

Perancangan Sistem Analisis Sentimen Berita Konflik Internasional Berbahasa Indonesia menggunakan *Naïve Bayes* dan *SVM*

Design of a Sentiment Analysis System for Indonesian-Language International Conflict News using Naïve Bayes and SVM

¹Fitri Anisa*, ²Asto Purwanto

^{1,2}Program Studi Sistem Informasi, STMIK-IM

^{1,2}Jl. Belitung No.7 Bandung, Indonesia

*e-mail: fitriannissaaa@gmail.com

(received: 4 May 2026, revised: 9 May 2026, accepted: 15 May 2026)

Abstrak

Lonjakan volume berita digital, khususnya yang meliputi konflik internasional, mendorong kebutuhan mendesak akan sistem klasifikasi sentimen yang mampu bekerja secara otomatis dan objektif. Studi ini merancang sebuah sistem analisis sentimen untuk berita konflik internasional berbahasa Indonesia dengan memanfaatkan dua algoritma klasifikasi, yaitu *Naïve Bayes* dan *Support Vector Machine (SVM)*, sekaligus membandingkan efektivitas keduanya. Sebanyak 339 teks berita yang bersumber dari Reuters, BBC, CNN, dan Al Jazeera digunakan sebagai data, terdiri atas 146 teks bersentimen negatif, 123 positif, dan 70 netral. Tahapan pengolahan data mencakup preprocessing (case folding, tokenizing, stopword removal, stemming), pembobotan fitur TF-IDF, dan klasifikasi dengan skema pembagian data 80:20. Sistem dibangun menggunakan Python dan di-deploy melalui framework Streamlit sebagai antarmuka web. Hasil pengujian menunjukkan SVM mencapai akurasi 82,35% dengan nilai precision, recall, dan F1-score yang lebih merata di setiap kelas, melampaui *Naïve Bayes* yang memperoleh akurasi 77,94%. Kebaruan utama studi ini adalah pemanfaatan teks berita lengkap dari sumber jurnalistik internasional terpercaya sebagai objek analisis — berbeda dari mayoritas studi terdahulu yang mengandalkan data media sosial berformat pendek. Studi ini menyimpulkan SVM lebih unggul untuk klasifikasi sentimen berita konflik internasional berbahasa Indonesia, dan merekomendasikan perluasan dataset serta eksplorasi metode representasi semantik seperti *IndoBERT* untuk peningkatan performa di masa mendatang.

Kata kunci: analisis sentimen, machine learning, naïve bayes, support vector machine, TF-IDF

Abstract

The rapid growth of digital news content, particularly coverage related to international conflicts, has created an urgent need for sentiment classification systems capable of operating automatically and objectively. This study designs a sentiment analysis system for Indonesian-language international conflict news using two classification algorithms, namely *Naïve Bayes* and *Support Vector Machine (SVM)*, while also comparing their effectiveness. A total of 339 news articles sourced from Reuters, BBC, CNN, and Al Jazeera were used as the dataset, consisting of 146 negative, 123 positive, and 70 neutral sentiment texts. The data processing stages included preprocessing (case folding, tokenizing, stopword removal, and stemming), TF-IDF feature weighting, and classification using an 80:20 train-test split scheme. The system was developed using Python and deployed through the Streamlit framework as a web-based interface. The experimental results indicate that SVM achieved an accuracy of 82.35%, with more balanced precision, recall, and F1-score values across all classes, outperforming *Naïve Bayes*, which achieved an accuracy of 77.94%. The main novelty of this study lies in the use of complete news articles from credible international journalistic sources as the object of analysis, unlike most previous studies that primarily relied on short-form social media data. The study concludes that SVM is more effective for sentiment classification of Indonesian-language international conflict news and recommends expanding the dataset and exploring semantic representation methods such as *IndoBERT* to improve performance in future research.

Keywords: machine learning, naïve bayes, sentiment analysis, support vector machine, TF-IDF

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

1 Pendahuluan

Penyebaran berita konflik internasional melalui media digital menghasilkan kumpulan teks berita yang terus bertambah dan membutuhkan proses analisis yang lebih sistematis. Pemberitaan tersebut tidak hanya memuat informasi faktual, tetapi juga dapat memperlihatkan kecenderungan sentimen tertentu dalam menggambarkan suatu peristiwa. Analisis terhadap kecenderungan tersebut penting karena dapat membantu pemerintah, organisasi, maupun media dalam memahami dinamika isu konflik yang berkembang [1]. Oleh sebab itu, diperlukan sistem otomatis yang mampu mengidentifikasi sentimen pada teks berita secara lebih efisien dan konsisten.

Berita konflik internasional yang dipublikasikan media global seperti CNN, BBC, Reuters, dan Al Jazeera menjadi salah satu sumber informasi yang banyak diikuti masyarakat. Jumlah teks berita konflik internasional yang terus meningkat membuat proses analisis manual menjadi kurang efektif serta berpotensi menghasilkan interpretasi yang subjektif [2]. Pemberitaan konflik internasional umumnya mengandung nuansa sentimen yang beragam, mulai dari laporan faktual yang netral hingga narasi yang dipengaruhi emosi positif maupun negatif. Variasi sentimen tersebut menjadikan analisis sentimen berbasis otomatis sebagai solusi yang relevan untuk mengolah dan memahami data berita dalam skala besar. Selain itu, meningkatnya jumlah sumber berita dan perkembangan isu konflik yang dinamis menyebabkan proses pemantauan sentimen secara konsisten dan menyeluruh sulit dilakukan tanpa dukungan metode otomatis yang andal.

Dalam ranah text mining, pendekatan yang bertujuan mendeteksi kecenderungan opini suatu teks ke dalam kategori tertentu — seperti positif, negatif, atau netral — dikenal sebagai analisis sentimen [1]. Untuk keperluan penelitian ini, dua algoritma dipilih berdasarkan kesesuaian karakteristiknya dengan domain teks berita berbasis TF-IDF, yaitu Naïve Bayes dan Support Vector Machine (SVM). Naïve Bayes unggul dalam hal kecepatan pelatihan dan efisiensi komputasi pada ruang fitur berdimensi besar, sementara SVM dikenal andal dalam membangun pemisah kelas yang optimal pada data dengan distribusi fitur yang kompleks [4][10]. Perpaduan kedua algoritma ini membuka peluang perbandingan yang komprehensif pada domain spesifik berita konflik internasional berbahasa Indonesia. TF-IDF dipilih sebagai skema pembobotan fitur karena kemampuannya merefleksikan tingkat kepentingan suatu kata secara proporsional terhadap seluruh korpus [5], dan kombinasinya dengan Naïve Bayes maupun SVM telah terbukti menghasilkan klasifikasi yang baik pada berbagai tipe teks, termasuk teks jurnalistik [4][12].

Sari et al. [6], Fitriyani et al. [7], dan Muliana et al. [8] menggunakan data media sosial seperti Twitter dan ulasan aplikasi sebagai objek analisis sentimen. Sementara itu, Andriawan dan Ernawati [14] telah membahas analisis sentimen terkait konflik internasional, namun masih menggunakan data Twitter dan bukan teks berita penuh. Studi yang secara khusus melakukan perbandingan langsung antara Naïve Bayes dan SVM pada berita konflik internasional berbahasa Indonesia serta mengimplementasikannya ke dalam sistem web interaktif masih jarang ditemukan [6][7][8][14]. Karakteristik data media sosial yang cenderung singkat dan tidak terstruktur juga berbeda dengan teks berita jurnalistik yang lebih panjang, terstruktur, dan kaya konteks [12][13]. Oleh karena itu, penelitian ini difokuskan pada penggunaan teks berita penuh sebagai objek analisis sentimen konflik internasional.

Oleh karena itu, penelitian ini bertujuan merancang sistem analisis sentimen pada berita konflik internasional berbahasa Indonesia menggunakan Naïve Bayes dan SVM, membandingkan performa kedua metode, serta mengimplementasikannya dalam bentuk aplikasi web menggunakan framework Streamlit. Penelitian ini memberikan tiga kontribusi utama: pertama, penggunaan teks berita penuh dari sumber jurnalistik internasional terpercaya (Reuters, BBC, CNN, Al Jazeera) sebagai domain analisis sentimen konflik yang lebih kaya konteks dibandingkan data media sosial [6][7][8]; kedua, perbandingan langsung Naïve Bayes dan SVM pada domain berita konflik internasional berbahasa Indonesia yang belum banyak diteliti [11][14]; dan ketiga, implementasi sistem analisis sentimen berbasis web interaktif menggunakan Streamlit yang memungkinkan penggunaan langsung oleh pengguna non-teknis tanpa memerlukan pengetahuan pemrograman [9]. Secara praktis, sistem ini dapat dimanfaatkan oleh jurnalis, analis kebijakan, maupun organisasi internasional untuk memantau dan memahami sentimen pemberitaan konflik secara otomatis dan efisien.

2 Tinjauan Literatur

Perkembangan analisis sentimen dalam beberapa tahun terakhir menunjukkan bahwa supervised machine learning masih banyak digunakan untuk mengolah data teks pada berbagai domain. Birjali et al. [1] dan Wankhade et al. [2] menekankan bahwa pendekatan ini tetap relevan karena mampu menghasilkan performa klasifikasi yang kompetitif. Dalam aspek representasi fitur, Agarwal et al. [3] membandingkan TF-IDF dan word embeddings, sedangkan Hasan et al. [5] mengkaji beberapa teknik ekstraksi fitur pada analisis sentimen berbasis machine learning. Temuan dari penelitian tersebut menunjukkan bahwa pemilihan representasi fitur perlu disesuaikan dengan karakteristik data dan tujuan klasifikasi.

Berbagai kajian telah menguji kemampuan Naïve Bayes dan SVM secara komparatif dalam tugas klasifikasi teks sentimen. Pada dataset berita BBC dengan representasi TF-IDF, Hai et al. [4] menguji empat algoritma klasifikasi teks dan mendapati bahwa SVM secara konsisten mengungguli metode lainnya dari sisi akurasi. Konsistensi keunggulan SVM juga tampak pada ranah ulasan dan media sosial: Sari et al. [6] pada ulasan PUBG Mobile, Fitriyani et al. [7] pada ulasan Weverse, serta Leandro dan Fianty [10] pada berbagai platform media sosial — seluruhnya membuktikan superioritas SVM atas Naïve Bayes. Akar dari keunggulan ini terletak pada kemampuan SVM mengonstruksi hyperplane pemisah yang optimal dalam ruang fitur berdimensi tinggi, kondisi yang lazim muncul ketika teks direpresentasikan menggunakan TF-IDF. Di sisi berlawanan, Naïve Bayes tetap menjadi pilihan yang kompetitif ketika kecepatan dan kesederhanaan menjadi prioritas — Muliana et al. [8] memperlihatkan algoritma ini sanggup mengklasifikasikan sentimen pemilu dari 2.500 cuitan platform X dengan akurasi rata-rata 66,90%, sementara Firda et al. [9] memanfaatkannya sebagai basis pembandingan dalam eksplorasi strategi pelabelan sentimen berbasis SVM.

Kendati riset analisis sentimen terus tumbuh pesat, dominasi data media sosial dan ulasan aplikasi masih menjadi keterbatasan utama dari studi-studi yang ada. Sejumlah penelitian yang memanfaatkan teks berita justru mengungkap perbedaan karakteristik yang signifikan: Li et al. [12] menguji 10 model machine learning pada 9.382 artikel berita lintas batas dan menemukan pola yang berbeda dari studi berbasis media sosial, sementara Rahman et al. [13] mengaplikasikan SVM dan Naïve Bayes dengan TF-IDF pada 2.225 artikel berita BBC. Kedua penelitian ini menunjukkan bahwa teks berita memiliki kerumitan struktural dan kedalaman konteks yang jauh melampaui unggahan media sosial yang umumnya pendek dan tidak terstruktur.

Pada domain konflik khususnya, Umer et al. [11] mengevaluasi sejumlah model supervised learning pada cuitan tentang konflik Rusia-Ukraina, yang mempertegas relevansi metode klasifikasi pada domain berita konflik. Adapun Andriawan dan Ernawati [14] membandingkan Naïve Bayes dan SVM untuk sentimen konflik Palestina-Israel, tetapi masih bertumpu pada data Twitter yang bersifat singkat dan tidak terstruktur, bukan teks berita jurnalistik yang komprehensif. Zhang [15] turut mengonfirmasi efektivitas gabungan Naïve Bayes dan TF-IDF untuk klasifikasi teks berita, namun kajiannya belum menyentuh domain konflik internasional berbahasa Indonesia.

Berdasarkan penelitian-penelitian sebelumnya, penggunaan teks berita penuh dari sumber jurnalistik internasional sebagai objek analisis sentimen konflik masih belum banyak dilakukan. Selain itu, penelitian yang membandingkan Naïve Bayes dan SVM secara langsung pada berita konflik internasional berbahasa Indonesia serta mengimplementasikannya dalam sistem berbasis web interaktif juga masih terbatas. Oleh karena itu, penelitian ini difokuskan pada tiga aspek utama, yaitu penggunaan teks berita penuh dari Reuters, BBC, CNN, dan Al Jazeera yang memiliki konteks lebih kaya dibandingkan data Twitter pada penelitian sebelumnya [11][14], perbandingan performa Naïve Bayes dan SVM pada domain berita konflik internasional berbahasa Indonesia, serta pengembangan sistem analisis sentimen berbasis web menggunakan Streamlit [6][7][8][9].

3 Metode Penelitian

Untuk menjawab pertanyaan penelitian, studi ini mengadopsi rancangan eksperimental berbasis pengukuran kuantitatif [12]. Korpus penelitian dibangun secara langsung melalui penelusuran mandiri pada empat portal berita internasional — Reuters, BBC, CNN, dan Al Jazeera — sepanjang tahun 2026. Setiap artikel dipilih berdasarkan relevansinya dengan topik konflik internasional, dan proses seleksi ini menghasilkan 339 teks berita yang terbagi atas 146 data bersentimen negatif, 123 data positif, dan 70 data netral. Data yang sebagian besar berbahasa Inggris diterjemahkan ke dalam bahasa

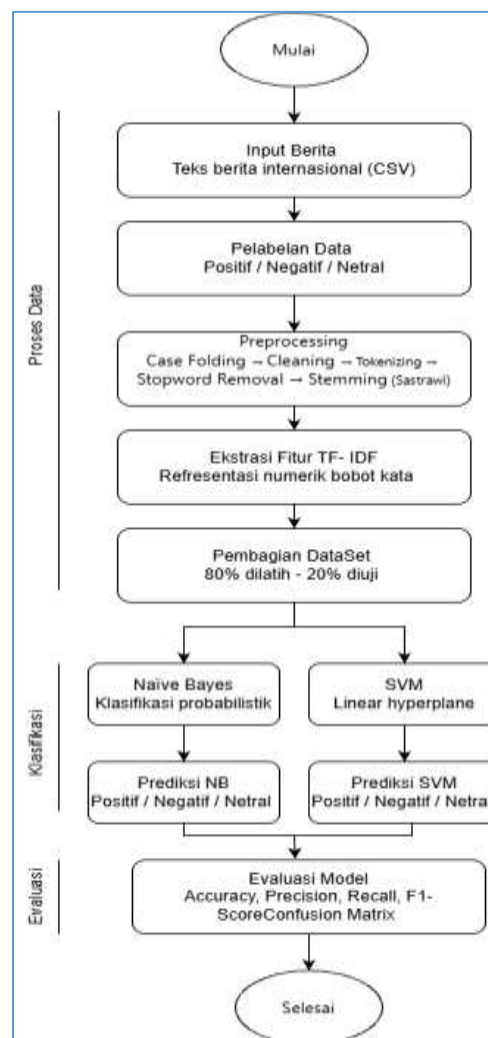
Indonesia menggunakan bantuan alat penerjemah (Google Translate), kemudian dilakukan pengecekan ulang secara manual untuk menjaga kesesuaian makna dan konteks berita. Penggunaan hasil terjemahan mesin dapat mempengaruhi nuansa sentimen pada teks, sehingga berpotensi menimbulkan pergeseran makna. Namun, proses pengecekan manual dilakukan untuk meminimalkan kesalahan interpretasi.

Proses pelabelan data dilakukan berdasarkan definisi operasional sentimen yang telah ditetapkan untuk setiap kelas. Guna menekan subjektivitas dalam proses anotasi, setiap data dilabeli mengacu pada kriteria yang sudah dirancang secara ketat, dan pada kasus yang meragukan dilakukan peninjauan ulang untuk menjaga konsistensi antar label. Keterbatasan penelitian ini adalah pelabelan dilakukan oleh anotator tunggal, sehingga evaluasi konsistensi antar anotator (inter-annotator agreement) belum dapat dilakukan dan menjadi agenda perbaikan pada penelitian selanjutnya. Kriteria pelabelan sentimen yang digunakan dalam penelitian ini ditunjukkan pada Tabel 1.

Tabel 1 Kriteria pelabelan sentimen

| Kelas Sentimen | Kriteria |
|----------------|--|
| Negatif | Mengandung konflik, kekerasan, atau ketegangan |
| Positif | Mengandung perdamaian, kerja sama, atau resolusi konflik |
| Netral | Informasi faktual tanpa kecenderungan emosi |

Kriteria ini diterapkan secara seragam pada seluruh data untuk menjaga objektivitas antar data. Pada kasus teks ambigu, pelabelan diutamakan berdasarkan konteks keseluruhan kalimat, bukan hanya kehadiran kata kunci tertentu. Alur sistem ditunjukkan pada Gambar 1.



Gambar 1 Alur sistem analisis sentimen

Alur proses sistem analisis sentimen pada penelitian ini ditunjukkan pada Gambar 1. Proses sistem diawali dengan memasukkan kumpulan teks berita internasional berformat CSV ke dalam sistem untuk dilakukan pelabelan sentimen. Teks selanjutnya melalui proses pembersihan dan normalisasi data berupa case folding, cleaning, tokenisasi, penghapusan stopword, serta stemming dengan bantuan library Sastrawi kemudian dilakukan pembobotan fitur menggunakan TF-IDF untuk mengubah data teks menjadi representasi numerik yang dapat diproses model klasifikasi.

Dataset kemudian dibagi ke dalam data pelatihan dan data pengujian dengan rasio 80:20, kemudian diklasifikasikan menggunakan Naïve Bayes dan Support Vector Machine (SVM) untuk memprediksi label sentimen teks ke dalam tiga kategori: positif, negatif, atau netral. Kinerja kedua algoritma kemudian dianalisis menggunakan metrik accuracy, precision, recall, F1-score, serta confusion matrix untuk melihat performa klasifikasi masing-masing model. Proses berakhir setelah diperoleh hasil evaluasi dari masing-masing model.

Ekstraksi fitur menggunakan TF-IDF [5] berdasarkan persamaan (1), (2), dan (3).

$$TF(w, d) = \frac{f(w, d)}{\sum f(t, d)} \quad (1)$$

$$IDF(w) = \log_{10} \left(\frac{N}{df(w)} \right) \quad (2)$$

$$TF - IDF(w, d) = TF(w, d) \times IDF(w) \quad (3)$$

Keterangan:

$f(w, d)$: menyatakan jumlah kemunculan kata w dalam dokumen d

$\sum f(t, d)$: menyatakan total kata dalam dokumen d

N : menyatakan jumlah seluruh dokumen

$df(w)$: menyatakan jumlah dokumen yang mengandung kata w

Klasifikasi menggunakan Multinomial Naïve Bayes [4] berdasarkan persamaan (4).

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (4)$$

Linear SVM [10] digunakan berdasarkan persamaan (5).

$$f(x) = w \cdot x + b \quad (5)$$

Penelitian ini menggunakan hold-out validation (80:20) sebagai skema evaluasi utama. Dataset dibagi dengan rasio 80:20 menghasilkan 271 data latih dan 68 data uji dari total 339 data. Pembagian ini dilakukan dengan pembulatan ke bawah dari hasil perhitungan $80\% \times 339 = 271,2$, sehingga diperoleh 271 data latih dan 68 data uji, dan umum digunakan dalam penelitian klasifikasi teks sebagai baseline evaluasi [4][6]. Pembagian data dilakukan secara acak dengan random state tetap (random_state=42) untuk memastikan hasil eksperimen dapat direplikasi. Model dilatih menggunakan parameter default dari library Scikit-learn sebagai baseline awal dan untuk menjaga kemudahan replikasi eksperimen sebelum optimasi lebih lanjut. Evaluasi dilakukan menggunakan metrik accuracy, precision, recall, F1-Score, dan confusion matrix dengan rumus yang ditunjukkan pada persamaan (6), (7), (8), dan (9).

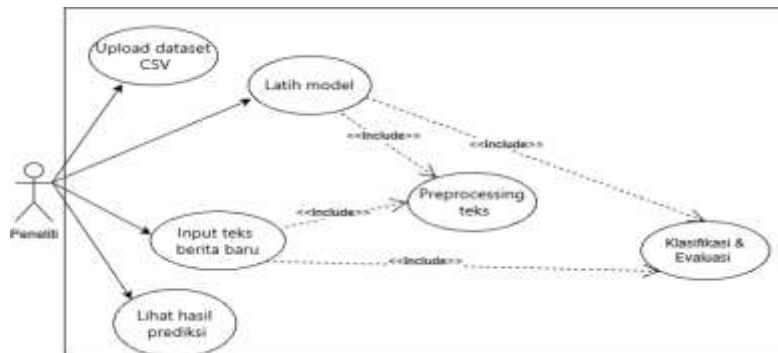
$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

$$Precision = TP / (TP + FP) \quad (7)$$

$$Recall = TP / (TP + FN) \quad (8)$$

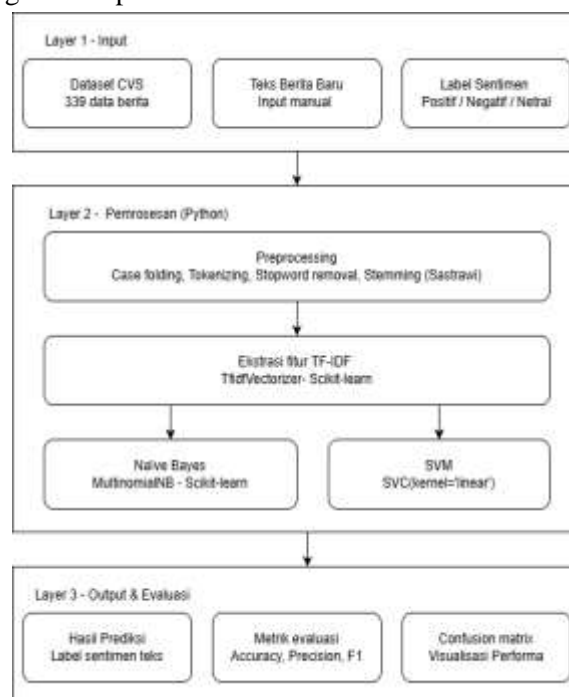
$$F1 - Score = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (9)$$

Pada klasifikasi multiclass, perhitungan TP, FP, dan FN dilakukan untuk setiap kelas menggunakan pendekatan one-vs-rest. Sistem diimplementasikan menggunakan Python dan Streamlit. Use case diagram dan arsitektur sistem ditunjukkan pada Gambar 2 dan Gambar 3.



Gambar 2 Use case diagram sistem

Gambar 2 memperlihatkan interaksi antara peneliti sebagai pengguna utama dengan sistem analisis sentimen yang dikembangkan. Melalui sistem ini, peneliti dapat mengunggah dataset dalam format CSV, menjalankan pelatihan model, memasukkan teks berita baru untuk dianalisis, serta melihat hasil prediksi sentimen. Pada proses pelatihan dan prediksi, sistem menjalankan preprocessing secara otomatis sebelum data diproses menggunakan algoritma Naïve Bayes dan Support Vector Machine. Relasi <<include>> pada diagram menunjukkan bahwa preprocessing, klasifikasi, dan evaluasi selalu menjadi bagian dari proses utama sistem.



Gambar 3 Arsitektur sistem

Gambar 3 memperlihatkan rancangan arsitektur sistem analisis sentimen yang terdiri atas tiga lapisan, yaitu layer input, layer pemrosesan, serta layer output dan evaluasi. Layer input berfungsi menerima dataset berita berformat CSV, teks berita baru yang dimasukkan pengguna, serta label sentimen positif, negatif, dan netral. Pada layer pemrosesan, data melewati tahap preprocessing menggunakan Sastrawi, kemudian direpresentasikan dengan TF-IDF melalui TfidfVectorizer. Setelah

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

itu, proses klasifikasi dilakukan menggunakan Multinomial Naïve Bayes dan SVM kernel linear. Hasil akhir sistem berupa prediksi sentimen, nilai evaluasi model, dan confusion matrix sebagai visualisasi performa klasifikasi.

4 Hasil dan Pembahasan

Bagian ini menguraikan temuan dari pengujian sistem, mencakup output preprocessing, hasil pembobotan TF-IDF, performa Naïve Bayes dan SVM, serta analisis komparatif kedua model pada korpus berita konflik internasional berbahasa Indonesia.

4.1 Hasil Preprocessing

Tahap preprocessing menghasilkan teks berita yang lebih konsisten dan terstruktur setelah dilakukan proses pembersihan serta normalisasi data. Tabel 2 menunjukkan contoh perbandingan teks asli dengan hasil preprocessing.

Tabel 2 Hasil preprocessing teks berita

| No | Teks Asli | Hasil Preprocessing |
|----|---|--|
| 1 | Harga minyak dunia melonjak di atas 100 dolar setelah rencana blokade Selat Hormuz oleh Amerika Serikat | harga minyak dunia lonjak dolar rencana blokade selat hormuz amerika serikat |
| 2 | Iran memperingatkan akan menyerang kapal militer yang mendekati Selat Hormuz | iran ingat serang kapal militer dekat selat hormuz |
| 3 | Perjanjian gencatan senjata berhasil disepakati oleh kedua belah pihak konflik | perjanjian gencat senjata sepakat pihak konflik |
| 4 | Ketegangan meningkat setelah kapal tanker berbalik arah di dekat Selat Hormuz | tegang tingkat kapal tanker balik arah selat hormuz |
| 5 | Selat Hormuz mengalami gangguan besar akibat konflik militer yang menghambat distribusi minyak | selat hormuz alami ganggu akibat konflik militer hambat distribusi minyak |

Berdasarkan Tabel 2, teks yang semula mengandung kata berimbuhan dan struktur kalimat kompleks telah disederhanakan menjadi bentuk kata dasar yang lebih seragam. Sebagai contoh, kata "melonjak" menjadi "lonjak", "rencana" tetap "rencana", dan "menghambat" menjadi "hambat" — menunjukkan bahwa stemming berjalan dengan baik pada sebagian besar kata. Namun pada beberapa kasus, proses stemming menghasilkan over-stemming, seperti "memperingatkan" menjadi "ingat" dan "meningkat" menjadi "tingkat" [15]. Hal ini merupakan karakteristik library Sastrawi berbasis rule morphological yang perlu dipertimbangkan dalam interpretasi hasil klasifikasi, terutama pada kelas netral yang karakteristik teksnya lebih halus.

4.2 Hasil Pembobotan TF-IDF

Setelah preprocessing selesai, TF-IDF menghasilkan sejumlah fitur kata unik dari seluruh 339 dokumen. Tabel 3 menunjukkan sampel pembobotan TF-IDF pada beberapa kata yang sering muncul dalam dataset.

Tabel 3 Sampel hasil pembobotan TF-IDF

| Dokumen | blokade | hormuz | konflik | damai | militer |
|---------|---------|--------|---------|-------|---------|
| D1 | 0.356 | 0.247 | 0.000 | 0.000 | 0.000 |
| D2 | 0.389 | 0.270 | 0.000 | 0.264 | 0.000 |
| D3 | 0.000 | 0.266 | 0.000 | 0.000 | 0.236 |
| D4 | 0.000 | 0.249 | 0.122 | 0.000 | 0.000 |

| | | | | | |
|----|-------|-------|-------|-------|-------|
| D5 | 0.000 | 0.267 | 0.131 | 0.000 | 0.237 |
|----|-------|-------|-------|-------|-------|

Bobot TF-IDF yang tinggi mengindikasikan bahwa suatu kata memiliki tingkat kepentingan lebih besar pada dokumen tertentu [5]. Berdasarkan Tabel 3, kata 'blokade' memiliki bobot tertinggi pada D1 (0.356) dan D2 (0.389), mengindikasikan kedua dokumen tersebut membahas topik blokade secara dominan. Pada D2, kata 'damai' (0.264) juga muncul bersama 'blokade', menunjukkan konteks pemberitaan yang membahas blokade dalam kaitannya dengan upaya perdamaian. Kata 'hormuz' muncul secara konsisten pada D1 hingga D5 dengan bobot yang bervariasi, mencerminkan kata tersebut relevan di banyak dokumen namun tidak terlalu unik untuk satu dokumen tertentu. Kata 'konflik' hanya muncul pada D4 (0.122) dan D5 (0.131), sedangkan 'militer' muncul pada D3 (0.236) dan D5 (0.237). Nilai 0 menunjukkan kata tersebut tidak muncul dalam dokumen bersangkutan, mencerminkan sifat sparse dari representasi TF-IDF pada data teks berita.

4.3 Hasil Evaluasi Model

Pengujian klasifikasi dijalankan pada 68 sampel data uji dari keseluruhan 339 data. Perbandingan kinerja kedua model ditunjukkan pada Tabel 4.

Tabel 4 Perbandingan kinerja model

| Model | Accuracy | Precision | Recall | F1-Score |
|-------------|----------|-----------|--------|----------|
| Naïve Bayes | 77,94% | 0,82 | 0,78 | 0,76 |
| SVM | 82,35% | 0,83 | 0,82 | 0,82 |

Tabel 4 menunjukkan bahwa SVM mencatatkan akurasi 82,35%, melampaui Naïve Bayes yang berada di angka 77,94%. Temuan ini selaras dengan sejumlah penelitian terdahulu yang secara konsisten menempatkan SVM di atas Naïve Bayes dalam tugas klasifikasi teks [6][10]. Classification report per kelas ditunjukkan pada Tabel 5 dan Tabel 6.

Tabel 5 Classification report naïve bayes

| Kelas | Precision | Recall | F1-Score | Support |
|--------------|-----------|--------|----------|---------|
| Negatif | 0,79 | 0,92 | 0,85 | 24 |
| Netral | 1,00 | 0,39 | 0,56 | 18 |
| Positif | 0,73 | 0,92 | 0,81 | 26 |
| Weighted Avg | 0,82 | 0,78 | 0,76 | 68 |

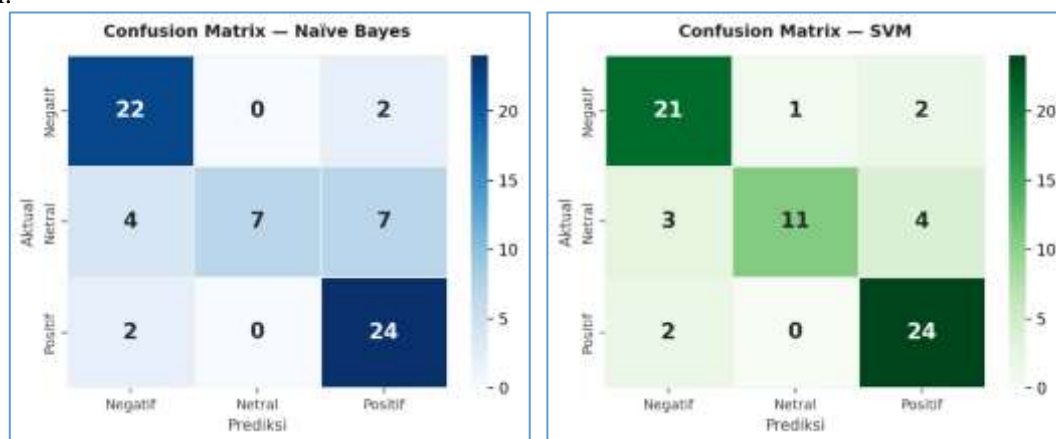
Tabel 6 Classification report SVM

| Kelas | Precision | Recall | F1-Score | Support |
|--------------|-----------|--------|----------|---------|
| Negatif | 0,81 | 0,88 | 0,84 | 24 |
| Netral | 0,92 | 0,61 | 0,73 | 18 |
| Positif | 0,80 | 0,92 | 0,86 | 26 |
| Weighted Avg | 0,83 | 0,82 | 0,82 | 68 |

Berdasarkan Tabel 5 dan Tabel 6, model Naïve Bayes dan SVM memiliki performa yang berbeda pada setiap kelas sentimen. Pada model Naïve Bayes, nilai recall pada kelas netral relatif rendah dibandingkan kelas lainnya, sedangkan model SVM menunjukkan performa yang lebih seimbang pada seluruh kelas, terutama pada peningkatan recall dan F1-Score pada kelas netral.

Rendahnya nilai recall pada kategori netral mengindikasikan bahwa model Naïve Bayes masih mengalami kendala dalam membedakan teks netral dari kategori lainnya. Hal ini tidak hanya disebabkan oleh distribusi data yang tidak seimbang, tetapi juga karena karakteristik teks netral yang cenderung mengandung kata-kata yang juga muncul pada kelas negatif maupun positif, misalnya kata

'konflik' yang berbobot tinggi namun dapat muncul dalam konteks faktual yang seharusnya berlabel netral.



Gambar 4 Confusion matrix naïve bayes (kiri) dan SVM (kanan)

Gambar 4 menunjukkan confusion matrix dari model Naïve Bayes dan SVM dalam mengklasifikasikan data uji. Pada model Naïve Bayes, masih terdapat kesalahan klasifikasi terutama pada kelas netral yang sering diprediksi sebagai kelas lain. Sementara itu, model SVM menunjukkan hasil yang lebih baik dengan jumlah prediksi yang benar lebih tinggi dan kesalahan yang lebih sedikit dibandingkan Naïve Bayes.

4.4 Pembahasan

Performa kedua algoritma menunjukkan perbedaan yang dipengaruhi oleh mekanisme klasifikasi masing-masing metode. Naïve Bayes mengasumsikan independensi antar fitur yang tidak sepenuhnya terpenuhi pada teks berita karena kata-kata memiliki keterkaitan kontekstual kuat [1]. SVM yang mencari hyperplane optimal memperoleh nilai akurasi lebih tinggi dalam menangani ruang vektor berdimensi tinggi hasil TF-IDF [4], konsisten dengan temuan Hai et al. [4] pada dataset berita BBC dan Li et al. [12] pada 9.382 artikel berita internasional. SVM menunjukkan performa lebih tinggi dibandingkan Naïve Bayes karena tidak bergantung pada asumsi independensi fitur, sehingga menghasilkan nilai evaluasi yang lebih besar dalam menangani data teks berita yang memiliki keterkaitan kontekstual antar kata. Kondisi tersebut turut dipengaruhi oleh karakteristik pembobotan TF-IDF yang memberikan bobot lebih tinggi pada kata-kata domain spesifik seperti 'konflik', 'serangan', dan 'gencatan senjata', sehingga SVM dapat memanfaatkan distribusi fitur tersebut secara lebih optimal dalam membentuk batas keputusan antar kelas [5].

Kategori sentimen netral menjadi bagian yang paling sulit diklasifikasikan oleh model. Naïve Bayes hanya mencapai recall 0,39 — dari 18 data netral pada data uji, hanya 7 yang teridentifikasi benar. Hal ini tidak lepas dari distribusi dataset yang tidak seimbang, di mana kelas netral merupakan kelas dengan jumlah data paling sedikit (70 data) dibandingkan kelas negatif (146 data) dan positif (123 data). Ketidakseimbangan jumlah data antar kelas menyebabkan model cenderung bias terhadap kelas mayoritas, sehingga performa pada kelas netral menjadi lebih rendah dibandingkan kelas lainnya. Selain faktor distribusi, karakteristik teks netral yang mengandung kata berkonotasi negatif dalam konteks faktual seperti "konflik" dan "ketegangan" turut mempersulit klasifikasi [11]. Contoh, kalimat "Dewan Keamanan PBB mengadakan sidang darurat terkait konflik Timur Tengah" seharusnya netral namun diklasifikasikan negatif oleh Naïve Bayes karena bobot tinggi kata "konflik". SVM meningkatkan recall netral menjadi 0,61 karena mempertimbangkan kombinasi keseluruhan fitur dalam membentuk keputusan klasifikasi [10].

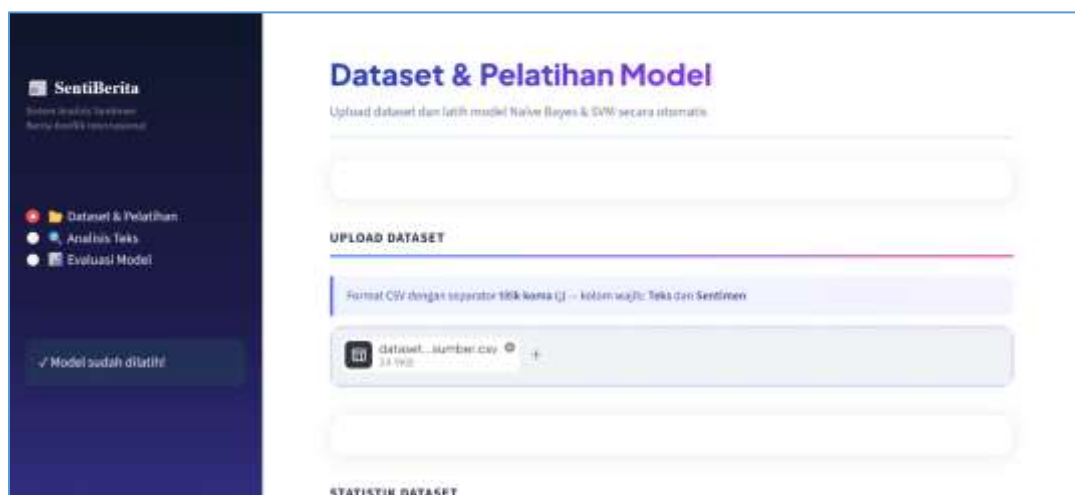
Kesalahan klasifikasi serupa juga ditemukan pada kelas positif, di mana teks seperti "Kedua negara sepakat melanjutkan negosiasi meskipun ketegangan masih berlanjut" diklasifikasikan sebagai negatif oleh Naïve Bayes akibat bobot tinggi kata "ketegangan", padahal konteks keseluruhan kalimat menunjukkan upaya perdamaian. Pola kesalahan ini mengindikasikan bahwa model berbasis frekuensi kata seperti Naïve Bayes rentan terhadap ambiguitas kontekstual, terutama pada teks berita yang mengandung kata bermuatan sentimen dalam konteks yang berlawanan.

Kelebihan penelitian ini dibandingkan studi sebelumnya [6][7][8][14]: (1) menggunakan teks berita penuh (full text news) dari 4 sumber jurnalistik internasional terpercaya dengan konteks lebih kaya dibandingkan data media sosial yang pendek dan tidak terstruktur; (2) membandingkan Naïve Bayes dan SVM secara langsung pada domain spesifik berita konflik internasional yang belum banyak dieksplorasi; (3) mengimplementasikan sistem web interaktif menggunakan Streamlit yang memungkinkan penggunaan langsung tanpa pengetahuan pemrograman. Kombinasi ketiga aspek ini dalam satu kerangka penelitian merupakan kontribusi yang membedakan penelitian ini dari studi terdahulu. Meskipun jumlah dataset relatif terbatas, kualitas teks berita utuh dari sumber terpercaya memberikan konteks informasi yang lebih kaya dan representatif dibandingkan data Twitter yang umumnya pendek dan bising.

Berdasarkan hasil evaluasi, pemilihan algoritma yang sesuai dengan karakteristik data teks berdimensi tinggi terbukti memberikan pengaruh signifikan terhadap performa sistem analisis sentimen. SVM dengan kernel linear terbukti memberikan dampak signifikan terhadap performa keseluruhan dibandingkan Naïve Bayes, terutama pada data dengan distribusi fitur yang tidak seimbang [10][12].

4.5 Tampilan Antarmuka Sistem

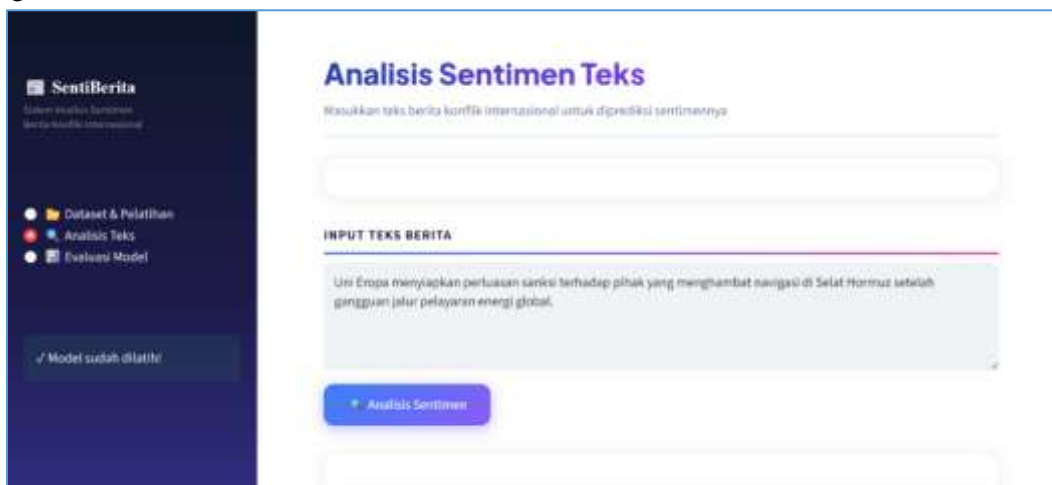
Sistem diimplementasikan menggunakan Streamlit dengan tiga halaman utama.



Gambar 5 Tampilan halaman upload dataset dan pelatihan model

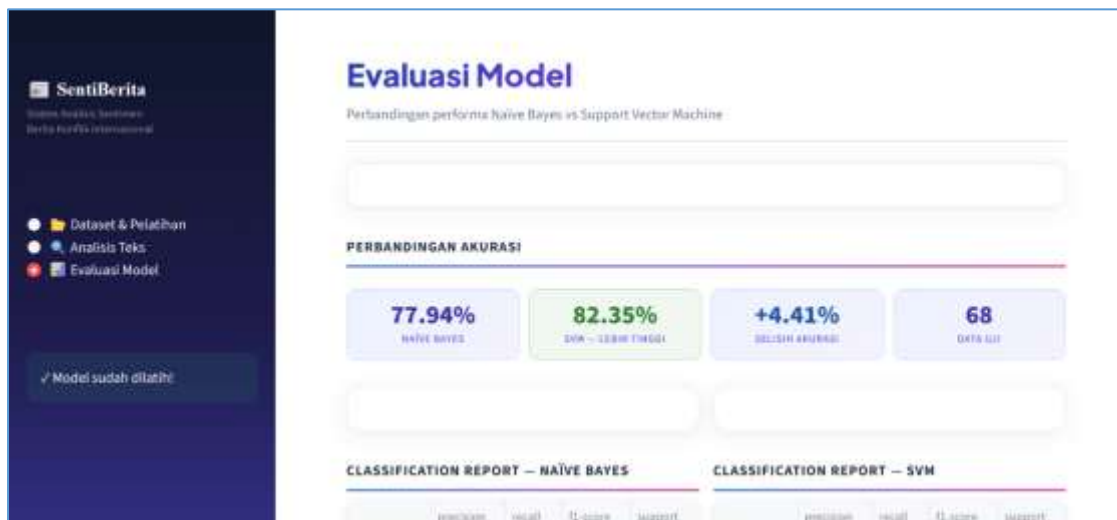
Gambar 5 menunjukkan tampilan halaman Dataset & Pelatihan Model pada sistem yang dikembangkan. Pada halaman ini, pengguna dapat mengunggah dataset dalam format CSV yang berisi teks berita dan label sentimen. Sistem menyediakan informasi format dataset yang diperlukan untuk memastikan data dapat diproses dengan benar.

Setelah dataset berhasil diunggah, sistem akan melakukan proses pelatihan model secara otomatis menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM). Status pelatihan model ditampilkan pada antarmuka untuk memberikan informasi kepada pengguna bahwa model telah siap digunakan.



Gambar 6 Tampilan halaman prediksi sentimen

Gambar 6 menunjukkan halaman Prediksi Sentimen. Pengguna memasukkan teks berita baru, sistem menjalankan preprocessing dan mengklasifikasikan sentimen menggunakan Naïve Bayes dan SVM secara bersamaan.



Gambar 7 Tampilan halaman evaluasi model

Gambar 7 menunjukkan halaman Evaluasi Model yang menampilkan akurasi, classification report, dan confusion matrix dari kedua model. Pengguna dapat mengunduh gambar confusion matrix untuk dokumentasi. Sistem ini memungkinkan pengguna melakukan analisis sentimen secara mandiri tanpa memerlukan pengetahuan pemrograman.

5 Kesimpulan

Penelitian ini berhasil membangun sistem analisis sentimen berita konflik internasional berbahasa Indonesia berbasis web yang mengintegrasikan metode Naïve Bayes dan SVM. Dari 68 data uji yang diambil dari total 339 dataset, SVM terbukti lebih unggul dengan akurasi 82,35%, precision 0,83, recall 0,82, dan F1-score 0,82, dibandingkan Naïve Bayes yang memperoleh akurasi 77,94%, precision 0,82, recall 0,78, dan F1-score 0,76. Performa SVM yang lebih tinggi dipengaruhi kemampuannya dalam membentuk hyperplane optimal pada ruang fitur berdimensi tinggi hasil TF-IDF tanpa bergantung pada asumsi independensi fitur. Secara praktis, sistem yang dikembangkan dapat digunakan oleh jurnalis, analis kebijakan, maupun organisasi internasional untuk membantu pemantauan sentimen pemberitaan konflik secara otomatis, sementara dari sisi akademik, penelitian ini memberikan kontribusi pada pengembangan analisis sentimen berbasis teks berita penuh dari sumber jurnalistik internasional terpercaya pada domain konflik berbahasa Indonesia yang membedakannya dari studi sebelumnya yang umumnya menggunakan data media sosial pendek. Penelitian ini masih memiliki beberapa keterbatasan, di antaranya jumlah dataset yang relatif terbatas, distribusi kelas yang belum seimbang, serta penggunaan parameter default tanpa proses hyperparameter tuning, serta pelabelan oleh anotator tunggal, sehingga untuk penelitian selanjutnya disarankan memperbesar dataset, menggunakan teknik oversampling SMOTE, menerapkan hyperparameter tuning dan cross-validation, serta mengeksplorasi representasi teks lebih canggih seperti IndoBERT. Pengembangan lanjutan masih dapat dilakukan pada domain berita lain seperti ekonomi, kesehatan, maupun politik domestik.

Referensi

- [1] M. Birjali, M. Kasri, dan A. Beni-Hssane, "A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends," Knowledge-Based Systems, Vol. 226, Art. No. 107134, 2021, Doi: 10.1016/j.knosys.2021.107134.
- [2] M. Wankhade, A. C. S. Rao, dan C. Kulkarni, "A Survey on Sentiment Analysis Methods, Applications, and Challenges," Artificial Intelligence Review, Vol. 55, pp. 5731–5780, 2022, DOI: 10.1007/s10462-022-10144-1. [Online]. Available: <https://link.springer.com/article/10.1007/s10462-022-10144-1>
- [3] K. Agarwal, R. Tayal, dan S. Gupta, "A Comparative Analysis of TF-IDF and Word Embeddings for Text Classification," International Journal of Advanced Computer Science and Applications, Vol. 14, No. 3, pp. 712–720, 2023, DOI: 10.14569/IJACSA.2023.0140381.
- [4] T. Hai, J. Zhou, S. A. Zadeh, dan O. A. Adetiloye, "Evaluation of Text Classification using Support Vector Machine Compare with Naive Bayes, Random Forest Decision Tree and K-NN," in Proc. ICACTCE 2023, Springer, 2023, DOI: 10.1007/978-3-031-37164-6_23.
- [5] M. U. Hasan, N. Islam, M. J. Islam, dan M. H. Kabir, "Enhancing Machine Learning-based Sentiment Analysis Through Feature Extraction Techniques," PLOS ONE, Vol. 19, No. 2, Art. No. e0297396, 2024, DOI: 10.1371/journal.pone.0297396. [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10866497>
- [6] I. K. Sari, P. D. Lestari, dan H. Fauzi, "Comparison of Naive Bayes and SVM Algorithms for Sentiment Analysis of PUBG Mobile on Google Play Store," Sistemasi: Jurnal Sistem Informasi, Vol. 13, No. 4, 2024, DOI: 10.32520/stmsi.v13i4.4814.
- [7] N. Fitriyani et al., "Performance Comparison of Naïve Bayes, Random Forest, and Support Vector Machine in Sentiment Analysis of the Weverse Application," Sistemasi: Jurnal Sistem Informasi, Vol. 14, No. 1, 2025, DOI: 10.32520/stmsi.v14i1.5893.
- [8] Muliana et al., "Analysis of Public Sentiment on Election Results using Naïve Bayes in Social Media X," Sistemasi: Jurnal Sistem Informasi, Vol. 13, No. 4, 2024, DOI: 10.32520/stmsi.v13i4.4592.
- [9] Firda et al., "Comparison of Rating-based and Inset Lexicon-based Labeling in Sentiment Analysis using SVM," Sistemasi: Jurnal Sistem Informasi, Vol. 14, No. 1, 2025, DOI: 10.32520/stmsi.v14i1.4795.
- [10] J. O. Leandro dan M. I. Fianty, "Evaluation of Sentiment Analysis Methods for Social Media Applications: A Comparison of SVM and Naïve Bayes," JOIV: International Journal on Informatics Visualization, Vol. 9, No. 2, pp. 796-807, 2025, DOI: 10.62527/joiv.9.2.2905.
- [11] M. Umer et al., "Sentiment Analysis and Comprehensive Evaluation of Supervised Machine Learning Models using Twitter Data on Russia-Ukraine War," SN Computer Science, Vol. 4, No. 3, 2023, DOI: 10.1007/s42979-023-01790-5.
- [12] C. Li, G. Zhan, dan Z. Li, "A Comparative Study of Machine Learning Models for Sentiment Analysis of Transboundary Rivers News Articles," Soft Computing, 2024, DOI: 10.1007/s00500-024-10357-2.
- [13] M. A. Rahman, M. S. Islam, M. S. Rahman, dan M. S. Hossain, "Leveraging Textual Information for Social Media News Categorization and Sentiment Analysis," PLOS ONE, Vol. 19, No. 3, Art. No. e0289719, 2024, DOI: 10.1371/journal.pone.0289719. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0289719>
- [14] M. G. Andriawan dan T. Ernawati, "Penggunaan Algoritma Naive Bayes dan Support Vector Machine untuk Analisis Sentimen Konflik Palestina dan Israel pada Platform X," Jurnal Informatika dan Teknik Elektro Terapan, Vol. 12, No. 3, 2024, DOI: 10.23960/jitet.v12i3.6250.
- [15] L. Zhang, "Features Extraction based on Naive Bayes Algorithm and TF-IDF for News Classification," PLOS ONE, Vol. 20, No. 4, Art. No. e0327347, 2025, DOI: 10.1371/journal.pone.0327347.