

PREDIKSI HASIL PEMILIHAN UMUM BERDASARKAN DATA MEDIA SOSIAL MENGGUNAKAN TEKNIK DATA MINING NAIVE BAYES

Rusli, Harmayani

Universitas Asahan

e-mail: ¹Ruslypsb255@gmail.com, ²mayong3010@gmail.com

Abstract: *In the digital era, social media has a significant influence in shaping public opinion, including in elections. This research aims to analyze and predict public opinion trends related to the 2024 Election using social media data. Data mining techniques, particularly the Naive Bayes method, are employed for classifying public sentiment. The analyzed data includes text from posts and comments relevant to the 2024 Election. This study is limited to a specific time frame to ensure the data's relevance. The analysis results are expected to provide insights into the patterns and trends of public opinion that may impact the election outcomes. Furthermore, this research can serve as a reference for political parties and candidates to design more effective campaign strategies. The study also aims to enhance understanding of the role of social media in the democratic process and encourage more active and informed public participation. This research contributes to the development of data mining techniques and sentiment classification in political data analysis.*

Keywords: *Social Media, Election, Data Mining, Naive Bayes, Sentiment Analysis.*

Abstrak: Di era digital, media sosial memiliki pengaruh signifikan dalam membentuk opini publik, termasuk dalam Pemilihan Umum (Pemilu). Penelitian ini bertujuan untuk menganalisis dan memprediksi tren opini masyarakat terkait Pemilu 2024 dengan menggunakan data dari media sosial. Teknik data mining, khususnya metode Naive Bayes, digunakan untuk klasifikasi sentimen opini publik. Data yang dianalisis berupa teks dari postingan dan komentar yang relevan dengan Pemilu 2024. Penelitian ini dibatasi pada periode waktu tertentu untuk memastikan data yang terkini. Hasil analisis diharapkan memberikan wawasan mengenai pola dan tren opini publik yang dapat mempengaruhi hasil Pemilu. Selain itu, penelitian ini dapat menjadi referensi bagi partai politik dan kandidat dalam merancang strategi kampanye yang lebih efektif. Studi ini juga diharapkan dapat meningkatkan pemahaman tentang peran media sosial dalam proses demokrasi, serta mendorong partisipasi masyarakat yang lebih aktif dan informatif. Penelitian ini berkontribusi pada pengembangan teknik data mining dan klasifikasi sentimen dalam analisis data politik.

Kata kunci: Media Sosial, Pemilu, Data Mining, Naive Bayes, Analisis Sentimen.

PENDAHULUAN

Di era digital saat ini, media sosial memainkan peran penting dalam membentuk opini publik, termasuk dalam bidang politik. Di Indonesia, platform seperti Twitter, Facebook, dan Instagram sering kali menjadi ruang bagi masyarakat untuk mengemukakan pendapat terkait isu-isu politik, termasuk Pemilihan Umum (Pemilu). Pemilu merupakan

momen krusial bagi suatu negara untuk menentukan pemimpin dan kebijakan masa depan. Partisipasi masyarakat dalam proses ini sering kali dipengaruhi oleh informasi yang mereka peroleh dari media sosial.

Data yang diperoleh dari media sosial memberikan gambaran yang lebih real-time tentang opini publik, menjadikannya sumber informasi yang berharga untuk analisis politik. Dengan

kemajuan teknik data mining, data besar yang terkumpul dari media sosial dapat dianalisis secara efektif untuk mengungkap pola dan tren yang tersembunyi. Salah satu metode yang digunakan dalam penelitian ini adalah Naive Bayes, sebuah teknik klasifikasi yang efisien untuk mengkategorikan sentimen publik.

Penelitian ini bertujuan untuk memahami bagaimana opini masyarakat terkait Pemilu 2024 terbentuk di media sosial, serta mengevaluasi dampaknya terhadap hasil Pemilu. Dengan menggunakan data mining, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai tren opini publik. Selain itu, penelitian ini diharapkan dapat menjadi referensi untuk perancangan strategi kampanye yang lebih efektif dan mendorong demokrasi yang lebih transparan. Studi ini juga berkontribusi pada pengembangan teknik analisis data dan penerapannya dalam dunia politik.

Berdasarkan Latar Belakang masalah diatas dirumuskan permasalahan yaitu Seberapa efektif teknik data mining, khususnya metode Naive Bayes, dalam mengklasifikasikan sentimen publik mengenai Pemilu 2024 yang diungkapkan di media sosial?.

Penelitian ini dibatasi pada beberapa hal berikut: Hanya menganalisis data teks dari media sosial X terkait Pemilu 2024, tanpa mempertimbangkan data gambar, video, atau multimedia lainnya. Data yang dianalisis dibatasi pada periode [Tanggal Awal] hingga [Tanggal Akhir] untuk menjaga relevansi. Metode analisis yang digunakan adalah Naive Bayes untuk klasifikasi sentimen, tanpa menerapkan metode lain seperti *decision tree*, *random forest*, atau *deep learning*.

Adapun tujuan penelitian ini adalah sebagai berikut : Menganalisis sentimen publik terkait Pemilu 2024 di media sosial dan dampaknya pada opini masyarakat, Menilai pengaruh opini di media sosial terhadap hasil Pemilu 2024.

Menggunakan data mining untuk menemukan pola opini publik di media sosial terkait Pemilu.

Manfaat Penelitian Dengan memahami bagaimana media sosial mempengaruhi opini publik dan hasil Pemilu, penelitian ini dapat meningkatkan kesadaran masyarakat tentang peran media sosial dalam proses demokrasi dan mendorong partisipasi yang lebih aktif dan informatif dalam Pemilu.

Hasil penelitian dapat memberikan informasi yang berguna bagi partai politik, kandidat, dan pengambil keputusan untuk merancang strategi kampanye yang lebih efektif berdasarkan sentimen publik dan tren opini yang terdeteksi.

Penelitian ini berkontribusi pada pengembangan pengetahuan dalam bidang data mining dan teknik klasifikasi, serta aplikasinya dalam analisis sentimen, yang bermanfaat bagi peneliti dan praktisi di bidang ilmu komputer dan informasi.

METODE

Dalam penelitian ini, fokus utama adalah pada Pemilu sebagai mekanisme demokratis di Indonesia, di mana warga negara memilih anggota legislatif dan presiden secara langsung. Berdasarkan Undang-Undang No. 7 Tahun 2017, Pemilu harus dilaksanakan secara langsung, umum, bebas, rahasia, jujur, dan adil sesuai dengan prinsip Pancasila dan UUD 1945. Metode ini akan meneliti pelaksanaan kampanye oleh peserta Pemilu serta bagaimana pemungutan suara dilakukan dan dihitung, untuk menentukan pemenang secara adil dan transparan (Wartoyo, 2021; Morris, 2017). Penelitian ini akan menggunakan data dari proses Pemilu untuk mengidentifikasi pola partisipasi warga negara dan hasil Pemilu.

Data mining digunakan sebagai metode untuk mengekstraksi pola yang berguna dari data besar yang terkait dengan hasil Pemilu. Menurut Pregibin, data mining menggabungkan ilmu

statistika, kecerdasan buatan, dan riset database untuk mengelompokkan tugas-tugas seperti deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi (Susilo Yuda Irawan, 2023). Dalam penelitian ini, data mining akan diterapkan untuk menganalisis data Pemilu dengan tujuan untuk memprediksi hasil berdasarkan data sosial media.

Algoritma Naive Bayes

Penelitian ini menggunakan algoritma Naive Bayes untuk mengklasifikasikan dan memprediksi hasil Pemilu. Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang dikembangkan oleh Thomas Bayes, di mana ia menggunakan informasi data sebelumnya untuk memprediksi kemungkinan di masa depan. Keunggulan dari algoritma ini termasuk kemudahan implementasi, akurasi tinggi, dan kemampuan untuk mengolah dataset besar dengan cepat (Amelia, 2017). Dalam konteks penelitian ini, algoritma Naive Bayes akan digunakan untuk memodelkan data hasil Pemilu berdasarkan data sosial media dan variabel lainnya.

Perangkat Lunak

Website: Situs web akan digunakan sebagai platform untuk menampilkan hasil penelitian dan informasi terkait. Menurut Arifin & Krisnadita (2017) serta Hastanti (2015), website adalah media untuk menyajikan konten multimedia yang dapat diakses melalui internet. Dalam penelitian ini, website dinamis akan dikembangkan untuk memfasilitasi pembaruan data secara berkala dan menampilkan analisis hasil Pemilu secara real-time. Python: Python dipilih sebagai bahasa pemrograman utama dalam penelitian ini karena kemampuannya yang kuat dalam pengolahan data dan sintaks yang sederhana. Python menyediakan pustaka yang luas yang memungkinkan analisis data dengan mudah, serta mendukung implementasi metode data mining dan *machine learning* (Perkovic, 2019).

Teknik Analisis Data

Pengumpulan Data Langkah 1: Ambil data dari platform media sosial menggunakan teknik web scraping atau API, Langkah 2: Kumpulkan data yang relevan, termasuk postingan, komentar, dan interaksi terkait dengan topik penelitian.

Pra-pemrosesan Data

Langkah 1: Hapus noise, data duplikat, dan informasi yang tidak relevan untuk membersihkan dataset.

Langkah 2: Lakukan tokenisasi, yaitu memecah teks menjadi unit-unit yang lebih kecil seperti kata-kata atau frasa.

Langkah 3: Normalisasi teks dengan mengubah huruf kapital menjadi huruf kecil dan menghapus tanda baca.

Langkah 4: Hapus stopwords, yaitu kata-kata umum yang tidak memberikan informasi penting, untuk memperjelas analisis.

Ekstraksi Fitur

Langkah 1: Ubah teks yang telah diproses menjadi fitur numerik dengan teknik seperti *Bag of Words* (BoW) atau *Term Frequency-Inverse Document Frequency* (TF-IDF).

Langkah 2: Pilih fitur yang relevan dan signifikan untuk analisis lebih lanjut.

Penerapan Model Naive Bayes

Langkah 1: Latih model Naive Bayes dengan menggunakan dataset pelatihan yang telah diproses dan diekstraksi fiturnya.

Langkah 2: Klasifikasikan data baru untuk menentukan sentimen atau opini publik terkait Pemilu.

Evaluasi Model

Langkah 1: Uji model dengan dataset pengujian untuk mengukur kinerja model.

Langkah 2: Gunakan metrik evaluasi seperti akurasi, precision, recall, dan F1-score untuk menilai kualitas prediksi.

Langkah 3: Bandingkan hasil prediksi dengan data hasil Pemilu yang nyata jika data tersebut tersedia untuk validasi lebih lanjut.

HASIL DAN PEMBAHASAN

Proses pengumpulan data dilakukan untuk mendapatkan informasi mengenai partai politik yang mencakup data jumlah *follower* di berbagai platform media sosial dan status verifikasi (*verified*) dari akun-akun tersebut. Data yang dikumpulkan terdiri dari:

Data Partai dan Jumlah Follower

Data ini mencakup jumlah *follower* dari beberapa partai politik pada platform media sosial seperti Facebook, Instagram, Twitter, dan TikTok.

Hasil pengumpulan data dapat dilihat pada Tabel 1 yang memperlihatkan jumlah *follower* masing-masing partai di platform-platform tersebut.

Tabel 1. Data Partai dan Follower di Semua Platform

Partai	Follower Facebook	Follower Instagram	Follower Twitter	Follower TikTok
PKB	89.000	566.322	126.058	17.000
Gerindra	3.400.000	549.534	675.761	6.751
PDIP	1.500.000	277.854	292.568	7.438
Golkar	26.000	80.954	3.191	32.000
NasDem	128.000	117.756	145.599	18.300
Partai Buruh	5.500	6.807	2.179	4.670
Partai Gelora	12.000	45.378	12.763	6.765
PKS	1.000.000	319.248	475.195	52.200
PKN	2.200	2.071	337	421
Hanura	491.000	6.495	118	343

Data Partai Berdasarkan Status Verifikasi

Setiap partai yang terdaftar pada penelitian ini telah terverifikasi 100%, yang berarti akun-akun mereka di media sosial dapat diandalkan untuk analisis data.

Data ini disajikan dalam Tabel 2 yang menampilkan nama partai, website, dan probabilitas verifikasi.

Tabel 2. Partai Berdasarkan Status Verifikasi

Id_partai	Nama_partai	Website	Probabilitas_verified	Rentang Nilai Probabilitas
P1	PKB	pkb.id	1 (100%)	0,90 – 1,00
P2	GERINDRA	gerindra.id	1 (100%)	0,90 – 1,00
P3	PDIP	pdiperjuangan.id	1 (100%)	0,90 – 1,00
P4	NASDEM	nasdem.id	1 (100%)	0,90 – 1,00
P5	PKS	pks.id	1 (100%)	0,90 – 1,00

Dari hasil analisis, rentang nilai probabilitas untuk semua partai yang terverifikasi adalah antara 0.90 hingga 1.00, yang menunjukkan tingkat kepercayaan yang tinggi terhadap kehadiran partai tersebut di media sosial. Rentang nilai ini dikategorikan menjadi tiga tingkatan:

Tinggi (0.90 - 1.00) : Menunjukkan verifikasi yang sangat baik.

Sedang (0.70 - 0.89) : Verifikasi yang cukup baik tetapi memerlukan peningkatan.

Rendah (0.00 - 0.69) : Memerlukan perbaikan signifikan untuk kehadiran digital.

Analisis Jumlah Follower Berdasarkan Verifikasi

Analisis ini dilakukan untuk melihat keterkaitan antara jumlah *follower* dan status verifikasi akun. Hasil analisis disajikan dalam Tabel 3.

Tabel 3. Jumlah Follower Berdasarkan Verifikasi

Id_media	Id_partai	Platform	Follower	Probabilitas_verified
M1	P1	Facebook	89.000	1 (100%)
M2	P1	Instagram	566.322	1 (100%)
M3	P1	Twitter	126.058	1 (100%)
M4	P2	Facebook	3.400.000	1 (100%)
--	--	--	--	--

Pembentukan Basis Pengetahuan

Dari data yang diperoleh, basis pengetahuan dibangun untuk menggambarkan keterkaitan antara partai politik dan jumlah *follower*-nya berdasarkan status verifikasi. Basis pengetahuan ini dapat digunakan untuk analisis lanjutan mengenai keterlibatan partai di media sosial.

Analisis Naive Bayes

Untuk memprediksi hasil Pemilu berdasarkan data jumlah *follower*, digunakan model Naive Bayes yang mengandalkan probabilitas dan hukum total probabilitas. Proses perhitungannya melibatkan langkah-langkah berikut:

Perhitungan Probabilitas Awal (Prior)

Sebagai contoh, untuk partai PKB:

$$P(H) = P(PKB) = \frac{\text{Jumlah PKB}}{\text{Total PKB}} = \frac{1}{10} = 0.1$$

1. Perhitungan Likelihood P(X|H)

Untuk *follower* Facebook PKB:

$$P(X/H) = P(\text{Follower Facebook PKB}) = \frac{\text{Follower Facebook PKB}}{\text{Total Follower Facebook}} = \frac{99000}{10000000} = 0.01$$

Perhitungan Probabilitas Total P(X)

Menghitung probabilitas total berdasarkan data semua partai:

$$P(X) = \sum_{i=1}^n P(X|H_i) \cdot P(H_i)$$

Perhitungan Probabilitas Posterior P(H|X)

Probabilitas untuk PKB diberikan data X:

$$P(PKB/X) = \frac{P(X|PKB) \cdot P(PKB)}{P(X)} = \frac{0.0134 \cdot 0.1}{0.07954} \approx \frac{0.00134}{0.07954} \approx 0.01684$$

Hasil ini menunjukkan bagaimana data jumlah *follower* dapat memberikan gambaran mengenai kemungkinan preferensi publik terhadap partai politik tertentu, yang dianalisis menggunakan model prediktif Naive Bayes.

SIMPULAN

Penelitian ini berhasil menerapkan teknik data mining Naive Bayes untuk memprediksi hasil pemilihan umum dengan menggunakan data dari berbagai platform media sosial. Hasil yang diperoleh menunjukkan adanya hubungan signifikan antara jumlah pengikut partai dan probabilitas kemenangan. Meskipun terdapat tantangan dalam pengolahan data yang tidak teratur dan bersifat informal, pendekatan yang digunakan telah memberikan gambaran yang jelas mengenai potensi setiap partai dalam konteks pemilihan umum. Temuan ini menegaskan bahwa data media sosial dapat menjadi alat yang efektif dalam meramalkan hasil pemilu.

DAFTAR PUSTAKA

Alkhairi & Widarto, d. (2019). ANALISIS

SENTIMEN MASYARAKAT PADA TWITTER TERHADAP PEMILIHAN UMUM 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES.

Amelia, L. &. (2017). JTAM (Jurnal Teori dan Aplikasi Matematika) Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur.

Andika. (2019). Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes. Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes.

Arifin & Krisnadita. (2017). Membangun Website MTS Negeri 01 OKU Timur Menggunakan Php dan Mysql.

Fitriani, A. S. (2019). JTAM (Jurnal Teori dan Aplikasi Matematika) Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur.

Hastanti. (2015). Membangun Website MTS Negeri 01 OKU Timur Menggunakan Php dan Mysql.

Hidayatullah & Kawistara. (2017). Membangun Website MTS Negeri 01 OKU Timur Menggunakan Php dan Mysql.

Juanita, & Hendy B, M. (2020). Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes. Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes.

Krisnadita, A. d. (2017). Yayasan Insan Cipta Medan APLIKASI BUKU TAMU MENGGUNAKAN FITUR KAMERA DAN AJAX BERBASIS WEBSITE PADA KANTOR DISPORA KOTA MEDAN.

kusnadi & putri, p. (2021). ANALISIS SENTIMEN MASYARAKAT PADA TWITTER TERHADAP

- PEMILIHAN UMUM 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES.
- Manullang, O. (2023). Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based dan Random Forest. Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based dan Random Forest.
- Morrisan. (2017). Pemilihan Umum berdasarkan undang-undang dasar 1945. BAB 2 KAJIAN PUSTAKA. Palit. (2015).
- Palit. (2015). Yayasan Insan Cipta Medan APLIKASI BUKU TAMU MENGGUNAKAN FITUR KAMERA DAN AJAX BERBASIS WEBSITE PADA KANTOR DISPORA KOTA MEDAN.
- Perkovic, L. (2019).
- Putra. (2021). Membangun Website MTS Negeri 01 OKU Timur Menggunakan Php dan Mysql.
- Rahmawati, E. S. (2022). Yayasan Insan Cipta Medan APLIKASI BUKU TAMU MENGGUNAKAN FITUR KAMERA DAN AJAX BERBASIS WEBSITE PADA KANTOR DISPORA KOTA MEDAN.
- RSipayung, M. (2024). Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes. Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes.
- Safitri, P. d. (2015). Yayasan Insan Cipta Medan APLIKASI BUKU TAMU MENGGUNAKAN FITUR KAMERA DAN AJAX BERBASIS WEBSITE PADA KANTOR DISPORA KOTA MEDAN.
- Salamah, U. G. (2021). Yayasan Insan Cipta Medan APLIKASI BUKU TAMU MENGGUNAKAN FITUR KAMERA DAN AJAX BERBASIS WEBSITE PADA KANTOR DISPORA KOTA MEDAN.
- Saleh. (2015). JTAM (Jurnal Teori dan Aplikasi Matematika) Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur.
- sarbaini. (2015). Kajian Pustaka.
- Satria M D, Z. Z. (2023). The Role of Social Media in Increasing Political Literacy of Beginning Voters. The Role of Social Media in Increasing Political Literacy of Beginning Voters.
- Soekanto, S. (2010). PENGARUH MEDIA SOSIAL TERHADAP PERUBAHAN SOSIAL MASYARAKAT DI INDONESIA. 79-Article Text-148-1-10-20170314.
- Supardi, M. &. (2017). Membangun Website MTS Negeri 01 OKU Timur Menggunakan Php dan Mysql.
- Suprpto, I. &. (2015). JTAM (Jurnal Teori dan Aplikasi Matematika) Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur.
- Susilo Yuda Irawan, A. (2023). ANALISIS SENTIMEN MASYARAKAT PADA TWITTER TERHADAP PEMILIHAN UMUM 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES. ANALISIS SENTIMEN MASYARAKAT PADA TWITTER TERHADAP PEMILIHAN UMUM 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES.
- Wartoyo, X. (2021). THE CONCEPT AND ITS IMPLEMENTATION OF INDONESIAN LEGISLATIVE ELECTIONS BASED ON THE PANCASILA DEMOCRACY PERSPECTIVE. THE CONCEPT AND ITS IMPLEMENTATION OF INDONESIAN LEGISLATIVE ELECTIONS BASED ON THE PANCASILA DEMOCRACY PERSPECTIVE.