



Tersedia Online : <http://e-journals.unmul.ac.id/>

ADOPSI TEKNOLOGI DAN SISTEM INFORMASI (ATASI)

Alamat Jurnal : <http://e-journals2.unmul.ac.id/index.php/atasi/index>



Analisis Sentimen Twitter Ibu Kota Negara Nusantara Menggunakan Algoritma Naive Bayes, Logistic Regression dan K-Nearest Neighbors

Alisya Nisrina Sativa ^{1)*}, Avinka Rizky ²⁾, Imelda Putri ³⁾, Juventia Adelia Putri ⁴⁾,
Akhmad Irsyad ⁵⁾, Islamiyah ⁶⁾

Program Studi Sistem Informasi, Fakultas Teknik, Universitas Mulawarman

E-Mail : alisya.nisrina@gmail.com¹⁾; avinkarizky242@gmail.com²⁾; imelptrii27@gmail.com³⁾;
adelia010604@gmail.com⁴⁾; akhmadirsyad@ft.unmul.ac.id⁵⁾; islamiyahunmul@gmail.com⁶⁾

ARTICLE INFO

Article history:

Received : 18 December 2024
Revised : 23 December 2024
Accepted : 29 December 2024
Available online : 30 December 2024

Keywords:

KNN
IKN
Nusantara
Naive Bayes
Logistic Regression

ABSTRACT

Indonesia plans to move its capital from Jakarta to eastern Kalimantan under the name Nusantara. On the other hand, the new capital-building plan of Nusantara in eastern Kalimantan is called a public policy expert that will squeeze the state's incomes and expenditure (APBN) budget post in the midst of massive spending on health care and economic recovery from the Covid-19 pandemic. The relocation plan harvested a number of citizens' opinions on Twitter. The sentiment analysis research was conducted to find out the response of tweets related to the transfer of the Indonesian capital after the passing of the ICT Act as well as the comparison of the number of positive and negative tweets. The analysis process involves processing text data, creating a word cloud, and applying regression algorithms like Logistic Regression, classifications like Naive Bayes and K-Nearest Neighbors to predict sentiment. The Naive Bayes model's accuracy is the best algorithm model achieving a precision of 0.9480, with precision 0.9399, recall 0.9609, and f1-score 0.9503. This research provides in-depth insight into public opinion related to the transfer of the state capital of Nusantara.

ABSTRAK

Indonesia berencana memindahkan ibu kota negara dari Jakarta ke Kalimantan Timur dengan nama Nusantara. Di sisi lain rencana pembangunan Ibu Kota Negara (IKN) baru Nusantara di Kalimantan Timur disebut pakar kebijakan publik akan memberatkan pos anggaran pendapatan dan belanja negara (APBN) yang terjepit di tengah pengeluaran besar untuk penanganan kesehatan dan pemulihan ekonomi akibat pandemi Covid-19. Rencana pemindahan tersebut menuai beberapa pendapat warga negara di Twitter. Penelitian *sentiment analysis* ini dilakukan untuk mengetahui respon tweet terkait pemindahan Ibu Kota Indonesia setelah disahkannya Undang-Undang IKN serta perbandingan jumlah tweet yang positif dan negatif. Proses analisis melibatkan pengolahan data teks, pembuatan *word cloud*, serta penerapan algoritma klasifikasi seperti *Logistic Regression*, seperti *Naive Bayes* dan *K-Nearest Neighbors* untuk memprediksi sentimen. Hasil penelitian menunjukkan distribusi sentimen dengan mayoritas tweet bersentimen positif, diikuti oleh netral dan negatif. Akurasi model *Naive Bayes* menjadi model algoritma terbaik mencapai akurasi 94.80%, dengan *precision* 93.99%, *recall* 96.09%, dan *f1-score* 95.03%. Penelitian ini memberikan wawasan yang mendalam tentang opini publik terkait pemindahan Ibu Kota Negara Nusantara.

Kata Kunci :

KNN
IKN
Nusantara
Naive Bayes
Logistic Regression

APA style in citing this article:

Sativa, A. N., Rizky, A.,
Imelda Putri, Putri, J. A.,
Akhmad Irsyad, & Islamiyah.
Analisis Sentimen Twitter Ibu
Kota Negara Nusantara
Menggunakan Algoritma
Naive Bayes, Logistic
Regression dan K-Nearest
Neighbors. *Adopsi Teknologi
Dan Sistem Informasi
(ATASI)*, 3(2).
<https://doi.org/10.30872/atasi.v3i2.1371>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

*) Corresponding Author

<https://doi.org/10.30872/atasi.v3i2.1371>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

1. PENDAHULUAN

Dalam upaya pemindahan Ibu Kota Nusantara telah lama menjadi wacana pemerintah, tetapi tepatnya pada tahun 2017 upaya tersebut kembali dimunculkan oleh kementerian. Mengenai pemindahan Ibu Kota Negara Nusantara, yakni pada masa Presiden Bapak Ir.H.Joko Widodo, Pembangunan Ibu Kota Nusantara pada pertengahan Maret 2022 akan direlokasi secara bertahap dari tahun 2024 hingga 2045. Tepat pada Senin, 26 Agustus 2019, Melalui siaran pers, lokasi baru IKN berada di Kabupaten Penajam Paser Utara, Provinsi Kalimantan Timur. Alasan pemindahan Ibu Kota dikarenakan beban kota Jakarta sebagai ibu kota sudah terlalu berat, bukan hanya dari sisi fungsi tetapi dari pertumbuhan penduduk yang juga kian meningkat, jumlah penduduk di Pulau Jawa berkisar 56% dari total seluruh warga di Indonesia. Menurut data dari Badan Pencatatan Statistik jumlah penduduk di Daerah Khusus Ibukota Jakarta pada tahun 2021 berjumlah 10.609.681 jiwa.

Dalam hal tersebut adanya pemindahan Ibu Kota Negara Indonesia, tentu mengundang berbagai reaksi, terutama bagi masyarakat Indonesia. Mengingat ibu kota negara baru di Indonesia merupakan hal yang sensitif sehingga ramai diperbincangkan di media sosial termasuk Twitter. Twitter merupakan sosial media yang memungkinkan pengguna dapat mengakses informasi yang sedang terjadi saat ini dengan cepat. Twitter digunakan seseorang untuk mengekspresikan pendapatnya dengan pesan yang biasa disebut tweet. Tweet yang dikirimkan dapat berupa teks, gambar maupun video. Pengguna Twitter bebas menyampaikan pendapat atau ekspresi mereka tentang layanan, fasilitas atau isu politik atau hal-hal yang sedang dibahas. Walaupun saat ini Undang-Undang IKN telah disahkan tetapi masih menuai beberapa pendapat warga negara di Twitter. Pendapat tersebut berupa pendapat positif dan negatif akan pemindahan Ibu Kota Indonesia. Untuk mengetahui respon tweet warga negara terkait pemindahan Ibu Kota Indonesia setelah disahkan Undang-Undang IKN maka dilakukan penelitian *sentiment analysis*.

Analisis Sentimen adalah suatu metode otomatis yang digunakan untuk memahami, mengekstrak, dan memproses data teks dengan tujuan untuk memperoleh wawasan atau informasi yang berharga dan bertujuan untuk mengklasifikasikan teks dalam sebuah kalimat. Teknik ini sering digunakan untuk mengevaluasi dan menganalisis seberapa puas pengguna atau pelanggan terhadap suatu produk atau kebijakan tertentu. Penambangan data atau proses menggunakan alat dan teknik tertentu untuk mencari pola atau wawasan berharga dalam data dalam jumlah besar. Penambangan data (data mining) dapat dilakukan dengan menggunakan sejumlah pendekatan, metode, dan algoritma yang berbeda.

Proses analisis sentimen melibatkan visualisasi data untuk memahami distribusi sentimen dan penggunaan algoritma klasifikasi untuk memprediksi sentimen dari teks yang tidak berlabel. Hasil dari penelitian ini menunjukkan distribusi emosi. Mayoritas tweet memiliki sentimen positif, disusul sentimen netral dan sentimen negatif. Oleh karena itu, penelitian ini tidak hanya memberikan gambaran opini masyarakat terhadap pemindahan Ibu Kota Negara, namun juga menyoroti keunggulan model *Logistic Regression* dibandingkan *Naive Bayes* dan juga *K-Nearest Neighbors* dalam tugas klasifikasi sentimen. Hasil-hasil ini dapat digunakan untuk lebih memahami preferensi dan kekhawatiran masyarakat, sehingga dapat mempermudah perencanaan dan pelaksanaan perpindahan Ibu Kota Negara yang lebih efektif.

2. TINJAUAN PUSTAKA

A. Analisis Sentimen

Analisis sentimen adalah proses ekstraksi, Secara otomatis memproses dan memahami data dalam bentuk teks tidak terstruktur untuk mengambil informasi sentimen yang terkandung dalam kalimat opini atau opini (Brah, Touahria, & Tari, 2019). Analisis sentimen untuk menilai opini dan kecenderungan terhadap suatu topik memiliki kecenderungan negatif dan positif (Rozi, Pramono & Dahlan, 2012). Media sosial Twitter ini membuka jendela bagi peneliti untuk mempelajari sentimen dan opini publik melalui analisis sentimen (Qiu, Lin, dan Shuai, 2019). Tugas dasar analisis sentimen adalah mengklasifikasikan teks ke dalam kalimat atau dokumen dan menentukan apakah pendapat yang diungkapkan dalam kalimat atau dokumen tersebut positif atau negatif.

B. Penelitian Terkait

Pada penelitian-penelitian yang telah dilakukan sebelumnya, telah menunjukkan bahwa Analisis Sentimen dalam kontes tentang pemindahan Ibu Kota Negara itu penting. Dalam penelitian yang berjudul "*Analisis Sentimen Masyarakat Indonesia terhadap Pemindahan Ibu Kota Negara Indonesia pada Twitter*" yang ditulis oleh Sri Lestari, Mupaat, Adhitia Erfina (2022), menunjukkan jika masyarakat Indonesia yang beranggapan positif terhadap Ibu Kota Negara baru Indonesia. Lalu pada penelitian lainnya yang berjudul "*Analisis Sentimen Twitter Terhadap Perpindahan Ibu Kota Negara Ke IKN Nusantara Menggunakan Orange Data Mining*" yang ditulis oleh Hafizh Dhery Al Assyam, Firman Noor Hasan menunjukkan jika yaitu (*surprise*) kaget sebanyak 80%, diikuti dengan bentuk emosi kedua (suka cita) sebanyak 50%, kemudian bentuk emosi ketiga (sedih) sebanyak 15%, kemudian bentuk emosi keempat (takut) sebanyak 10%. Penelitian ke-3 dengan judul "*PENERAPAN ALGORITMA KLASIFIKASI NAÏVE BAYES UNTUK ANALISIS SENTIMEN TENTANG PEMINDAHAN IBU KOTA NEGARA*" yang ditulis oleh Abdul Halim Hasugian, Raissa Amanda Putri, Mulya Alfian Simatupang, menunjukkan jika rincian data dalam kelas positif sebanyak 66% yaitu 400 data, dan negatif sebanyak 34% yaitu 206 data.

*) Corresponding Author

<https://doi.org/10.30872/atasi.v3i2.1371>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

3. METODE PENELITIAN

A. Algoritma Klasifikasi

1. Logistic Regression

Algoritma yang digunakan adalah *logistic regression*. *Logistic regression* adalah teknik analisis data dalam statistika yang dirancang untuk mengetahui hubungan antar variabel, dimana variabel respon bersifat kategorikal, baik nominal maupun ordinal, dan variabel penjelas bersifat kategoris atau kontinu (Ramli et al., 2013).

2. Naive Bayes

Naive Bayes merupakan klasifikasi yang bersifat *supervised learning* karena memiliki supervisor (manusia melakukan klasifikasi secara manual pada data yang digunakan dalam pelatihan) selaku pengajar dalam proses belajar atau *learning*. Selain itu, performansi *Naive Bayes* memiliki waktu klasifikasi yang singkat sehingga mempercepat proses sistem analisis sentimen.

3. K-Nearest Neighbors

K-Nearest Neighbor (KNN) adalah suatu metode yang menggunakan algoritma *supervised* di mana hasil dari *query instance* yang baru diklasifikasi berdasarkan mayoritas dari kategori pada KNN (Sikki, 2009). Tujuan dari algoritma KNN adalah untuk mengklasifikasi objek baru berdasarkan atribut dan *training samples*. Di mana hasil dari sampel uji yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN.

B. Pengumpulan Data

Kegiatan ini merupakan langkah awal dalam melakukan penelitian atau analisis data untuk memperoleh informasi yang relevan. Pengumpulan data merupakan langkah penting dalam siklus penelitian dan seringkali memerlukan perencanaan yang matang untuk menjamin kualitas data yang diperoleh. Dalam hal ini, kami mengambil dataset dari situs :

https://github.com/varaah/analisis-sentimen_ikn/tree/525aa102a23bff3eb46be40f240cb43c2841746a/data

1. Pre-Processing

Pra-pemrosesan adalah proses menyiapkan data mentah untuk dianalisis atau digunakan dalam model pembelajaran mesin. Ini adalah langkah penting dalam proses pembelajaran mesin, karena kualitas dan karakteristik data dapat mempengaruhi performa model secara signifikan. Untuk memaksimalkan output preprocessing dari penelitian sebelumnya, penelitian ini melakukan beberapa tahapan *pre-processing* yaitu:

- Case folding*: Sebuah proses untuk mengedit teks dokumen ke dalam bentuk lower case
- Cleaning*: Data yang digunakan perlu adanya proses pembersihan data seperti symbol tautan URL, angka. Data Twitter sendiri tentunya banyak data kotor seperti tagar, angka, nama pengguna, URL dan teks retweet
- Tokenize*: Proses data yang sebelumnya kalimat kemudian dipecah menjadi kata per kata
- Stopword Removal*: Kata yang terdapat di dalam stoplist akan melalui tahap pembersihan
- Weighting*: dalam proses ini untuk pembobotan kata dengan TF-IDF.

2. TF-IDF

Pembobotan kata adalah proses pemberian bobot untuk setiap kata yang terdapat dalam sebuah dokumen. Dalam pencarian informasi peringkat berdasarkan frekuensi kata, salah satu metode yang paling populer adalah metode TFIDF (*Term Frequency - Inversed Document Frequency*). Dalam metode TF-IDF, *Term Frequency* lebih berfokus pada istilah yang sering muncul dalam suatu dokumen sedangkan *Inverse Document Frequency* lebih berfokus pada pemberian bobot rendah untuk istilah yang muncul dalam banyak dokumen.

Rumus dari TF-IDF:

$$TF - IDF = TF \times IDF$$

Keterangan : d = dokumen ke-d, t = kata ke-t dari kata kunci, W = bobot dokumen ke-d terhadap kata ke-t, tf = term frekuensi/frekuensi kata

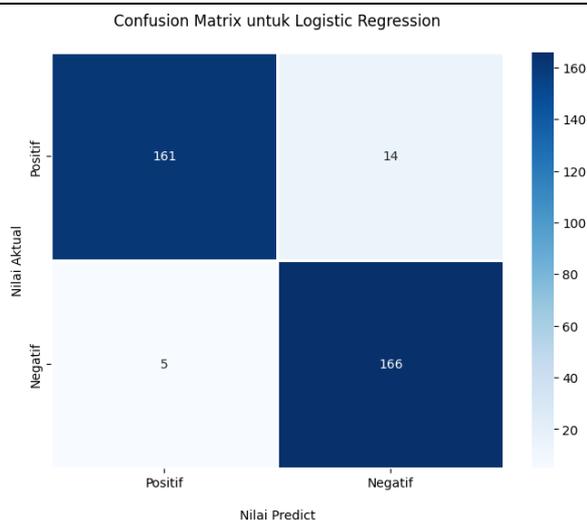
4. HASIL DAN PEMBAHASAN

Penelitian ini melakukan analisis sentimen terhadap tweet yang berkaitan dengan pemindahan Ibu Kota Negara. Hasil distribusi sentimen dari dataset yang telah dianalisis: Sentimen Positif: 50.14%, dan Sentimen Negatif: 49.86%. Menunjukkan bahwa mayoritas tweet memiliki sentimen positif terhadap pemindahan Ibu Kota Negara dibanding dengan sentimen negatif yang memiliki persentase yang lebih rendah dibanding sentimen positif.

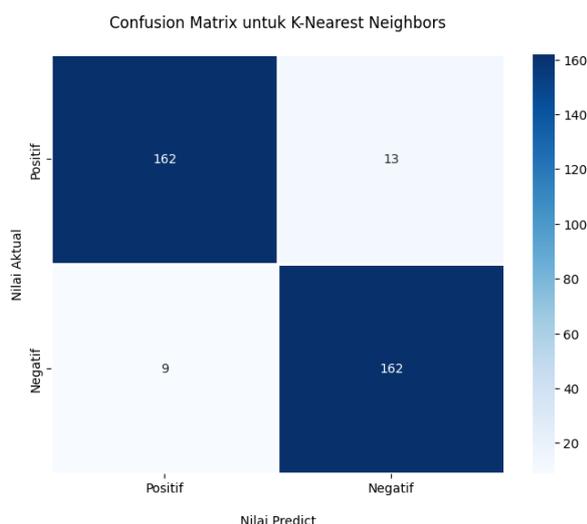
*) Corresponding Author

<https://doi.org/10.30872/atasi.v3i2.1371>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.



Gambar 5. Confusion Matrix Logistic Regression



Gambar 6. Confusion Matrix K-Nearest Neighbors

5. KESIMPULAN

Analisis sentimen opini terhadap Ibu Kota Nusantara di media sosial, khususnya Twitter, memiliki relevansi signifikan dalam memahami pandangan publik terkait dengan aspek-aspek tertentu dari ibu kota. Ini membuktikan bahwa pemahaman sentimen masyarakat dapat memberikan wawasan berharga untuk perbaikan dan pengembangan. Hasil analisis menunjukkan bahwa mayoritas tweet mengandung sentimen positif terhadap pemindahan Ibu Kota Negara, diikuti oleh sentimen negatif. Distribusi sentimen ini mencerminkan adanya dukungan yang signifikan dari masyarakat di media sosial.

Dalam Penelitian ini setelah menyelesaikan tahapan analisis data sentimen Twitter terhadap pembangunan Ibu Kota Negara Nusantara, dapat disimpulkan, diantaranya data tweet yang diperoleh dianalisis menggunakan metode *Naive Bayes*, *Logistic Regression*, dan *K-Nearest Neighbors*. Dengan algoritma terbaik adalah *Naive Bayes* dengan persentase Akurasi nya adalah 94.80%, *Precision* 93.99%, *Recall* 96.09%, dan *F1-Score* nya adalah 95.03%.

6. DAFTAR PUSTAKA

- Al Assyam, H. D., & Hasan, F. N. (2023). Analisis sentimen Twitter terhadap perpindahan ibu kota negara ke IKN nusantara menggunakan orange data mining. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(1), 341-349
- Aripiyanto, S., Tukino, T., Sufyan, A., & Nandaputra, R. (2022). Sentimen analisis Twitter Ibu Kota Negara Nusantara menggunakan long short-term memory dan lexicon based. *EXPERT: Jurnal Manajemen Sistem Informasi Dan Teknologi*, 12(2), 119-125.
- Gunawan, B., Sastypratiwi, H., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 4(2), 113-118.
- Hasugian, A. H., Putri, R. A., & Simatupang, M. A. (2024). PENERAPAN ALGORITMA KLASIFIKASI NAÏVE BAYES UNTUK ANALISIS SENTIMEN TENTANG PEMINDAHAN IBU KOTA NEGARA.

*) Corresponding Author

<https://doi.org/10.30872/atasi.v3i2.1371>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

JOURNAL OF SCIENCE AND SOCIAL RESEARCH, 7(2), 635-644.

- Kurniawan, A., & Waluyo, S. (2022, September). Penerapan Algoritma Naive Bayes Dalam Analisis Sentimen Pindahan Ibukota Pada Twitter. In *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)* (Vol. 1, No. 1, pp. 455-461).
- Lestari, S., Mupaat, M., & Erfina, A. (2022). Analisis Sentimen Masyarakat Indonesia terhadap Pindahan Ibu Kota Negara Indonesia pada Twitter. *JUSIFO (Jurnal Sistem Informasi)*, 8(1), 13-22.
- Novantika, A., & Sugiman, S. (2022, February). Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunakan Metode SVM dan Logistic Regression. In *PRISMA, Prosiding Seminar Nasional Matematika* (Vol. 5, pp. 808-813).
- Prasetyo, S. D., Hilabi, S. S., & Nurapriani, F. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN. *Jurnal KomtekInfo*, 1-7.
- Safra, I. A., & Zuliarso, E. (2020). Analisa sentimen persepsi masyarakat terhadap pemindahan ibukota baru di kalimantan timur pada media sosial twitter.

*) Corresponding Author

<https://doi.org/10.30872/atasi.v3i2.1371>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.