

Analisis Sentimen Publik terhadap Hasil Pemilu dengan Naïve Bayes pada Media Sosial X

Analysis of Public Sentiment on Election Results using Naïve Bayes in Social Media X

¹Ahmad Syakir Muliana, ²Dinda Lestarini*, ³Sarifah Putri Raflesia

^{1,2,3}Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya

^{1,2,3}Jl. Palembang- Indralaya KM.33, Indralaya, Sumatra Selatan, Indonesia

*e-mail: dinda.lestarini@gmail.com

(received: 9 September 2024, revised: 7 October 2024, accepted: 9 October 2024)

Abstrak

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap hasil Pemilu Indonesia 2024 menggunakan Naïve Bayes pada data media sosial dari platform X (sebelumnya Twitter). Sebanyak 2.500 tweet terkait pemilu dikumpulkan melalui teknik web scraping dan diproses menggunakan tokenisasi, penghapusan stopword, stemming, serta pembobotan TF-IDF. Untuk menangani ketidakseimbangan kelas, diterapkan Teknik Oversampling Minoritas Sintetis (SMOTE). Model Naïve Bayes dievaluasi dengan Stratified K-Fold Cross-Validation, menghasilkan akurasi rata-rata sebesar 66,90% pada data uji dan 80% selama cross-validation. Hasil penelitian menunjukkan efektivitas klasifikasi sentimen positif, namun terdapat tantangan dalam mengidentifikasi sentimen negatif dan netral secara akurat. Temuan ini memberikan implikasi penting bagi pembuat kebijakan dan partai politik dalam merumuskan strategi komunikasi yang responsif. Penelitian lanjutan direkomendasikan untuk mengeksplorasi algoritma yang lebih canggih guna meningkatkan akurasi klasifikasi sentimen, khususnya dalam mengidentifikasi sentimen netral.

Kata kunci: Sentiment analysis, naïve bayes, social media, hasil pemilu, natural language processing (NLP)

Abstract

The objective of the research is to examine the public opinion regarding the 2024 Indonesian election results by applying Naïve Bayes to social media data obtained from platform X of Twitter. A dataset comprising 2,500 election-related tweets was obtained by web scraping and then subjected to tokenization, stopword elimination, stemming, and TF-IDF weighting for preprocessing. The application of the Synthetic Minority Oversampling Technique (SMOTE) was attempted to mitigate class imbalance. The performance of the Naïve Bayes model was assessed using Stratified K-Fold Cross-Validation. The model achieved an average accuracy of 66.90% on the test set and 80% during cross-validation. The results demonstrate successful categorization of positive sentiment, although the model encountered difficulties in precisely detection of negative and neutral sentiments. The results underscore significant consequences for policymakers and political parties in formulating effective communication strategies. Further study is advised to investigate sophisticated algorithms to improve the accuracy of sentiment classification, namely in detecting neutral sentiments.

Keywords: Sentiment analysis, naïve bayes, social media, election results, natural language processing (NLP).

1 Pendahuluan

Pemilihan umum merupakan salah satu pilar utama demokrasi yang memiliki dampak signifikan terhadap stabilitas politik dan sosial suatu negara. Hasil pemilu tidak hanya menentukan arah kebijakan pemerintahan, tetapi juga mencerminkan persepsi dan sentimen masyarakat terhadap proses politik. Dalam era digital saat ini, media sosial telah berkembang menjadi platform utama bagi masyarakat untuk mengekspresikan opini dan perasaan mereka terkait hasil pemilu. Oleh karena itu,

analisis sentimen terhadap hasil pemilu menjadi sangat penting dalam memahami reaksi publik serta implikasinya terhadap dinamika politik dan sosial[1].

Teknik analisis sentimen dalam penelitian pengolahan bahasa alami (Natural Language Processing/NLP) yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini dalam teks sebagai positif, negatif, atau netral[2]. Salah satu metode yang efektif dalam analisis sentimen adalah Naïve Bayes, sebuah algoritma klasifikasi probabilistik yang sederhana namun kuat. Algoritma ini menggunakan teorema Bayes dengan asumsi independensi antar fitur untuk memprediksi kategori dari teks yang dianalisis[3].

Penelitian ini memberikan kontribusi kebaruan dalam beberapa aspek. Pertama, penelitian ini menggabungkan teknik NLP dan algoritma Naïve Bayes untuk menganalisis sentimen masyarakat terhadap hasil Pemilu Indonesia, yang merupakan konteks politik lokal yang masih jarang dieksplorasi dalam literatur sebelumnya[1]. Kedua, penelitian ini memanfaatkan berbagai platform media sosial untuk mengumpulkan data, memberikan perspektif yang lebih komprehensif dibandingkan penelitian sebelumnya yang mungkin hanya terfokus pada satu sumber data.

Masalah yang diangkat dalam konteks penelitian ini menyangkut “Bagaimana sentimen masyarakat terhadap hasil pemilu di Indonesia dapat diidentifikasi dan diklasifikasikan menggunakan metode Naïve Bayes, dan apa implikasinya terhadap stabilitas politik dan sosial?” Hipotesis yang diuji akan menunjukkan bahwa dominasi sentimen negatif atau positif terhadap hasil pemilu dapat berfungsi sebagai indikator awal dari potensi ketidakstabilan atau kepuasan politik di masyarakat[4].

Penelitian ini terdiri dari beberapa tahap. Pertama, data teks dikumpulkan dari media sosial dan diproses melalui tokenisasi menggunakan NLTK (Natural Language Toolkit), penghapusan stopword, dan stemming dengan Sastrawi. Setelah itu, teks yang telah diproses diterjemahkan dari Bahasa Indonesia ke Bahasa Inggris menggunakan Deep Translator[5]. Kemudian, sentimen dianalisis menggunakan VADER (Valence Aware Dictionary and Sentiment Reasoner)[2].

Penelitian ini diharapkan dapat memberikan wawasan mendalam tentang bagaimana masyarakat merespons hasil pemilu, serta bagaimana analisis sentimen dapat digunakan sebagai alat untuk memprediksi kecenderungan politik dan sosial[6]. Dengan memahami pola sentimen masyarakat, pembuat kebijakan dapat mengambil keputusan yang lebih bijaksana dan responsif terhadap kebutuhan dan aspirasi publik[7].

Kontribusi utama dari penelitian ini adalah pengembangan model analisis sentimen berbasis Naïve Bayes yang dapat diterapkan dalam konteks politik, terutama dalam mengukur reaksi masyarakat terhadap hasil pemilu. Selain itu, penelitian ini menekankan pentingnya penggunaan teknik NLP dalam menganalisis data teks dari media sosial yang beragam dan kompleks[8]. Oleh karena itu, hasil penelitian ini diharapkan dapat memberikan kontribusi signifikan bagi literatur akademis serta praktik dalam bidang ilmu politik dan data sains.

2 Tinjauan Literatur

Analisis sentimen menjadi salah satu topik penting dalam studi terkait pemilu di era digital, terutama dalam konteks media sosial yang semakin mendominasi sebagai platform utama bagi masyarakat untuk mengekspresikan opini. Dalam konteks ini, beberapa penelitian terbaru telah mengkaji efektivitas berbagai algoritma pembelajaran mesin dalam klasifikasi sentimen, termasuk Naïve Bayes, yang sering dipilih karena sifatnya yang sederhana namun andal dalam menangani teks pendek.

Penelitian yang dilakukan oleh Maharani et al[7] menunjukkan bahwa Naïve Bayes tetap menjadi salah satu algoritma yang efisien untuk klasifikasi sentimen di media sosial, terutama untuk mengidentifikasi sentimen positif dan negatif. Namun, penelitian ini menemukan bahwa algoritma tersebut memiliki keterbatasan dalam membedakan sentimen netral yang merupakan tantangan signifikan karena model sering bias terhadap sentimen positif atau negatif. Jeet et al[2] mendukung temuan ini dengan menekankan pentingnya teknik praproses data seperti Term Frequency-Inverse Document Frequency (TF-IDF) yang dapat memperbaiki akurasi klasifikasi sentimen. Mereka menemukan bahwa penggunaan TF-IDF secara signifikan meningkatkan kualitas fitur teks, terutama ketika digunakan untuk teks pendek seperti tweet. Teknik ini memprioritaskan kata-kata yang lebih relevan sehingga dapat mengidentifikasi sentimen dengan lebih akurat.

Dalam hal ketidakseimbangan data, yang merupakan tantangan besar dalam analisis sentimen, penelitian Gede dan Arya Budaya et al[9] menunjukkan bahwa Synthetic Minority Over-sampling Technique (SMOTE) memberikan solusi yang efektif. SMOTE membantu mengatasi ketidakseimbangan kelas dengan menciptakan sampel sintetis dari kelas minoritas, sehingga memperbaiki distribusi data yang tidak seimbang sehingga membantu meningkatkan performa model secara keseluruhan.

Meskipun banyak penelitian telah berhasil menerapkan algoritma Naïve Bayes dan teknik NLP seperti TF-IDF dan SMOTE, masih terdapat beberapa celah yang belum terisi. Salah satunya adalah kurangnya penelitian yang secara spesifik menggunakan data berbahasa Indonesia dalam konteks politik lokal, seperti analisis sentimen terkait pemilu di Indonesia. Bahasa Indonesia memiliki karakteristik linguistik yang berbeda dari bahasa Inggris, yang membuat hasil dari penelitian berbasis bahasa Inggris tidak sepenuhnya dapat diaplikasikan pada konteks lokal terutama dalam konteks politik lokal seperti pemilu. Selain itu, klasifikasi sentimen netral masih menjadi tantangan, karena model cenderung bias terhadap sentimen positif atau negatif, dengan tingkat akurasi yang lebih rendah dalam mendeteksi sentimen netral.

Berdasarkan analisis literatur ini, hipotesis yang dikembangkan dalam penelitian ini adalah bahwa dengan mengombinasikan Naïve Bayes, TF-IDF, dan SMOTE[10] pada data media sosial X yang berbahasa Indonesia, model ini akan mampu meningkatkan akurasi dalam mengklasifikasikan sentimen, terutama dalam mendeteksi sentimen netral. Penelitian ini juga berhipotesis bahwa penggunaan teknik praproses yang sesuai, seperti tokenisasi, stemming, dan normalisasi untuk bahasa Indonesia, akan lebih meningkatkan kinerja model dalam konteks politik Indonesia.

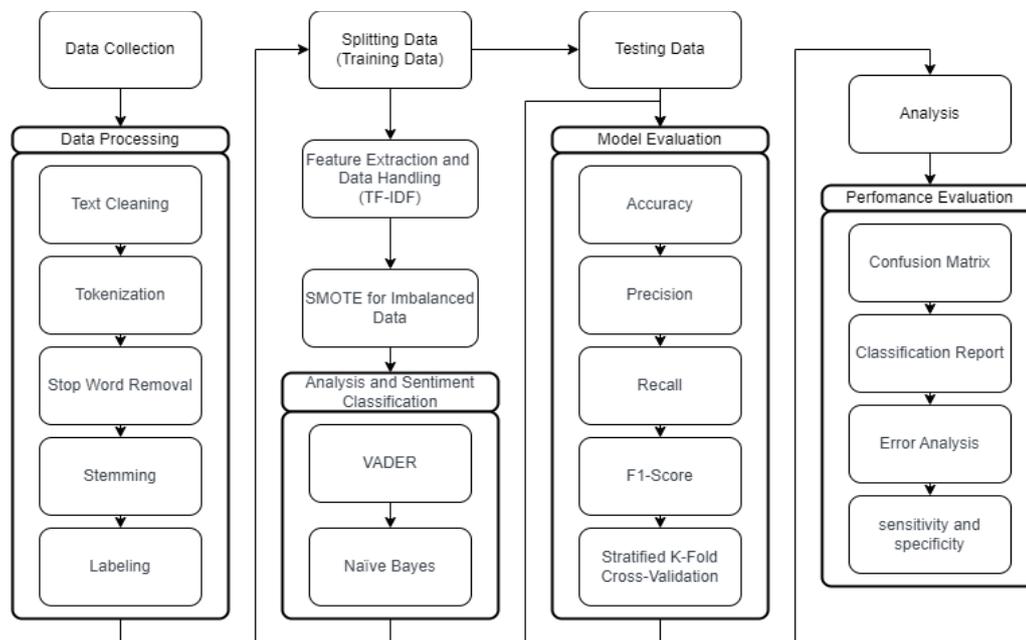
Artikel ini berfokus untuk mengisi kesenjangan tersebut dengan menawarkan pendekatan baru yang belum banyak dieksplorasi dalam literatur sebelumnya. Pertama, penelitian ini secara khusus menggunakan data berbahasa Indonesia yang relevan dengan konteks politik lokal, seperti tweet terkait hasil Pemilu Indonesia 2024. Kedua, artikel ini berupaya memperbaiki kelemahan model Naïve Bayes dalam mendeteksi sentimen netral melalui pengoptimalan metode SMOTE dan TF-IDF yang disesuaikan untuk menangani karakteristik bahasa Indonesia di media sosial. Dengan demikian, penelitian ini memberikan kontribusi yang signifikan dalam meningkatkan pemahaman tentang pola sentimen masyarakat terhadap hasil pemilu di Indonesia, sekaligus menawarkan solusi teknis untuk permasalahan ketidakseimbangan data dan klasifikasi sentimen netral yang belum diatasi oleh penelitian sebelumnya.

Pada bagian akhir tinjauan literatur, penelitian ini menunjukkan novelty dengan menggabungkan berbagai teknik NLP dalam analisis sentimen berbahasa Indonesia, sesuatu yang masih jarang dibahas dalam penelitian terdahulu. Dengan memfokuskan pada sentimen netral yang sering diabaikan, artikel ini menawarkan pendekatan yang lebih menyeluruh dalam memahami dinamika sosial-politik di Indonesia, serta memberikan kontribusi baru dalam penggunaan Naïve Bayes dan teknik SMOTE pada data yang tidak seimbang.

3 Metode Penelitian

Bagian ini menjelaskan tentang tahapan yang dilakukan dalam penelitian ini. Penelitian ini menggunakan pendekatan kuantitatif untuk menganalisis sentimen publik terhadap hasil Pemilu Indonesia 2024 dengan menerapkan algoritma Naïve Bayes.

3.1 Desain Penelitian



Gambar 1. Metodologi penelitian

Gambar 1 merupakan tahapan penelitian untuk menganalisis sentimen publik terhadap hasil Pemilu Indonesia 2024 dengan menerapkan algoritma Naïve Bayes. Penelitian ini dimulai dengan pengumpulan data dari platform media sosial X (sebelumnya Twitter)[11], diikuti oleh serangkaian proses praproses data yang mencakup tokenisasi, penghapusan stopword, stemming, dan pembobotan kata menggunakan Term Frequency-Inverse Document Frequency (TF-IDF)[12]. Data yang telah diproses kemudian digunakan untuk melatih dan menguji model Naïve Bayes guna mengklasifikasikan sentimen menjadi tiga kategori: positif, negatif, dan netral. Pembagian data latih dan uji dilakukan dengan rasio 80:20, menggunakan teknik stratifikasi untuk memastikan distribusi kelas sentimen yang seimbang dalam kedua subset[13].

3.2 Pengumpulan dan Persiapan Data

created_at	full_text	username
Wed May 01 11:30:44 +0000 2024	Aku sebagai pendukung Prabowo pendapat pribadi sebaiknya Anies Muhaimin jangan masuk kabinet BH4RATA_JA	
Wed May 01 01:30:55 +0000 2024	Bubarrrrrrr https://t.co/ugFyySOy36 Hermawi mengatakan Surya Paloh bangga dengan kinerja Timn: maximilan_angel	
Tue Apr 30 10:53:31 +0000 2024	@aniesbaswedan @cakimiNOW @feryfarhati Sygkuuuuu u did well. Terima kasih cerita di Novembe ddaifuku	
Tue Apr 30 08:40:38 +0000 2024	@cakimiNOW Pak Anies dan Gus Muhaimin benar2 telah memberi kita contoh berpolitik yang baik	Agusmal11
Tue Apr 30 05:11:47 +0000 2024	@gatauyaudahlah amin nih anies muhaimin apa gimana?	S4pickTank
Tue Apr 30 01:13:09 +0000 2024	@abu_waras Coba ja u jual foto Anies dan muhaimin	abinandar2003
Mon Apr 29 17:46:02 +0000 2024	@didik_lp Sepertinya pak Anies memang buru2 dideklarasikan dgn tujuan agar kalah lalu dipilih Gus ! QiuQiu_Official	
Mon Apr 29 12:45:56 +0000 2024	Ganjar Anies Airlangga Muhaimin dan Jokowi semuanya alumni UGM. FYI.	Opcux
Mon Apr 29 12:42:03 +0000 2024	@Jelli_cent Coba ganti foto Anies - Muhaimin atau Ganjar -Mahfudz ada yg beli ndak ? Terus diband khusus_mub	

Gambar 2. Sampel dataset

Data dikumpulkan melalui teknik web scraping, dengan fokus pada tweet yang mengandung kata kunci terkait hasil pemilu di Indonesia seperti "hasil pemilu," "pilpres," "Capres," dan "KPU." Setelah

pengumpulan, data dibersihkan dari elemen tidak relevan seperti tautan, mention, dan karakter khusus[1][10]. Gambar 2 menunjukkan contoh dataset yang berhasil dikumpulkan. Dataset awal yang terdiri dari 946 contoh sentimen positif, 479 contoh sentimen negatif, dan 288 contoh sentimen netral menunjukkan adanya ketidakseimbangan kelas yang signifikan. Untuk mengatasi hal ini, diterapkan Synthetic Minority Over-sampling Technique (SMOTE) yang menghasilkan distribusi yang lebih seimbang di seluruh kelas sentimen[9].

Original dataset shape : Positive: 946, Negative: 479, Neutral: 288.

Resampled dataset shape : Positive: 946, Negative: 946, Neutral: 946.

3.3 Praproses Data

full_text	cleaned_text	processed_text
Aku sebagai pendukung Prabowo pendapat pribadi saya sebagai pendukung prabowo pendapat pribadi dukung prabowo dapat pribadi baik anies muhaimin jan	Bubarrrrrrr https://t.co/ugFyySOy36 Hermawi me bubarrrrrrr hermawi mengatakan surya paloh ba	bubarrrrrrr hermawi kata surya paloh bangga kerja timn
@aniesbaswedan @cakimiNOW @feryfarhati Syg sygkuuuuu kamu did well terima kasih cerita di n	@aniesbaswedan @cakimiNOW @feryfarhati Syg sygkuuuuu kamu did well terima kasih cerita november a	@aniesbaswedan @cakimiNOW @feryfarhati Syg sygkuuuuu kamu did well terima kasih cerita november a
@cakimiNOW Pak Anies dan Gus Muhaimin benar pak anies dan gus muhaimin benar telah membe	@cakimiNOW Pak Anies dan Gus Muhaimin benar pak anies dan gus muhaimin benar telah membe	@aniesbaswedan @cakimiNOW @feryfarhati Syg sygkuuuuu kamu did well terima kasih cerita november a
@gatauyaudahlah amin nih anies muhaimin apa g amin nih anies muhaimin apa gimana	@gatauyaudahlah amin nih anies muhaimin apa g amin nih anies muhaimin apa gimana	amin nih anies muhaimin apa gimana
@abu_waras Coba ja u jual foto Anies dan muhai coba saja kamu jual foto anies dan muhaimin	@abu_waras Coba ja u jual foto Anies dan muhai coba saja kamu jual foto anies dan muhaimin	coba kamu jual foto anies muhaimin
@didik_lp Sepertinya pak Anies memang buru2 di seperti pak anies memang buru dideklarasika	@didik_lp Sepertinya pak Anies memang buru2 di seperti pak anies memang buru dideklarasika	seperti pak anies memang buru deklarasi tuju kalah lalu

Gambar 3. Sampel sentimen yang diproses

Gambar 3 menunjukkan contoh dataset yang telah melalui tahap praproses. Tahap praproses data melibatkan beberapa langkah penting untuk memastikan kualitas data yang akan dianalisis:

1. **Casefolding:** Menghapus URL, mention (@), hashtag (#), dan karakter khusus lainnya yang tidak relevan dengan analisis sentimen menggunakan metode regex.
2. **Penghapusan Stop Words:** Menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis sentimen seperti "dan," "atau," "yang," dan sebagainya menggunakan library Sastrawi[6].
3. **Tokenisasi:** Memecah teks menjadi unit-unit kata atau token untuk memudahkan analisis menggunakan library NLTK[2].
4. **Stemming:** Mengubah kata-kata ke bentuk dasarnya untuk mengkonsolidasikan variasi kata yang memiliki makna sama menggunakan library Sastrawi[6].
5. **Normalisasi Teks:** Mengganti kata-kata informal atau slang dengan kata-kata baku sesuai kamus normalisasi yang telah disiapkan.
6. **Translasi Teks:** Untuk memudahkan analisis dan labeling, teks yang telah diproses diterjemahkan dari bahasa Indonesia ke bahasa Inggris menggunakan library Deep Translator[5].
7. **Labeling:** Pemberian label sentimen secara manual pada subset data, mengklasifikasikannya sebagai positif, negatif, atau netral.

3.4 Stratified K-Fold Cross-Validation

Untuk memastikan bahwa model Naïve Bayes yang dikembangkan mampu menggeneralisasi dengan baik pada data yang tidak terlihat, digunakan teknik Stratified K-Fold Cross-Validation dengan lima kali pemisahan. Teknik ini memastikan bahwa setiap subset data latih dan uji memiliki proporsi kelas yang seimbang, sehingga evaluasi model mencerminkan performa yang konsisten dan andal [14].

3.5 Pembobotan Kata

Metode **Term Frequency-Inverse Document Frequency (TF-IDF)** digunakan dalam model pembelajaran mesin. TF-IDF mengukur kepentingan suatu kata dalam sebuah dokumen dengan mempertimbangkan frekuensinya dalam dokumen tersebut relatif terhadap seluruh koleksi dokumen. Metode ini dipilih karena kemampuannya dalam memberikan bobot lebih tinggi pada kata-kata yang jarang muncul tetapi signifikan dalam analisis[15]. Perhitungan TF-IDF dapat dilakukan dengan menggunakan rumus (1) dan (2).

$$TF(t, d) = \frac{n_{t,d}}{\sum_{t \in d} n_{t,d}} \quad (1)$$

$$IDF(t, d) = \log\left(\frac{D}{|\{d \in D: t \in d\}|}\right) \quad (2)$$

Proses pembobotan ini dilakukan menggunakan pustaka scikit-learn, yang menghasilkan matriks fitur yang merepresentasikan bobot kata-kata dalam teks yang telah diproses[16]. Perhitungan untuk pembobotan dapat dilakukan dengan menggunakan rumus (3).

$$TFIDF(t, d, D) = TF(t, d) \cdot IDF(t, D) \quad (3)$$

3.6 Klasifikasi Sentimen

Model Naïve Bayes digunakan untuk mengklasifikasikan sentimen teks menjadi tiga kategori: positif, negatif, dan netral. Model ini dibangun berdasarkan data pelatihan yang telah diberi label sentimen. Data pelatihan digunakan untuk melatih model agar dapat mengidentifikasi pola sentimen dalam data teks[3]. Ketidakseimbangan kelas dalam data dapat menyebabkan bias terhadap kelas mayoritas. Untuk mengatasi ini, diterapkan Synthetic Minority Over-sampling Technique (SMOTE) menghasilkan sampel sintetis untuk kelas minoritas dengan menciptakan titik data baru di antara sampel minoritas dan tetangga terdekatnya[9]. Teknik ini membantu menyeimbangkan distribusi kelas, sehingga model dapat belajar lebih efektif dari semua kelas, dan meningkatkan akurasi dalam klasifikasi sentimen[17].

3.7 Analisis Sentimen

Setelah teks diterjemahkan, analisis sentimen dilakukan menggunakan VADER (Valence Aware Dictionary and sEntiment Reasoner). VADER adalah alat yang digunakan untuk analisis sentimen berbasis aturan, khususnya efektif untuk teks pendek seperti tweet[5][11].

3.8 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja model Naïve Bayes dalam mengklasifikasikan sentimen. Metode evaluasi yang digunakan meliputi:

1. **Akurasi:** Mengukur persentase prediksi yang benar dari seluruh prediksi yang dibuat oleh model. Pengukuran nilai akurasi dapat dilakukan dengan menggunakan rumus (4).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Observations} \quad (4)$$

2. **Presisi:** Mengukur jumlah prediksi positif yang benar dibagi dengan jumlah keseluruhan prediksi positif. Pengukuran nilai presisi dapat dilakukan dengan menggunakan rumus (5)

$$Accuracy = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (5)$$

3. **Recall:** Mengukur jumlah prediksi positif yang benar dibagi dengan jumlah keseluruhan data positif yang sebenarnya. Pengukuran nilai recall dapat dilakukan dengan menggunakan rumus (6).

$$Accuracy = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (6)$$

4. **F1-Score:** Menggabungkan presisi dan recall menjadi satu metrik harmonis. Pengukuran nilai F1 Score dapat dilakukan dengan menggunakan rumus (7).

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (7)$$

5. **Stratified K-Fold Cross-Validation:** Digunakan untuk menguji model pada berbagai subset data guna memastikan konsistensi kinerja model[14].

3.9 Analisis Hasil

Hasil klasifikasi sentimen dianalisis menggunakan **Confusion Matrix** untuk mengevaluasi performa model[12]. Confusion Matrix menampilkan jumlah prediksi benar dan salah untuk setiap

kategori sentimen. Confusion matrix yang digunakan untuk menilai kinerja model dapat dilihat pada Tabel 1.

Tabel 1. Confusion matrix

Class		Predicted Class		
		Negative	Neutral	Positive
Actual Class	Negative	True Negative	False Neutral	False Positive
	Neutral	False Negative	True Neutral	False Positive
	Positive	False Negative	False Neutral	True Positive

- True Positives (TP):** Kasus positif yang diprediksi benar.
- True Neutral (TNet):** Kasus Neutral diprediksi benar kelas Netral oleh model.
- True Negatives (TN):** Kasus negatif yang diprediksi benar.
- False Positives (FP):** Kasus negatif yang salah diprediksi sebagai positif.
- False Neutral (FNet):** Kasus Positif atau Negatif, salah diprediksi sebagai kelas Netral oleh model.
- False Negatives (FN):** Kasus positif yang salah diprediksi sebagai negatif.

Dari Table 1 Confusion Matrix diatas ini, menjelaskan bahwa metrik seperti akurasi, presisi, recall, dan F1-score dihitung untuk menilai kinerja model. Akurasi mencerminkan persentase keseluruhan prediksi yang benar, presisi mengukur ketepatan prediksi positif, recall menilai kemampuan model dalam mendeteksi semua kasus positif, dan F1-score menggabungkan presisi serta recall untuk memberikan gambaran keseluruhan performa model[18].

Data yang digunakan diproses ulang untuk menangani nilai-nilai yang hilang dan diubah menjadi fitur numerik melalui tokenisasi dan TF-IDF. Model Naïve Bayes kemudian digunakan untuk menganalisis data ini, yang memberikan wawasan mengenai distribusi sentimen publik dan faktor-faktor yang mempengaruhinya terhadap hasil pemilu[19].

Setelah pelatihan, model diuji pada data uji terpisah untuk menilai kinerjanya dalam kondisi nyata. Evaluasi ini mencakup perhitungan akurasi, pembuatan Confusion Matrix, laporan klasifikasi, dan Error Analysis. Confusion Matrix digunakan untuk mengidentifikasi distribusi prediksi yang benar dan salah di antara berbagai kelas sentimen, yang divisualisasikan dalam bentuk heatmap untuk memudahkan interpretasi.

Error Analysis klasifikasi dianalisis lebih lanjut melalui heatmap distribusi Error Analysis, yang mengungkap hubungan antara sentimen aktual dan prediksi[20]. Analisis ini membantu memahami kelemahan model dalam mengklasifikasikan sentimen yang kompleks atau ambigu. Selain itu, Sensitivity dan Specificity dievaluasi untuk setiap kelas sentimen, dengan hasil yang divisualisasikan guna memberikan wawasan mengenai kemampuan model dalam deteksi dan menghindari kesalahan[21]. Laporan klasifikasi yang mencakup presisi, recall, dan F1-score juga divisualisasikan untuk menyajikan gambaran menyeluruh mengenai performa model dalam setiap kelas sentimen.

3.10. Visualisasi dan Interpretasi

Semua hasil evaluasi dari model Naïve Bayes ini kemudian divisualisasikan untuk memfasilitasi interpretasi dan penarikan kesimpulan. Visualisasi mencakup Confusion Matrix, Error Analysis Heatmap, Sensitivity and Specificity per Class, serta grafik metrik laporan klasifikasi. Visualisasi ini tidak hanya membantu dalam memahami kinerja model, tetapi juga mengidentifikasi area yang memerlukan peningkatan lebih lanjut

4 Hasil dan Pembahasan

Bagian ini menjelaskan hasil penelitian dan pembahasan untuk menganalisis sentimen publik terhadap hasil Pemilu Indonesia 2024 yang dikumpulkan dari platform media sosial X. Hasil penelitian disajikan dalam bentuk tabel dan grafik untuk menggambarkan distribusi sentimen positif, negatif, dan netral, serta efektivitas model Naïve Bayes mengklasifikasikan sentimen tersebut. Pembahasan dilakukan dengan menginterpretasikan hasil yang diperoleh serta mengaitkannya dengan

studi terdahulu untuk menunjukkan kelebihan dan kontribusi penelitian ini dibandingkan penelitian sebelumnya. Analisis dilakukan secara logis dengan memperhatikan keakuratan model dalam mengklasifikasikan sentimen, serta mengidentifikasi tantangan yang dihadapi dalam membedakan sentimen negatif dan netral.

4.1 Hasil Penelitian

Hasil penelitian ini mencakup klasifikasi sentimen publik terhadap hasil Pemilu Indonesia 2024 dengan menggunakan model Naïve Bayes. Model Naïve Bayes dievaluasi menggunakan data pelatihan yang telah disampling ulang. Untuk memvalidasi kinerja model memastikan setiap pemisahan mempertahankan proporsi kelas dataset yang seimbang, menghasilkan skor berikut:

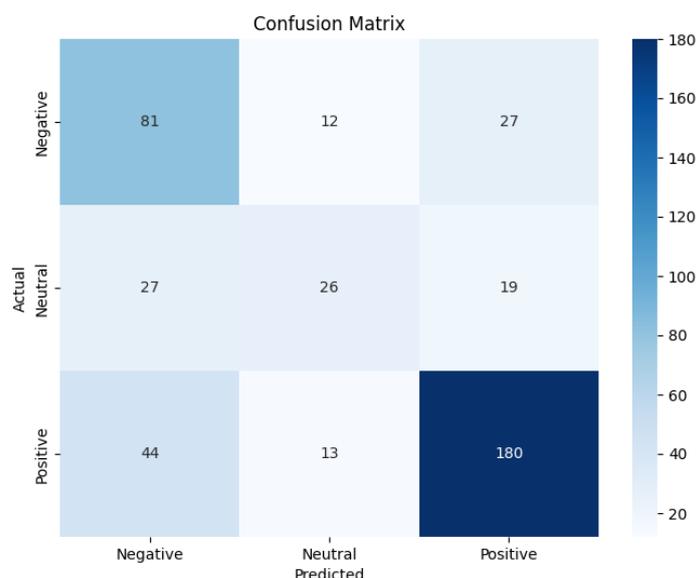
Cross-Validation Scores : 0.746, 0.787, 0.764, 0.866, 0.829
Mean Cross-Validation Score : 0.798
Standar Deviasi : 0.04358936627979872

Hasil ini menunjukkan kinerja yang konsisten dan andal, dengan akurasi rata-rata 79,8% mendekati 80%, menyoroti ketahanan model di seluruh pembagian data pelatihan yang berbeda. Model dilatih dan diuji pada data yang telah diproses dengan rasio 80:20 untuk data latih dan uji. Hasil dari model dievaluasi menggunakan Confusion Matrix, termasuk Classification Report terdiri dari akurasi, presisi, recall, F1-score dan Support.

Tabel 2. Confusion matrix

	Predicted Negative	Predicted Neutral	Predicted Positive
Actual Negative	81	12	27
Actual Neutral	27	26	19
Actual Positive	44	13	180

Tabel 2 menunjukkan hasil evaluasi model pada data uji setelah menerapkan teknik **Synthetic Minority Over-sampling Technique (SMOTE)** untuk menangani ketidakseimbangan kelas



Gambar 4. Heatmap confusion matrix

Pada Gambar 4 **Heatmap Confusion Matrix** berikut menunjukkan distribusi kesalahan dan prediksi benar yang dihasilkan oleh model dari hasil visualisasi ini, dapat dilihat bahwa prediksi untuk sentimen **positif** cenderung lebih akurat, sedangkan sentimen **netral** dan **negatif** sering salah

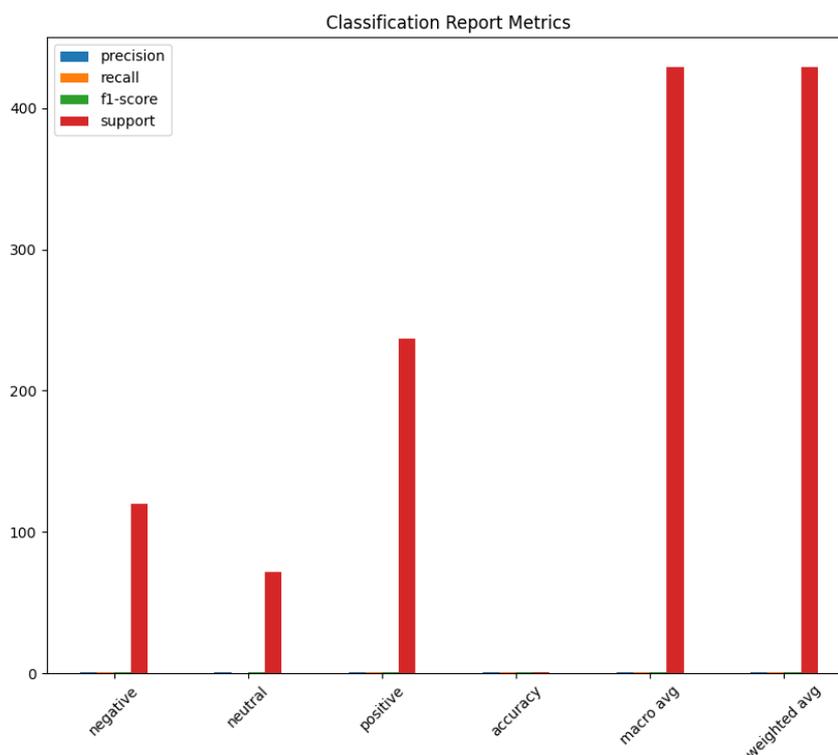
diklasifikasikan.

Tabel 3. Classification report

Sentiment	Precision	Recall	F1-Score	Support
Negative	0.53	0.68	0.60	120
Neutral	0.51	0.36	0.42	72
Positive	0.80	0.76	0.78	237
Accuracy			0.67	429
Macro Avg	0.61	0.60	0.60	429
Weighted Avg	0.67	0.67	0.67	429

Akurasi Keseluruhan: 0.668997668997669

Pada Table 3 di atas terlihat bahwa model secara keseluruhan mencapai akurasi sebesar 66,90%, dengan nilai presisi dan recall di berbagai sentimen menyoroti berbagai tingkat efektivitas model Rata-rata berbobot untuk precision, recall, dan f1-score adalah 0.67.

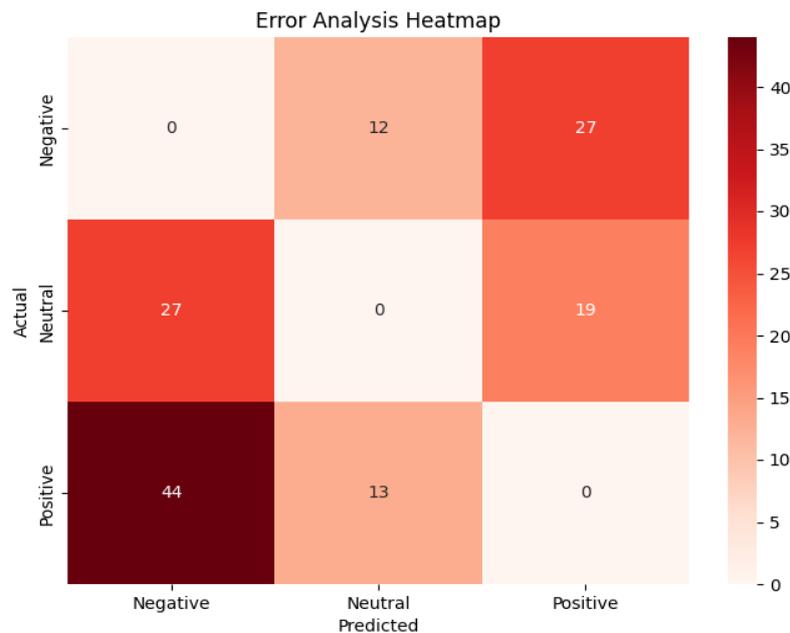


Gambar 5. Classification report metrics

Pada Gambar 5 **Classification Report Metrics** model memiliki kinerja terbaik dalam klasifikasi sentimen positif, dengan nilai F1-Score tertinggi. Dimana model memiliki f1-score terbaik untuk kelas sentimen positif (0.78), namun kinerja untuk sentimen negatif (0.60) dan netral (0.42) masih memerlukan peningkatan kinerja model dalam membedakan sentimen negatif dan netral masih perlu ditingkatkan.

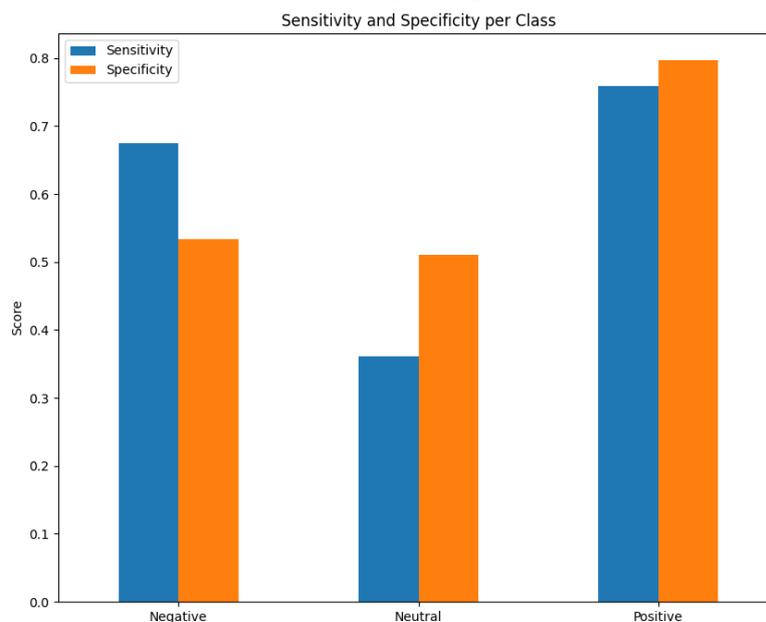
4.2 Error analysis and Sensitivity and Specificity

Untuk lebih memahami area di mana model gagal melakukan Error analysis yang divisualisasikan dalam bentuk heatmap.



Gambar 6. Error analysis

Gambar 6 memvisualisasikan distribusi Error analysis klasifikasi. Dapat dilihat bahwa sebagian besar kesalahan terjadi ketika model salah mengklasifikasikan sentimen netral sebagai negatif, dan sentimen positif sebagai negatif[20]. Ini menunjukkan bahwa model mungkin mengalami kesulitan dalam membedakan konteks yang lebih halus, yang sering terjadi pada sentimen netral.



Gambar 7. Sensitivity and specificity

Sensitivity and Specificity dihitung untuk setiap kelas sentimen untuk memberikan wawasan tambahan tentang bagaimana model ini menangani deteksi kelas tertentu versus menghindari kesalahan pada kelas lainnya, memperlihatkan sensitivitas dan spesifisitas untuk masing-masing kelas sentimen. Sensitivitas tertinggi tercapai pada kelas positif (0.759), sementara spesifisitas terbaik tercapai juga pada kelas positif (0.796). Ini mengindikasikan bahwa model lebih efisien dalam

<http://sistemasi.ftik.unisi.ac.id>

mendeteksi dan mengklasifikasikan sentimen positif dibandingkan dengan sentimen negatif atau netral[21].

5 Kesimpulan

Penelitian ini berhasil menunjukkan bahwa penggunaan algoritma Naïve Bayes untuk analisis sentimen publik terhadap hasil Pemilu Indonesia 2024, melalui data yang dikumpulkan dari media sosial X, memberikan hasil yang signifikan, khususnya dalam klasifikasi sentimen positif. Dengan akurasi keseluruhan sebesar 66,90% dan rata-rata cross-validation sebesar 80%, model menunjukkan kemampuan yang baik dalam mendeteksi sentimen positif. Namun, tantangan tetap ada dalam membedakan antara sentimen negatif dan netral, yang memerlukan perhatian lebih lanjut dalam pengembangan model di masa depan. Penerapan SMOTE untuk mengatasi ketidakseimbangan kelas terbukti membantu meningkatkan kinerja model, terutama dalam menangani data dengan distribusi yang tidak seimbang. Teknik TF-IDF juga berkontribusi dalam memberi bobot yang tepat pada kata-kata penting dalam proses analisis sentimen. Temuan penelitian ini relevan bagi pembuat kebijakan dan pihak-pihak yang tertarik pada reaksi publik terhadap hasil pemilu, karena dapat memberikan wawasan yang lebih mendalam mengenai persepsi masyarakat dan potensi ketidakstabilan politik. Selain itu, penelitian ini juga menyarankan untuk menerapkan algoritma yang lebih kompleks, seperti transformer atau deep learning, guna meningkatkan akurasi dalam klasifikasi sentimen negatif dan netral di masa mendatang.

Referensi

- [1] A. R. Abdillah and F. N. Hasan, "Analisis Sentimen terhadap Kandidat Calon Presiden berdasarkan Tweets Di Sosial Media Menggunakan Naive Bayes Classifier," *SMATIKA J.*, vol. 13, no. 01, pp. 117–130, Jul. 2023, doi: 10.32664/smatika.v13i01.750.
- [2] Nehal, D. Jeet, V. Sharma, S. Mishra, C. Iwendi, and J. Osamor, "Twitter Sentiment Analysis and Emotion Detection using NLTK and TextBlob," *2023 4th Int. Conf. Comput. Autom. Knowl. Manag. ICCAKM 2023*, 2023, doi: 10.1109/ICCAKM58659.2023.10449540.
- [3] J. Sihombing, "Klasifikasi Data Antropometri Individu menggunakan Algoritma Naïve Bayes Classifier," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 1–10, Mar. 2021, doi: 10.37148/bios.v2i1.15.
- [4] A. Risma Alifvia, U. Saprudin, S. Dharma Wacana Metro, J. Kenanga No, K. Metro Barat, and K. Metro, "Analisis Sentimen Review Data Twitter Komisi Pemilihan Umum (Kpu) Menggunakan Metode Naïve Bayes," *J. Inf. dan Komput.*, vol. 11, no. 1, p. 2023, 2023.
- [5] H. Yuliansyah, S. A. Mulasari, S. Sulistyawati, F. A. Ghozali, and B. Sudarsono, "Sentiment Analysis of the Waste Problem based on YouTube comments using VADER and Deep Translator," *J. Media Inform. Budidarma*, vol. 8, pp. 663–673, 2024, doi: 10.30865/mib.v8i1.6918.
- [6] M. D. Purbolaksono, F. D. Reskyadita, and A. A. Suryani, "Indonesian-text-classification-using-back-propagation-and-sastrawi-stemming-analysis-with-information-gain-for-selection-feature2020International-Journal-on-Advanced-Science-Engineering-and-Information-TechnologyOp.pdf," vol. 10, no. 1, pp. 234–238, 2020.
- [7] Z. N. Maharani, A. Luthfiarta, and N. Z. Farsya, "Sentiment Analysis of the 2024 Indonesian Presidential Dispute Trial Election using SVM and Naïve Bayes on Platform X," vol. 6, no. 1, pp. 440–449, 2024, doi: 10.47065/bits.v6i1.5380
- [8] P. Meilina, "Penerapan Data Mining dengan Metode Kalsifikasi menggunakan," *Teknologi*, vol. 7, no. March, pp. 10–20, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>

- [9] J. Homepage *et al.*, “Comparison of Sentiment Analysis Algorithms with SMOTE Oversampling and TF-IDF Implementation on Google Reviews for Public Health Centers,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 3, pp. 1077–1086, 2024, [Online]. Available: <https://journal.irpi.or.id/index.php/malcom/article/view/1459>
- [10] H. H. Limbong, “Optimasi Analisis Sentimen Ulasan Aplikasi Amikom One menggunakan SMOTE pada Algoritma Artificial Neural Network Optimization of Sentiment Analysis for Amikom One Application Reviews Using SMOTE with Artificial Neural Network Algorithm,” vol. 13, pp. 2048–2059, 2024.
- [11] J. Kimani, A. Karanjah, and P. Kihara, “Sentiment Classification of Safaricom PLC Social Media Sentiments on X (Formerly Twitter),” *Asian J. Probab. Stat.*, vol. 26, no. 6, pp. 31–40, 2024, doi: 10.9734/ajpas/2024/v26i6622.
- [12] D. Lestarini, T. Rusdy, S. Iriyani, and S. P. Raflesia, “Detection of Indonesian wildlife sales and promotion through social media using machine learning approach,” *Bull. Electr. Eng. Informatics*, vol. 12, no. 6, pp. 3798–3804, 2023, doi: 10.11591/eei.v12i6.5418.
- [13] M. I. Syafii, “Sentimen analisis Pada Media Sosial Twitter menggunakan Metode Naive Bayes Classifier (NBC),” *J. Teknol. Pint.*, vol. 2, no. 7, pp. 1–11, 2022, [Online]. Available: <http://teknologipintar.org/index.php/teknologipintar/article/view/207%0Ahttp://teknologipintar.org/index.php/teknologipintar/article/download/207/197>
- [14] S. Prusty, S. Patnaik, and S. K. Dash, “SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer,” *Front. Nanotechnol.*, vol. 4, no. August, pp. 1–12, 2022, doi: 10.3389/fnano.2022.972421.
- [15] G. Popoola, K. K. Abdullah, G. S. Fuhnwi, and J. Agbaje, “Sentiment Analysis of Financial News Data using TF-IDF and Machine Learning Algorithms,” *2024 IEEE 3rd Int. Conf. AI Cybersecurity, ICAIC 2024*, no. February, 2024, doi: 10.1109/ICAIC60265.2024.10433843.
- [16] A. Apriani, H. Zakiyudin, and K. Marzuki, “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta,” *J. Bumigora Inf. Technol.*, vol. 3, no. 1, pp. 19–27, Jul. 2021, doi: 10.30812/bite.v3i1.1110.
- [17] Y. A. Singgalen, “KLIK: Kajian Ilmiah Informatika dan Komputer Comparative Analysis of DT and SVM Model Performance with SMOTE in Sentiment Classification,” *Media Online*, vol. 4, no. 5, pp. 2485–2494, 2024, doi: 10.30865/klik.v4i5.1828.
- [18] Y. Liu *et al.*, “RoBERTa: A Robustly Optimized BERT Pretraining Approach,” no. 1, 2019, [Online]. Available: <http://arxiv.org/abs/1907.11692>
- [19] A. Deolika, K. Kusriani, and E. T. Luthfi, “Analisis Pembobotan Kata pada Klasifikasi Text Mining,” *J. Teknol. Inf.*, vol. 3, no. 2, p. 179, 2019, doi: 10.36294/jurti.v3i2.1077.
- [20] L. He, T. Yin, and K. Zheng, “They May Not Work! an evaluation of eleven sentiment analysis tools on seven social media datasets,” *J. Biomed. Inform.*, vol. 132, no. February, 2022, doi: 10.1016/j.jbi.2022.104142.
- [21] M. Iqbal, A. Karim, and F. Kamiran, “Balancing prediction errors for robust sentiment classification,” *ACM Trans. Knowl. Discov. Data*, vol. 13, no. 3, 2019, doi: 10.1145/3328795.