

# Generative AI Image Sentiment Analysis on Social Media X using TF-IDF and FastText

Rahman Saputra <sup>1\*</sup>, Yoga Pristyanto <sup>2\*</sup>, Ika Nur Fajri <sup>3\*</sup>

\* Sistem Informasi, Universitas Amikom Yogyakarta

[rhmansptra@students.ac.id](mailto:rhmansptra@students.ac.id) <sup>1</sup>, [yoga.pristyanto@amikom.ac.id](mailto:yoga.pristyanto@amikom.ac.id) <sup>2</sup>, [fajri@amikom.ac.id](mailto:fajri@amikom.ac.id) <sup>3</sup>

## Article Info

### Article history:

Received 2025-08-01

Revised 2025-08-28

Accepted 2025-09-03

### Keyword:

*Sentiment Analysis,  
AI-Generated Images,  
Social Media X,  
Complement Naïve Bayes,  
Support Vector Machine,  
TF-IDF,  
FastText,  
SMOTE.*

## ABSTRACT

This research investigates public opinion on AI-generated images on Social Media X using machine learning-driven text classification. Three classification models were evaluated: Complement Naïve Bayes (CNB) utilizing TF-IDF features, Support Vector Machine (SVM) merging TF-IDF with FastText embeddings, and IndoBERT as a modern transformer-based baseline. A total of 1,958 Indonesian tweets were collected via web scraping with relevant keywords, followed by a pipeline involving text cleaning, manual labeling into positive, negative, and neutral categories, and data balancing using the Synthetic Minority Over-sampling Technique (SMOTE) for the classical models (with class weighting applied for IndoBERT). Results show that the SVM model outperformed the others, achieving 68.7% accuracy with average precision, recall, and F1-score of 0.69, 0.69, and 0.68, respectively; CNB attained 64.1% accuracy with average metrics of 0.64; while IndoBERT recorded 58.2% accuracy with average precision, recall, and F1-score of 0.58, 0.58, and 0.57. Confusion matrix analysis revealed SVM's superior ability to distinguish positive and neutral sentiments in casual language, though IndoBERT demonstrated potential for capturing deeper semantic nuances despite underperforming due to dataset size and informal text. The findings highlight the efficacy of integrating statistical and semantic representations for improved sentiment analysis on unstructured, noisy social media data related to AI-generated imagery, while suggesting that transformer models like IndoBERT may benefit from larger datasets for optimal performance.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

## I. PENDAHULUAN

Dengan perkembangan pesat teknologi kecerdasan buatan (AI), banyak aspek kehidupan telah berubah, salah satunya bidang seni visual. Gambar AI generatif adalah salah satu teknologi AI terkini yang dapat menghasilkan gambar secara otomatis dengan kualitas yang hampir sama dengan karya buatan manusia. Teknologi ini bekerja dengan mempelajari pola dari jutaan gambar yang sudah ada dan kemudian membuat gambar baru secara real-time berdasarkan pola tersebut[1]. Dengan munculnya teknologi gambar AI generatif, masyarakat umum kini sangat tertarik padanya, terutama di berbagai platform media sosial, salah satunya adalah "media sosial X", yang memungkinkan pengguna berinteraksi secara langsung dan berbagi pendapat mereka

tentang topik yang sedang berkembang, seperti gambar AI generatif[2].

Analisis emosi pengguna terhadap konten yang diunggah menjadi penting untuk memahami sikap masyarakat karena minat publik terhadap teknologi ini meningkat. Analisis sentimen adalah jenis text mining yang digunakan untuk mengekstrak, mengolah, dan membagi pendapat atau perasaan pengguna media sosial tentang suatu subjek ke dalam kategori positif, negatif, atau netral[3]. Menurut penelitian sebelumnya, analisis sentimen telah banyak digunakan dalam berbagai konteks media sosial. Ini termasuk sentimen terhadap program vaksinasi COVID-19[2], opini publik tentang ChatGPT[1], aplikasi kesehatan[4], produk pasar[4], [5], dan analisis sentimen terkait program pemerintah[6]. Namun, tidak banyak penelitian yang secara

husus menyelidiki persepsi masyarakat terhadap teknologi gambar AI generatif, terutama di media sosial X.

Karena algoritma berbasis pembelajaran mesin seperti Naive Bayes dan Support Vector Machine (SVM) dikenal sangat baik untuk mengklasifikasikan sentimen dengan tingkat akurasi yang tinggi, sebagian besar penelitian sebelumnya tentang analisis sentimen menggunakan algoritma ini[7]. Studi sebelumnya menunjukkan bahwa keunggulan metode Naive Bayes dan SVM bergantung pada jenis dataset yang digunakan serta parameter tertentu, seperti teknik ekstraksi fitur dan representasi kata[8]. Selain itu, telah terbukti bahwa penggunaan FastText sebagai metode penerapan dapat meningkatkan efisiensi klasifikasi sentimen[9], terutama untuk data teks dari platform media sosial seperti X, yang biasanya tidak terstruktur dan informal[10].

Meskipun demikian, masih ada kekurangan penelitian yang jelas, karena belum ada penelitian komprehensif yang menghubungkan penggunaan metode Naive Bayes dan SVM dengan sentimen terhadap gambar AI generatif di media sosial X. Penelitian ini sangat penting karena akan membantu pengembang, pembuat kebijakan, dan pemasar memahami bagaimana teknologi baru ini dilihat oleh masyarakat. Ini dapat membantu mereka mengantisipasi dampak AI generatif yang akan datang[11].

Ulasan produk lokal di pasar, masalah politik, dan kebijakan publik dan fiskal adalah fokus penelitian sebelumnya tentang analisis sentimen di Indonesia [12],[13],[14]. Namun, hingga saat ini belum ada penelitian yang secara eksplisit menyelidiki tanggapan masyarakat Indonesia terhadap gambar AI generatif di media sosial X. Padahal, citra yang diciptakan oleh AI semakin luas dan memiliki kemampuan untuk memengaruhi persepsi publik, termasuk dalam konteks politik [12], [14]. Penelitian ini mengeksplorasi tanggapan publik terhadap gambar AI generatif di media sosial X, serta membandingkan efektivitas metode representasi teks tradisional seperti TF-IDF dan FastText [12].

Perkembangan terbaru menunjukkan bahwa model berbasis transformer seperti IndoBERT lebih baik daripada metode tradisional dalam mengklasifikasikan sentimen berbahasa Indonesia. IndoBERT telah ditunjukkan efektif dalam berbagai bidang, termasuk e-commerce, masalah kebijakan publik, masalah politik, dan mitigasi bencana[14],[15],[16]. Model menunjukkan akurasi antara 90% dan 93% pada berbagai dataset, melampaui metode konvensional seperti Naive Bayes[15],[16]. Oleh karena itu, untuk memberikan perbandingan yang lebih akurat terhadap kinerja TF-IDF dan FastText, IndoBERT dianggap sebagai baseline modern dalam penelitian ini.

Dengan demikian, tujuan dari penelitian ini adalah untuk melakukan analisis sentimen terhadap pandangan masyarakat tentang teknologi gambar AI generatif, seperti yang terlihat dalam percakapan media sosial X menggunakan pendekatan algoritma Naive Bayes dan SVM. Hasil penelitian ini diharapkan dapat memberikan gambaran yang jelas tentang persepsi publik dan menjadi referensi akademik yang berguna

untuk studi lanjutan tentang penerimaan dan dampak teknologi gambar AI generatif.

## II. METODE

Metode pembelajaran supervised berbasis klasifikasi teks digunakan dalam penelitian kuantitatif ini. Membangun model klasifikasi sentimen adalah tujuan utama dari penelitian ini, yang dilakukan melalui berbagai tahapan terstruktur, termasuk pengumpulan data, pra-pemrosesan, ekstraksi fitur, penyeimbangan data, pelatihan model, dan evaluasi kinerja model.

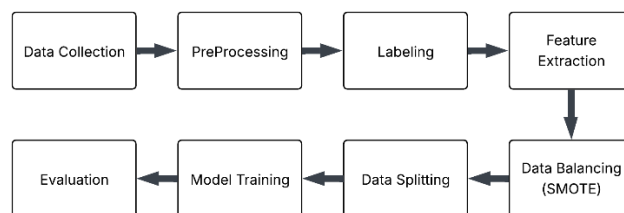
Pilihan algoritma Naive Bayes dan Support Vector Machine (SVM) didasarkan pada kemampuan mereka untuk mengklasifikasikan teks dan menangani data besar[17],[18]. Untuk menampilkan karakteristik semantik teks secara menyeluruh, kombinasi TF-IDF dan FastText digunakan. Penelitian sebelumnya menunjukkan bahwa kombinasi ini dapat meningkatkan akurasi[19], [20].

Selain itu, data yang digunakan dalam penelitian ini tidak seimbang antara kelas sentimen, sehingga teknik penyeimbangan data diperlukan untuk mencegah bias model. Oleh karena itu, teknik over-sampling minoritas sintetis (SMOTE) digunakan, yang telah terbukti berhasil meningkatkan kualitas klasifikasi pada dataset yang tidak seimbang[21].

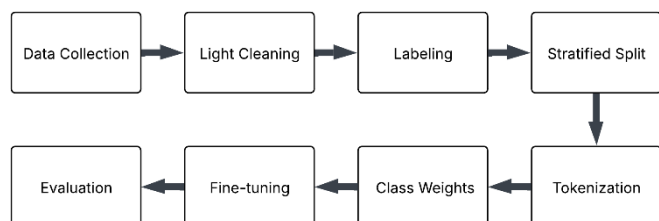
Untuk menilai konsistensi hasil klasifikasi, evaluasi performa model dilakukan dengan metrik akurasi, ketepatan, recall, dan skor F1. Selain itu, teknik cross-validation k-fold juga digunakan[22].

Penelitian ini menambahkan model transformer IndoBERT berbahasa Indonesia sebagai pembanding modern (baseline) karena telah ditunjukkan unggul dalam analisis sentimen dalam berbagai domain (e-commerce, kebijakan publik, politik, dan kebencanaan)[14],[15],[16]. Jalur IndoBERT berbeda dari jalur konvensional dengan menggunakan pra-pemrosesan minimal non-destruktif (pembersihan ringan tanpa stemming atau penghapusan kata akhir), representasi end-to-end melalui tokenisasi sub-kata, kehilangan kelas yang berat (tanpa SMOTE), dan skema pembagian data 80/10/10 (latihan, validasi, dan uji). Agar hasil kedua metode dapat dibandingkan secara wajar, pelatihan dilakukan menggunakan AdamW dengan rasio pembelajaran  $2e-5$ , ukuran batch 16/32, penundaan dini berbasis Macro-F1, dan evaluasi dilakukan dengan metrik yang sama [14], [15],[16].

Diagram alur berikut menunjukkan langkah-langkah utama yang diambil dalam penelitian ini:



Gambar 1. Tahapan penelitian menggunakan metode klasik (CNB dan SVM)



Gambar 2. Tahapan penelitian menggunakan baseline modern IndoBERT.

### A. Data Collection

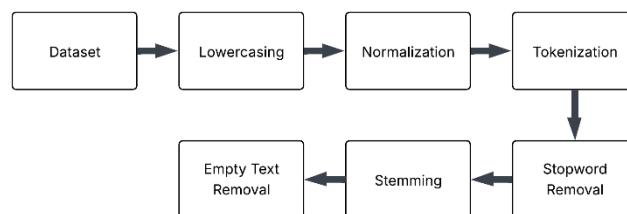
Data dalam penelitian ini diambil dari media sosial X (dulu disebut Twitter) melalui metode web scraping dengan skrip Python yang dibuat secara independen tanpa menggunakan pustaka pihak ketiga seperti SNScrape. Metode ini memungkinkan penyesuaian terhadap struktur HTML terkini dari situs web dan memungkinkan pengambilan data yang lebih luwes berdasarkan kata kunci spesifik dalam bahasa Indonesia serta jangka waktu tertentu. Kata-kata kunci yang diterapkan dalam proses ini mencakup: “gambar oleh ai”, “karya ciptaan ai”, dan “gambar yang diproduksi ai”. Dari kegiatan pengambilan data ini, berhasil terkumpul sebanyak 2.000 tweet.

Selanjutnya, dilakukan proses pembersihan awal dengan menghapus duplikasi, konten yang kosong, serta tweet yang mengandung spam atau tidak relevan. Setelah langkah ini, total tweet yang dapat digunakan berjumlah 1.958. Tweet-tweet ini selanjutnya diberi label secara manual oleh peneliti ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Proses pelabelan mempertimbangkan makna kontekstual dan ekspresi emosional yang terdapat dalam teks, sesuai dengan pendekatan leksikal yang sering dipakai dalam penelitian sejenis[23].

Komposisi data menunjukkan bahwa sentimen netral mendominasi di dalam dataset. Hasil ini sejalan dengan sejumlah penelitian sebelumnya yang mengindikasikan kecenderungan umum orang untuk menyampaikan pandangan secara netral di media sosial X[24]. Untuk menangani ketidakseimbangan jumlah data di antara kelas sentimen, penelitian ini menggunakan metode Synthetic Minority Over-sampling Technique (SMOTE) selama proses pelatihan model untuk meningkatkan kinerja klasifikasi pada data minoritas[22].

### B. Preprocessing

Karena kualitas input teks sangat mempengaruhi kinerja model klasifikasi, pra-pemrosesan adalah langkah penting dalam pipeline analisis sentimen. Data yang dikumpulkan dari X, misalnya, biasanya tidak sistematis dan informal, dan banyak mengandung ekspresi emosi yang tidak baku[20]. Akibatnya, untuk mengubah teks mentah menjadi representasi yang lebih bersih dan informatif, diperlukan beberapa tahapan pra-pemrosesan. Ini mencakup langkah-langkah berikut:



Gambar 3. Tahapan Preprocessing.

1) *Case Folding*: Teks menjadi huruf kecil. Ini sangat penting agar model tidak memperlakukan istilah seperti "AI", "Ai", atau "ai" sebagai token yang berbeda. Metode ini biasa digunakan dalam banyak studi analisis sentimen berbasis teks[20], [24].

2) *Normalization*: Tujuan normalisasi adalah untuk mengubah kata-kata yang tidak formal, slang, atau singkatan yang umum digunakan dalam percakapan online menjadi bentuk baku yang lebih tepat. Misalnya, kata "gk", "ga", dan "gak" berarti "tidak", dan "bgt" berarti "banget". Berdasarkan temuan penelitian sebelumnya [23],[25], dan analisis frekuensi kata dalam kumpulan data, kamus normalisasi disusun secara manual. Karena mayoritas tweet menggunakan bahasa tidak formal, langkah ini penting dalam konteks media sosial X.

3) *Tokenization*: Menggunakan pustaka NLTK, tokenisasi memecah kalimat menjadi unit kata atau token. Ini penting untuk memudahkan penghapusan stopwords dan menjaga kualitas vektor representasi dengan menghapus token yang sangat pendek atau tidak bermakna[26].

4) *Stopword Removal*: Stopword merujuk pada kata-kata umum seperti “yang”, “dan”, “itu”, yang sering muncul tetapi tidak memberikan kontribusi makna dalam analisis sentimen. Pustaka stopwords dari Sastrawi diterapkan dalam penelitian ini dan disesuaikan dengan ciri bahasa di media sosial. Langkah-langkah ini juga digunakan dalam studi oleh [21], [25].

5) *Stemming*: Stemming mengembalikan setiap kata ke bentuk aslinya (root word). Penelitian ini menggunakan algoritma stemming Sastrawi untuk bahasa Indonesia. Contohnya, kata "menampilkan", "ditampilkan", dan "penampilan" kembali ke kata dasar "tampil". Studi [17] dan [17] juga menunjukkan bahwa stemming sangat berfungsi untuk mengurangi dimensi serta meningkatkan akurasi klasifikasi sentimen dalam bahasa Indonesia.

6) *Empty Text Removal*: Setelah semua tahapan pra-pemrosesan, dokumen yang kosong atau hanya berisi kata-kata tanpa makna akan dihapus. Langkah ini krusial untuk mempertahankan mutu data latihan, sebagaimana diterapkan dalam kajian analisis RUU Kesehatan oleh [18]. Semua tahap di atas dirancang dengan memperhatikan karakteristik unik dari data media sosial, yang sangat berbeda dari data formal. Penggabungan teknik ini telah terbukti berhasil dalam penelitian sebelumnya yang juga memanfaatkan algoritma TF-IDF dan FastText untuk representasi fitur [17],[19], [21].

Berbeda dengan model konvensional, IndoBERT tidak memerlukan banyak pra-pemrosesan untuk mempertahankan struktur kata. Proses pembersihan sederhana dilakukan, menghapus spasi ganda, emoji, hashtag, mention, dan URL. Karena IndoBERT menggunakan tokenisasi sub-word, yang mampu menangani variasi morfologi kata, tidak ada stemming atau penghapusan stopword. Metode ini sejalan dengan temuan studi sebelumnya yang menggunakan IndoBERT untuk masalah sosial dan kebijakan publik [14], [15].

### C. Labeling

Setelah menyelesaikan tahap pre-pemrosesan data, langkah berikutnya adalah pemberian label (labeling) pada data. Pada penelitian ini, peneliti melakukan pelabelan secara manual terhadap setiap tweet yang telah dibersihkan. Proses pelabelan ini membagi tweet ke dalam tiga kategori sentimen: positif, negatif, dan netral, berdasarkan makna konteks dan ekspresi emosional dari tweet itu. Penandaan manual dipilih karena memberikan interpretasi yang lebih tepat terhadap konteks bahasa santai yang digunakan oleh pengguna di media sosial X. Setiap tweet dianalisis secara individual untuk mengenali nuansa emosionalnya, terutama karena susunan kalimat dalam tweet seringkali singkat, menggunakan bahasa gaul, atau mengandung ironi dan sarkasme yang sulit dipahami oleh metode otomatis [23], [26].

Penelitian ini menggunakan format klasifikasi multi-class, dengan tiga label sentimen utama (positif, netral, negatif). Pemilihan format multi-class dianggap lebih sesuai dibanding binary classification, karena opini publik terhadap gambar AI generatif sering kali bersifat netral atau ambigu dan tidak dapat direduksi hanya menjadi positif atau negatif.

Pendekatan ini sejalan dengan studi sebelumnya yang juga mengandalkan proses pelabelan manual untuk memastikan kualitas data pelatihan dan pengujian dalam domain analisis sentimen berbasis media sosial [23]. Di samping itu, hasil pelabelan secara manual berperan krusial dalam menciptakan dataset yang tepat, terutama saat diterapkan dalam proses pembelajaran mesin dengan algoritma klasifikasi seperti SVM dan Naïve Bayes [26].

Distribusi data dari proses pelabelan menunjukkan bahwa sentimen netral mendominasi jumlah tweet dalam kumpulan data. Pernyataan ini sejalan dengan hasil penelitian [19] yang menunjukkan bahwa sebagian besar pengguna media sosial X cenderung berbagi pendapat dengan cara informatif tanpa mengekspresikan emosi secara langsung.

Studi IndoBERT tentang masalah PPN [14] dan ulasan produk e-commerce [15] juga menggunakan strategi ini.

### D. Feature Extraction

Ekstraksi fitur adalah langkah krusial dalam analisis sentimen karena menentukan cara teks diubah menjadi vektor angka yang bisa diproses oleh algoritma pengklasifikasi. Representasi yang akurat akan berdampak pada ketepatan model, terutama pada data teks singkat dan santai seperti tweet [17], [19], [20]. Dalam studi ini,

diterapkan dua metode representasi teks, yaitu Term Frequency–Inverse Document Frequency (TF-IDF) dan FastText, yang masing-masing memiliki keunggulan dalam menangkap ciri-ciri teks untuk tujuan klasifikasi sentimen [21], [26].

TF-IDF merupakan teknik statistik yang menilai seberapa signifikan suatu kata dalam sebuah dokumen dibandingkan dengan keseluruhan korpus. Nilai TF merepresentasikan frekuensi kemunculan kata dalam dokumen, sedangkan IDF menggambarkan seberapa langka kata itu muncul dalam seluruh dokumen. TF-IDF kerap dipakai dalam analisis sentimen karena dapat menyeimbangkan istilah yang umum dan spesifik dalam teks, sehingga menciptakan representasi yang lebih relevan bagi model klasifikasi [19], [21]. Dalam penerapannya, fitur TF-IDF digunakan untuk menggambarkan tweet dalam format matriks vektor berdimensi tinggi. Setiap tweet diubah menjadi vektor berdasarkan berat setiap kata yang berhubungan. Studi sebelumnya mengungkapkan bahwa TF-IDF sangat efektif dalam memodelkan data teks singkat seperti tweet, terutama bila digabungkan dengan algoritma Naïve Bayes [19], [22].

Nilai TF-IDF dapat dihitung secara matematis dengan menggunakan rumus berikut:

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \cdot \text{IDF}(t, D)$$

dengan definisi masing-masing komponen sebagai berikut:

Term-Frequency:

$$\text{TF}(t, d) = \frac{\text{freq}(t, d)}{\sum_{t' \in d} \text{freq}(t', d)}$$

Inverse Document Frequency:

$$\text{IDF}(t, D) = \log \left( \frac{|D|}{|\{d \in D: t \in d\}|} \right) + 1$$

Tidak seperti TF-IDF yang bergantung pada frekuensi kata, FastText merupakan metode word embedding yang diciptakan oleh Facebook AI Research. FastText menghasilkan representasi kata sebagai vektor berdimensi rendah yang menjaga makna semantik. Keunggulan FastText ada pada kemampuannya mengenali sub-kata dan menangani kata baru (out-of-vocabulary), yang sangat berharga dalam konteks bahasa tidak formal seperti tweet [17]. Dalam penelitian ini, FastText digunakan untuk menciptakan representasi vektor dari setiap kata di dalam tweet, yang selanjutnya dirata-ratakan untuk membentuk representasi keseluruhan tweet. Representasi ini dianggap lebih efektif dalam menangkap konteks makna dibandingkan pendekatan yang hanya mengandalkan frekuensi murni [9], [26]. Kedua metode ini digunakan untuk membandingkan efektivitas pendekatan statistik (TF-IDF) dan pendekatan semantik (FastText) dalam pembangunan model klasifikasi sentimen dengan algoritma Naïve Bayes dan SVM. Kombinasi seperti ini telah digunakan dalam beberapa penelitian sebelumnya dan terbukti efektif dalam meningkatkan kinerja klasifikasi [17], [19], [20].

Dalam penelitian ini, representasi vektor dari setiap kata dalam tweet dibuat menggunakan FastText. Vektor teks akhir dihitung sebagai rata-rata terbobot dari vektor kata

secara keseluruhan, dengan bobot berasal dari nilai IDF yang dihasilkan dari ekstraksi TF-IDF. Formula yang digunakan adalah sebagai berikut:

$$v_{tx} = \frac{\sum_{t \in \text{tokens}} \text{IDF}(t) \cdot v_t}{\sum_{t \in \text{tokens}} \text{IDF}(t)}$$

Jika model FastText atau kamus TF-IDF tidak menemukan token dalam tweet, digunakan vektor nol berdimensi 300:

$$v_{tx} = 0 \in R^{300}$$

Berbeda dengan metode klasik, pada IndoBERT tidak dilakukan ekstraksi fitur manual. Tokenisasi dilakukan menggunakan tokenizer indobenchmark/indobert-base-pl, yang memungkinkan pemrosesan teks secara langsung hingga panjang maksimum 128 token. Selanjutnya, arsitektur transformer mempelajari representasi sub-word tersebut secara end-to-end selama proses pelatihan. Pendekatan ini terbukti mampu menghasilkan representasi teks sosial yang lebih kaya dan kontekstual dalam berbagai domain aplikasi [14],[15], [16].

Dalam penelitian ini, embedding FastText digunakan untuk menggunakan model bahasa Indonesia yang telah dilatih sebelumnya oleh AI Facebook (cc.id.300.vec). Model ini tidak dilatih ulang, tetapi digunakan secara langsung untuk menghasilkan representasi vektor kata. Oleh karena itu, informasi semantik bahasa Indonesia sudah sangat luas, sehingga dapat membantu klasifikasi tanpa memerlukan pengetahuan tambahan tentang korpus yang terbatas [17].

#### E. Data Balancing

Ketidakseimbangan kelas dalam dataset sentimen adalah masalah umum dalam penelitian analisis sentimen, terutama saat jumlah data di kelas mayoritas (misalnya netral) sangat lebih banyak dibandingkan kelas minoritas (positif atau negatif). Kesenjangan ini dapat membuat model klasifikasi cenderung pada kelas mayoritas dan mengurangi kemampuan generalisasi dari model tersebut. Oleh sebab itu, diperlukan teknik untuk menyeimbangkan data agar kinerja model pada data minoritas lebih baik [21], [22].

Dalam studi ini, digunakan metode Synthetic Minority Over-sampling Technique (SMOTE) untuk menangani ketidakseimbangan antara kelas. SMOTE berfungsi dengan menghasilkan sampel sintetis baru dari kelas minoritas melalui interpolasi antara sampel yang sudah ada, sehingga dapat memperbaiki representasi kelas itu dalam data pelatihan. Metode ini telah terbukti berhasil dalam memperbaiki distribusi kelas dan telah diterapkan dalam banyak penelitian sebelumnya untuk meningkatkan kinerja model klasifikasi di bidang teks [21].

Penerapan SMOTE dilaksanakan sesudah proses ekstraksi fitur, baik pada representasi TF-IDF maupun kombinasi TF-IDF dan FastText. Dataset yang telah di-oversampling selanjutnya digunakan sebagai input untuk melatih model Naïve Bayes dan SVM. Tujuan langkah ini adalah untuk memastikan model dapat mempelajari distribusi fitur secara merata dari seluruh kelas sentimen dan menghindari dominasi kelas utama selama proses pembelajaran [22].

SMOTE menghasilkan sampel sintetis dari kelas minoritas melalui interpolasi linier antara satu titik data dan titik data yang berdekatan:

$$x_{nw} = x_i + \lambda \cdot (x_{nn} - x_i)$$

Penelitian [21] secara tegas menunjukkan bahwa penggabungan SVM dengan teknik penyeimbangan SMOTE-Tomek Link dapat meningkatkan akurasi klasifikasi dalam konteks ulasan pelanggan, dan prinsip yang serupa diterapkan dalam studi ini. Oleh karena itu, penerapan SMOTE menjadi elemen krusial dalam rangkaian klasifikasi untuk membangun model yang lebih tepat dan setara terhadap semua tipe sentimen.

Pada IndoBERT, SMOTE tidak digunakan untuk menyeimbangkan distribusi kelas, tetapi bobot kelas juga dikenal sebagai bobot kelas disesuaikan pada fungsi penurunan cross-entropy untuk menangani ketidakseimbangan. Akibatnya, kelas minoritas dihukum lebih berat daripada kelas mayoritas. Metode ini dipilih karena tidak mengubah distribusi data asli; sebaliknya, kontribusi tiap kelas dalam proses pelatihan hanya diubah. Oleh karena itu, pendekatan ini lebih cocok untuk model berbasis transformer yang peka terhadap objek sintetis pada tingkat data [16].

#### F. Data Splitting

Setelah melakukan pra-pemrosesan, ekstraksi fitur, dan penyeimbangan data, tahap berikutnya adalah membagi dataset menjadi dua bagian: data pelatihan dan data pengujian. Dalam penelitian ini, proporsi distribusi data yang diterapkan adalah 80% untuk melatih model dan 20% untuk menguji kinerja model. Pemilihan rasio ini merujuk pada praktik standar dalam studi analisis sentimen karena dapat memberikan hasil pelatihan yang cukup mewakili tanpa mengorbankan data uji yang digunakan untuk penilaian [19], [22].

Pembagian ini dilakukan pada dua jenis representasi fitur, yaitu hanya TF-IDF dan kombinasi TF-IDF + FastText, yang telah diseimbangkan dengan metode SMOTE. Data yang direpresentasikan menggunakan TF-IDF diaplikasikan untuk melatih model Complement Naïve Bayes (CNB), sementara data yang digabungkan dimanfaatkan untuk model Support Vector Machine (SVM). Setelah data dibagi, model dilatih dengan menggunakan data pelatihan dan dievaluasi dengan data pengujian. Evaluasi dilakukan dengan menggunakan metrik umum dalam klasifikasi, seperti akurasi, presisi, recall, F1-score, dan matriks kebingungan. Agar hasil evaluasi yang diperoleh valid dan kemungkinan overfitting berkurang, teknik validasi silang k-fold cross-validation diterapkan, dengan k = 5 untuk model CNB dan k = 10 untuk model SVM [18], [22].

Penelitian ini menggunakan train-test split 80:20 (stratified) untuk CNB dan SVM, sedangkan untuk IndoBERT menggunakan 80:10:10 (train, validasi, test).

Complement Naïve Bayes (CNB) dipilih karena kemampuannya untuk mengatasi distribusi data teks yang

tidak seimbang dan telah digunakan luas dalam berbagai penelitian klasifikasi teks [22], [23].

Support Vector Machine (SVM) digunakan karena efektif dalam menangani data berdimensi tinggi dan memiliki performa baik dalam memisahkan kelas secara optimal, terutama pada representasi fitur gabungan seperti TF-IDF dan FastText [17], [19], [26].

Untuk jalur IndoBERT, dataset dibagi menjadi 80% data latihan, 10% data validasi, dan 10% data uji. Ini memastikan bahwa distribusi kelas sentimen pada masing-masing subset tetap proporsional. Selanjutnya, untuk menerapkan early stopping dan memilih model terbaik berdasarkan skor Macro-F1, subset validasi digunakan. Strategi ini sejalan dengan praktik umum dalam penelitian berbasis IndoBERT di bidang kebijakan publik dan kebencanaan, yang menekankan pentingnya subset validasi untuk proses fine-tuning model transformer [14], [16]. Selain itu, Macro-F1 sebagai metrik seleksi dianggap lebih representatif daripada metrik seleksi sederhana karena mampu memberikan penilaian seimbang pada dataset dengan distribusi kelas tidak merata, yang sering terjadi pada data media sosial.

#### G. Model Training

Pelatihan dilakukan setelah dataset dibagi dan disusun sesuai dengan representasi fitur yang ada. Dalam studi ini, dua algoritma diterapkan untuk mengembangkan model klasifikasi sentimen, yaitu Complement Naïve Bayes (CNB) dan Support Vector Machine (SVM). Model CNB dilatih dengan representasi fitur TF-IDF, sementara model SVM dilatih menggunakan representasi fitur gabungan antara TF-IDF dan FastText.

Untuk dokumen berikut, Complement Naïve Bayes (CNB) menghitung probabilitas kelas:

$$P(c|d) \propto P(c) \cdot \prod_{t \in d} P(t|c)^{TF(t,d)}$$

Dengan mengetahui kemungkinan kata \$t\$ terhadap komplemen kelas \$c\$, hasilnya dapat dihitung sebagai berikut:

$$P(t|c) = \frac{\sum_{d' \in c} TF(t, d') + \alpha}{\sum_{d' \in c} \sum_{t'} TF(t', d') + \alpha \cdot |V|}$$

Kelas akhir dipilih berdasarkan:

$$\hat{c} = \arg \max_c P(c|d)$$

Hyperplane digunakan oleh Support Vector Machine (SVM) untuk memisahkan kelas:

$$w \cdot x + b = 0$$

Fungsi keputusan untuk klasifikasi adalah:

$$f(x) = \text{sign}(w \cdot x + b)$$

Untuk menangani data non-linear, digunakan kernel RBF dengan fungsi:

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2)$$

Dalam implementasi praktis, algoritma SVM menggunakan kernel Radial Basis Function (RBF) dengan parameter regulasi \$C = 1.0\$ dan nilai gamma = scale sesuai konfigurasi standar scikit-learn. Pemilihan kernel RBF dilakukan karena sifat data teks yang non-linear, sementara

nilai \$C\$ dipilih untuk menjaga keseimbangan antara margin maksimal dan kesalahan klasifikasi.

Sementara itu, algoritma Complement Naïve Bayes (CNB) dilatih dengan representasi TF-IDF menggunakan smoothing parameter \$\alpha = 1.0\$ (Laplace smoothing). Parameter ini digunakan untuk menghindari probabilitas nol pada kata yang jarang muncul dalam suatu kelas dan telah terbukti efektif dalam klasifikasi teks tidak seimbang.

Kedua model diuji menggunakan data uji yang telah dipisahkan sebelumnya, dan kinerjanya dievaluasi dengan sejumlah metrik klasifikasi, yaitu:

Akurasi: menghitung rasio prediksi yang tepat dibandingkan total data pengujian.

Presisi: menilai sejauh mana prediksi positif yang dibuat benar-benar sesuai.

Recall: menilai kemampuan model untuk menangkap seluruh data yang relevan dari kelas yang sebenarnya.

F1-score: adalah rata-rata harmonis dari presisi serta recall, mencerminkan keseimbangan di antara keduanya.

Confusion Matrix: menampilkan distribusi klasifikasi yang benar dan yang salah untuk tiap kelas.

GridSearchCV (cv=10, 24 kandidat, seleksi Macro-F1) mengatur SVM pada ruang {kernel: linear/rbf; C: 1,10,100,1000; gamma: scale/auto (RBF); class\_weight: None/balanced}. Konfigurasi terbaik adalah RBF dengan C=1000, gamma=scale, dan class\_weight=balanced. CNB dituning pada \$\alpha \in \{0.1, 0.5, 1.0, 2.0\}\$, dengan nilai terbaik adalah 1.0. Stabilitas kinerja ditunjukkan melalui CV lima bagian yang berbeda.

Selain evaluasi menggunakan data uji, juga dilakukan validasi silang (k-fold cross-validation) untuk menguji ketahanan model pada data lainnya. Pada model CNB diterapkan validasi silang dengan \$k = 5\$, demikian juga pada model SVM menggunakan \$k = 5\$. Penentuan nilai \$k\$ ini sejalan dengan praktik standar dalam literatur untuk memelihara keseimbangan antara bias dan variansi [18], [22].

Hasil validasi silang menunjukkan bahwa model CNB memperoleh akurasi sebesar 0.5835, 0.5640, 0.5207, 0.6467, dan 0.6756, dengan rata-rata 0.5981 dan standar deviasi 0.0561. Sementara itu, model SVM dengan representasi gabungan TF-IDF dan FastText memperoleh akurasi sebesar 0.6124, 0.6074, 0.5930, 0.6653, dan 0.8079, dengan rata-rata 0.6572 dan standar deviasi 0.0816. Hasil ini menunjukkan bahwa meskipun SVM memiliki akurasi lebih tinggi dibanding CNB, variasi antar fold yang lebih besar membuat kinerjanya cenderung lebih fluktuatif.

Model CNB menunjukkan kinerja yang konsisten dalam klasifikasi teks yang menggunakan TF-IDF, khususnya dalam mengatasi distribusi kata yang tidak merata. Di sisi lain, model SVM yang dilatih dengan kombinasi TF-IDF dan FastText dapat memanfaatkan informasi semantik secara lebih mendalam, sehingga menghasilkan prediksi yang lebih tepat dan konsisten, sebagaimana ditemukan dalam studi sebelumnya [17], [19].

Selain CNB dan SVM, penelitian ini juga memakai baseline modern IndoBERT sebagai perbandingan.



IndoBERT dilatih dengan arsitektur transformer indobenchmark/indobert-base-p1 untuk pengklasifikasian tiga jenis sentimen (negatif, netral, positif).

Sebelum pelatihan, dataset dibagi secara stratifikasi menjadi 80% untuk pelatihan, 10% untuk validasi, dan 10% untuk pengujian. Subset validasi digunakan untuk menerapkan penghentian awal serta memilih model terbaik berdasar nilai Macro-F1, yang dianggap lebih mencerminkan pada data yang tidak seimbang [24],[27].

Pelatihan IndoBERT dilaksanakan dengan pengaturan sebagai berikut:

Optimisasi: Pengoptimal AdamW dengan laju pembelajaran  $2e-5$ , pengurangan bobot 0.01, rasio pemanasan 0.1, dan pemotongan gradien 1.0.

Ukuran batch: 16 (pelatihan) dan 32 (penilaian).

Epoch: 3 siklus pelatihan, dengan penghentian awal (kesabaran=2) untuk menghindari overfitting.

Fungsi kerugian: Cross-Entropy dengan bobot kelas untuk menyeimbangkan kontribusi setiap kelas, menggantikan penerapan SMOTE pada metode konvensional [24].

Tokenisasi: diterapkan menggunakan tokenizer IndoBERT dengan panjang maksimum 128 token, truncation aktif, dan padding yang dinamis.

Selama proses pelatihan, model dievaluasi dengan menggunakan metrik validasi utama yang berupa Macro-F1. Model yang berkinerja terbaik pada data validasi selanjutnya diuji pada set pengujian. Evaluasi akhir meliputi metrik Accuracy, Macro-F1, Precision, Recall, dan analisis menggunakan classification report serta confusion matrix.

Hasil pengujian menunjukkan bahwa IndoBERT dapat menangkap representasi semantik yang lebih rumit dibandingkan metode klasik, meskipun kinerjanya masih dipengaruhi oleh ukuran dataset dan adanya bahasa informal. Hasil ini konsisten dengan studi sebelumnya yang mengungkapkan keunggulan IndoBERT dalam ranah e-commerce [15] dan bencana [16].

#### H. Model Evaluation

Setelah tahap pelatihan, dilakukan penilaian terhadap performa kedua model klasifikasi, yaitu Complement Naïve Bayes (CNB) dan Support Vector Machine (SVM). Evaluasi dilaksanakan berdasarkan hasil perkiraan pada data uji, dengan memanfaatkan metrik akurasi, presisi, recall, F1-score, serta matriks kebingungan. Selain itu, dilakukan juga validasi silang (k-fold cross-validation) untuk menilai stabilitas kinerja model secara keseluruhan. Tabel di bawah ini menyajikan ikhtisar hasil penilaian model:

TABEL I  
PERBANDINGAN MODEL

| Model                          | Akurasi | Presisi (avg) | F1-Score (avg) |
|--------------------------------|---------|---------------|----------------|
| <b>CNB + TF-IDF</b>            | 0.641   | 0.64          | 0.64           |
| <b>SVM + TF-IDF + FastText</b> | 0.687   | 0.69          | 0.68           |

Model Complement Naïve Bayes menunjukkan tingkat akurasi 64,1% dengan rata-rata F1-score sebesar 0,64, yang mengindikasikan bahwa model ini cukup efektif dalam mengenali kelas-kelas sentimen secara seimbang, meskipun masih ada peluang untuk perbaikan. Hasil dari validasi silang dengan rata-rata nilai 59,8% menunjukkan bahwa kinerja model ini masih cenderung berbeda-beda pada data yang beragam. Di sisi lain, model SVM yang dilatih dengan kombinasi fitur TF-IDF dan FastText mencapai hasil yang lebih baik, dengan akurasi mencapai 68,7% dan rata-rata F1-score 0,68. Nilai cross-validation sebesar 65,7% menunjukkan bahwa model SVM memiliki stabilitas performa yang lebih unggul daripada CNB. Ini menunjukkan bahwa penggabungan representasi statistik dan semantik dalam fitur teks dapat meningkatkan ketepatan prediksi pada tugas klasifikasi sentimen, terutama pada data media sosial yang kompleks dan tidak formal.

Model IndoBERT dievaluasi sebagai baseline modern bersama dengan CNB dan SVM. Data yang diuji dievaluasi dengan metrik akurasi, ketepatan, ulang, dan macro-F1, serta laporan klasifikasi per kelas dan matrix confusion. Karena proses fine-tuning transformer biasanya menggunakan pembagian data eksplisit (train, val, atau split test), tidak dilakukan validasi silang k-fold [14], [15], [16].

TABEL 2  
HASIL EVALUASI MODEL INDOBERT

| Model           | Akurasi | Presisi (avg) | Recall (avg) | F1-Score (avg) |
|-----------------|---------|---------------|--------------|----------------|
| <b>IndoBERT</b> | 0.582   | 0.581         | 0.578        | 0.566          |

Model IndoBERT mencapai akurasi 58,2% dengan rata-rata Macro-F1 sebesar 0,566. Dibandingkan dengan model klasik, performa IndoBERT pada dataset ini masih tergolong belum maksimal. Analisis matriks kebingungan menunjukkan bahwa model masih sering salah dalam membedakan kelas positif dan netral, sedangkan kelas negatif terklasifikasi dengan lebih konsisten.

Kinerja IndoBERT yang tidak optimal ini mungkin diakibatkan oleh ukuran dataset yang kecil, ketidakseimbangan distribusi kelas, serta sifat bahasa nonformal pada platform media sosial. Hasil ini konsisten dengan studi sebelumnya yang menekankan bahwa transformer membutuhkan volume data yang besar untuk mencapai kinerja terbaik [14],[16].

Walaupun demikian, IndoBERT masih memperlihatkan kemampuan untuk menangkap representasi semantik yang lebih mendalam dibandingkan metode frekuensi (TF-IDF) atau embedding yang bersifat statis (FastText). Dengan data yang lebih banyak atau metode augmentasi tambahan, performa IndoBERT bisa melebihi model-model klasik, seperti yang ditunjukkan dalam penelitian di bidang e-commerce [15] dan juga kebencanaan [16].

### III. HASIL DAN PEMBAHASAN

#### A. Distribusi Sentimen

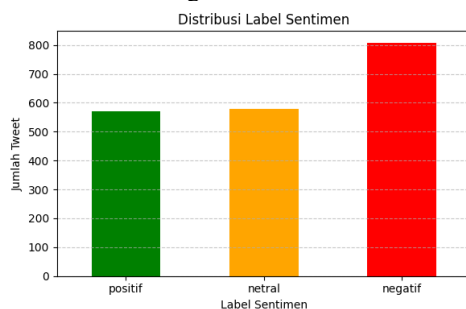
Setelah melakukan pra-pemrosesan, ekstraksi fitur, dan penyeimbangan data, tahap berikutnya adalah membagi dataset menjadi dua bagian: data pelatihan dan data pengujian. Dalam penelitian ini, proporsi distribusi data yang diterapkan adalah 80% untuk melatih model dan 20% untuk menguji kinerja model. Pemilihan rasio ini merujuk pada praktik standar dalam studi analisis sentimen karena dapat memberikan hasil pelatihan yang cukup mewakili tanpa mengorbankan data uji yang digunakan untuk penilaian [19], [22].

Setelah proses pelabelan manual selesai pada 1.958 tweet yang diperoleh dari media sosial X, data tersebut dibagi menjadi tiga kategori sentimen: positif, negatif, dan netral. Distribusi sentimen ditampilkan pada Tabel 3 berikut ini:

TABEL 3  
DISTRIBUSI SENTIMEN

| Sentimen       | Jumlah tweet | Presentase (%) |
|----------------|--------------|----------------|
| <i>Negatif</i> | 808          | 41.26%         |
| <i>Netral</i>  | 579          | 29.57%         |
| <i>Positif</i> | 571          | 29.17%         |
| <i>Total</i>   | <b>1.958</b> | <b>100%</b>    |

Dari tabel tersebut, tampak bahwa sentimen negatif mendominasi dengan proporsi mencapai 41,26%, diikuti oleh sentimen netral dan positif yang memiliki persentase hampir seimbang, yaitu masing-masing 29,57% dan 29,17%. Hasil penelitian ini menunjukkan bahwa masyarakat cenderung bersikap kritis terhadap teknologi gambar AI generatif. Ini sejalan dengan penelitian [19], yang menunjukkan bahwa pengguna media sosial X cenderung lebih berani dalam menyatakan ketidakpuasan terhadap isu-isu kontroversial atau inovasi teknologi baru.



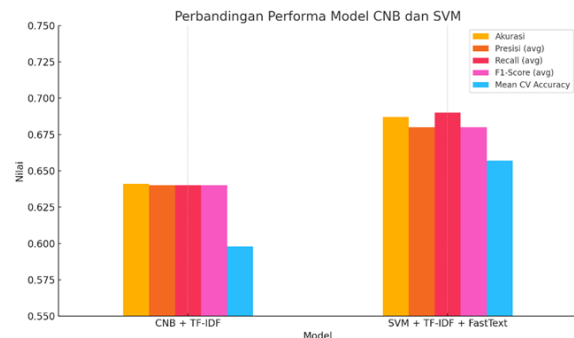
Gambar 4. Distribusi label sentimen.

#### B. Analisis Hasil Klasifikasi

Pada tahap ini, dilakukan evaluasi terhadap hasil klasifikasi dari dua model yang diterapkan, yaitu Complement Naïve Bayes (CNB) dan Support Vector Machine (SVM). Evaluasi dilakukan dengan menggunakan data uji dan validasi silang (cross-validation) pada data yang sudah diproses dan disesuaikan keseimbangannya. Tabel di bawah ini menyajikan ringkasan hasil penilaian kedua model:

TABEL 4  
PERBANDINGAN MODEL KLASIK

| Model                          | Akurasi | Presisi ( avg ) | F1-Score ( avg ) |
|--------------------------------|---------|-----------------|------------------|
| <i>CNB + TF-IDF</i>            | 0.641   | 0.64            | 0.64             |
| <i>SVM + TF-IDF + FastText</i> | 0.687   | 0.69            | 0.68             |



Gambar 5. Perbandingan metrik evaluasi antara model CNB dan SVM.

Umumnya, model SVM menunjukkan hasil yang lebih baik dibandingkan CNB pada semua metrik evaluasi. Hal ini menunjukkan bahwa penggabungan antara TF-IDF dan FastText sebagai representasi fitur dapat memahami konteks semantik dengan lebih baik dibandingkan hanya menggunakan TF-IDF, terutama pada teks tidak formal seperti tweet. Model CNB masih menunjukkan hasil yang sangat kompetitif, terutama dalam aspek kecepatan pelatihan dan efisiensi sumber daya. Akan tetapi, SVM memiliki keunggulan dalam stabilitas prediksi, terlihat dari nilai rata-rata cross-validation yang lebih tinggi (65,7% dibandingkan 59,8%). Penemuan ini mendukung studi sebelumnya [12], [14], [16], yang menunjukkan bahwa representasi berdasarkan FastText dapat meningkatkan presisi dan ketahanan model klasifikasi berbasis SVM dalam mengelola data teks dari platform media sosial.

TABEL 5  
HASIL EVALUASI MODEL INDOBERT

| Model           | Akurasi | Presisi ( avg ) | Recall ( avg ) | F1-Score ( avg ) |
|-----------------|---------|-----------------|----------------|------------------|
| <i>IndoBERT</i> | 0.582   | 0.581           | 0.578          | 0.566            |

Akurasi IndoBERT pada dataset ini belum melampaui metode SVM konvensional; dengan Macro-F1 sebesar 0,566, IndoBERT mencapai 58,2%, lebih rendah dari SVM (68,7%, 0,68). Ini menunjukkan bahwa, seperti yang ditunjukkan oleh penelitian sebelumnya [14], [15], dan [16], dominasi bahasa informal dan ukuran dataset yang terbatas sangat mempengaruhi kinerja transformer. Namun, IndoBERT masih memiliki keunggulan dalam menampilkan makna semantik yang lebih kompleks, sehingga model ini dapat melampaui metode konvensional untuk data yang lebih besar.

#### C. Analisis Confusion Matrix



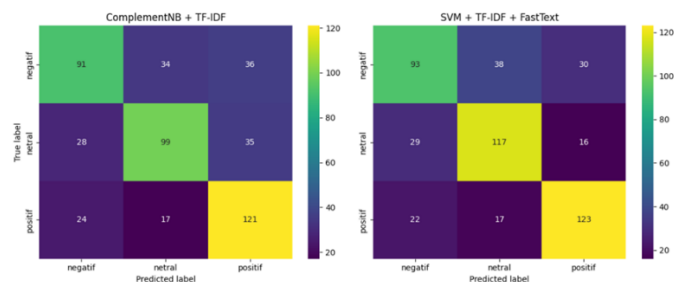
Analisis matriks kebingungan dilakukan untuk menggali lebih dalam kinerja model dalam mengkategorikan setiap jenis sentimen: positif, netral, dan negatif. Matriks ini menunjukkan jumlah prediksi yang benar (true positive) dan salah (false positive/false negative) untuk setiap kelas. Berdasarkan hasil dari pengujian data, berikut adalah matriks kebingungan dari kedua model:

TABEL 6  
CONFUSION MATRIX – CNB + TF-IDF

|               | Pred: Negatif | Pred: Netral | Pred: Positif |
|---------------|---------------|--------------|---------------|
| Asli: Negatif | 91            | 34           | 36            |
| Asli: Netral  | 28            | 99           | 35            |
| Asli: Positif | 24            | 17           | 121           |

TABEL 7  
CONFUSION MATRIX – SVM + TF-IDF + FastText

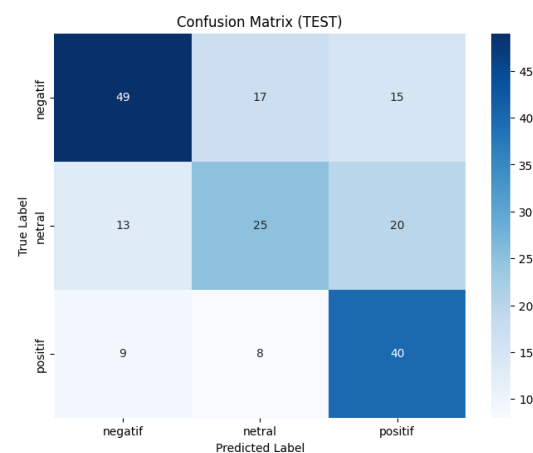
|               | Pred: Negatif | Pred: Netral | Pred: Positif |
|---------------|---------------|--------------|---------------|
| Asli: Negatif | 49            | 17           | 15            |
| Asli: Netral  | 13            | 25           | 20            |
| Asli: Positif | 9             | 8            | 40            |



Gambar 6. Confusion matrix model CNB dan SVM.

TABEL 8  
CONFUSION MATRIX INDOBERT

|               | Pred: Negatif | Pred: Netral | Pred: Positif |
|---------------|---------------|--------------|---------------|
| Asli: Negatif | 93            | 38           | 30            |
| Asli: Netral  | 29            | 117          | 16            |
| Asli: Positif | 22            | 17           | 123           |



Gambar 7. Confusion matrix model IndoBERT.

#### D. Format Teks

Agar dapat memahami lebih mendalam tentang karakteristik kesalahan model klasifikasi, analisis dilakukan terhadap beberapa contoh tweet yang diklasifikasikan salah (misclassification) oleh tiap model. Tujuan analisis ini adalah untuk menemukan pola umum dari kesalahan dan menilai kekurangan model dalam memahami konteks tweet.

TABEL 9  
CONTOH KASUS KESALAHAN PADA COMPLEMENT NAÏVE BAYES (TF-IDF)

| Clean Text   | Label Asli | Prediksi | Catatan  |
|--|------------|----------|--|
| <i>menyebalkan bagian humas yang mulai menggunakan gambar AI</i> | Negatif    | Positif  | Complement NB cenderung peka terhadap frekuensi kata dan tidak bisa menangkap konteks sarkasme atau ironi yang terkandung dalam kalimat          |
| <i>video ilustrasi ai yg menjadi kebanggaan</i>                  | Positif    | Netral   | Model tidak mampu menangkap kata-kata yang memiliki makna positif karena TF-IDF hanya memperhatikan bobot kata, bukan keseluruhan makna semantik |
| <i>kata lucu gambar ai minta hak cipta</i>                       | Negatif    | Netral   | Model kesulitan dalam memahami makna ironi; kata “lucu” secara statistik dapat dianggap netral atau positif meskipun konteksnya negatif          |

TABEL 10  
CONTOH KASUS KESALAHAN PADA (SVM + TF-IDF + FastText)

| Clean Text                                 | Label Asli | Prediksi | Catatan   |
|--|------------|----------|---|
| <i>menyebalkan bagian humas yang mulai</i> | Negatif    | Netral   | Walaupun FastText menyertakan representasi semantis, kalimat yang ambigu atau sarkastik tetap susah untuk |

|   |         |        |   |
|---|---------|--------|---|
|   |         |        | ditafsirkan dengan akurat.  |
| <i>video ilustrasi ai yg menjadi kebanggaan</i> | Positif | Netral | Model tidak dapat mendeteksi kata "kebanggaan" sebagai indikator sentimen positif; mungkin disebabkan oleh sebaran kata-kata sejenis di kategori netral.    |
| <i>yg kiri gambar ai min harus jangan pakai</i> | Negatif | Netral | SVM menghadapi tantangan dengan data yang mengandung noise atau kalimat yang tidak formal; konteks larangan tidak teridentifikasi sebagai sentimen negatif. |

TABEL 11  
CONTOH KASUS KESALAHAN PADA INDOBERT

| Clean Text   | Label Asli | Prediksi | Catatan   |
|--|------------|----------|---|
| <i>ai generate gemini hasil emang ralistic</i>   | Positif    | Netral   | Model gagal menangkap nuansa positif.             |
| <i>kalo gw pake ai buat bikin meme brainrot gpp emang yg open comission buat gambar brainrot art</i> | Netral     | Positif  | Bahasa gaul salah diinterpretasi sebagai positif. |
| <i>gambar ai banyak gaya</i>   | Negatif    | Positif  | Sarkasme tidak dikenali sebagai negatif.          |

Berbagai pola kesalahan yang ditemukan pada kedua model dapat dijelaskan sebagai berikut. Pertama, model masih mengalami kesulitan dalam menangani sarkasme dan ironi, karena baik representasi TF-IDF maupun FastText belum mampu memahami nuansa emosional secara kontekstual. Kedua, kata-kata yang memiliki makna ganda seperti "lucu", "gila", atau "keren" sering disalahartikan jika tidak dianalisis dalam konteks keseluruhan kalimat.

Ketiga, Complement Naïve Bayes (CNB) cenderung terlalu bergantung pada frekuensi kata, sehingga mudah salah klasifikasi jika teks tidak mengandung kata-kata yang sering muncul. Sementara itu, Support Vector Machine (SVM) meskipun lebih unggul dalam menangkap makna semantik, tetap mengalami kesulitan dalam memahami kalimat ambigu atau gaya bahasa informal.

Keempat, kalimat-kalimat yang sangat singkat atau tidak memiliki kata kunci yang kuat membuat model tidak memiliki cukup informasi untuk menentukan polaritas sentimen secara akurat.

Hasil analisis ini mengindikasikan bahwa pendekatan konvensional masih terbatas dalam memahami konteks emosional yang kompleks, dan bahkan model berbasis transformer seperti IndoBERT pun membutuhkan dataset yang lebih besar, lebih seimbang, serta normalisasi bahasa gaul yang lebih baik untuk mencapai performa optimal. Oleh karena itu, penelitian selanjutnya dapat mengintegrasikan teknik augmentasi data, memperluas kamus slang media sosial, dan mengeksplorasi model pembelajaran mendalam lainnya seperti LSTM atau RoBERTa yang dirancang khusus untuk menangani teks informal.

#### E. Analisis Signifikansi Perbedaan Performa Model

Pengujian signifikansi berpasang-pasang dilakukan pada set uji yang sama berukuran  $N = 485$  untuk memastikan bahwa perbedaan kinerja antara Complement Naïve Bayes (CNB) dan Support Vector Machine (SVM) bukanlah hasil dari variasi acak. Dua tes digunakan: McNemar (berbasis pasangan benar atau salah untuk masing-masing instansi uji) dan Wilcoxon yang ditandatangani pada skor cross-validation ( $k = 5$ ). Ada ambang signifikansi  $\alpha = 0,05$ .

Hasil McNemar (test set identik, CNB versus SVM). Tabel kontingensi  $2 \times 2$  berikut terdiri dari pasangan keputusan model benar atau salah kedua untuk setiap instansi uji:

TABEL 12  
KONTINGENSI MCNEMAR

|           | SVM Benar | SVM Salah |
|-----------|-----------|-----------|
| CNB Benar | 263       | 48        |
| CNB Salah | 70        | 104       |

Hasil uji McNemar (tepat) adalah  $p = 0.0527$ . Pada test set ini, perbedaan performa CNB dan SVM tidak signifikan secara statistik karena  $p \geq 0,05$  dan  $p \geq 0,05$ . Namun, dalam kasus diskordan SVM yang unggul 70 dibandingkan 48 ( $\approx 59,3\%$  dibandingkan  $40,7\%$ ), tren SVM menunjukkan peningkatan, tetapi tidak cukup kuat secara statistik dengan ukuran dan karakter data saat ini.

Verifikasinya menggunakan Wilcoxon (skor CV  $k = 5$ ). Wilcoxon signed-rank atas skor cross-validation CNB versus SVM digunakan sebagai cross-check. Hasilnya menunjukkan kembali tidak signifikan pada  $\alpha = 0,05$  dan  $p = 0,0625$ . Hasil ini konsisten dengan McNemar dan mendukung kesimpulan bahwa kedua model memiliki kinerja yang sama pada data ini.

TABEL 13  
RINGKASAN P-VALUE UJI SIGNIFIKANSI

| Uji Statistik         | Model 1 | Model 2 | p-value | Signifikan ( $\alpha = 0,05$ )? |
|-----------------------|---------|---------|---------|---------------------------------|
| McNemar (test set)    | CNB     | SVM     | 0.0527  | Tidak                           |
| Wilcoxon (CV, $k=5$ ) | CNB     | SVM     | 0.0625  | Tidak                           |

Karena dijalankan pada jalur eksekusi yang berbeda, IndoBERT disajikan secara deskriptif sebagai baseline modern pada bagian hasil yang berbeda. Apabila prediksi IndoBERT dievaluasi pada test set yang identik, uji

signifikansi pairwise terhadap model klasik akan dilaporkan. Temuan "tidak signifikan" CNB dan SVM tetap informatif: karena ukuran data yang terbatas, ketidakseimbangan kelas, noise bahasa (slang, variasi ejaan, atau campuran bahasa), dan selisih kinerja antarmodel yang tipis, sampel ini sulit mendapatkan bukti statistik yang kuat.

#### F. Diskusi dan Keterbatasan Data

Hasil evaluasi menunjukkan bahwa model SVM yang menggunakan kombinasi TF-IDF dan FastText hanya mendapatkan akurasi sebesar 68,7%, sementara model CNB memiliki akurasi yang lebih rendah. Nilai ini tergolong moderat dibandingkan dengan standar umum penelitian analisis sentimen, yang umumnya dapat melebihi 70–80% pada dataset yang seimbang dan bersih.

Akurasi yang rendah bisa dijelaskan oleh berbagai faktor. Pertama, ukuran dataset tersebut terbatas, yaitu hanya 1.958 tweet setelah proses pra-pemrosesan. Jumlah ini tergolong sedikit untuk melatih model pembelajaran mesin, terutama dalam kasus klasifikasi multi-class dengan tiga kategori emosi. Kedua, ciri-ciri bahasa di media sosial yang ada dalam dataset ini menghasilkan banyak slang, singkatan, emotikon, dan campuran bahasa, sehingga representasi teks menjadi lebih bising (noise) dan susah dipahami oleh model yang berbasis statistik maupun embedding. Ketiga, terdapat ketidakseimbangan dalam distribusi kelas (class imbalance), di mana jumlah data didominasi oleh sentimen netral. Hal ini berdampak pada kinerja model dalam mengidentifikasi kelas minoritas (positif dan negatif), yang terlihat pada matriks kebingungan di mana banyak kesalahan klasifikasi muncul pada kelas netral.

Faktor-faktor ini menunjukkan bahwa mutu data sangat berpengaruh pada kesuksesan klasifikasi. Dalam penelitian ini, meskipun SVM dapat memanfaatkan representasi semantik dari FastText, keterbatasan dalam dataset membuat model belum mencapai optimalisasi. Pembahasan ini sejalan dengan hasil penelitian sebelumnya yang menyoroti bahwa kebisingan bahasa gaul dan data yang tidak seimbang adalah tantangan utama dalam analisis sentimen yang berbasis media sosial [24], [18].

#### IV. KESIMPULAN

Penelitian ini menganalisis sentimen mengenai teknologi gambar AI generatif di platform media sosial X dengan menggunakan metode klasifikasi teks. Tiga model klasifikasi diterapkan, yakni Complement Naïve Bayes (CNB) dengan representasi TF-IDF, Support Vector Machine (SVM) yang menggabungkan TF-IDF dan FastText, serta baseline modern IndoBERT sebagai perwakilan pendekatan berbasis transformer.

Evaluasi menunjukkan bahwa model SVM yang menggabungkan TF-IDF dan FastText menunjukkan kinerja paling unggul dibandingkan model-model lainnya. SVM menghasilkan akurasi sebesar 68,7%, dengan nilai rata-rata presisi, recall, dan F1-score masing-masing mencapai 0,69, 0,68, dan 0,68. Validasi silang menunjukkan konsistensi hasil

dengan rerata cross-validation sebesar 65,7%. Model CNB memperoleh akurasi 64,1%, dengan nilai rata-rata presisi, recall, dan F1-score sebesar 0,64, serta rata-rata cross-validation sebesar 59,8%.

Sementara itu, IndoBERT mencapai akurasi 58,2% dan nilai Macro-F1 sebesar 0,566, yang lebih rendah daripada SVM dan CNB dalam dataset terbatas ini. Analisis matriks kebingungan menunjukkan bahwa IndoBERT cukup konsisten dalam mendeteksi sentimen negatif, namun masih sering keliru dalam membedakan antara kelas netral dan positif. Temuan ini menunjukkan bahwa meskipun IndoBERT secara teori lebih baik dalam merepresentasikan makna semantik, kinerjanya sangat dipengaruhi oleh ukuran dataset, ketidakseimbangan distribusi kelas, dan keberadaan bahasa informal yang dominan.

Secara keseluruhan, studi ini mengindikasikan bahwa penggabungan representasi frekuensi (TF-IDF) dan semantik (FastText) pada SVM lebih berhasil untuk dataset kecil dan tidak formal, sedangkan IndoBERT mungkin lebih unggul dengan adanya dataset yang lebih luas dan seimbang.

Di masa depan, disarankan untuk melakukan penelitian lebih lanjut dengan menggunakan dataset yang lebih besar, seimbang, dan beragam, agar model dapat belajar dari distribusi data yang lebih mencerminkan. Di samping itu, penerapan teknik pelabelan semi-otomatis dapat meningkatkan efisiensi proses anotasi, terutama pada data yang berukuran besar. Penelitian selanjutnya juga bisa menyelidiki metode data augmentation yang dibuat khusus untuk mengatasi bahasa gaul atau slang di platform media sosial. Selain itu, algoritma berbasis deep learning yang lain, seperti LSTM atau varian transformer (contohnya IndoBERTweet dan IndoRoBERTa), perlu diuji untuk meningkatkan akurasi dan kemampuan generalisasi dalam analisis sentimen terhadap teks yang tidak formal.

#### DAFTAR PUSTAKA

- [1] A. D. Pratama and H. Hendry, "Analisa Sentimen Masyarakat Terhadap Penggunaan Chatgpt Menggunakan Metode Support Vector Machine (SVM)," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 1, pp. 327–338, Feb. 2024, doi: 10.29100/jipi.v9i1.4285.
- [2] Mukti M Kusairi and Surya Agustian, "SVM Method with FastText Representation Feature for Classification of Twitter Sentiments Regarding the Covid-19 Vaccination Program," *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, Nov. 2022, doi: <https://doi.org/10.31849/digitalzone.v13i2.11531>.
- [3] Y. K. Sari, F. Rozi, S. Muhyiddin, and F. Sukmana, "Sentiment Analysis Of Public Opinion On Application X (Twitter) In Indonesia Against Chatgpt Using Naïve Bayes Algorithm," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 4, pp. 2473–2484, Dec. 2024, doi: 10.29100/jipi.v9i4.7052.
- [4] F. Matheos Sarimole and K. Kudrat, "Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine," *Jurnal Sains dan Teknologi*, vol. 5, no. 3, pp. 783–790, Feb. 2024, doi: 10.55338/saintek.v5i3.2702.
- [5] R. F. Rahmanda, Y. Sibaroni, and S. S. Prasetyowati, "Effectiveness of Bi-GRU and FastText in Sentiment Analysis of

- Shopee App Reviews,” *Sinkron*, vol. 9, no. 1, pp. 444–454, Feb. 2025, doi: 10.33395/sinkron.v9i1.14474.
- [6] Muhammad Afif Raihan and Erwin Budi Setiawan, “Aspect Based Sentiment Analysis with FastText Feature Expansion and Support Vector Machine Method on Twitter,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 591–598, Aug. 2022, doi: 10.29207/resti.v6i4.4187.
- [7] Moh Hadi Subowo and Farikh Alzami, “Using 2024 election twitter data, sentiment analysis based on TF-IDF and Naïve Bayes,” *MONETER: Jurnal Keuangan dan Perbankan*, vol. 12, Jul. 2024, doi: <https://doi.org/10.32832/moneter.v12i2.789>.
- [8] T. Safitri, Y. Umaidah, and I. Maulana, “Analisis Sentimen Pengguna Twitter Terhadap Grup Musik BTS Menggunakan Algoritma Support Vector Machine,” *Journal of Applied Informatics and Computing*, vol. 7, no. 1, pp. 28–35, Jul. 2023, doi: 10.30871/jaic.v7i1.5039.
- [9] D. R. Wijaya, G. M. A. Sasmitha, and W. O. Vihikan, “Sentiment Analysis of Indonesian Citizens on Electric Vehicle Using FastText and BERT Method,” *Journal of Information Systems and Informatics*, vol. 6, no. 3, pp. 1360–1372, Sep. 2024, doi: 10.51519/journalisi.v6i3.784.
- [10] Inggit Restu Illahi and E. B. Setiawan, “Sentiment Analysis on Social Media Using Fasttext Feature Expansion and Recurrent Neural Network (RNN) with Genetic Algorithm Optimization,” *International Journal on Information and Communication Technology (IJoICT)*, vol. 10, no. 1, pp. 78–89, Jun. 2024, doi: 10.21108/ijoict.v10i1.905.
- [11] N. Z. B. Jannah and K. Kusnawi, “Comparison of Naïve Bayes and SVM in Sentiment Analysis of Product Reviews on Marketplaces,” *Sinkron*, vol. 8, no. 2, pp. 727–733, Mar. 2024, doi: 10.33395/sinkron.v8i2.13559.
- [12] I. Hendrawan Rifky, E. Utami, and A. Hartanto Dwi, “Analisis Perbandingan Metode Tf-Idf dan Word2vec pada Klasifikasi Teks Sentimen Masyarakat Terhadap Produk Lokal di Indonesia,” *Smart Comp: Jurnalnya Orang Pintar Komputer*, vol. 11, no. 3, Jul. 2022, doi: 10.30591/smartcomp.v11i3.3902.
- [13] A. Putri, S. Agustian, I. Afrianty, and F. Sains dan Teknologi, “Eksplorasi Fitur Fasttext, Tf-Idf Dan Indobert Pada Metode K-Nearest Neighbor Untuk Klasifikasi Sentimen,” 2025.
- [14] M. R. Manoppo *et al.*, “Analisis Sentimen Publik Di Media Sosial Terhadap Kenaikan Ppn 12% Di Indonesia Menggunakan Indobert,” *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 4, no. 2, pp. 152–163, May 2025, doi: 10.69916/jkbt.v4i2.322.
- [15] S. Aras, M. Yusuf, R. Y. Ruimassa, E. A. B. Wambrau, and E. B. Pala’langan, “Sentiment Analysis on Shopee Product Reviews Using IndoBERT,” *Journal of Information Systems and Informatics*, vol. 6, no. 3, pp. 1616–1627, Sep. 2024, doi: 10.51519/journalisi.v6i3.814.
- [16] S. M. Anugerah, R. Wijaya, and M. A. Bijaksana, “Sentimen Analysis Social Media for Disaster using Naïve Bayes and IndoBERT,” *INTEK: Jurnal Penelitian*, vol. 11, no. 1, pp. 51–58, Apr. 2024, doi: 10.31963/intek.v11i1.4771.
- [17] M. Mustasaruddin, E. Budianita, M. Fikry, and F. Yanto, “Klasifikasi Sentiment Review Aplikasi MyPertamina Menggunakan Word Embedding FastText dan SVM (Support Vector Machine),” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 3, p. 526, Mar. 2023, doi: 10.30865/json.v4i3.5695.
- [18] T. Widyanto, I. Ristiana, and A. Wibowo, “Komparasi Naïve Bayes dan SVM Analisis Sentimen RUU Kesehatan di Twitter,” *SINTECH (Science and Information Technology) Journal*, vol. 6, no. 3, pp. 147–161, Dec. 2023, doi: 10.31598/sintechjournal.v6i3.1433.
- [19] U. I. Arsyah, M. Pratiwi, and A. Muhammad, “Twitter Sentiment Analysis of Public Space Opinions using SVM and TF-IDF Methods,” *Indonesian Journal of Computer Science*, vol. 13, no. 1, Feb. 2024, doi: 10.33022/ijcs.v13i1.3594.
- [20] D. A. Kristiyanti and Sri Hardani, “Sentiment Analysis of Public Acceptance of Covid-19 Vaccines Types in Indonesia using Naïve Bayes, Support Vector Machine, and Long Short-Term Memory (LSTM),” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 3, pp. 722–732, Jun. 2023, doi: 10.29207/resti.v7i3.4737.
- [21] D. I. Sumantiawan, J. E. Suseno, and W. A. Syafei, “Sentiment Analysis of Customer Reviews Using Support Vector Machine and Smote-Tomek Links For Identify Customer Satisfaction,” *J. Sistem Info. Bisnis*, vol. 13, no. 1, pp. 1–9, Jun. 2023, doi: 10.21456/vol13iss1pp1-9.
- [22] M. A. Saddam, E. Kurniawan D, and I. Indra, “Analisis Sentimen Fenomena PHK Massal Menggunakan Naive Bayes dan Support Vector Machine,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 8, no. 3, pp. 226–233, Sep. 2023, doi: 10.30591/jpit.v8i3.4884.
- [23] R. A. Sitorus and I. Zufria, “Application of the Naïve Bayes Algorithm in Sentiment Analysis of Using the Shopee Application on the Play Store,” *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 15, no. 1, pp. 53–66, May 2024, doi: 10.31849/digitalzone.v15i1.19828.
- [24] N. A. Maulana and D. Darwis, “Perbandingan Metode SVM dan Naïve Bayes untuk Analisis Sentimen pada Twitter tentang Obesitas di Kalangan Gen Z,” *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 3, pp. 655–666, Mar. 2025, doi: 10.52436/1.jpti.691.
- [25] S. A. Amira, S. Utama, and M. H. Fahmi, “Penerapan Metode Support Vector Machine untuk Analisis Sentimen pada Review Pelanggan Hotel,” *Edu Komputika Journal*, vol. 7, no. 2, pp. 40–48, Dec. 2020, doi: 10.15294/edukomputika.v7i2.42608.
- [26] M. F. Madjid, D. E. Ratnawati, and B. Rahayudi, “Sentiment Analysis on App Reviews Using Support Vector Machine and Naïve Bayes Classification,” *Sinkron*, vol. 8, no. 1, pp. 556–562, Feb. 2023, doi: 10.33395/sinkron.v8i1.12161.