TECHNOVATAR Jurnal Teknologi, Industri dan Informasi
Vol. 2, No. 4, Oktober 2024, pp. 1 – 12
Open Access: https://journal.awatarapublisher.com/index.php/technovatar
ISSN: 3031-8874, DOI: 10.61434/technovatar.v2i4.385

ANALISIS SENTIMEN PENGGUNA X TERHADAP PUTUSAN MK PEMILU 2024 DENGAN NAIVE BAYES DAN GRADIENT BOOSTING

Indra Surya Permana[1], Christina Juliane[2]
[1],[2] STMIK LIKMI, Bandung, Indonesia
e-mail: indrajamz@gmail.com[1], christina.juliane@likmi.ac.id[2]

INFO ARTIKEL
Sejarah Artikel:
Diterima, September 2024
Direvisi, Oktober 2024
Disetujui, Oktober 2024
KEYWORDS
Sentiment analysis
Naive Bayes
Gradient Boosted Machines
X/Twitter
TF-IDF

ABSTRACT
This study analyzes Twitter reactions to the Constitutional Court's decision on Indonesia's 2024 election using Naive Bayes and Gradient Boosted Machines (GBM) with TF-IDF features. The evaluation indicates that Naive Bayes performs strongly for positive sentiment and relatively better for negative sentiment, while GBM provides stronger recall and F1-score for neutral sentiment. Based on the reported confusion matrices, Naive Bayes achieved 94.33% accuracy and 70.14% macro-F1, whereas GBM achieved 93.51% accuracy and 69.64% macro-F1. Because the dataset is strongly imbalanced and several data-collection, labeling, and validation parameters are not documented in the available working file, the findings should be interpreted within this experimental context. The results support model selection based on the sentiment class of interest rather than a single overall claim of superiority.

This work is licensed under a Creative Commons Attribution 4.0 International License.

KATA KUNCI
Analisis sentimen
Naive Bayes
Gradient Boosted Machines
X/Twitter
TF-IDF

CORRESPONDING AUTHOR
Indra Surya Permana
STMIK LIKMI
Bandung
indrajamz@gmail.com

ABSTRAK
Penelitian ini menganalisis reaksi pengguna Twitter terhadap putusan Mahkamah Konstitusi mengenai Pemilu Indonesia 2024 menggunakan Naive Bayes dan Gradient Boosted Machines (GBM) dengan fitur TF-IDF. Evaluasi menunjukkan bahwa Naive Bayes berkinerja kuat pada sentimen positif dan relatif lebih baik pada sentimen negatif, sedangkan GBM memberikan recall dan F1-score yang lebih baik pada sentimen netral. Berdasarkan confusion matrix, Naive Bayes mencapai accuracy 94,33% dan macro-F1 70,14%, sedangkan GBM mencapai accuracy 93,51% dan macro-F1 69,64%. Karena dataset sangat tidak seimbang dan beberapa parameter pengumpulan data, pelabelan, serta validasi tidak terdokumentasi pada berkas kerja, temuan perlu ditafsirkan dalam konteks eksperimen ini. Hasil mendukung pemilihan model berdasarkan kelas sentimen yang diprioritaskan, bukan klaim keunggulan tunggal secara keseluruhan.

PENDAHULUAN

Pemilihan Umum 2024 di Indonesia merupakan peristiwa penting yang menarik perhatian besar dari masyarakat dan menjadi sorotan utama. Dalam konteks ini, beberapa dinamika penting telah terjadi. Pertama, teknologi, khususnya media sosial, telah menjadi platform utama bagi calon presiden dan pendukungnya untuk menyampaikan visi, misi, dan program kerja (Afnira, 2023). Kedua, pendidikan politik dan pemahaman akan dampak keputusan politik menjadi semakin penting dalam konteks partisipasi pemilih. Ketiga, isu-isu global seperti perubahan iklim, keamanan energi, dan kesehatan global menjadi fokus utama dalam perdebatan kandidat (BBC News Indonesia, 2024). Keempat, Pemilu khususnya Pilpres 2024 mencerminkan semangat inklusivitas dengan banyaknya kandidat dari berbagai latar belakang dan kelompok masyarakat. Kelima, polarisasi politik, potensi penyalahgunaan wewenang, dan ancaman suburnya politik dinasti menjadi isu yang mewarnai

jalannya Pilpres 2024. Keenam, generasi muda berusia 22-30 tahun akan mendominasi pemilih secara nasional, dengan porsi 56%, atau sekitar 114 juta. Ketujuh, misinformasi dan disinformasi di media sosial masih menjadi isu utama menjelang pemilu 2024 (Hutabarat, n.d.), (Nur, n.d.).

Pemilu 2024 telah selesai pada saat hasil rekapitulasi diumumkan pada bulan Maret 2024, di mana pasangan Capres dan Cawapres pasangan nomor urut 2 mendapatkan 96.214.691 suara atau lebih dari 57 % dari total suara dan meraih kemenangan yang terpaut jauh dari pasangan nomor urut 1 dan 2. namun demikian hasil Pemilu 2024 di Indonesia ini telah menimbulkan berbagai kontroversi dan perdebatan, diantaranya adalah adanya 718 sengketa yang tercatat di Daftar Sengketa Pemilu Presiden & Wakil Presiden Ri 2024 (Komisi Pemilihan Umum Indonesia, 2024). Mahkamah Konstitusi (MK) memutuskan menolak seluruh gugatan hasil Pemilihan Presiden 2024. Meskipun demikian, banyak pihak yang merasa kecewa dan meragukan keputusan tersebut.

Saat ini platform X atau Twitter seringkali dianggap menjadi sumber informasi dalam analisis pendapat masyarakat (Dhiman & Toshniwal, 2022). Hal tersebut dikarenakan twitter menjadi tempat bagi masyarakat untuk berbagi pendapat, mengungkapkan kondisi emosi seperti kebahagiaan, kekecewaan, kemarahan, dan lain sebagainya, termasuk juga pandangan politik seseorang maupun kelompok (Devi & Chingangbam, 2021), (Vindua & Zailani, 2023). Dalam pelaporan, identitas akun pengguna perlu dilindungi dan kutipan yang berpotensi mengarah pada identifikasi individu tidak ditampilkan.

Pendapat pengguna media sosial X (Twitter) terkait hasil putusan Mahkamah Konstitusi (MK) terhadap Pemilu 2024 sangat beragam dan mencerminkan dinamika politik yang ada. Beberapa pengguna Twitter menunjukkan dukungan mereka terhadap keputusan MK, sementara yang lainnya mengekspresikan kekecewaan dan keraguan. Ada juga sejumlah pengguna yang mempertanyakan proses dan dasar hukum dari keputusan tersebut. Misalnya, ada yang mempertanyakan independensi lembaga penyelenggaraan Pemilu dan menyoroti adanya dugaan penyalahgunaan kekuasaan oleh pemerintah. Perilaku pengguna media sosial ini dibahas pada penelitian yang dilakukan oleh Septian,dkk (Septian et al., 2023) dan Vindua dan Zailani (Vindua & Zailani, 2023) yang membahas tentang sentimen pengguna twitter terhadap suatu putusan hukum di Indonesia yang berskala nasional.

Dalam era digital, data sentimen yang melimpah dari media sosial dan platform daring memberikan sumber informasi yang berharga untuk menganalisis pandangan publik untuk melihat sentimen atau opini yang terbentuk. Namun, kompleksitas dan volume data ini menuntut pendekatan yang canggih dalam analisisnya. Di sinilah model Naive Bayes dan Gradient Boosted Machines memainkan peran penting dalam opinion mining atau analisis sentimen. Naive Bayes, dengan keunggulannya dalam klasifikasi teks, telah menjadi salah satu pendekatan yang populer dalam analisis sentimen seperti penelitian sebelumnya yang dilakukan oleh Isnain, dkk (Isnain et al., 2021), Zahra (Zahra, 2023), dan Ansari (Ansari et al., 2018). Sementara model Gradient Boosted Machines (GBM) adalah metode yang kuat dalam machine learning yang dapat menghasilkan model prediksi yang akurat (Duan et al., 2023), (Nair et al., 2021). Dalam konteks analisis sentimen, gradient boosted machines dapat digunakan untuk memodelkan hubungan non-linear antara fitur-fitur dalam teks dengan sentimen yang ada(Dang et al., 2020), (Mehmood et al., 2019).

Dengan menggunakan model-model ini, opinion mining atau analisis sentimen terhadap putusan MK terkait Pemilu 2024 dapat dilakukan untuk mengidentifikasi pola dan tren opini masyarakat. Analisis sentimen tersebut dapat memberikan gambaran tentang bagaimana masyarakat merespons keputusan MK, apakah mayoritas mendukung atau mempertanyakan keputusan tersebut, serta faktor apa yang memengaruhi pandangan mereka.

Rumusan Masalah

Bagaimana Opinion Mining Pengguna X (Twitter) terhadap Putusan MK Tentang Pemilu 2024 dengan model Naive Bayes dan Gradient Boosted Machines ?

Tujuan Penelitian

Untuk mengetahui Opinion Mining Pengguna X (Twitter) terhadap Putusan MK Tentang Pemilu 2024 dengan model Naive Bayes dan Gradient Boosted Machines.

TINJAUAN PUSTAKA

Analisis Text (Text Analysis) dalam Pemrosesan bahasa alami /NLP (Natural Language Processing) adalah salah satu metode dalam pembelajaran mesin (machine learning) yang digunakan untuk memahami dan mengekstrak informasi dari teks (Gogia, 2021), (Kuckartz, 2014). Metode ini melibatkan pemrosesan dan pembacaan teks dengan menggunakan algoritma komputer untuk mengenali pola dan mencari makna dari teks tersebut. Metode analisis teks ini bisa mencakup identifikasi entitas seperti nama orang atau tempat, klasifikasi teks ke dalam kategori tertentu, ekstraksi informasi, analisis sentimen, dan lainnya.

Opion mining atau sering disebut Analisis sentimen adalah metode yang digunakan dalam analisis teks untuk mengidentifikasi, mengekstrak, dan menganalisis sentimen atau opini . Opion mining penting karena dengan meningkatnya kompleksitas dan volume data yang dihasilkan oleh media sosial dan platform daring, maka perlu untuk memahami pendapat publik dengan cepat dan akurat dengan bantuan pembelajaran mesin (Mohan Kumar & Nandkumar, 2020), (Osmani & Mohasefi, 2023).

Beberapa metode dalam machine learning ada yang masuk ke dalam metode sederhana dan juga deep learning atau pembelajaran mendalam. Salah satu metode dalam opinion mining yang sering dilakukan adalah dengan menggunakan metode Naïve Bayes Classifier (NBC). Metode NBC sendiri adalah metode yang sederhana dan efisien yang dapat digunakan untuk mengklasifikasikan teks ke dalam kategori sentimen yang berbeda (Isnain et al., 2021), (Foster & Wagner, 2021). Dalam metode Naive Bayes, probabilistik digunakan untuk mengklasifikasikan teks ke dalam kategori sentimenberdasarkan kemungkinan terjadinya suatu sentimen berdasarkan fitur-fitur linguistik yang ada dalam teks tersebut.

Gradient Boosted Machines (GBM) adalah teknik pembelajaran mesin yang menggunakan serangkaian model prediktif yang lemah (biasanya pohon keputusan) untuk membangun model prediktif yang kuat secara bertahap (Alzamzami et al., 2020). Algoritma ini meningkatkan prediksi dengan secara iteratif menyesuaikan model baru yang berfokus pada kasus-kasus yang salah diprediksi oleh model sebelumnya dalam seri. Model Gradient Boosted Machines (GBM) adalah teknik pembelajaran mesin yang canggih untuk regresi dan klasifikasi. Ini termasuk dalam kategori algoritma 'ensemble learning', di mana beberapa model prediktif dikombinasikan untuk meningkatkan keakuratan dan kekuatan prediksi (Khalid et al., 2020) .

Dengan menggunakan model-model analisis sentimen seperti Naïve Bayes dan Gradient Boosted Machines, diharapkan dapat menghasilkan opinion mining atau analisis sentimen yang lebih akurat terkait pendapat pengguna platform X atau Twitter terhadap keputusan Mahkamah Konstitusi terkait Pemilu 2024.

Berikut ini adalah review penelitian terdahulu terkait penelitian ini, khususnya penelitian yang menggunakan pendekatan NBC dan GBM.

Balshettar, Avilash, dan Jarmisa meneliti deteksi berita palsu di media sosial melalui analisis sentimen dengan pendekatan lexicon-based, Naive Bayes, passive-aggressive, dan Deep Neural Network (DNN) (Balshetwar et al., 2023). Penelitian tersebut mengintegrasikan analisis sentimen dan beberapa strategi imputasi untuk data yang hilang serta menggunakan unstructured K-fold untuk mengurangi overfitting. Metode yang diusulkan mencapai akurasi 99,8%, tetapi teknik analisisnya memiliki tingkat kerumitan tinggi.

Huang, Wu, Wu, dan Wang mengembangkan algoritma analisis sentimen untuk rekomendasi point of interest dengan contrastive learning dan adversarial training (Huang et al., 2023). Pendekatan ini ditujukan untuk mengatasi kesulitan anotasi, keterbatasan akurasi anotasi

manual, dan ekstraksi informasi dari beragam data. Model tersebut meningkatkan F1-score sebesar 12% dibandingkan metode umum, meskipun implementasinya pada kondisi dunia nyata masih menjadi tantangan.

Lu, Sun, dan Panjang menyajikan tinjauan komprehensif mengenai kemajuan dan tantangan terbuka dalam analisis sentimen (Lu et al., 2024). Kajian tersebut menawarkan kerangka kerja berdasarkan modalitas dan taksonomi untuk metode analisis sentimen multimodal, termasuk potensi penggunaan Generative Pre-trained Transformer untuk meningkatkan model pralatih. Kajian ini memberikan wawasan luas, tetapi tidak melaporkan tingkat akurasi dan pendekatan tinjauannya dinilai belum sistematis.

Singh et al. menerapkan Gradient Boosting untuk analisis sentimen dalam rekomendasi pekerjaan dan profil kandidat menggunakan data Twitter (Singh et al., 2022). Gradient Boosting Classifier menghasilkan akurasi 78,08% dan skor AUC 0,819. Namun, penggunaan satu model Gradient Boosting saja tetap menyisakan risiko overfitting yang perlu dikendalikan dalam proses pembelajaran.

Osmani dan Mohasefi mengusulkan Enriched Joint Sentiment-Topic Model untuk meningkatkan akurasi opinion mining (Osmani & Mohasefi, 2023). Model tersebut mencapai akurasi klasifikasi sentimen hingga 95%. Kelebihannya terletak pada parameter yang mampu meningkatkan akurasi, sedangkan keterbatasannya adalah parameter emosi yang belum memadai sehingga faktor tambahan masih diperlukan untuk meningkatkan kinerja.

Chen, Peng, dan Zihua mengembangkan CharAs-CBert untuk memperbaiki representasi kalimat pada klasifikasi sentimen (Chen et al., 2022). Metode ini menggabungkan karakter, kata, dan vektor konstruksi untuk mengatasi hilangnya konteks dan kesulitan membangun dependensi jangka panjang. Hasilnya mencapai F1-score dan akurasi 87,54%, serta 92,88% pada ACL14, tetapi kompleksitas, redundansi data, dan kegagalan dalam mengestimasi opini pada kondisi tertentu masih menjadi keterbatasan.

Adamu, Muhammad, dan Jingi membahas penerapan algoritma Gradient Boosting dalam pemodelan statistik (Adamu et al., 2019). Jumlah iterasi optimal ditentukan menggunakan Akaike Information Criterion dan cross-validation. Cross-validation menurunkan jumlah iterasi pendorong dari 111 menjadi 71 dengan hasil yang lebih baik, sementara pengendalian iterasi diperlukan untuk mengurangi risiko overfitting pada algoritma boosting.

Karimi, Rossi, dan Prati menerapkan adversarial training pada BERT untuk aspect-based sentiment analysis (Karimi et al., 2021). Model BERT Adversarial Training yang dipadukan dengan Post-Trained BERT menghasilkan F1-score hingga 84,8% dan mengungguli beberapa model BERT pembanding. Walaupun demikian, pendekatan tersebut masih perlu diuji pada beragam konteks analisis sentimen untuk menilai konsistensi ketepatannya.

Duan, Vasarhelyi, Codesso, dan Alzamil menggabungkan informasi media sosial dengan sistem informasi akuntansi pemerintah melalui text mining dan machine learning (Duan et al., 2023). Penelitian tersebut menggunakan Naive Bayes, Random Forest, dan XGBoost untuk menghubungkan data Twitter dan Facebook dengan dataset pemerintah. Pendekatan multimodel ini menawarkan inovasi dalam pengambilan keputusan, tetapi tidak menyajikan angka akurasi secara eksplisit.

Nicolas, Desroziers, Caccioli, dan Aste mengkaji risiko reputasi Environmental, Social, and Governance berdasarkan sekitar 114 juta tweet mengenai perusahaan S&P 100 sepanjang 2016–2022 (Nicolas et al., 2024). Analisis berbasis leksikon tersebut menggambarkan sentimen masyarakat dan investor terhadap isu ESG dengan cakupan data yang besar. Keterbatasannya adalah analisis masih berfokus pada text mining dan belum menggunakan model deep learning.

METODE PENELITIAN

Penelitian ini melakukan opinion mining menggunakan Naive Bayes Classifier (NBC) dan Gradient Boosted Machines (GBM). NBC merupakan metode probabilistik yang efisien untuk klasifikasi teks, sedangkan GBM membangun model prediktif secara bertahap melalui teknik boosting. Kedua model diterapkan pada representasi fitur TF-IDF dan dibandingkan menggunakan confusion matrix, precision, recall, dan F1-score.

Berikut ini adalah tahapan yang akan dilakukan dalam analisis sentimen pada penelitian ini

Pengumpulan Data

Proses mengumpulkan data yang relevan dengan penelitian. Dalam konteks ini, ini melibatkan pengumpulan tweets atau teks dari platform X yang terkait dengan topik penelitian.

Dataset

Dataset awal terdiri atas 9.952 komentar pada platform X (Twitter) yang dikumpulkan menggunakan tagar terkait keputusan Mahkamah Konstitusi, dengan cakupan lima hari menjelang hingga hari putusan final mengenai perselisihan hasil Pemilu 2024. Berkas kerja yang tersedia tidak mencatat tanggal kalender, daftar tagar/kueri, metode akses data, kriteria inklusi–eksklusi, prosedur deduplikasi, atau identifikasi akun otomatis secara rinci. Karena itu, unsur-unsur tersebut diperlakukan sebagai keterbatasan dokumentasi dan tidak direkonstruksi dengan asumsi baru.

Pra-processing Data

Tahap mempersiapkan teks mentah untuk analisis dengan membersihkannya dari unsur yang tidak perlu atau tidak memberikan informasi penting. Ini meliputi:

1. Case Folding: Mengkonversi semua teks ke huruf kecil untuk standarisasi.
2. Filtering: Menghapus karakter tidak penting seperti angka dan simbol.
3. Stopword Removal: Menghilangkan kata-kata umum (seperti “dan”, “di”, “yang”) yang tidak memiliki nilai analitis.
4. Tokenizing: Memecah teks menjadi token atau kata-kata individu.
5. Stemming: Mengurangi kata ke bentuk dasar mereka.

Labelling

Pelabelan membagi teks ke dalam kategori positif, negatif, dan netral. Berkas kerja tidak mendokumentasikan secara memadai apakah seluruh label ditetapkan secara manual atau dengan bantuan algoritma, jumlah anotator, pedoman anotasi, penyelesaian perbedaan, maupun reliabilitas antar-anotator. Kekurangan dokumentasi ini membatasi reproduktibilitas dan menjadi salah satu keterbatasan validitas penelitian.

TF-IDF

TF-IDF atau Term Frequency-Inverse Document Frequency, sebuah metode untuk mengubah teks menjadi vektor angka yang menggambarkan pentingnya kata dalam dokumen yang berkaitan dengan corpus (kumpulan dokumen).

Pelatihan Model Naive Bayes

Proses pelatihan algoritma klasifikasi Naive Bayes dengan dataset yang telah dilabeli. Model ini menggunakan probabilitas dan statistik untuk memprediksi label sentimen dari teks, dilatih dengan dataset yang telah dilengkapi dengan label sentimen pada langkah sebelumnya

Pelatihan Model GBM

Proses melatih model Gradient Boosted Machines, yang merupakan algoritma klasifikasi lanjutan yang menggunakan teknik boosting untuk menghasilkan model prediktif. Penjelasan: GBM membangun serangkaian model prediktif yang lemah secara bertahap, dengan setiap model baru berusaha memperbaiki kesalahan dari model sebelumnya.

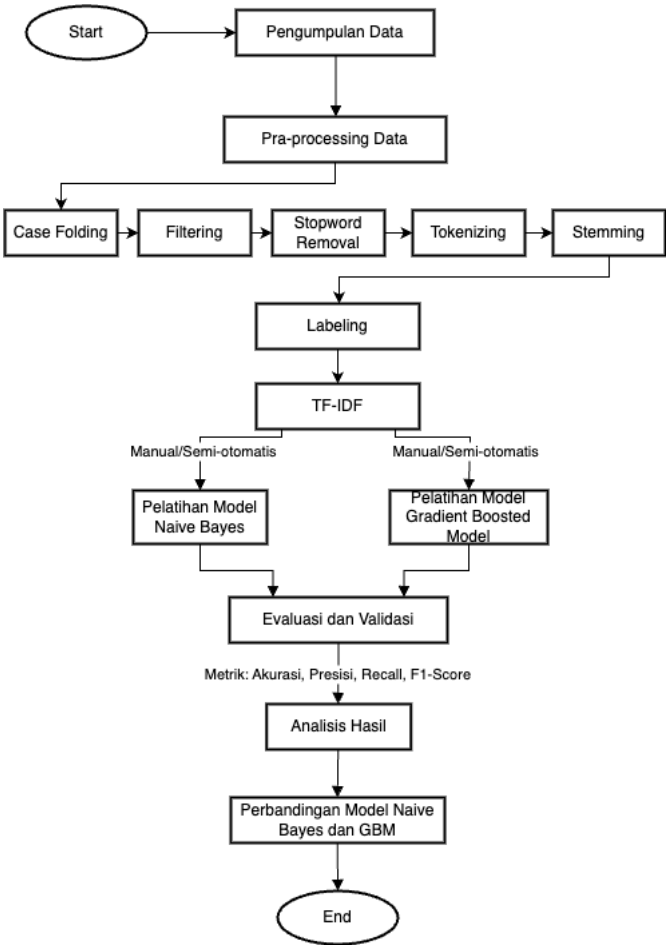
Evaluasi Model

Evaluasi dilakukan menggunakan cross-validation dan confusion matrix dengan metrik accuracy, precision, recall, dan F1-score. Namun, berkas kerja tidak mencatat jumlah fold, skema

stratifikasi, random seed, konfigurasi hyperparameter, maupun urutan pemodelan TF-IDF pada setiap fold. Oleh karena itu, hasil dapat digunakan untuk membandingkan performa pada eksperimen ini, tetapi reproduksi persis membutuhkan catatan konfigurasi tambahan dan penerapan TF-IDF hanya pada data latih setiap fold untuk mencegah data leakage.

Analisis Hasil

Proses ini melakukan pemeriksaan hasil klasifikasi, mencari keterkaitan, hubungan dan pola dalam data yang diprediksi dari model yang dibuat



Gambar 1 Diagram Alur Penelitian

Tools Analisis

Penelitian ini menggunakan RapidMiner sebagai alat utama untuk analisis data. RapidMiner adalah platform analisis data yang canggih dan user-friendly, yang memungkinkan pengguna untuk melakukan berbagai proses analitik tanpa perlu pengetahuan pemrograman yang mendalam (Makhi & Fauzan, 2022). Hal ini sudah termasuk preprocessing data, pemodelan, hingga evaluasi model.

DISKUSI DAN PEMBAHASAN

Tahap pra-pemrosesan data, dimulai dengan membaca file CSV yang mengandung 9,952 entry yang merupakan komentar dari Twitter. Langkah ini penting untuk memastikan bahwa semua data yang diperlukan terakses dan siap untuk proses analisis selanjutnya. Setelah itu, dilanjutkan dengan tahap pelabelan sentimen di mana masing-masing komentar dikategorikan menjadi salah satu dari tiga label sentimen: positif, negatif, atau netral. Proses ini sangat penting untuk opinion mining atau analisis sentimen lebih lanjut, memungkinkan untuk klasifikasi yang berbasis data aktual dan situasi yang sedang berlangsung. Confusion matrix pada tahap evaluasi masing-masing

memuat 1.464 observasi, sedangkan dataset awal berjumlah 9.952 entri. Catatan rinci mengenai jumlah data yang tersisih pada setiap tahap pembersihan, validasi, atau pemilihan sampel tidak tersedia dalam berkas kerja sehingga selisih tersebut dilaporkan sebagai keterbatasan dokumentasi.

Proses selanjutnya yang dilakukan adalah proses konversi dari data nominal ke teks untuk memfasilitasi tokenisasi. Tokenisasi berhasil memecah teks menjadi unit yang lebih kecil, atau token, yang memungkinkan analisis lebih rinci terhadap komponen individual dari teks tersebut. Transformasi kasus kemudian dilakukan, mengkonversi semua teks ke bentuk huruf kecil untuk menghindari duplikasi token yang tidak perlu berdasarkan perbedaan kapitalisasi.

Langkah berikutnya adalah penghapusan stopword menggunakan kumpulan stopword bahasa Indonesia yang diunduh dari situs Kaggle, yang mengeliminasi kata-kata yang tidak membawa beban sentimen signifikan dan membantu menyederhanakan dataset. Terakhir, filter tokens diterapkan berdasarkan panjang untuk menyingkirkan token yang terlalu pendek atau terlalu panjang yang mungkin tidak relevan untuk analisis sentimen. Proses pra-pemrosesan ini menghasilkan dataset yang bersih dan terstruktur yang siap untuk tahap pelatihan model menggunakan Naive Bayes dan Gradient Boosted Machines.

Pelatihan Model

Pelatihan model yang dilakukan mencakup pelatihan model Naive Bayes dan Gradient Boosted Machines. Dataset yang digunakan untuk pelatihan model Naive Bayes telah diproses menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency), yang umumnya digunakan untuk mengkonversi teks menjadi format yang bisa dipahami oleh model machine learning. Dataset ini terdiri dari 8,477 contoh dengan 11,510 atribut reguler.

Evaluasi

Analisis Evaluasi model menggunakan metode cross validation. Ini adalah metode validasi yang sangat umum dan robust, di mana data dibagi menjadi beberapa "folds" atau bagian. Dalam setiap iterasi, satu fold digunakan sebagai data tes dan sisanya sebagai data pelatihan. Berikut ini adalah hasil cross validation untuk masing-masing model

Naïve Bayes:

Tabel 1 Evaluasi Model Naive Bayes

	true positif	true negatif	true netral	class precision
pred. positif	1310	13	12	98.13%
pred. negatif	19	32	12	50.79%
pred. netral	12	15	39	59.09%
class recall	97.69%	53.33%	61.90%	

Hasil dari cross validation untuk model Naïve Bayes menunjukkan performa klasifikasi yang cukup baik pada kelas 'positif' dengan presisi 98.13% dan recall 97.69%, yang menandakan bahwa model sangat efektif dalam mengidentifikasi dan memprediksi sentimen positif dengan tingkat kesalahan yang sangat rendah. Untuk kelas 'negatif' dan 'netral', presisi dan recall menunjukkan performa yang lebih rendah. Presisi untuk 'negatif' adalah 50.79% dan untuk 'netral' 59.09%, sedangkan recall untuk kelas tersebut adalah 53.33% dan 61.90% berturut-turut. Presisi yang lebih rendah di kedua kelas ini menunjukkan bahwa model tersebut memiliki kesulitan dalam secara konsisten mengidentifikasi komentar negatif dan netral, sering kali mengklasifikasikan komentar tersebut secara tidak tepat. Recall yang moderat mengindikasikan bahwa sementara model bisa mengidentifikasi sebagian besar komentar negatif dan netral, masih ada jumlah signifikan yang terlewatkan. Distribusi kelas aktual pada confusion matrix sangat timpang, yaitu 1.341 positif, 60 negatif, dan 63 netral. Karena dominasi kelas positif, interpretasi tidak hanya didasarkan pada accuracy, tetapi juga pada macro-F1 dan kinerja setiap kelas.

Gradient Boosted Machines

Tabel 2 Evaluasi Model GBM

	true positif	true negatif	true netral	class precision
pred. positif	1298	29	11	97.01%

pred. negatif	30	23	4	40.35%
pred. netral	13	8	48	69.57%
class recall	96.79%	38.33%	76.19%	

Hasil dari cross validation untuk model Gradient Boosted Machines (GBM) menunjukkan performa klasifikasi yang sangat baik pada kelas 'positif' dengan presisi 97.01% dan recall 96.79%. Performa ini mengindikasikan bahwa model sangat efektif dalam mengidentifikasi dan memprediksi sentimen positif dengan tingkat kesalahan yang minimal. Namun, untuk kelas 'negatif' dan 'netral', hasilnya menunjukkan tingkat keefektifan yang lebih rendah. Presisi untuk kelas 'negatif' adalah 40.35% dan untuk 'netral' adalah 69.57%, sementara recall untuk kelas-kelas tersebut adalah 38.33% dan 76.19%, berturut-turut.

Presisi yang relatif rendah untuk kelas 'negatif' mengindikasikan bahwa model seringkali salah dalam mengidentifikasi komentar negatif, cenderung mengklasifikasikan komentar yang sebenarnya positif atau netral sebagai negatif. Hal ini juga diikuti dengan recall yang rendah, menunjukkan bahwa model ini gagal menangkap sejumlah besar komentar negatif yang sebenarnya. Dalam kelas 'netral', walaupun recall relatif lebih tinggi menunjukkan bahwa model ini lebih baik dalam menangkap komentar netral, presisi yang lebih rendah dari 70% tetap menandakan adanya kesalahan dalam mengklasifikasikan komentar dari kategori lain sebagai netral.

Secara keseluruhan, GBM mempertahankan performa tinggi pada kelas positif dan memberikan recall serta F1-score yang lebih baik untuk kelas netral. Namun, precision dan recall kelas negatif pada GBM lebih rendah daripada Naive Bayes. Dengan demikian, GBM tidak dapat dinyatakan lebih seimbang untuk seluruh kelas; keunggulannya pada eksperimen ini terutama terletak pada pengenalan sentimen netral, sedangkan Naive Bayes relatif lebih baik untuk sentimen negatif dan sedikit lebih baik pada kelas positif.

Nilai F1 Score

Untuk menambahkan skor F1 ke dalam tabel perbandingan untuk model Naive Bayes dan Gradient Boosted Machines (GBM), dapat dihitung melalui rumus sebagai berikut:

$$F1\ Score = 2x\left(\frac{Precission\ x\ Recall}{Precission\ x\ Recall}\right)$$

Hasil perhitungan F1 score dari kedua model adalah sebagai berikut:

Berdasarkan confusion matrix, Naive Bayes menghasilkan accuracy 94,33% dan macro-F1 70,14%, sedangkan GBM menghasilkan accuracy 93,51% dan macro-F1 69,64%. Nilai accuracy yang tinggi perlu dibaca bersama macro-F1 karena distribusi kelas didominasi sentimen positif.

Tabel 3 Nilai F1 Score dari Model Prediksi

Kategori Sentimen	Metrik	Naive Bayes	Gradient Boosted Machines
Positif	F1 Score	97.91%	96.90%
Negatif	F1 Score	52.03%	39.31%
Netral	F1 Score	60.47%	72.71%

PEMBAHASAN

1. Evaluasi Performa Model

Model Naive Bayes dan GBM menunjukkan keefektivitasan yang berbeda dalam mengklasifikasikan sentimen dari tweet. Secara khusus, model Naive Bayes menunjukkan performa yang sangat baik dalam mengidentifikasi sentimen positif dengan presisi dan recall di atas 95%. Hal tersebut mengindikasikan kekuatan model dalam mengenali pola kata yang secara konsisten terkait dengan sentimen positif. Namun, model ini kurang efektif dalam mengklasifikasikan sentimen negatif dan netral, dengan presisi dan recall yang jauh lebih rendah. Sebaliknya, GBM, meskipun menunjukkan penurunan presisi dalam klasifikasi sentimen positif dibandingkan Naive Bayes, memberikan hasil yang lebih seimbang di semua kategori sentimen. Terutama, performa model ini dalam mengidentifikasi sentimen netral lebih baik, yang menunjukkan kemampuannya dalam menangani kategori yang lebih ambigu dan kurang jelas.

2. Analisis Metodologi

Penggunaan metode TF-IDF untuk preprocessing data telah terbukti efektif dalam menyiapkan teks untuk analisis. Namun, jumlah atribut yang sangat besar (11,510 atribut) mungkin telah menyebabkan beberapa masalah dalam generalisasi model, khususnya bagi Naive Bayes. Kedepannya, teknik reduksi dimensi atau seleksi fitur mungkin dapat membantu meningkatkan efektivitas model tanpa mengorbankan informasi penting.

3. Pengaruh dan Implikasi Hasil Penelitian

Temuan penelitian menunjukkan bahwa pemilihan model perlu disesuaikan dengan kelas sentimen yang menjadi prioritas. Naive Bayes unggul pada sentimen positif dan relatif lebih baik pada sentimen negatif, sedangkan GBM lebih kuat dalam mendeteksi sentimen netral. Hasil ini bersifat kontekstual terhadap dataset dan prosedur yang digunakan sehingga tidak boleh digeneralisasi tanpa mempertimbangkan ketimpangan kelas dan keterbatasan dokumentasi eksperimen.

Keterbatasan Penelitian

Penelitian ini dibatasi oleh tidak lengkapnya catatan tanggal dan kueri pengumpulan data, kriteria seleksi, alur penyusutan dari 9.952 entri menjadi 1.464 observasi evaluasi, prosedur pelabelan, konfigurasi cross-validation, random seed, dan hyperparameter. Dataset juga sangat tidak seimbang serta hanya mencakup satu platform dan satu peristiwa politik, sehingga generalisasi ke periode, isu, dan populasi lain harus dilakukan dengan hati-hati. Data media sosial perlu dilaporkan dengan melindungi identitas pengguna dan tidak menampilkan nama akun atau kutipan yang memungkinkan identifikasi; berkas yang tersedia tidak memuat nomor persetujuan etik sehingga naskah tidak membuat klaim mengenai persetujuan yang tidak terdokumentasi.

REKOMENDASI

Berdasarkan tantangan yang dihadapi, penelitian berikutnya dapat fokus pada eksplorasi metode advanced preprocessing lainnya, seperti penggunaan word embeddings yang mungkin menangkap konteks kata dengan lebih baik dibandingkan TF-IDF. Selain itu, eksplorasi lebih lanjut terhadap teknik balanced training dataset atau penyesuaian parameter model dapat dilakukan untuk meningkatkan akurasi klasifikasi pada sentimen yang kurang representatif. Model deep mining juga dapat dijadikan rekomendasi pada penelitian mendatang untuk meningkatkan akurasi dalam opinion mining atau analisis sentimen

KESIMPULAN

Penelitian ini menganalisis opini pengguna Twitter mengenai Putusan Mahkamah Konstitusi tentang Pemilu 2024 menggunakan Naive Bayes dan Gradient Boosted Machines. Naive Bayes mencapai accuracy 94,33% dan macro-F1 70,14%, sedangkan GBM mencapai accuracy 93,51% dan macro-F1 69,64%. Naive Bayes lebih kuat pada sentimen positif dan relatif lebih baik pada sentimen negatif, sementara GBM unggul dalam mendeteksi sentimen netral. Hasil tersebut perlu dibaca dengan mempertimbangkan ketimpangan kelas dan keterbatasan dokumentasi pengumpulan, pelabelan, serta konfigurasi evaluasi. Dengan demikian, pemilihan model sebaiknya disesuaikan dengan kelas sentimen yang menjadi prioritas dan divalidasi kembali pada dataset yang lebih seimbang serta terdokumentasi lengkap.

KONTRIBUSI PENULIS

Seluruh penulis berkontribusi sesuai perannya dalam konseptualisasi, metodologi, pelaksanaan penelitian, analisis data, penulisan draf awal, serta penelaahan dan penyuntingan naskah. Semua penulis telah membaca dan menyetujui versi akhir naskah ini.

PERNYATAAN KETERSEDIAAN DATA

Data yang mendukung temuan dalam penelitian ini tersedia dari penulis atas permintaan yang wajar, dengan mempertimbangkan persetujuan etik dan peraturan institusi yang berlaku.

PERNYATAAN BEBAS KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak terdapat konflik kepentingan finansial maupun hubungan pribadi yang dapat memengaruhi hasil penelitian yang dilaporkan dalam naskah ini.

PERNYATAAN ETIKA PENELITIAN DAN PUBLIKASI

Para penulis menyatakan bahwa penelitian dan penulisan naskah ini telah dilaksanakan sesuai dengan prinsip etika penelitian dan publikasi yang berlaku, serta bebas dari plagiasi, fabrikasi, falsifikasi, dan publikasi ganda.

PERNYATAAN PEMANFAATAN TEKNOLOGI ASISTIF

Dalam proses penyiapan naskah, teknologi kecerdasan buatan generatif dan/atau teknologi asistif digunakan secara terbatas untuk membantu penyuntingan bahasa, konsistensi istilah, dan penataan naskah. Seluruh keluaran telah ditinjau dan disunting oleh para penulis. Para penulis bertanggung jawab penuh atas keakuratan, integritas, dan versi akhir naskah.

REFERENSI

E. Afnira, "Optimalisasi Media Sosial sebagai Sarana Publikasi Pengawasan Pemilu 2024: Kasus Bawaslu Kota Tanjungpinang," J. Mhs. Komun. Cantrik, vol. 3, no. 1, Art. no. 1, May 2023, doi: 10.20885/cantrik.vol3.iss1.art4.

BBC News Indonesia, "Putusan MK: Reaksi setelah MK tolak gugatan hasil Pilpres 2024, dari Prabowo hingga pengunjuk rasa," BBC News Indonesia. Accessed: Apr. 23, 2024. [Online]. Available: https://www.bbc.com/indonesia/articles/ce437lrlzwlo

J. Hutabarat, "Menjinakkan Polarisasi dan Politik Identitas Jelang Pemilu 2024," tirto.id. Accessed: Apr. 23, 2024. [Online]. Available: https://tirto.id/menjinakkan-polarisasi-dan-politik-identitas-jelang-pemilu-2024-gtQ8

M. F. Nur, "Potensi Polarisasi & Mencegah Konflik Pendukung di Pemilu 2024," tirto.id. Accessed: Apr. 23, 2024. [Online]. Available: https://tirto.id/potensi-polarisasi-mencegah-konflik-pendukung-di-pemilu-2024-gRPS

Komisi Pemilihan Umum Indonesia, "Daftar Sengketa Pemilu Presiden & Wakil Presiden Ri 2024 : Daftar Daerah Yang Mengalami Pemungutan Dan Penghitungan Suara Ulang/Susulan/Lanjutan (Baik Seluruhnya Maupun Sebagian)," Komisi Pemilihan Umum. Accessed: Apr. 23, 2024. [Online]. Available: https://pemilu2024.kpu.go.id/pilpres/daftar-sengketa

A. Dhiman and D. Toshniwal, "AI-based Twitter framework for assessing the involvement of government schemes in electoral campaigns," Expert Systems with Applications, vol. 203, p. 117338, Oct. 2022, doi: 10.1016/j.eswa.2022.117338.

W. R. Devi and C. Chingangbam, "Sentiment Analysis for Electoral Prediction Using Twitter Data," in Emerging Technologies in Data Mining and Information Security, vol. 1300, A. E. Hassanien, S. Bhattacharyya, S. Chakrabati, A. Bhattacharya, and S. Dutta, Eds., in Advances in Intelligent Systems and Computing, vol. 1300. , Singapore: Springer Nature Singapore, 2021, pp. 25–33. doi: 10.1007/978-981-33-4367-2_3.

R. Vindua and A. U. Zailani, "Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python," Jur. Ris. Kom., vol. 10, no. 2, p. 479, Apr. 2023, doi: 10.30865/jurikom.v10i2.5945.

L. Septian, T. Aljauza, and C. Juliane, "Analisis Sentimen Putusan Mahkamah Konstitusi terhadap Batas Usia Capres dan Cawapres Menggunakan IndoBERT," Indonesian Journal of Computer Science, vol. 12, no. 6, Art. no. 6, 2023, doi: 10.33022/ijcs.v12i6.3614.

A. R. Isnain, N. S. Marga, and D. Alita, "Sentiment Analysis Of Government Policy On Corona Case Using Naive Bayes Algorithm," Indonesian J. Comput. Cybern. Syst., vol. 15, no. 1, p. 55, Jan. 2021, doi: 10.22146/ijccs.60718.

D. F. Zahra, “Analisis Sentimen penggunaan Mypertamina untuk Pembelian BBM Bersubsidi mengggunakan Algoritma Naive Bayes,” Explore. jurnal. sistem. inf. dan. telematika, vol. 14, no. 1, p. 94, Jul. 2023, doi: 10.36448/jsit.v14i1.3098.

S. Ansari, K. Hans, and S. K. Khatri, “A Naive Bayes classifier approach for detecting hypervisor attacks in virtual machines,” 2nd International Conference on Telecommunication and Networks, TEL-NET 2017, vol. 2018-January, pp. 1–6, Apr. 2018, doi: 10.1109/TEL-NET.2017.8343551.

H. K. Duan, M. A. Vasarhelyi, M. Codesso, and Z. Alzamil, “Enhancing the government accounting information systems using social media information: An application of text mining and machine learning,” International Journal of Accounting Information Systems, vol. 48, p. 100600, Mar. 2023, doi: 10.1016/j.accinf.2022.100600.

A. Nair, C. Paralkar, J. Pandya, Y. Chopra, and D. Krishnan, “Comparative Review on Sentiment analysis-based Recommendation system,” in 2021 6th International Conference for Convergence in Technology (I2CT), Apr. 2021, pp. 1–6. doi: 10.1109/I2CT51068.2021.9418222.

N. C. Dang, M. N. Moreno-García, and F. De la Prieta, “Sentiment Analysis Based on Deep Learning: A Comparative Study,” Electronics, vol. 9, no. 3, Art. no. 3, Mar. 2020, doi: 10.3390/electronics9030483.

K. Mehmood, D. Essam, K. Shafi, and M. K. Malik, “Sentiment Analysis for a Resource Poor Language—Roman Urdu,” ACM Trans. Asian Low-Resour. Lang. Inf. Process., vol. 19, no. 1, p. 10:1-10:15, Aug. 2019, doi: 10.1145/3329709.

E. Gogia, “A Review on Sentiment Analysis techniques and applications.,” GRET, pp. 24–29, Dec. 2021, doi: 10.47893/GRET.2021.1041.

U. Kuckartz, Qualitative text analysis: a guide to methods, practice & using software. Los Angeles London New Delhi Singapore Washington, DC: Sage, 2014.

A. V. Mohan Kumar and A. N. Nandkumar, “A Survey on Challenges and Research Opportunities in Opinion Mining,” SN COMPUT. SCI., vol. 1, no. 3, p. 171, May 2020, doi: 10.1007/s42979-020-00149-4.

A. Osmani and J. B. Mohasefi, “Opinion Mining Using Enriched Joint Sentiment-Topic Model,” Int. J. Info. Tech. Dec. Mak., vol. 22, no. 01, pp. 313–375, Jan. 2023, doi: 10.1142/S0219622022500584.

J. Foster and J. Wagner, “Naive Bayes versus BERT: Jupyter notebook assignments for an introductory NLP course,” in Proceedings of the Fifth Workshop on Teaching NLP, Online: Association for Computational Linguistics, 2021, pp. 112–114. doi: 10.18653/v1/2021.teachingnlp-1.20.

F. Alzamzami, M. Hoda, and A. E. Saddik, “Light Gradient Boosting Machine for General Sentiment Classification on Short Texts: A Comparative Evaluation,” IEEE Access, vol. 8, pp. 101840–101858, 2020, doi: 10.1109/ACCESS.2020.2997330.

M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad, and G. S. Choi, “GBSVM: Sentiment Classification from Unstructured Reviews Using Ensemble Classifier,” Applied Sciences, vol. 10, no. 8, p. 2788, Apr. 2020, doi: 10.3390/app10082788.

S. V. Balshetwar, A. Rs, and D. J. R, “Fake news detection in social media based on sentiment analysis using classifier techniques,” Multimed Tools Appl, vol. 82, no. 23, pp. 35781–35811, Sep. 2023, doi: 10.1007/s11042-023-14883-3.

S. Huang, X. Wu, X. Wu, and K. Wang, “Sentiment analysis algorithm using contrastive learning and adversarial training for POI recommendation,” Soc. Netw. Anal. Min., vol. 13, no. 1, p. 75, Apr. 2023, doi: 10.1007/s13278-023-01076-x.

Q. Lu, X. Sun, Y. Long, Z. Gao, J. Feng, and T. Sun, “Sentiment Analysis: Comprehensive Reviews, Recent Advances, and Open Challenges,” IEEE Trans. Neural Netw. Learning Syst., pp. 1–21, 2024, doi: 10.1109/TNNLS.2023.3294810.

S. Singh, D. Krishnan, P. Sehgal, H. Sharma, T. Surani, and J. Singh, “Gradient Boosting Approach for Sentiment Analysis for Job Recommendation and Candidate Profiling,” in 2022 IEEE Bombay Section Signature Conference (IBSSC), Dec. 2022, pp. 1–6. doi: 10.1109/IBSSC56953.2022.10037443.

B. Chen, W. Peng, and J. Song, “CharAs-CBert: Character Assist Construction-Bert Sentence Representation Improving Sentiment Classification,” Sensors, vol. 22, no. 13, p. 5024, Jul. 2022, doi: 10.3390/s22135024.

H. A. Adamu, M. Muhammad, and A. M. Jingi, “Application of Gradient Boosting Algorithm In Statistical Modelling,” 2019. Accessed: Apr. 26, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Application-of-Gradient-Boosting-Algorithm-In-Adamu-Muhammad/5dbf9b52f1afdaba9f6c858c02b84227193f4310

A. Karimi, L. Rossi, and A. Prati, “Adversarial Training for Aspect-Based Sentiment Analysis with BERT,” in 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy: IEEE, Jan. 2021, pp. 8797–8803. doi: 10.1109/ICPR48806.2021.9412167.

M. L. D. Nicolas, A. Desroziers, F. Caccioli, and T. Aste, “ESG reputation risk matters: An event study based on social media data,” Finance Research Letters, vol. 59, p. 104712, Jan. 2024, doi: 10.1016/j.frl.2023.104712.

A. Makhi and Abd. C. Fauzan, “SISTEM INFORMASI MANAJEMEN ADMINISTRASI PADA LEMBAGA SERTIFIKASI PROFESI P1 UNIVERSITAS NAHDLATUL ULAMA BLITAR MENGGUNAKAN MODEL RAPID APPLICATION DEVELOPMENT,” jatim, vol. 3, no. 1, pp. 21–34, Apr. 2022, doi: 10.31102/jatim.v3i1.1423.