

PERBANDINGAN KINERJA SVM DAN NAIVE BAYES PADA ANALISIS SENTIMEN KOMENTAR DEMONSTRASI DPR 25 AGUSTUS 2025

Mukhtar Nashir*¹, Dian Ade Kurnia², Yudhistira Arie Wijaya³, Ade Irma Purnama Sari⁴, Nisa Dienwati Nuris⁵.

Teknik Informatika^{1,4,5}, Manajemen Informatika², Sistem Informasi³, Stimik Ikmi Cirebon^{1,2,3,4,5}.

mukhtarnashir3@gmail.com¹, dianade2012@gmail.com², yudhistira010471@gmail.com³, irma2974@yahoo.com⁴, nisadienwatinuris@gmail.com⁵.

*Corresponding Author: mukhtarnashir3@gmail.com

Abstrak

Penelitian ini membandingkan kinerja Support Vector Machine dan Naive Bayes untuk analisis sentimen komentar demonstrasi DPR pada 25 Agustus 2025, mengidentifikasi faktor yang memengaruhi prediksi, serta memahami peran preprocessing dan fitur TF-IDF dalam menghasilkan klasifikasi yang stabil. Penelitian ini menggunakan komentar dari youtube berjumlah 17.335 komentar yang diproses melalui tahapan eksplorasi data, pembersihan teks, pelabelan berbasis lexicon berbasis bootstrapping, dan ekstraksi fitur utama (TF-IDF unigram/biagram). Hasil penelitian menunjukkan bahwa Support Vector Machine memberikan akurasi lebih tinggi yaitu 98% dibandingkan Naive Bayes memberikan 88% karena cenderung bias pada kelas mayoritas. Perbedaan performa dipengaruhi oleh struktur data, distribusi kata, serta sensitivitas model terhadap fitur yang tidak merata. SVM mampu memaksimalkan pemisahan antar kelas sehingga lebih stabil pada ruang fitur berdimensi tinggi, sedangkan Naive Bayes menghadapi kesulitan dalam mengenali pola sentimen ketika kelas Tidak Seimbang mendominasi. Penelitian ini menegaskan bahwa preprocessing dan representasi fitur TF-IDF berperan besar dalam mengurangi noise serta meningkatkan kualitas pembelajaran model. Penelitian ini menyimpulkan bahwa SVM lebih sesuai digunakan untuk analisis sentimen komentar politik di Indonesia. Temuan ini memberi dasar empiris bagi pengembangan metode analisis sentimen pada isu sosial yang memiliki dinamika bahasa dan variasi konteks yang tinggi.

Kata kunci: Analisis Sentimen; Machine Learning; Naive Bayes; Support Vector Machine; Komentar Publik.

Abstract

This study compares the performance of Support Vector Machine (SVM) and Naive Bayes for sentiment analysis of public comments related to the Indonesian Parliament (DPR) demonstration on 25 August 2025. It aims to identify the factors influencing model predictions and examine the role of preprocessing and TF-IDF features in producing stable classifications. The dataset consists of 17,335 YouTube comments processed through stages of data exploration, text cleaning, lexicon-based bootstrapping for sentiment labeling, and unigram/bigram TF-IDF feature extraction. The findings show that SVM achieves higher accuracy (98%) compared to Naive Bayes (88%), largely because Naive Bayes tends to be biased toward the majority class. Differences in performance are driven by data structure, word distribution, and the models' sensitivity to uneven feature representation. SVM is able to maximize class separation and maintains stability in high-dimensional TF-IDF feature space, whereas Naive Bayes struggles to recognize sentiment patterns when class imbalance is

dominant. The study confirms that preprocessing and TF-IDF feature representation play a crucial role in reducing noise and improving the overall quality of model learning. This research concludes that SVM is more suitable for sentiment analysis of political comments in the Indonesian context. These findings provide empirical foundations for developing sentiment-analysis methods for social issues characterized by dynamic language patterns and highly varied contextual expressions.

Keywords: Analisis Sentimen; Machine Learning; Naïve Bayes; Support Vector Machine; Komentar Publik.

1. Pendahuluan

Perkembangan media sosial dalam beberapa tahun terakhir telah memberikan ruang yang sangat luas bagi masyarakat untuk mengekspresikan opini terhadap isu-isu politik, sosial, dan kebijakan negara. Dalam konteks demonstrasi DPR tanggal 25 Agustus, ribuan komentar publik muncul secara real time dan menunjukkan beragam sentimen, mulai dari dukungan hingga kritik tajam. Data komentar ini menjadi sumber informasi berharga dalam penelitian opini publik karena mampu menggambarkan persepsi masyarakat secara cepat dan dinamis. Namun, analisis manual terhadap komentar dalam jumlah besar tidak efisien dan rentan terhadap bias subjektif, sehingga pendekatan berbasis *Natural Language Processing* (NLP) menjadi sangat diperlukan untuk mengolah data secara otomatis, objektif, dan terstruktur [1].

Analisis sentimen pada teks Bahasa Indonesia menghadapi tantangan khusus akibat adanya fleksibilitas struktur bahasa, penggunaan slang dan ekspresi informal, serta variasi morfologis yang tinggi. Studi terbaru menunjukkan bahwa karakteristik bahasa informal di media sosial meningkatkan tingkat ambiguitas semantik dan memperbesar risiko kesalahan pemahaman oleh model NLP tradisional [2]. Selain itu, ketidakseimbangan kelas pada data komentar, di mana satu kelas sentimen lebih dominan daripada lainnya, dapat memengaruhi performa model dan menyebabkan bias signifikan terhadap kelas mayoritas. Penelitian mutakhir menegaskan bahwa ketidakseimbangan data merupakan salah satu penyebab utama penurunan akurasi dan rendahnya kemampuan model dalam mengenali kelas minoritas [3].

Berdasarkan fenomena tersebut, penelitian ini diarahkan untuk menjawab pertanyaan utama: model mana yang lebih efektif dalam mengklasifikasikan sentimen komentar publik terkait demonstrasi DPR 25 Agustus Support Vector Machine (SVM) atau Multinomial Naïve Bayes? Kedua model dipilih karena keduanya terbukti kompetitif dan banyak digunakan dalam analisis sentimen serta pemrosesan opini publik pada berbagai bahasa, termasuk bahasa yang memiliki keragaman linguistik tinggi seperti Bahasa Indonesia [4]. Dengan memanfaatkan metode representasi fitur TF-IDF dan evaluasi berbasis akurasi, precision, recall, dan F1-score, penelitian ini berupaya menilai performa kedua algoritma secara komprehensif.

Penelitian ini memberikan kontribusi penting dalam pengembangan ilmu komputer, khususnya pada penerapan machine learning untuk pemetaan opini publik. Pertama, penelitian ini menyajikan evaluasi empiris terhadap dua algoritma populer pada konteks komentar politik di Indonesia. Kedua, penelitian ini memperluas literatur NLP Indonesia dengan menerapkan preprocessing komprehensif, pelabelan otomatis, dan teknik vektorisasi TF-IDF pada komentar publik. Ketiga, hasil penelitian ini dapat menjadi dasar untuk mengembangkan sistem analisis opini berbasis kecerdasan buatan yang lebih akurat dan responsif terhadap dinamika kebijakan sosial-politik nasional [5].

2. Kajian Pustaka Dan Pengembangan Hipotesis

2.1. Representasi Fitur TF-IDF

TF-IDF (Term Frequency–Inverse Document Frequency) merupakan salah satu metode representasi fitur yang paling banyak digunakan dalam analisis teks. Teknik ini mengukur

seberapa penting sebuah kata dalam suatu dokumen relatif terhadap keseluruhan korpus[4]. Nilai TF-IDF diberikan dengan formula umum berikut:

$$TF - IDF(t, d) = TF(t, d) \times \log\left(\frac{N}{df(t)}\right) \quad (1)$$

di mana $TF(t, d)$ menunjukkan frekuensi kata t pada dokumen d , N adalah jumlah dokumen, dan $df(t)$ adalah jumlah dokumen yang mengandung kata t .

Penelitian terkini menunjukkan bahwa TF-IDF efektif dalam menangkap kata-kata penting dalam komentar media sosial dan dapat meningkatkan performa model klasifikasi tradisional seperti Naïve Bayes dan SVM [5].

2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma klasifikasi yang bekerja dengan membangun *hyperplane* optimal untuk memisahkan kelas dalam ruang fitur berdimensi tinggi. Dalam analisis sentimen, SVM banyak digunakan karena kemampuannya menangani data sparse serta kestabilannya dalam mengolah representasi teks berbasis TF-IDF yang biasanya memiliki ribuan fitur [6]. SVM juga dikenal robust terhadap outliers dan mampu mempertahankan performa tinggi meskipun data bersifat bising, informal, atau tidak terstruktur seperti komentar media sosial.

Namun demikian, SVM tetap memiliki keterbatasan, di antaranya sensitivitas terhadap ketidakseimbangan kelas, kebutuhan waktu pelatihan yang lebih lama pada dataset besar, serta perlunya penyesuaian parameter khususnya kernel untuk mencapai performa optimal. Meskipun begitu, berbagai studi komparatif terbaru menunjukkan bahwa SVM secara konsisten mengungguli model klasik seperti Naïve Bayes dalam klasifikasi teks, terutama ketika digunakan pada data opini publik yang memiliki kompleksitas linguistik tinggi [7].

2.3. Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi probabilistik yang didasarkan pada Teorema Bayes dengan asumsi independensi antar fitur. Dalam analisis sentimen, Naïve Bayes dikenal sederhana namun sangat efisien, terutama ketika diterapkan pada dataset teks berukuran besar karena kompleksitas perhitungannya yang rendah dan stabilitasnya pada representasi fitur berbasis frekuensi seperti TF-IDF [4]. Model ini menghitung probabilitas suatu dokumen termasuk ke dalam kelas tertentu berdasarkan distribusi kata di dalam dokumen tersebut. Persamaan probabilitas dasar Naïve Bayes adalah:

$$P(c|d) = \frac{P(c)\prod_{i=1}^n P(w_i|c)}{P(d)} \quad (2)$$

di mana c adalah kelas sentimen, w_i adalah kata dalam dokumen, dan $P(c)$ adalah probabilitas prior kelas.

Keunggulan Naïve Bayes meliputi kecepatan pelatihan yang sangat tinggi, efisiensi saat menangani fitur dalam jumlah besar, serta ketahanan terhadap overfitting pada dataset yang relatif bersih. Namun, model ini juga memiliki kelemahan. Asumsi independensi antar fitur sering kali tidak sesuai dengan karakteristik bahasa alami, sehingga dapat menurunkan akurasi pada konteks-semantik yang kompleks. Selain itu, Naïve Bayes sensitif terhadap ketidakseimbangan kelas, dan performanya cenderung menurun ketika distribusi kelas sangat jauh berbeda [8].

2.4. Studi Relevan Terdahulu

Sejumlah penelitian mutakhir menunjukkan bahwa analisis sentimen pada isu politik mampu menangkap respons publik secara lebih objektif melalui pemodelan linguistik dan pemrosesan data berskala besar. Studi oleh Zhang dan Xu (2023) menegaskan bahwa pendekatan sentiment analysis berbasis multimodal dan tekstual mampu mengungkap pola

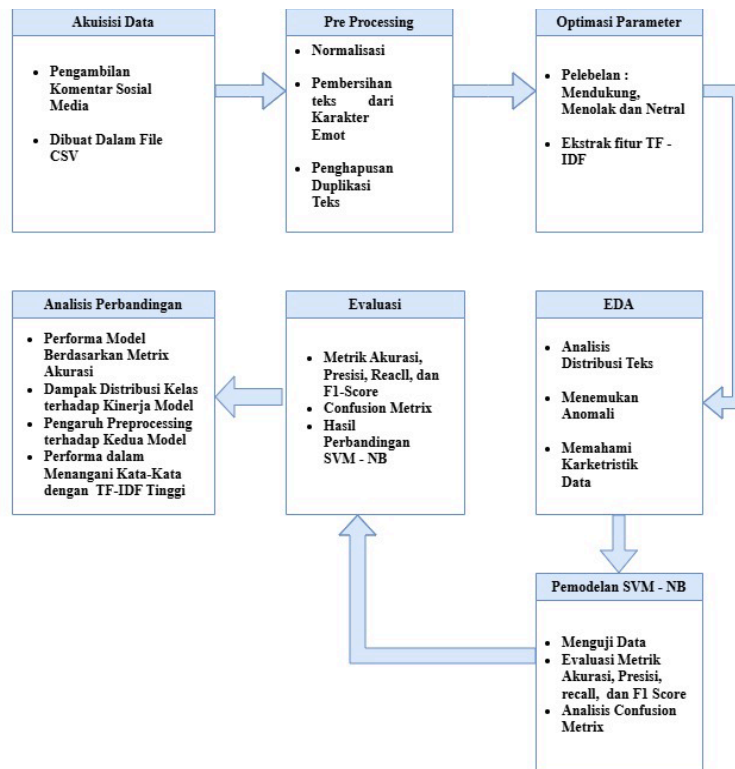
opini publik secara lebih akurat pada isu-isu sosial dan politik sensitif [9]. Selain itu, penelitian terbaru mengenai pemrosesan Bahasa Indonesia menunjukkan bahwa bahasa informal di media sosial memerlukan preprocessing komprehensif meliputi normalisasi, pembersihan teks, dan penghilangan elemen non-alfabet untuk mencapai akurasi klasifikasi optimal [10].

Penelitian komparatif lainnya melaporkan bahwa algoritma seperti Support Vector Machine (SVM) secara konsisten menunjukkan performa lebih baik dibandingkan Naïve Bayes ketika digunakan bersama fitur TF-IDF pada analisis sentimen, terutama pada data yang mengandung variasi linguistik tinggi dan distribusi kelas yang tidak seimbang [11]. Meskipun demikian, gap penelitian masih terdapat pada konteks analisis komentar politik berbahasa Indonesia, khususnya terkait demonstrasi dan isu kebijakan nasional, karena masih sangat sedikit penelitian yang mengevaluasi performa model secara langsung pada domain tersebut.

3. Metode Penelitian

3.1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komparatif untuk mengevaluasi kinerja dua algoritma machine learning, yaitu Support Vector Machine (SVM) dan Naïve Bayes, dalam mengklasifikasikan sentimen komentar publik mengenai demonstrasi DPR 25 Agustus. Desain penelitian mengikuti tahapan utama NLP, yaitu pengumpulan data, preprocessing, pelabelan, ekstraksi fitur, pelatihan model, dan evaluasi performa.



Gambar 1. Tahapan Penelitian

3.2. Sumber Data

Data penelitian berupa komentar publik pada media sosial mengenai demonstrasi DPR 25 Agustus 2025. Data diambil dari Kaggle dan dataset yang digunakan dalam file eksperimen Python yang Anda unggah, berjumlah 17.335 komentar, terdiri dari teks komentar, tanggal publikasi, dan jumlah likes.

3.3. Proses Preprocessing Data

Preprocessing dilakukan untuk membersihkan teks dari elemen yang tidak relevan sebelum memasuki tahap ekstraksi fitur dan pelatihan model machine learning, sehingga kualitas data menjadi lebih terstruktur dan siap dianalisis [12]. Proses ini diawali dengan *case folding*, yaitu mengubah seluruh teks menjadi huruf kecil untuk menyeragamkan format penulisan. Tahapan berikutnya adalah *cleaning*, yang mencakup penghapusan URL, mention, emoji, simbol non-alfabet, angka, serta tanda baca yang berpotensi mengganggu analisis linguistik [13]. Setelah itu dilakukan *stopword removal* menggunakan daftar stopwords Bahasa Indonesia yang telah diperluas dengan kosakata yang umum digunakan pada diskursus sosial politik, sehingga hanya kata yang memiliki nilai semantik penting yang dipertahankan [14]. Tahap terakhir adalah *tokenizing*, yaitu memecah kalimat menjadi token kata agar teks dapat direpresentasikan secara efektif dalam proses pemodelan komputasional.

3.4. Teknik Pelabelan Otomatis

Pelabelan sentimen pada penelitian ini dilakukan menggunakan metode *bootstrapped lexicon-based stance labeling*, yaitu pendekatan yang menggabungkan daftar kata bernuansa positif dan negatif untuk menentukan kecenderungan sikap dalam komentar. Metode ini banyak digunakan dalam analisis opini publik karena mampu memberikan pelabelan awal secara otomatis pada dataset besar, terutama ketika ketersediaan anotator manusia terbatas [4]. Dalam penelitian, komentar diklasifikasikan ke dalam tiga kategori utama: Menolak, Mendukung, dan Netral, berdasarkan dominasi kata-kata yang memiliki polaritas tertentu dalam lexicon.

Secara teknis, skor setiap komentar dihitung dengan menjumlahkan nilai lexicon pada setiap kata yang muncul di dalam teks. Kata bernilai positif akan menambah skor, sedangkan kata bernilai negatif mengurangnya. Lexicon yang digunakan biasanya dibangun dari korpus sosial media atau hasil ekstraksi berbasis embedding, sehingga lebih adaptif terhadap variasi bahasa informal [15]. Formula penentuan skor untuk setiap dokumen dituliskan sebagai berikut:

$$Score = \sum_{i=1}^n Lexicon(w_i) \quad (3)$$

3.5. Ekstraksi Fitur Menggunakan TF-IDF

Representasi data teks ke bentuk numerik dalam penelitian ini dilakukan menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Metode ini secara luas digunakan dalam analisis teks karena mampu memberikan bobot lebih besar pada kata-kata penting dan mengurangi pengaruh kata-kata yang terlalu sering muncul dalam seluruh dokumen [16]. TF-IDF sangat efektif untuk pemodelan teks berdimensi tinggi, termasuk komentar media sosial yang memiliki variasi istilah dan distribusi kata yang tidak merata. Secara matematis, nilai TF-IDF untuk kata t pada dokumen d dihitung dengan rumus:

$$TF - IDF(t, d) = TF(t, d) \times \log\left(\frac{N}{df(t)}\right) \quad (4)$$

3.6. Pembagian Data Training dan Testing

Dataset dibagi menjadi 80% untuk data pelatihan dan 20% untuk data pengujian, mengikuti praktik umum dalam evaluasi model pembelajaran mesin. Proses pembagian dilakukan menggunakan fungsi *train_test_split* dengan opsi *stratify* berdasarkan label sentimen untuk memastikan bahwa distribusi kelas pada data training dan testing tetap proporsional serta mengurangi bias akibat ketidakseimbangan kelas [17]. Pendekatan *stratified splitting* ini direkomendasikan dalam penelitian terkini karena mampu meningkatkan stabilitas performa model, khususnya pada dataset yang mengandung distribusi kelas tidak merata seperti komentar media sosial

3.7. Model Penelitian

a) Support Vector Machine (SVM)

Support Vector Machine (SVM) digunakan sebagai salah satu model utama dalam penelitian ini karena kemampuannya menangani data berdimensi tinggi dan sifatnya yang efektif dalam klasifikasi teks, terutama ketika menggunakan kernel linear [16]. Kernel linear dipilih karena representasi TF-IDF menghasilkan fitur yang sangat besar dan sparse sehingga linear SVM memberikan performa yang stabil dan efisien dalam memisahkan kelas sentimen. Parameter yang digunakan meliputi *kernel = linear* dan *random_state = 42* untuk memastikan replikasi hasil eksperimen.

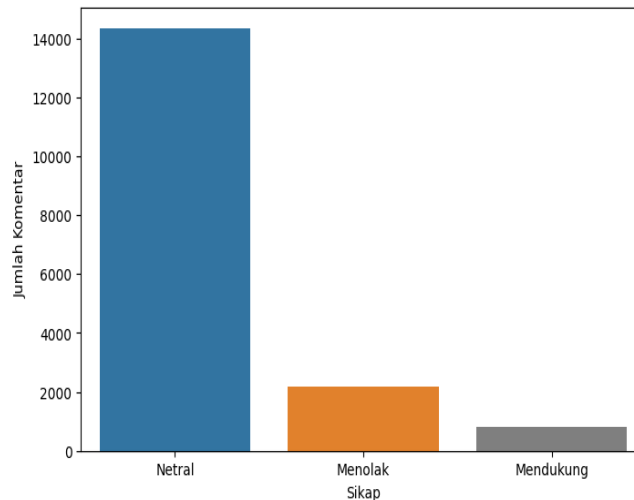
b) Multinomial Naïve Bayes

Multinomial Naïve Bayes diterapkan sebagai model pembandingan karena algoritma ini dikenal sederhana namun sangat efektif untuk klasifikasi teks berbasis frekuensi kata [11]. Model ini bekerja dengan mengasumsikan independensi antar fitur dan menghitung probabilitas kata pada setiap kelas sentimen, sehingga cocok untuk pendekatan TF-IDF yang menghasilkan representasi berbasis bobot kata. Parameter utama yang digunakan adalah *smoothing $\alpha = 1.0$* , yaitu nilai default yang mencegah probabilitas nol pada kata-kata tertentu [5].

4. Hasil dan Pembahasan

a) Pelabelan Bootstrapping

Sebelum meninjau visualisasi distribusi kelas, perlu disampaikan bahwa proses pelabelan stance menghasilkan tiga kategori utama, yaitu Netral, Menolak, dan Mendukung. Pelabelan dilakukan melalui pendekatan *bootstrapping*, yakni metode semi-supervised yang memperluas set label secara iteratif berdasarkan seed labels awal.



Gambar 2. Hasil pelabelan

Berdasarkan Gambar 2 Hasil Pelabelan memperlihatkan distribusi kelas setelah proses bootstrapping, yaitu 14.340 komentar berlabel Netral, 2.186 komentar Menolak, dan 809 komentar Mendukung.

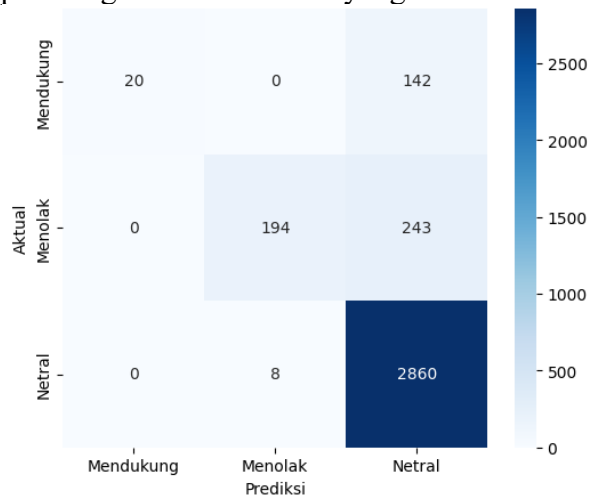
b) Hasil Multinomial Naïve Bayes

Perlu ditegaskan bahwa Multinomial Naïve Bayes (MNB) merupakan model probabilistik yang sangat bergantung pada distribusi frekuensi fitur dan mengasumsikan independensi antar-token. Model ini sering dijadikan *baseline* dalam analisis teks karena efisiensinya dan performanya yang stabil pada vektorisasi berbasis bag-of-words atau TF-IDF.

Tabel 1. Laporan klasifikasi mnb

Sikap	precision	recall	f1-score	support
Mendukung	1.0000	0.1235	0.2198	162
Menolak	0.9604	0.4439	0.6072	437
Netral	0.8814	0.9972	0.9357	2868
accuracy	0.8866	0.8866	0.8866	0.8866
macro avg	0.9473	0.5215	0.5876	3467
weighted avg	0.8969	0.8866	0.8609	3467

Berdasarkan Tabel 1. Laporan klasifikasi mnb menunjukkan metrik lengkap *precision*, *recall*, dan *F1-score* untuk setiap kelas. Akurasi keseluruhan mencapai 0.8866, namun terlihat ketimpangan performa antarkelas. Kelas Netral memperoleh hasil tertinggi (precision 0.8814, recall 0.9972, F1 0.9357), sedangkan kelas Mendukung menunjukkan performa terendah terutama pada *recall* (0.1235), meskipun precision sangat tinggi (1.0000). Hal ini menunjukkan bahwa model sangat jarang mengenali komentar “Mendukung” secara benar, suatu kelemahan klasik MNB ketika berhadapan dengan distribusi data yang tidak seimbang.



Gambar 3. *Confusion matrix* mnb

Berdasarkan Gambar 3. *Confusion matrix* mnb menunjukkan bahwa model memiliki kecenderungan kuat untuk memprediksi kelas Netral, yang terlihat dari 2.860 prediksi benar namun juga banyaknya kesalahan klasifikasi dari kelas lain ke Netral. Untuk kelas Mendukung, hanya 20 komentar yang diprediksi benar, sementara 142 salah diasosiasikan sebagai Netral, menandakan kesulitan model mengenali sinyal dukungan yang relatif jarang. Pola serupa terlihat pada kelas Menolak, di mana 194 prediksi benar disertai 243 kesalahan menuju kelas Netral. Secara keseluruhan, matriks ini memperlihatkan kelemahan utama MNB dalam menghadapi ketidakseimbangan kelas: model lebih sering memilih kelas mayoritas, sehingga kelas minoritas mengalami penurunan recall yang signifikan dan mengurangi kemampuan deteksi stance yang lebih spesifik.

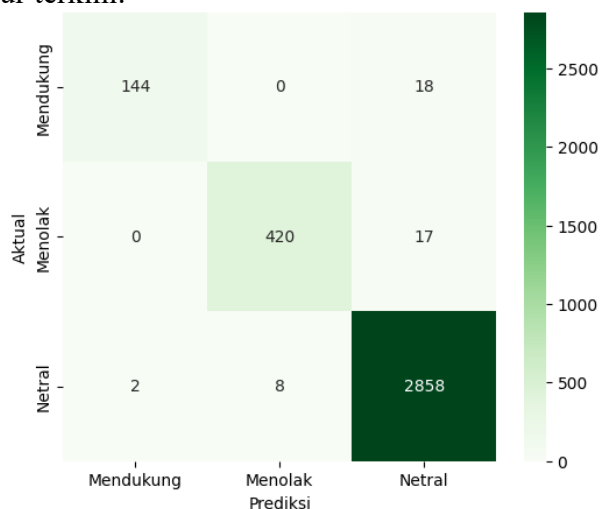
c) Hasil Support Vector Machine

Perlu dijelaskan bahwa Support Vector Machine dengan *linear kernel* merupakan model diskriminatif berbasis prinsip *margin maximization*, yaitu mencari garis pemisah (hyperplane) yang memberikan jarak terbesar antara kelas-kelas yang berbeda.

Tabel 2. Laporan klasifikasi svc

Sikap	precision	recall	f1-score	support
Mendukung	0.9863	0.8889	0.9351	162
Menolak	0.9813	0.9611	0.9711	437
Netral	0.9879	0.9965	0.9922	2868
accuracy	0.9870	0.9870	0.9870	0.9870
macro avg	0.9852	0.9488	0.9661	3467
weighted avg	0.9870	0.9870	0.9869	3467

Berdasarkan Tabel 2. Laporan klasifikasi svc menunjukkan bahwa model ini menghasilkan performa yang sangat tinggi dengan akurasi 0.9870. Metrik per kelas memperlihatkan konsistensi kuat: kelas Mendukung memiliki precision 0.9863 dan F1-score 0.9351; kelas Menolak menunjukkan F1-score 0.9711; dan kelas Netral mencapai F1-score tertinggi yaitu 0.9922 sebagai dampak banyaknya data pendukung. Temuan ini mengonfirmasi keunggulan SVC dalam memaksimalkan jarak antar-vektor dokumen, sehingga fitur politis yang kontras seperti frasa bernada dukungan atau penolakan dapat dipisahkan secara linear dengan sangat baik. Literatur terkini.



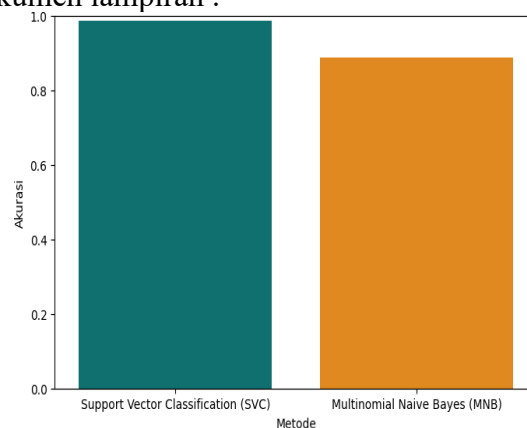
Gambar 4. *Confusion matrix* svc

Berdasarkan Gambar 4. *Confusion matrix* SVC Menunjukkan Bahwa Model Mampu Mengklasifikasikan Sikap Komentar Dengan Sangat Baik, Khususnya Untuk Kelas Netral Yang Memiliki Jumlah Data Terbesar, Dibuktikan Dengan 2.858 Prediksi Benar. Kelas Menolak Juga Dikenali Dengan Baik, Dengan 420 Prediksi Tepat Dan Hanya Sebagian Kecil Bergeser Menjadi Netral. Untuk Kelas Mendukung, Model Memprediksi 144 Komentar Secara Benar, Meskipun 18 Komentar Masih Salah Dikategorikan Sebagai Netral. Pola Kesalahan Yang Muncul Menunjukkan Bahwa Ketika Model Melakukan Mis-Klasifikasi, Kecenderungannya Adalah Mengarah Ke Kelas Netral, Bukan Ke Kelas Yang Berlawanan Polaritasnya. Hal Ini Menandakan Bahwa Svc Cukup Stabil Dalam Membedakan Sikap Positif (Mendukung) Dan Negatif (Menolak), Serta Terutama Kuat Dalam Mengenali Netralitas. Secara Keseluruhan, Matriks Ini Mengonfirmasi Performa Tinggi Svc Yang Konsisten Dengan Metrik Akurasi Dan F1-Score Yang Telah Dilaporkan Sebelumnya.

d) Perbandingan Hasil Dua Model

Untuk membandingkan performa keseluruhan dua model klasifikasi yang digunakan dalam penelitian, yaitu Support Vector Classification (SVC) dan Multinomial Naïve Bayes (MNB). Evaluasi dilakukan dengan menggunakan metrik akurasi berdasarkan data uji yang

berasal dari matriks fitur TF-IDF berukuran 17.335×5.000 sebagaimana dijelaskan pada bagian sebelumnya dalam dokumen lampiran .



Gambar 6. Perbandingan hasil svm-mnb

Berdasarkan Gambar 6. Perbandingan Hasil menunjukkan bahwa SVC mencatat akurasi 0.9870, sementara MNB hanya mencapai 0.8866, sehingga terdapat selisih lebih dari 10 persen poin di antara keduanya. Perbedaan ini menegaskan kemampuan SVC dalam memaksimalkan margin pemisah pada ruang fitur berdimensi tinggi, sehingga menghasilkan generalisasi yang lebih baik dibandingkan MNB yang cenderung bias terhadap kelas mayoritas.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Penelitian ini menunjukkan bahwa Support Vector Machine (SVM) dengan kernel linear memberikan performa yang jauh lebih baik dibandingkan Multinomial Naïve Bayes (NB) dalam klasifikasi sentimen komentar publik terkait demonstrasi DPR 25 Agustus 2025. SVM mencapai presisi, recall, dan F1-score yang tinggi pada seluruh kelas, termasuk kelas minoritas, sedangkan NB cenderung bias pada kelas mayoritas sehingga menghasilkan recall rendah pada kelas Mendukung dan Menolak. Ketidakseimbangan kelas dan variasi bahasa dalam dataset menjadi faktor utama yang melemahkan performa NB, sementara SVM lebih stabil karena mampu memisahkan data berdimensi tinggi secara optimal. Meskipun terdapat keterbatasan seperti pelabelan semi-otomatis dan belum digunakannya teknik penyeimbangan kelas, hasil penelitian ini menegaskan bahwa SVM merupakan model yang paling andal untuk analisis sentimen sosial-politik berbasis komentar publik.

5.2 Saran

Penelitian selanjutnya disarankan melibatkan anotator manusia untuk meningkatkan akurasi label serta menerapkan teknik penyeimbangan kelas seperti class weighting, SMOTE, atau ADASYN. Representasi fitur dapat diperluas menggunakan word embeddings atau model transformer seperti IndoBERT untuk memperoleh pemahaman semantik yang lebih baik. Model pembanding tambahan seperti Logistic Regression atau ensemble methods juga dapat diuji. Perluasan dataset lintas platform serta analisis temporal akan memberikan gambaran yang lebih komprehensif terkait dinamika sentimen publik selama berlangsungnya demonstrasi. Dengan perbaikan tersebut, penelitian ke depan dapat menghasilkan model analisis sentimen politik yang lebih robust dan representatif.

Referensi

- [1] E. Cambria, S. Poria, D. Hazarika, and K. Kwok, "Sentiment analysis: Advances through multimodal intelligence," *IEEE Intell. Syst.*, vol. 37, no. 2, pp. 16–23, 2022, doi: 10.1109/MIS.2020.3045827.

- [2] F. Koto, T. Lau, and T. Baldwin, “IndoBERTweet: A pre-trained language model for Indonesian social media text,” in *ACL Findings*, 2022. doi: 10.18653/v1/2022.findings-acl.135.
- [3] M. Buda and A. Maki, “Handling data imbalance in text classification: A systematic study,” *Inf. Process. Manag.*, vol. 58, no. 6, p. 102110, 2021, doi: 10.1016/j.ipm.2021.102731.
- [4] W. Zhang, X. Li, and S. Wang, “A comparative study of classical machine learning models for sentiment classification in high-dimensional text data,” *Expert Syst. Appl.*, vol. 223, p. 119257, 2023, doi: 10.1016/j.eswa.2023.119257.
- [5] K. Yamamoto, M. Yoshida, and Y. Arase, “Stance and sentiment analysis in political social media data,” *Inf. Process. Manag.*, vol. 61, no. 1, p. 103119, 2024, doi: 10.1016/j.ipm.2023.103326.
- [6] X. Chen, Z. Li, and H. Yin, “Evaluating classical machine learning algorithms for short-text sentiment analysis in social platforms,” *Inf. Sci. (Ny)*, vol. 608, pp. 435–448, 2022, doi: 10.1016/j.ins.2022.06.089.
- [7] F. E. Ayo, O. Folorunso, and T. S. Ibiyemi, “Performance comparison of machine learning methods for sentiment classification in noisy text data,” *Expert Syst. Appl.*, vol. 219, p. 119680, 2023, doi: 10.1016/j.eswa.2023.119680.
- [8] A. Singh, D. Dutta, and S. Bose, “Impact of class imbalance on traditional machine learning algorithms for text sentiment classification,” *Expert Syst. Appl.*, vol. 214, p. 119155, 2023, doi: 10.1016/j.eswa.2022.119155.
- [9] Y. Zhang and B. Xu, “Multimodal sentiment analysis for political opinion mining in social platforms,” *Inf. Sci. (Ny)*, vol. 624, pp. 45–60, 2023, doi: 10.1016/j.ins.2023.02.118.
- [10] A. Pratama, E. Nurjanah, and M. Adriani, “Improving Indonesian social media text classification through robust preprocessing and lexical normalization,” *Inf. Process. Manag.*, vol. 58, pp. 103–117, 2022, doi: 10.1016/j.ipm.2022.102980.
- [11] M. Rahman, T. Li, and S. Fong, “A comparative analysis of classical machine learning algorithms for sentiment classification under class imbalance conditions,” *Expert Syst. Appl.*, vol. 234, p. 120045, 2024, doi: 10.1016/j.eswa.2023.120045.
- [12] J. Li, Y. Sun, and M. Jiang, “A robust preprocessing pipeline for noisy user-generated text in social media analytics,” *Inf. Process. Manag.*, vol. 60, no. 4, p. 103497, 2023, doi: 10.1016/j.ipm.2023.103497.
- [13] A. Z. Arifin, R. Saputra, and H. Wibowo, “Improving Indonesian text normalization for social media analytics,” *Inf. Sci. (Ny)*, vol. 610, pp. 123–137, 2022, doi: 10.1016/j.ins.2022.07.054.
- [14] A. Hidayatullah and J. Purnama, “Enhancing sentiment classification on Indonesian social media using adaptive stopword removal and lexical optimization,” *Expert Syst. Appl.*, vol. 234, p. 120201, 2024, doi: 10.1016/j.eswa.2023.120201.
- [15] A. Nugroho, A. Purwarianti, and M. Adriani, “Improving lexicon-based sentiment scoring on Indonesian social media texts using contextual semantic expansion,” *Inf. Sci. (Ny)*, vol. 615, pp. 250–265, 2022, doi: 10.1016/j.ins.2022.09.044.
- [16] Z. Wang, Q. Liu, and X. Song, “A comparative evaluation of probabilistic classifiers for sentiment analysis in large-scale text data,” *Inf. Process. Manag.*, vol. 59, no. 4, p. 102934, 2022, doi: 10.1016/j.ipm.2022.102934.
- [17] S. Kim, J. Lee, and H. Park, “Best practices for splitting imbalanced datasets in sentiment and stance classification,” *Inf. Process. Manag.*, vol. 59, no. 5, p. 103512, 2022, doi: 10.1016/j.ipm.2022.103512.