikaji ulang untuk pemindahan ibukota, trus skrg bgini. Yg bner yg mna?	_ArryOn_
5	@tempodotco Klo pemindahan ibukota tidak urgent kenapa ide pemindahan itu muncul dr zaman presiden soekarno??	immanuel_padang

Gambar. 2 Hasil Crawling

B. Preprocessing Data

Preprocessing dilakukan pada data yang telah dikumpulkan untuk mengatasi gangguan atau ketidakkonsistenan sebelum data tersebut siap untuk dikelola. Tahapan *preprocessing* ini meliputi:

1. *Cleansing*

Tahapan *cleansing* bertujuan untuk menghilangkan karakter atau simbol tertentu seperti tautan *URL*, *username* atau *mention*, *hashtag*, *retweet*, dan emoji. Hasil Data yang telah melalui proses *cleansing* bisa ditinjau pada Tabel III.

TABEL III CLEANSING

Before Cleansing	After Cleansing
@MataNajwa Mirip program Transmigrasi, harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru	Mirip program Transmigrasi harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru
@putrasetyawan@kalotwdiem@barubikinl ol seiring waktu tu brp lma bang realita paling jls dah, pulau jawa itu, masih byk jg yg jalan & fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg? puluhan tahun smp jalan & fasilitasnya terbangun jg? knp ga lgsg aja bangun fasilitas di kalimantan?	seiring waktu tu brp lma bang realita paling jls dah pulau jawa itu masih byk jg yg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas di kalimantan
@Alessandra_gy@mimih6mei@jokowi@bengkeldodo@alextham878@Minietweets_yosephrosario_digifeeder@Raden_Wijaya008 @rojeoz @th_sitie @gtobing2903 lihat desain IKN ini emang uda layak ibukota pindah.karena lebih strategis dan nyaman	lihat desain IKN ini emang uda layak ibukota pindahkarena lebih strategis dan nyaman

2. *Case Folding*

Case folding adalah langkah dalam pengolahan teks yang bertujuan untuk membuat teks menjadi standar dengan mengonversi semua huruf dalam teks menjadi huruf kecil. Hasil dari proses ini dapat diamati dalam Tabel IV.

TABEL IV CASE FOLDING

Before Case Folding	After Case Folding
Mirip program Transmigrasi harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru	mirip program transmigrasi harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru
seiring waktu tu brp lma bang realita paling jls dah pulau jawa itu masih byk jg yg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas di kalimantan	seiring waktu tu brp lma bang realita paling jls dah pulau jawa itu masih byk jg yg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas di kalimantan
lihat desain IKN ini emang uda layak ibukota pindah karena lebih strategis dan nyaman	lihat desain ikn ini emang uda layak ibukota pindahkarena lebih strategis dan nyaman

3. *Tokenizing*

Tokenizing merupakan proses yang memungkinkan analisis teks dengan memecah teks jadi unit-unit kecil yang dikatakan token. Token ini bisa berupa kata-kata, frasa, atau bahkan karakter tunggal, tergantung pada teknik tokenisasi yang digunakan. Hasil dapat dilihat pada Tabel V.

TABEL V TOKENIZING

Before Tokenizing	After Tokenizing
mirip program transmigrasi harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru	mirip program transmigrasi harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru

seiring waktu tu brp lma bang realita paling jls dah pulau jawa itu masih byk jg yg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas di kalimantan	seiring waktu tu brp lma bang realita paling jls dah pulau jawa itu masih byk jg yg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas di kalimantan
lihat desain ikn ini emang uda layak ibukota pindah karena lebih strategis dan nyaman	lihat desain ikn ini emang uda layak ibukota pindah karena lebih strategis dan nyaman

4. Stopword Removal

Stopword Removal merupakan tahap dalam proses pembersihan teks dalam menghapus kata-kata yang sering muncul dalam teks, namun tidak mempunyai arti penting pada analisa sentimen. Hasil dari tahapan ini ditampilkan pada Tabel VI.

TABEL VI STOPWORD REMOVAL

Before Stopword Removal	After Stopword Removal
mirip program transmigrasi harus fight kalau mau bertahan hidup sebagai pembuka ladang garapan ibukota baru	mirip program transmigrasi fight mau bertahan hidup pembuka ladang garapan ibukota baru
seiring waktu tu brp lma bang realita paling jls dah pulau jawa itu masih byk jg yg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota apakah bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas di kalimantan	seiring waktu tu brp lma bang realita paling jls dah pulau jawa byk jg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas kalimantan
lihat desain ikn ini emang uda layak ibukota pindah karena lebih strategis dan nyaman	lihat desain ikn emang uda layak ibukota pindah karena strategis nyaman

5. Stemming

Stemming merupakan tahapan di mana kata-kata dikembalikan ke bentuk dasarnya atau kata dasar. Hasil dari langkah ini bisa ditinjau pada Tabel VII.

TABEL VII STEMMING

Before Stemming	After Stemming
mirip program transmigrasi fight mau bertahan hidup pembuka ladang garapan ibukota baru	mirip program transmigrasi fight mau tahan hidup buka ladang garap ibukota baru
seiring waktu tu brp lma bang realita paling jls dah pulau jawa byk jg jalan amp fasilitas daerahnya jelek pdhl dah puluhan thn jkt jd ibukota bakal sm jg puluhan tahun smp jalan amp fasilitasnya terbangun jg knp ga lgsg aja bangun fasilitas kalimantan	iring waktu tu brp lma bang realita paling jls dah pulau jawa byk jg jalan amp fasilitas daerah jelek pdhl dah puluh thn jkt jd ibukota bakal sm jg puluh tahun smp jalan amp fasilitas bangun jg knp ga lgsg aja bangun fasilitas kalimantan
lihat desain ikn emang uda layak ibukota pindah karena strategis nyaman	lihat desain ikn emang uda layak ibukota pindah karena strategis nyaman

C. Labeling Data

Sebelum memasuki tahap klasifikasi model, peneliti melakukan pelabelan data. Dalam penelitian ini, data diberi label menjadi dua kategori, yaitu Positif dan Negatif. Contoh sentimen dapat dilihat pada Tabel VIII.

TABEL VIII LABELING DATA

Komentar	Sentimen
bangun ikn tetap lanjut pilih pemimpin tepat dukung lanjut pembangunan ikn merata bangun	Positif
gimana mau ngurus didik wong duit habis makan gratis ikn aja ragu lanjut tangan aki gemoy	Negatif

D. Klasifikasi Model *Support Vector Machine*

Dalam penelitian ini, *TF-IDF* digunakan untuk menghitung bobot setiap kata dalam hasil data. Kemudian, data dipisahkan menjadi data training dan testing dengan perbandingan 80% untuk data training dan 20% untuk data testing. Ini berarti terdapat 3069 data training dan 768 data testing. Proses ini dilakukan menggunakan *Platform Google Colab* dengan bahasa pemrograman *Python*. Hasil dari proses tersebut dapat dilihat dalam Tabel IX di bawah ini.

TABEL IX HASIL AKURASI KERNEL

Kernel Linear	Precision	Recall	F1-Score
Negatif	93%	100%	96%
Positif	80%	7%	12%
Accuracy	92.70%		

Berdasarkan hasil eksperimen, dapat disimpulkan bahwa penggunaan kernel linear dalam analisis sentimen terkait pemindahan Ibu Kota Negara memberikan tingkat akurasi tertinggi sebesar 92.70%. Ini menunjukkan bahwa model dengan *kernel linear* memberikan performa yang lebih baik dalam analisis sentimen suatu topik. Setelah itu dilakukan evaluasi model dengan *confusion matrix*. Hasil dapat dilihat pada Tabel X.

TABEL X HASIL CONFUSION MATRIX

Nilai Aktual	Nilai Prediksi	
	Negatif	Positif
Negatif	708	4
Positif	1	55

Setelah dilakukan evaluasi model dengan *confusion matrix* maka di validasi dengan *Stratified K-Fold Cross Validation*. Dalam menilai performa model dengan presisi yang lebih spesifik, maka Model klasifikasi divalidasi menggunakan *Stratified K-Fold Cross Validation* dengan data latih dari pengujian awal. Dalam penelitian ini, dilakukan 5 iterasi *Stratified K-Fold Cross Validation*. Selama setiap iterasi, kombinasi data latih dan data uji evaluasi oleh *Stratified K-Fold Cross Validation*. Meskipun terdapat ketidakseimbangan dalam jumlah sampel dari semua jenis label data yang digunakan, langkah ini bertujuan untuk memperkuat bukti dan kesimpulan yang dihasilkan dari evaluasi model. Hasil evaluasi ditemukan dalam Tabel XI.

TABEL XI HASIL K-FOLD

Iterasi Ke -	Accuracy
1	100%
2	99.92%
3	99.92%
4	99.92%
5	100%

Kemudian, dilakukan visualisasi data dengan *Wordcloud* untuk setiap kelas guna menampilkan kata-kata dominan dalam kelas positif dan negatif. Gambar. 3 menampilkan *Wordcloud* untuk kelas positif, sementara Gambar. 4 menampilkan *Wordcloud* untuk kelas negatif. *Wordcloud* memberikan representasi visual dari kata-kata yang paling menonjol atau sering muncul dalam setiap kategori sentimen ini dapat membantu pemahaman pola lebih lanjut

membantu dalam menyoroti kata-kata yang sering muncul dalam dataset. Dalam penelitian berikutnya, Dapat digunakan beberapa metode termasuk XGBoost untuk memaksimalkan mendapatkan akurasi lebih baik dari metode yang digunakan dalam penelitian ini.

UCAPAN TERIMAKASIH

Terimakasih kepada Universitas Teknokrat Indonesia, Dosen Pembimbing yang telah mendukung dan membimbing dalam penelitian ini. Serta orangtua dan keluarga yang selalu memberikan semangat dan do'a sampai menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- [1] A. Ernawati, J. Wahyudi, A. D. Astuti, and S. Q. Aini, "Analisis Kelayakan Pendirian Usaha Pengolahan Limbah Medis untuk Meningkatkan Pendapatan Asli Daerah," *Jurnal Ekonomi dan Kebijakan Publik*, vol. 13, no. 1, pp. 57–70, 2022, doi: 10.22212/jekp.v13i1.2155.
- [2] D. Ocsanda, C. Ilham Wijaya, M. Azzam Al Haq, J. Dwi Efendi, and F. Prihantoro, "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, 2023, doi: 10.47441/jkp.v18i1.291.
- [3] S. Hasan and H. Haeran, "Strategi Perpindahan Ibu Kota Negara (IKN) Ditinjau dari Perspektif Pertahanan Maritim," *SOSIO DIALEKTIKA*, vol. 8, no. 1, p. 117, Jul. 2023, doi: 10.31942/sd.v8i1.8697.
- [4] S. Aripriyanto, T. Tukino, A. Sufyan, and R. Nandaputra, "Sentimen Analisis Twitter Ibu Kota Negara Nusantara Menggunakan Long Short-Term Memory dan Lexicon Based," *EXPERT: Jurnal Manajemen Sistem Informasi dan Teknologi*, vol. 12, no. 2, p. 119, Dec. 2022, doi: 10.36448/expert.v12i2.2821.
- [5] D. J. Putnarubun and C. F. Palembang, "PENERAPAN METODE MACHINE LEARNING-NAIVE BAYES PADA ANALISIS SENTIMEN PEMINDAHAN IBU KOTA NEGARA BARU Implementation of Machine Learning-Naïve Bayes Method to Sentiment Analysis of New Country Capitals".
- [6] R. Nugroho and E. Setijaningrum, "Indonesia's New Capital: An Analysis Of Transparency, Public Participation, And Deliberative Democracy," *Journal of Namibian Studies*, vol. 35, pp. 1842–1874, 2023.
- [7] E. Sukmana and A. Azizah, "New Capital City of Indonesia, an Opportunity or Threat for Ecotourism Resilience in East Borneo," *Journal of Disaster Research*, vol. 19, no. 1, pp. 214–225, 2024, doi: 10.20965/jdr.2024.p0214.
- [8] A. Kurniawan and S. Waluyo, "Penerapan Algoritma Naive Bayes Dalam Analisis Sentimen Pemindahan Ibukota Pada Twitter Application Of Naive Bayes Algorithm In Capital Movement Sentiment Analysis On Twitter," 2022. [Online]. Available: <https://senafiti.budiluhur.ac.id/index.php>
- [9] D. Ambarwati Puspitasari, Y. Karlina, and B. M. Mulyo, "Jurnal Penelitian Pendidikan Bahasa dan Sastra Indonesia," vol. 9, pp. 239–257, 2023, doi: 10.22202/jg.v9i2.6947.
- [10] H. Hamjen and L. R. Said, "Exploring Research Trends in the Indonesian Archipelago's New National Capital," *Journal of International Conference Proceedings*, vol. 6, no. 7, pp. 125–135, Dec. 2023, doi: 10.32535/jicp.v6i7.2767.
- [11] A. Witanti, B. Yogyakarta Jl Raya Wates-Jogjakarta, K. Sedayu, K. Bantul, and D. Istimewa Yogyakarta, "ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)," *Jurnal Sistem Informasi dan Informatika (Simika) P-ISSN*, vol. 5, pp. 2622–6901, 2022.
- [12] A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, "SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM," *JDMISI*, vol. 2, no. 1, pp. 31–37, 2021, [Online]. Available: <https://t.co/NfhfMjtXw>
- [13] D. Alita, Y. Fernando, and H. Sulistiani, "IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER," *Jurnal TEKNOKOMPAK*, vol. 14, no. 2, p. 86, 2020.
- [14] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 3, p. 650, Jul. 2020, doi: 10.30865/mib.v4i3.2181.
- [15] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *Jambura Journal of Electrical and Electronics Engineering*, vol. 5, no. 1, pp. 32–35, 2023, doi: 10.37905/jjee.v5i1.16830.
- [16] D. Darwis, E. Shintya Pratiwi, A. Ferico, and O. Pasaribu, "PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA," 2020.
- [17] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," vol. 6, no. 3, 2021, [Online]. Available: <http://situs.com>

- [18] P. Kumala Sari and R. Randy Suryono, "KOMPARASI ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN METAVERSE," 2024.
- [19] A. Rahman *et al.*, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen".
- [20] E. Febriyani and H. Februariyanti, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter," vol. 17, no. 1.