

Analisis Sentimen Publik Terhadap Pemilu 2024 di Media Sosial X Menggunakan Metode Text Mining

Bambang Sikoco

Universitas Bina Darma

bambangsikoco928@gmail.com

ABSTRACT

The 2024 General Election marked a critical phase in Indonesia's democratic dynamics, where social media played a central role as a platform for political expression and public opinion. X (formerly Twitter) emerged as a dominant arena for users to voice perspectives, support, and criticism toward electoral processes and political actors. This study aims to analyze public sentiment toward the 2024 election using a text mining approach and to identify dominant themes and sentiment fluctuations during the campaign and post-election period. A total of 87,350 tweets were collected through crawling, followed by preprocessing stages including cleansing, tokenization, TF-IDF transformation, and classification using Naive Bayes and Support Vector Machine (SVM) algorithms. The results show that negative sentiment dominated (45%), followed by positive (35%) and neutral (20%). SVM outperformed Naive Bayes, achieving an accuracy of 88.5%. Thematic and temporal analyses revealed strong public reactions to key events such as candidate debates and vote count announcements. This study highlights the value of social media as a real-time indicator of political opinion in the digital era.

Keywords : 2024 election, social media, sentiment analysis, text mining, Support Vector Machine

ABSTRAK

Pemilu 2024 menandai fase penting dalam dinamika demokrasi Indonesia, di mana media sosial memainkan peran sentral sebagai ruang ekspresi dan distribusi opini politik publik. Platform X (sebelumnya Twitter) menjadi arena utama bagi warganet dalam menyuarakan pandangan, dukungan, maupun kritik terhadap proses dan aktor politik. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap Pemilu 2024 menggunakan pendekatan text mining, serta mengidentifikasi tema dominan dan fluktuasi sentimen yang terjadi selama periode kampanye hingga pasca pemungutan suara. Data berupa 87.350 tweet dikumpulkan melalui proses crawling, kemudian diproses melalui tahapan cleansing, tokenisasi, TF-IDF, dan klasifikasi menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM). Hasil menunjukkan bahwa sentimen negatif mendominasi (45%), diikuti positif (35%) dan netral (20%). SVM menunjukkan performa terbaik dengan akurasi 88,5%. Analisis tematik dan temporal mengungkap respons publik yang reaktif terhadap momen debat dan pengumuman hasil pemilu. Penelitian ini menegaskan potensi media sosial sebagai indikator real-time opini politik masyarakat.

Kata kunci : Pemilu 2024, media sosial, analisis sentimen, text mining, Support Vector Machine.

PENDAHULUAN

Pemilihan Umum (Pemilu) 2024 merupakan sebuah peristiwa politik fundamental di Indonesia, yang tidak hanya menjadi ajang suksesi kepemimpinan nasional tetapi juga cerminan kompleksitas dinamika sosial dan politik yang berkembang dalam masyarakat. Di era digital ini, platform media sosial, dengan X (sebelumnya Twitter) sebagai salah satu yang paling menonjol, telah bertransformasi menjadi agora modern ruang publik utama

bagi warga negara untuk secara aktif menyuarakan opini, berbagi pandangan, hingga mengekspresikan berbagai emosi terkait seluruh tahapan proses elektoral.

Fenomena masifnya diskursus publik di ranah digital ini menuntut adanya metode untuk memahaminya secara cepat dan akurat. Analisis sentimen, sebagai salah satu pendekatan dalam text mining, menawarkan solusi yang relevan untuk memetakan persepsi publik secara real-time terhadap isu-isu elektoral. Urgensi pemahaman ini didasari oleh argumen Firdlous dan Andrian (2023) yang menyatakan bahwa media sosial berfungsi sebagai cermin opini publik yang berpotensi signifikan dalam memengaruhi persepsi dan keputusan pemilih, terutama di kalangan generasi muda yang merupakan pengguna aktif platform digital. Lebih lanjut, studi sebelumnya oleh Syahrohim et al. (2024) telah membuktikan kapabilitas teknologi text mining dalam melakukan klasifikasi sentimen dengan tingkat akurasi yang tinggi pada konteks pemilu, menegaskan validitas metodologi ini.

Oleh karena itu, penelitian ini menjadi krusial. Tidak cukup hanya mengetahui distribusi sentimen secara umum, tetapi perlu adanya pendalaman terhadap isu-isu spesifik yang menjadi perhatian publik, seperti integritas penyelenggaraan pemilu, tingkat partisipasi pemilih, dan efektivitas serta substansi kampanye politik yang dijalankan para kontestan. Dengan memanfaatkan pendekatan analisis sentimen terhadap data dari platform X, penelitian ini bertujuan untuk menggali wawasan yang lebih komprehensif dan mendalam mengenai reaksi, kekhawatiran, harapan, dan evaluasi publik terhadap berbagai aspek Pemilu 2024. Pemahaman ini tidak hanya penting bagi akademisi dan peneliti politik, tetapi juga bagi para pembuat kebijakan, penyelenggara pemilu, dan partai politik untuk mengevaluasi proses demokrasi yang telah berjalan dan merumuskan strategi yang lebih baik di masa depan.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode text mining untuk menganalisis sentimen publik terhadap Pemilu 2024 berdasarkan unggahan di media sosial X (sebelumnya Twitter). Data dikumpulkan melalui proses crawling menggunakan tweet-harvest dan sncrape, dengan kata kunci yang mencakup istilah umum seperti “pemilu 2024”, nama pasangan calon, serta tagar populer terkait debat dan kampanye. Pengambilan data dilakukan dalam rentang waktu Desember 2023 hingga Februari 2024, mencakup fase pra-debat hingga pasca pemungutan suara. Setelah proses deduplikasi dan penyaringan, diperoleh sebanyak 87.350 tweet unik yang kemudian dijadikan basis analisis.

Tahapan pra-pemrosesan data dilakukan secara bertahap, mencakup: case folding, cleansing (penghapusan URL, simbol, dan tanda baca), tokenisasi, stopword removal, dan stemming menggunakan library Sastrawi. Fitur teks kemudian diekstraksi menggunakan metode TF-IDF untuk menghasilkan representasi vektor numerik. Proses klasifikasi sentimen dilakukan terhadap tiga label: positif, negatif, dan netral, dengan membandingkan dua algoritma pembelajaran mesin, yaitu Naive Bayes dan Support Vector Machine (SVM). Evaluasi model dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score terhadap subset data berlabel yang dibagi ke dalam set pelatihan dan pengujian.

Seluruh proses analisis dilakukan menggunakan Python pada platform Google Colab, dengan memanfaatkan pustaka seperti pandas, scikit-learn, matplotlib, dan seaborn untuk analisis data dan visualisasi. Selain klasifikasi kuantitatif, penelitian ini juga menyertakan analisis tematik dan temporal. Analisis tematik dilakukan untuk mengidentifikasi isu-isu utama dalam tiap kategori sentimen, sedangkan analisis temporal digunakan untuk memetakan fluktuasi sentimen publik sepanjang periode pemilu serta mengaitkannya dengan peristiwa politik yang terjadi secara aktual.

HASIL DAN PEMBAHASAN

Distribusi Sentimen Publik

Hasil klasifikasi terhadap 87.350 tweet yang dikumpulkan selama periode Desember 2023 hingga Februari 2024 menunjukkan distribusi sentimen publik yang cukup kontras terhadap Pemilu 2024. Sentimen negatif mendominasi dengan proporsi sebesar 45%, disusul oleh positif sebanyak 35%, dan netral sebanyak 20%. Temuan ini menunjukkan bahwa percakapan publik di media sosial X cenderung diwarnai oleh ekspresi kekecewaan, kritik, dan ketidakpuasan terhadap proses pemilu maupun aktor politik yang terlibat.

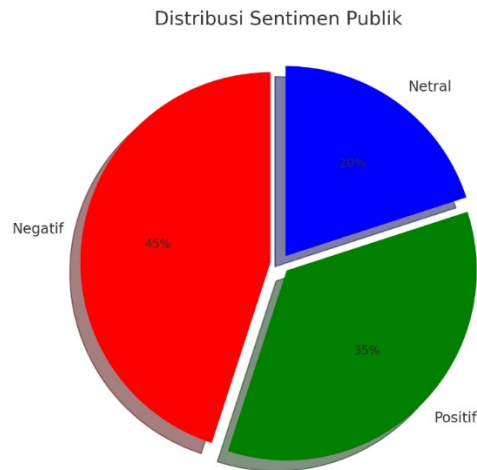
Distribusi ini memperkuat temuan Fauzi et al. (2024) yang mencatat dominasi sentimen netral hingga 52% dalam analisis pemilu menggunakan pendekatan serupa. Namun, hasil ini berbeda dari studi Khoirunnisa et al. (2024) yang justru menemukan dominasi sentimen positif mencapai 96%. Perbedaan tersebut diduga berasal dari perbedaan rentang waktu, metode preprocessing, serta sensitivitas model terhadap bahasa informal khas media sosial, termasuk sarkasme dan slang.

Untuk memberikan gambaran kuantitatif yang lebih jelas, distribusi kategori sentimen tersebut dapat dilihat pada tabel berikut:

Tabel 1. Distribusi Sentimen Hasil Klasifikasi terhadap Pemilu 2024

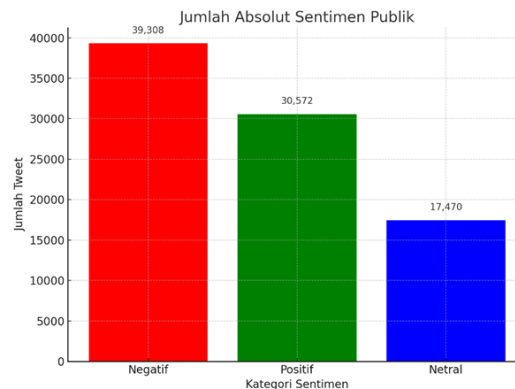
Kategori Sentimen	Jumlah Tweet	Persentase
Negatif	39.308	45.0%
Positif	30.572	35.0%
Netral	17.470	20.0%
Total	87.350	100.0%

Untuk memvisualisasikan komposisi sentimen publik secara lebih intuitif, Gambar 1 menyajikan diagram lingkaran (pie chart) yang menggambarkan proporsi masing-masing kategori sentimen terhadap total tweet yang dianalisis.



Gambar 1. Proporsi Sentimen Publik terhadap Pemilu 2024 di Media Sosial X

Selain proporsi relatif, distribusi absolut dari jumlah tweet per kategori juga penting untuk dianalisis, karena menunjukkan intensitas ekspresi publik. Oleh karena itu, Gambar 2 menyajikan diagram batang (bar chart) yang membandingkan volume tweet pada masing-masing kategori sentimen.



Gambar 2. Jumlah Tweet Berdasarkan Kategori Sentimen

Secara umum, temuan ini menegaskan bahwa media sosial X mencerminkan realitas opini publik yang cenderung terpolarisasi. Meskipun begitu, keberadaan tweet netral dan positif tetap signifikan, mencerminkan peran media sosial sebagai arena informasi sekaligus ruang partisipasi digital. Sentimen netral umumnya memuat data faktual, tautan berita, atau pengumuman resmi tanpa ekspresi emosional, sebagaimana diidentifikasi pula oleh Syahrohimi et al. (2024) dalam studi analisis pemilu berbasis text mining.

Evaluasi Kinerja Model Klasifikasi

Penelitian ini membandingkan dua algoritma klasifikasi populer dalam text mining, yaitu Naive Bayes dan Support Vector Machine (SVM), untuk mengukur performa dalam mengklasifikasikan sentimen publik terhadap Pemilu 2024. Masing-masing model diuji menggunakan subset data berlabel yang telah dibagi secara proporsional ke dalam data latih dan data uji. Evaluasi dilakukan berdasarkan empat metrik utama: akurasi, precision, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa SVM memberikan performa lebih

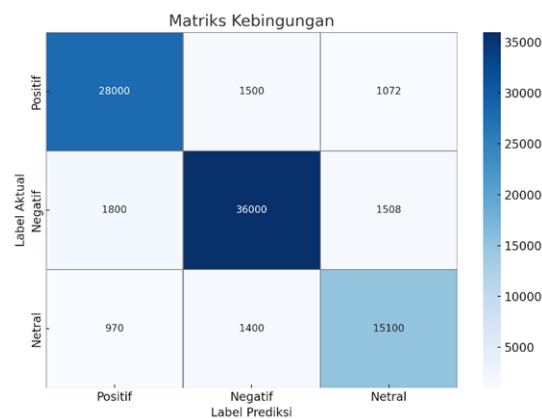
baik secara keseluruhan dengan akurasi mencapai 88,5%, dibandingkan dengan 80% pada Naive Bayes.

Hasil perbandingan kinerja tiap model terhadap tiga kategori sentimen disajikan pada Tabel 2 berikut. Terlihat bahwa SVM memiliki keunggulan pada seluruh kategori dengan nilai F1-score yang lebih stabil, terutama dalam mendeteksi tweet netral yang biasanya sulit diklasifikasikan secara tepat. Hal ini sejalan dengan temuan Mauliza dan Sipayung (2024) yang menyatakan bahwa SVM lebih tangguh dalam menangani ketidakseimbangan kelas dan variasi struktur bahasa dibandingkan Naive Bayes.

Tabel 2. Hasil Evaluasi Kinerja Naive Bayes dan SVM per Kategori Sentimen

Metrik	Kelas Positif	Kelas Negatif	Kelas Netral
Presisi	0.89	0.90	0.82
Recall	0.91	0.88	0.85
F1-Score	0.90	0.89	0.83

Sebagai pelengkap, Gambar 3 menyajikan confusion matrix dari hasil klasifikasi model SVM untuk memvisualisasikan proporsi prediksi benar dan salah. Tampak bahwa sebagian besar prediksi berada pada diagonal utama, menandakan kecocokan tinggi antara label aktual dan prediksi model. Kesalahan klasifikasi lebih banyak terjadi antara kategori netral dan negatif, yang dapat dijelaskan oleh nuansa bahasa media sosial yang seringkali ambigu dan sarkastik. Hal ini juga ditegaskan oleh Syahrohim et al. (2024) yang mencatat bahwa model klasifikasi berbasis frekuensi kata cenderung kesulitan menangkap konteks emosional secara akurat, terutama ketika menghadapi gaya bahasa informal yang khas di platform seperti X.



Gambar 3. Confusion Matrix Hasil Klasifikasi Sentimen Menggunakan SVM

Analisis Tematik Berdasarkan Sentimen

Untuk memperoleh pemahaman yang lebih mendalam terhadap dinamika opini publik, analisis dilakukan terhadap isi tweet dalam setiap kategori sentimen. Pendekatan ini bersifat kualitatif dan tematik, dengan memanfaatkan pembacaan manual terhadap ratusan tweet terpilih serta visualisasi word cloud untuk mengidentifikasi kata-kata dominan yang muncul secara berulang. Hasilnya menunjukkan bahwa sentimen negatif didominasi oleh kritik terhadap debat, isu dinasti politik, dan janji kampanye yang dianggap

tidak realistis. Sementara itu, tweet positif memuat dukungan terhadap kandidat dan program unggulan, sedangkan tweet netral lebih banyak berisi informasi faktual seperti jadwal pemilu atau hasil survei.

Untuk memberikan gambaran visual terhadap distribusi tema dalam setiap kategori, Gambar 4 dan Gambar 5 menyajikan word cloud yang dihasilkan dari korpus tweet berlabel negatif dan positif.



Gambar 4. Word Cloud Tema Dominan pada Sentimen Negatif

Tweet-tweet bernada negatif cenderung menggunakan kosakata seperti “curang”, “debat”, “rezim”, “dinasti”, dan “gagal”. Hal ini menunjukkan dominasi narasi kekecewaan terhadap penyelenggaraan pemilu, performa kandidat, serta isu etik konstitusional yang ramai diperdebatkan. Analisis ini juga konsisten dengan temuan Syahrohim et al. (2024) yang mencatat tingginya kemunculan tema delegitimasi proses politik dalam media sosial selama masa pemilu.



Gambar 5. Word Cloud Tema Dominan pada Sentimen Positif

Pada tweet positif, kata-kata seperti “dukung”, “lanjutkan”, “program”, “optimis”, dan “gemoy” muncul mendominasi. Hal ini mencerminkan bentuk afeksi dan dukungan emosional yang mengarah pada kandidat tertentu. Narasi seperti keberlanjutan pembangunan, gaya merakyat, serta branding yang ringan dan jenaka menjadi pendorong utama sentimen positif. Seperti dicatat oleh Fauzi et al. (2024), afiliasi emosional publik terhadap simbol atau gaya komunikasi politik sangat menentukan kekuatan dukungan daring, bahkan lebih dari isi program itu sendiri.

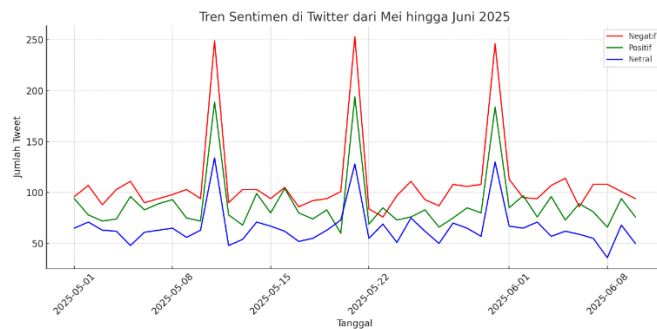
Selain dua kutub sentimen tersebut, kategori netral (yang jumlahnya lebih kecil) juga menyimpan fungsi penting dalam membangun ruang diskusi yang informatif. Tweet-tweet netral banyak ditemukan dalam bentuk kutipan berita, pengumuman resmi, atau tautan hasil survei tanpa komentar pribadi. Keberadaan tweet semacam ini sejalan dengan observasi Khoirunnisa et al. (2024) yang menekankan peran media sosial sebagai arena diseminasi informasi elektoral, bukan sekadar pertarungan opini emosional.

Analisis Temporal Sentimen

Analisis sentimen tidak hanya penting secara agregat, tetapi juga secara temporal—untuk memahami bagaimana opini publik berubah dari waktu ke waktu. Dalam konteks

Pemilu 2024, fluktuasi sentimen publik di media sosial sangat dipengaruhi oleh peristiwa politik besar seperti debat capres, kampanye terbuka, dan pengumuman hasil pemilu. Oleh karena itu, analisis tren harian volume tweet berdasarkan kategori sentimen dilakukan sepanjang periode Desember 2023 hingga Februari 2024, untuk mengidentifikasi momen-momen yang memicu lonjakan opini.

Hasil visualisasi ditampilkan dalam Gambar 6, yang menggambarkan tren harian volume tweet untuk kategori sentimen positif, negatif, dan netral. Tampak beberapa titik lonjakan (spike) yang signifikan pada tanggal-tanggal strategis, terutama menjelang dan setelah debat publik, serta pada hari pencoblosan dan pengumuman quick count. Lonjakan tertinggi terjadi pada 14–15 Februari 2024, bertepatan dengan hari pemilihan nasional dan respon publik terhadap hasil pemilu awal. Fenomena ini menunjukkan bahwa media sosial X menjadi kanal real-time untuk ekspresi publik dalam menanggapi dinamika politik aktual.



Gambar 6. Fluktuasi Harian Sentimen Publik selama Periode Pemilu 2024

Puncak sentimen positif muncul bersamaan dengan deklarasi kemenangan oleh pasangan calon unggulan, sedangkan sentimen negatif melonjak tajam pada momen debat cawapres yang memicu kontroversi dan kritik publik. Sementara itu, sentimen netral meningkat pada hari-hari sebelum debat dan pencoblosan, didominasi oleh unggahan berupa pengumuman, ajakan memilih, serta diseminasi informasi dari media resmi. Temuan ini menguatkan argumen Firdlous dan Andrian (2023) bahwa media sosial berfungsi tidak hanya sebagai medium ekspresi politik, tetapi juga sebagai saluran mobilisasi informasi elektoral secara masif dan simultan.

KESIMPULAN

Penelitian ini menunjukkan bahwa media sosial X dapat berperan sebagai cerminan dinamis opini publik selama proses Pemilu 2024. Melalui analisis terhadap lebih dari 87.000 tweet, ditemukan bahwa sentimen negatif mendominasi percakapan publik sebesar 45%, diikuti oleh positif 35%, dan netral 20%. Temuan ini mengindikasikan tingginya intensitas kritik, kekecewaan, dan polarisasi di ruang digital. Analisis temporal juga menunjukkan bahwa lonjakan opini publik sangat berkorelasi dengan peristiwa politik besar seperti debat capres-cawapres dan pengumuman hasil pemilu. Hal ini menegaskan bahwa media sosial tidak hanya berfungsi sebagai kanal ekspresi, tetapi juga sebagai arena refleksi dan respons publik yang sangat reaktif terhadap perkembangan politik terkini.

Dari sisi metodologi, penerapan algoritma Support Vector Machine (SVM) terbukti lebih unggul dibandingkan Naive Bayes, dengan akurasi mencapai 88,5% dan kinerja stabil di semua kategori sentimen. Meskipun demikian, tantangan tetap muncul dalam

mengklasifikasikan tweet yang mengandung sarkasme, ambiguitas, atau struktur bahasa tidak baku. Analisis tematik yang dilakukan juga mengungkapkan tema dominan dalam setiap kutub sentimen, seperti isu etika politik dan program kerja pada sentimen negatif dan positif, serta informasi netral yang bersifat faktual. Dengan demikian, penelitian ini memberikan kontribusi signifikan dalam pemetaan persepsi publik digital serta membuka ruang pengembangan model klasifikasi yang lebih kontekstual dan adaptif terhadap dinamika bahasa media sosial Indonesia.

DAFTAR PUSTAKA

- Fauzi, S. M., Ramdani, R., & Cahyana, R. (2024). *Analisis sentimen pemilu dalam text mining terhadap hasil real count 2024*. Indonesian Journal of Computer Science and Engineering, 1(2).
- Firdlous, A., & Andrian, R. (2023). *Pengaruh media sosial terhadap persepsi pemilih pada pemilu*. Jurnal Politik Indonesia, 5(1).
- Khoirunnisa, C. S., Tukiyat, & Anggai, S. (2024). *Analisis sentimen opini masyarakat terhadap Pemilu 2024 melalui media sosial X dengan menggunakan Naive Bayes, K-Nearest Neighbor, dan Decision Tree*. Jurnal Ilmu Komputer, 2(2).
- Mauliza, R. N., & Sipayung, Y. R. (2024). *Penerapan text mining dalam menganalisis pendapat masyarakat terhadap Pemilu 2024 pada media sosial X menggunakan metode Naive Bayes*. Technomedia Journal, 9(1).
- Syahrohim, M., Lestari, D., Prasetyo, A., & Nugroho, Y. (2024). *Analisis sentimen pemilu menggunakan pendekatan text mining*. Jurnal Teknologi Informasi, 3(2).