

**ANALISIS SENTIMEN MEDIA SOSIAL
TERKAIT ISU PEMBANGUNAN IBU KOTA NEGARA (IKN)
MENGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE**

Muhammad Nur Alif Haqiqi¹, Zaehol Fatah²

Universitas Ibrahimy

unliff12@gmail.com, zaeholfatah@gmail.com

ABSTRAK

Penelitian ini dilatarbelakangi oleh beragamnya opini publik di media sosial terkait isu pembangunan Ibu Kota Negara (IKN), yang menimbulkan sentimen positif, negatif, dan netral. Untuk mengklasifikasikan data komentar yang tidak terstruktur secara sistematis, penelitian ini bertujuan untuk mengetahui kecenderungan opini masyarakat serta membandingkan performa metode “Naive Bayes” dan “Support Vector Machine” (SVM). Metodologi yang digunakan adalah kuantitatif dengan pendekatan analisis sentimen berbasis “Machine Learning”. Data bersumber dari komentar YouTube tentang pembangunan IKN yang dikumpulkan melalui “Web Scraping” dan diseleksi menggunakan teknik “Purposive Sampling”. Tahapan penelitian meliputi “Preprocessing”, pembobotan TF-IDF, klasifikasi menggunakan algoritme “Naive Bayes” dan SVM, serta evaluasi model menggunakan akurasi, “Precision”, “recall”, dan “F1-Score”. Hasil penelitian menunjukkan bahwa metode “Naive Bayes” memperoleh akurasi sebesar 79,8%. Sementara itu, metode SVM menghasilkan performa yang jauh lebih unggul dengan tingkat akurasi mencapai 94,2%. Kesimpulan dari penelitian ini membuktikan bahwa metode SVM lebih efektif dibandingkan “Naive Bayes” dalam mengklasifikasikan sentimen publik terhadap isu pembangunan IKN. Penelitian ini diharapkan dapat menjadi acuan dalam memahami kecenderungan opini masyarakat serta pengembangan analisis sentimen berbahasa Indonesia.

Kata Kunci: Analisis Sentimen, IKN, Naive Bayes, Support Vector Machine, Media Sosial.

ABSTRACT

This research is based on public opinion on social media related to the development of the National Capital (IKN), which gives rise to positive, negative, and neutral sentiments. In order to systematically classify unstructured comment data, this research aims to find out the tendency of public opinion and compare the performance of "Naive Bayes" and "Support Vector Machine" (SVM) methods. The methodology used is quantitative with a sentiment analysis approach based on "Machine Learning". Data sourced from YouTube comments about the development of IKN collected through "Web Scraping" and using the "Purposive Sampling" technique. The research stages include "Preprocessing", TF-IDF weighting, classification using "Naive Bayes" and SVM algorithms, as well as evaluation models using precision, "Precision", "recall", and "F1-Score". Research results show that the "Naive Bayes" method obtains an accuracy of 79.8%. Meanwhile, the SVM method produces far superior performance with an achievement level of 94.2%. The conclusion of this research proves that the SVM method is more effective than "Naive Bayes" in classifying public sentiment towards the issue of IKN development. This research is expected to be a reference in understanding the tendency of opinion as well as the development of sentiment analysis in Indonesian.

Keywords: Sentiment Analysis, IKN, Naive Bayes, Support Vector Machine, Social Media.

Pendahuluan

Perkembangan teknologi informasi dan komunikasi membawa perubahan besar terhadap bentuk komunikasi masyarakat, terutama melalui adanya media sosial sebagai ruang publik digital. Media sosial tidak hanya digunakan sebagai sarana komunikasi pribadi, tetapi kian menjadi tempat masyarakat menyampaikan opini, kritik, dan juga tanggapan pada berbagai isu sosial serta kebijakan pemerintah secara cepat dan luas[1]. Dalam hal kebijakan publik, opini yang berkembang di media sosial bisa mencerminkan tingkat penerimaan maupun penolakan masyarakat pada suatu kebijakan, sehingga penting untuk dianalisis secara sistematis[2].

Salah satu kebijakan strategis nasional yang menimbulkan bermacam respons publik ialah pembangunan Ibu Kota Negara (IKN) di Samarinda, Kalimantan Timur. Kebijakan ini dilihat sebagai langkah pemerintah dalam memberikan pemerataan pembangunan dan pusat perkembangan ekonomi baru di Indonesia. Namun, pembangunan IKN juga menimbulkan perdebatan di masyarakat terkait persiapan infrastruktur, dampak lingkungan, hingga penggunaan keuangan negara[3]. Perbedaan pandangan ini banyak disampaikan masyarakat melalui media sosial, khususnya YouTube, sehingga membentuk sentimen positif, negatif, maupun netral terkait isu pembangunan IKN[4].

Banyaknya data opini yang dihasilkan pada media sosial menimbulkan tantangan dalam proses analisis jika dilakukan secara manual karena data berbentuk teks tidak terstruktur dan jumlahnya sangat besar. Dengan demikian, diperlukan pendekatan komputasional yang bisa membentuk data teks secara baik dan efisien. Analisis sentimen adalah salah satu cara pada text mining yang digunakan untuk menentukan serta mengelompokkan opini masyarakat ke dalam bentuk sentimen tertentu, seperti positif, negatif, dan netral[4].

Penelitian ini menerapkan strategi machine learning beserta menerapkan metode Naive Bayes dan Support Vector Machine (SVM). Naive Bayes dikenal dengan algoritma klasifikasi yang sangat sederhana dan efisien dalam pengolahan data teks, sedangkan Support Vector Machine mempunyai kemampuan yang baik saat menangani data beraspek tinggi dan menghasilkan performa klasifikasi yang sangat stabil[5]. Penelitian terdahulu membuktikan bahwa kedua metode tersebut mampu dan bisa memberikan fase akurasi yang baik dalam analisis sentimen berbasis media sosial, khususnya pada data teks berbahasa Indonesia.[6].

Materi dan Metode

Penelitian ini menerapkan bentuk pengkajian kuantitatif pada pendekatan analisis sentimen berbentuk machine learning. Penelitian dilakukan dengan cara mengolah data teks dari media sosial menjadi data statistik yang diklasifikasikan pada bentuk sentimen positif, negatif, dan netral. Maka dari itu, penelitian ini juga menerapkan pengujian terhadap dua algoritma, yaitu Naïve Bayes dan SVM, untuk menganalogikan tingkat akurasi kinerja kedua metode dalam mengklasifikasikan sentimen masyarakat terkait isu pembangunan Ibu Kota Negara (IKN).

Populasi pada penelitian ini yaitu semua komentar pengguna media sosial YouTube yang berkaitan pada isu pembangunan Ibu Kota Negara (IKN). Data komentar didapatkan melalui proses web scraping berdasarkan kata kunci yang berhubungan dengan pembangunan IKN.

Sampel penelitian terdiri dari 3.541 komentar YouTube terkait pembangunan IKN yang sesuai dengan topik penelitian ini serta telah diproses serta preprocessing data, seperti cleaning, stemming, stopword removal, dan tokenizing. Data yang digunakan kemudian dikelompokkan ke dalam kategori sentimen positif, negatif, dan netral. Teknik pemilihan sampel yang dipakai adalah purposive sampling, yaitu memilih komentar yang tepat pada tujuan pengkajian untuk dianalisis

mendayagunakan metode Naïve Bayes dan Support Vector Machine (SVM).

Hasil dan Pembahasan

Data penelitian diperoleh melewati proses web scraping pada media sosial YouTube terkait isu pembangunan Ibu Kota Negara (IKN). Total data yang berhasil dikumpulkan sebanyak 3.541 komentar.

Table 1. Perbandingan Performa Metode

Metode	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	79.8%	79%	80%	73%
SVM	94.2%	94%	94%	94%

Proses preprocessing merupakan tahapan awal yang dilakukan sebelum data digunakan pada proses klasifikasi sentimen. Tahapan ini bertujuan untuk membersihkan, menyederhanakan, dan menyeragamkan data teks sehingga menghasilkan data yang lebih berkualitas serta mampu meningkatkan akurasi proses klasifikasi. Pada penelitian ini, proses preprocessing terdiri atas lima tahapan, yaitu cleaning, case folding, tokenizing, stopword removal, dan stemming.

Tahap pertama adalah cleaning, yaitu proses pembersihan data teks dari karakter atau simbol yang tidak diperlukan, seperti tanda baca, angka, emoji, URL, hashtag, mention, serta karakter khusus lainnya. Tahapan ini dilakukan agar komentar menjadi lebih bersih dan hanya menyisakan informasi yang relevan untuk dianalisis.

Selanjutnya dilakukan case folding, yaitu proses mengubah seluruh huruf pada teks menjadi huruf kecil (lowercase). Proses ini bertujuan untuk menyeragamkan penulisan kata sehingga tidak terdapat perbedaan antara huruf besar dan huruf kecil yang dapat memengaruhi hasil klasifikasi.

Tahap berikutnya adalah tokenizing, yaitu proses memecah kalimat menjadi potongan kata atau token. Dengan proses ini, setiap kata dapat dianalisis secara individual sehingga

memudahkan pengolahan data pada tahap selanjutnya.

Setelah proses tokenizing, dilakukan stopword removal, yaitu proses menghapus kata-kata umum yang sering muncul tetapi tidak memiliki makna penting dalam analisis sentimen, seperti "dan", "yang", "di", "ke", dan kata-kata sejenis. Penghapusan stopword bertujuan untuk meningkatkan kualitas fitur dengan mempertahankan kata-kata yang lebih merepresentasikan isi komentar

Tahap terakhir adalah stemming, yaitu proses mengubah kata yang memiliki imbuhan menjadi bentuk kata dasar. Dengan proses ini, kata-kata yang memiliki makna sama tetapi berbeda bentuk penulisannya dapat diseragamkan sehingga membantu meningkatkan efektivitas proses klasifikasi sentimen.

Hasil pengujian menggunakan metode Naïve Bayes memperoleh tingkat akurasi sebesar 79.8%. Berdasarkan hasil klasifikasi yang diperlihatkan, model memiliki weighted precision sebesar 79%, recall sebesar 80%, dan f1-score sebesar 73%.

Pengujian dengan menggunakan metode SVM memperoleh fase akurasi sebanyak 94.2% dengan weighted precision sebanyak 94%, recall sebesar 94%, dan f1-score sebesar 94%.

Hasil penelitian kedua metode menunjukkan bahwa metode Naive Bayes memperoleh tingkat akurasi sebesar 79.8% dengan nilai weighted precision sebesar 79%, recall sebesar 80%, dan f1-score sebesar 73%. Sementara itu, metode Support Vector Machine (SVM) menghasilkan performa yang lebih baik dengan tingkat akurasi sebesar 94.2%, weighted precision sebesar 94%, recall sebesar 94%, dan f1-score sebesar 94%. Hasil tersebut menunjukkan bahwa metode SVM lebih efektif dibandingkan Naive Bayes dalam mengklasifikasikan sentimen positif, negatif, dan netral pada komentar media sosial terkait isu pembangunan Ibu Kota Negara (IKN).

Tabel 2. Distribusi Sentimen Komentar Youtube

Kategori Sentimen	Jumlah	Persentase
Positif	630	17.8%
Negatif	190	5.4%
Netral	2721	76.8%

AS (Analisis Sentimen) adalah pendekatan pada text mining yang diaplikasikan untuk memahami, mengekstraksi, serta menjadikan data tekstual secara otomatis agar mengidentifikasi kecenderungan sikap atau polaritas opini yang ada di dalamnya. Analisis ini memiliki tujuan untuk mengetahui bagaimana suatu tema, isu, atau objek dipersepsikan melalui teks tertulis, baik yang bersumber dari dokumen ilmiah, artikel, maupun publikasi lainnya. Data yang telah terhimpun selanjutnya dianalisis untuk mendapatkan gambaran kecenderungan sentimen secara lebih sistematis dan terukur, sehingga bisa digunakan sebagai acuan dalam penarikan kesimpulan. Hasil analisis sentimen biasanya diklasifikasikan menjadi tiga golongan utama, yaitu sentimen positif, negatif, dan netral.

Metode Naive Bayes digunakan pada pengkajian klasifikasi agar sentimen komentar masyarakat terkait isu pembangunan Ibu Kota Negara (IKN) ke dalam golongan negatif, positif, dan netral. Metode ini bekerja mengaplikasikan alur probabilistik berdasarkan Teorema Bayes pada anggapan independensi antar fitur. Naïve Bayes mempunyai keunggulan dalam langkah komputasi yang sederhana dan efisien sehingga banyak diimplementasikan dalam pengolahan data teks dan analisis sentimen berbasis media sosial.

Support Vector Machine (SVM) diaplikasikan pada pengkajian ini agar bisa mengklasifikasikan data sentimen dengan mengetahui hyperplane optimal yang mampu dan bisa memisahkan setiap kelas sentimen. Metode ini mempunyai kemampuan yang baik saat mengatasi data teks berdimensi tinggi sehingga kerap kali diterapkan dalam proses klasifikasi dalam bidang machine learning dan analisis sentimen. Penelitian ini, SVM

digunakan untuk mengidentifikasi sentimen masyarakat terkait isu pembangunan Ibu Kota Negara (IKN) berdasarkan komentar pada media sosial YouTube.

Metode TF-IDF adalah cara pembobotan kata yang diterapkan dalam mengubah data yang semula teks menjadi data numerik atau angka saat memberikan bobot dalam tiap kata beralaskan fase keperluannya dalam suatu dokumen. TF-IDF sendiri memiliki dua elemen, yaitu Term Frequency (TF) yang memperkirakan frekuensi munculnya kata pada sebuah berkas dan Inverse Document Frequency (IDF) yang memiliki fungsi menurunkan bobot kata yang sering keluar di seluruh dokumen serta menaikkan bobot kata yang jarang keluar. Hasil TF-IDF diperoleh dari hasil perkalian TF dan IDF, sehingga kata pada nilai TF-IDF tinggi dianggap lebih representatif dalam menentukan isi serta sentimen suatu dokumen.

Media Sosial menurut Ilmu Komunikasi, media sosial ialah platform digital yang memfasilitasi individu bisa berkomunikasi, mencari informasi, bergaul, bahkan membangun hubungan sosial tanpa batas ruang dan waktu. Media sosial bahkan menjadi sarana interaksi utama di era digital karena pengguna dapat berkomunikasi melalui teks, gambar, video, serta bentuk komunikasi daring lainnya secara cepat dan luas. Hadirnya media sosial banyak mengubah cara masyarakat berkomunikasi dan berinteraksi dalam kehidupan keseharian melalui berbagai platform digital seperti Instagram, Facebook, WhatsApp, TikTok, dan YouTube.

Kesimpulan

Saat penelitian telah dilakukan, hasil analisis sentimen terhadap komentar masyarakat mengenai pembangunan Ibu Kota Negara (IKN) pada media sosial YouTube bisa dilakukan dengan efektif menerapkan Naïve Bayes serta SVM. Alur analisis dilakukan melalui langkah preprocessing data dan pembobotan TF-IDF agar membuat data teks menjadi informasi yang dapat serta bisa dikelompokkan ke tiga sentimen utama yaitu

positif, negatif, dan netral. Hasil pengkajiannya memperlihatkan bahwa kedua metode bisa digunakan dalam klasifikasi sentimen, sehingga analisis sentimen berbasis machine learning bisa membantu memahami kecenderungan opini masyarakat terkait isu pembangunan IKN di media sosial.

Ucapan terima kasih

Penulis mengucapkan terima kasih kepada Universitas Ibrahimy atas dukungan dan fasilitas yang diberikan selama pelaksanaan penelitian, kepada Universitas Muhammadiyah Muara Bungo atas kesempatan serta dukungan dalam penyelenggaraan kegiatan ilmiah, dan kepada Panitia atas bantuan, koordinasi, serta dukungan selama pelaksanaan kegiatan sehingga penelitian dan penyusunan artikel ini dapat terselesaikan dengan baik.

Daftar Pustaka

- [1] H. Tuhuteru *et al.*, “Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier,” vol. 03, no. 03, pp. 394–401, 2018, doi: 10.30591/jpit.v3i3.977.
- [2] U. P. Triputra, F. Teknologi, and U. S. Hartono, “JITE (Journal of Informatics and Telecommunication Engineering) Surakarta City Residents Using the BiLSTM Model,” vol. 9, no. July, pp. 229–239, 2025.
- [3] C. A. Prianggodo, “PEMBANGUNAN IKN SEBAGAI MERCUSUAR STRATEGIS (KAJIAN KRITIS DALAM PERSPEKTIF SUSTAINABLE DEVELOPMENT),” vol. 8, no. 1, 2023.
- [4] A. Aningtyas, S. Ika, A. Wardhani, A. Fawwaz, and F. I. Hamida, “Pengambilan Keputusan Pembangunan Ibu Kota Negara (IKN): Analisis Perspektif Masyarakat di Luar Daerah Kalimantan Ditinjau dari Aspek Kemanusiaan,” vol. 2, no. 01, pp. 36–42, 2024, doi: 10.58812/sish.v2.i01.
- [5] B. A. Maulana and M. J. Fahmi, “Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM),” vol. 4, no. April, pp. 375–384, 2024.
- [6] U. Mercu and B. Yogyakarta, “PUBLIC SENTIMENT ANALYSIS OF DANANTARA POLICY THROUGH SOCIAL MEDIA X USING SVM AND RANDOM FOREST ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN DANANTARA MELALUI MEDIA SOSIAL X DENGAN METODE SVM DAN RANDOM FOREST,” vol. 10, no. 2, pp. 1208–1217, 2025.