



Perbandingan Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Untuk Analisis Sentimen Komentar Politik Youtube

Bayu Aji Santoso^{*}, Bangkit Indarmawan Nugroho, Dzurrotu Umi Asyfiya

Sistem Informasi, STMIK YMI Tegal, Tegal, Indonesia

Email: ^{1,*}bayuajisantoso28@gmail.com, ²efbeterang@gmail.com, ³dzurrotuumi13@gmail.com

Email Penulis Korespondensi: bayuajisantoso28@gmail.com

Abstrak—Analisis sentimen merupakan salah satu bidang penting dalam pemrosesan bahasa alami yang banyak digunakan untuk memahami opini publik di media sosial. Penelitian ini membandingkan kinerja tiga algoritma machine learning, yaitu Naïve Bayes, Support Vector Machine (SVM), dan Random Forest, dalam menganalisis sentimen komentar YouTube. Dataset terdiri dari 15.257 komentar yang diperoleh dari kanal Indonesia Lawyers Club (ILC) dan Rakyat Bersuara. Proses penelitian meliputi tahapan preprocessing (cleaning, case folding, tokenizing, normalisasi dengan kamus slang, stopword removal, dan stemming), pelabelan data dengan pendekatan Lexicon-based menggunakan InSet Lexicon, pembagian data dengan rasio 80% data latih dan 20% data uji, serta evaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score, dilengkapi dengan uji K-fold cross validation. Hasil distribusi sentimen menunjukkan dominasi sentimen negatif 43,2%, disusul netral 34,9%, dan positif 21,9%. Evaluasi model memperlihatkan bahwa SVM unggul dengan akurasi 83,52%, presisi 83,55%, recall 83,52%, dan F1-score 83,52%, diikuti oleh Random Forest dengan akurasi 77,20%, sedangkan Naïve Bayes memperoleh hasil terendah 64,71%. Uji K-Fold semakin memperkuat hasil tersebut dengan akurasi terbaik 84,14% pada SVM. Dengan demikian, SVM dapat disimpulkan sebagai algoritma paling efektif untuk analisis sentimen komentar politik di YouTube.

Kata Kunci: Analisis Sentimen; Lexicon-based; Naive Bayes; Support Vector Machine, Random Forest

Abstract—Sentiment analysis is an important field in natural language processing that is widely used to understand public opinion on social media. This study compares the performance of three machine learning algorithms, namely Naïve Bayes, Support Vector Machine (SVM), and Random Forest, in analyzing YouTube comment sentiment. The dataset consists of 15,257 comments obtained from the Indonesia Lawyers Club (ILC) and Rakyat Bersuara channels. The research process includes preprocessing stages (cleaning, case folding, tokenizing, normalization with a slang dictionary, stopword removal, and stemming), data labeling with a Lexicon-based approach using InSet Lexicon, data division with a ratio of 80% training data and 20% test data, and evaluation using accuracy, precision, recall, and F1-score metrics, complemented by K-fold cross validation tests. The results of the sentiment distribution show a dominance of negative sentiment at 43.2%, followed by neutral at 34.9%, and positive at 21.9%. Model evaluation showed that SVM excelled with 83.52% accuracy, 83.55% precision, 83.52% recall, and 83.52% F1-score, followed by Random Forest with 77.20% accuracy, while Naïve Bayes achieved the lowest result at 64.71%. The K-Fold test further strengthened these results, with the best accuracy of 84.14% for SVM. Thus, SVM can be concluded as the most effective algorithm for analyzing political comment sentiment on YouTube.

Keywords: Analisis Sentimen; Lexicon-based; Naive Bayes; Support Vector Machine; Random Forest

1. PENDAHULUAN

Perkembangan teknologi informasi dalam satu dekade terakhir telah membawa transformasi signifikan dalam pola komunikasi masyarakat. Media sosial kini menjadi ruang diskusi utama bagi isu-isu sosial, politik, ekonomi, hingga budaya. Platform seperti YouTube, Twitter, Facebook, dan Instagram memungkinkan siapa pun menyuarakan pendapat, berbagi informasi, serta membentuk opini publik dalam skala yang masif. Komentar pengguna pada media sosial tidak hanya berfungsi sebagai respon terhadap suatu konten, melainkan juga sebagai refleksi dari emosi, persepsi, dan sikap masyarakat terhadap isu tertentu. Hal ini menjadikan media sosial sebagai sumber data yang sangat berharga untuk memahami opini publik secara lebih mendalam (Prasetya et al., 2021).

Salah satu pendekatan yang banyak digunakan untuk mengekstraksi informasi dari data teks di media sosial adalah analisis sentimen. Analisis sentimen berfokus pada pengklasifikasian opini ke dalam kategori positif, negatif, atau netral. Dengan teknik ini, peneliti dan praktisi dapat memperoleh wawasan yang lebih terstruktur mengenai persepsi masyarakat, yang kemudian dapat dimanfaatkan untuk mendukung pengambilan keputusan, baik dalam ranah bisnis, pemerintahan, maupun politik. Dua pendekatan utama dalam analisis sentimen adalah lexicon-based dan machine learning. Pendekatan lexicon-based memanfaatkan kamus sentimen, seperti InSet atau VADER, yang berisi daftar kata dengan skor polaritas tertentu. Kelebihan pendekatan ini terletak pada kemampuannya melakukan pelabelan otomatis tanpa memerlukan anotasi manual, sehingga lebih efisien dari sisi waktu dan biaya (Ardinata et al., 2024) (Arif Widiyan Subagio et al., 2024).

Di sisi lain, pendekatan machine learning digunakan untuk membangun model klasifikasi yang lebih adaptif terhadap variasi data. Beberapa algoritma yang paling populer dan banyak digunakan adalah Naïve Bayes, Support Vector Machine (SVM), dan Random Forest. Naïve Bayes dikenal sederhana, efisien, serta memiliki performa yang baik pada dataset teks yang besar meskipun kompleksitasnya rendah (Hudha et al., 2022). Sementara itu di penelitian lain, SVM sering dianggap unggul dalam tugas klasifikasi karena sifatnya yang mampu memaksimalkan margin pemisah antar kelas, sehingga lebih tahan terhadap noise dan data yang sulit dipisahkan (Munandar et al., 2023) (Ridwansyah, 2022) Adapun Random Forest dikenal stabil karena menggabungkan banyak pohon keputusan dan mampu

menangani data berdimensi tinggi, meskipun terkadang performanya menurun jika data bersifat tidak seimbang (Busrayan & Andrianingsih, 2025).

Sejumlah penelitian sebelumnya menunjukkan hasil yang beragam terkait efektivitas algoritma-algoritma tersebut (Hudha et al., 2022) melaporkan bahwa Naïve Bayes dapat mencapai akurasi 90,36% pada analisis komentar YouTube. (Fathoni et al., 2024) membandingkan pendekatan lexicon-based (InSet dan VADER) dengan SVM dan menemukan bahwa kombinasi keduanya mampu menghasilkan akurasi di atas 85%. Penelitian lain juga memperlihatkan performa tinggi, misalnya pada ulasan aplikasi dengan Naïve Bayes yang mencapai akurasi lebih dari 82% pada Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (Duei Putri et al., 2022). Sementara itu, (Munandar et al., 2023) dalam kajian pada aplikasi belajar online melaporkan bahwa SVM dapat mencapai akurasi hingga 99%, menunjukkan keunggulannya dalam klasifikasi teks. Pada domain berbeda, (Marga, 2022) meneliti kebijakan pemerintah terkait Covid-19 dengan Naïve Bayes dan ekstraksi fitur N-Gram, yang menghasilkan akurasi 84%.

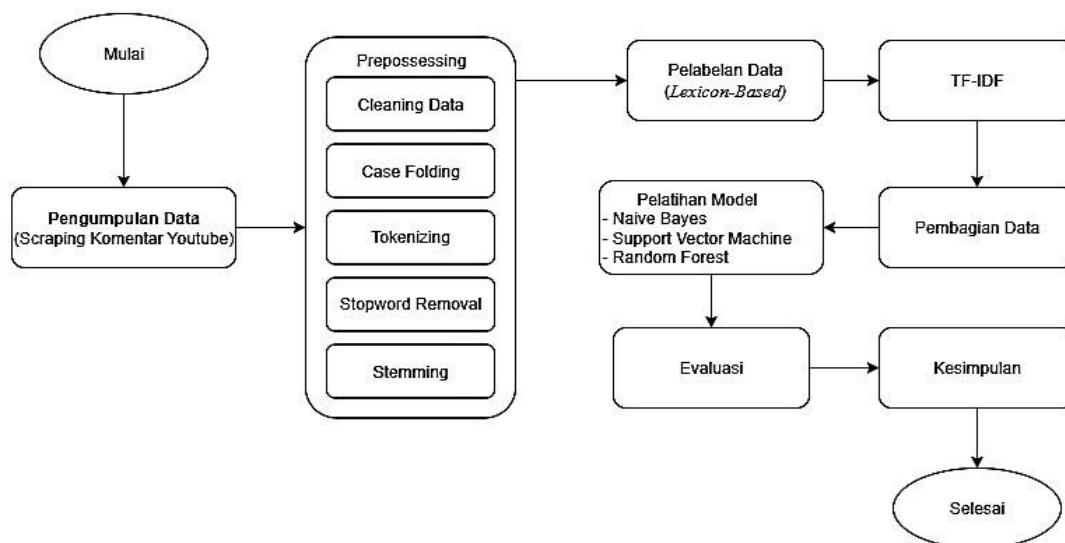
Namun demikian, sebagian besar penelitian terdahulu masih terbatas pada satu algoritma atau berfokus pada domain tertentu seperti e-commerce, ulasan aplikasi, dan isu kesehatan (Fauzi et al., 2023). Kajian mengenai sentimen politik, khususnya di YouTube, relatif masih jarang dilakukan. Padahal, komentar politik memiliki karakteristik yang khas: penggunaan bahasa informal, adanya singkatan, slang, hingga ekspresi emosi yang tinggi. Karakteristik ini menimbulkan tantangan tambahan, karena algoritma yang biasanya efektif pada data terstruktur atau semi-formal belum tentu optimal ketika diterapkan pada komentar politik di media sosial (Busrayan & Andrianingsih, 2025).

Untuk mengisi celah penelitian tersebut, studi ini difokuskan pada analisis sentimen komentar politik di YouTube, khususnya pada kanal Indonesia Lawyers Club (ILC) dan Rakyat Bersuara. Kedua kanal ini dipilih karena sering mengangkat topik politik yang sensitif dan kontroversial, sehingga menghasilkan volume komentar yang besar serta memuat opini publik yang beragam. Data yang diperoleh diberi label menggunakan pendekatan lexicon-based dengan kamus InSet, kemudian dibandingkan performa tiga algoritma machine learning, yaitu Naïve Bayes, SVM, dan Random Forest. Evaluasi dilakukan dengan menggunakan metrik akurasi, presisi, recall, F1-score, serta validasi silang dengan K-Fold Cross Validation.

Tujuan utama penelitian ini adalah untuk memberikan gambaran komparatif mengenai efektivitas tiga algoritma machine learning dalam menganalisis sentimen komentar politik berbahasa Indonesia. Selain itu, penelitian ini diharapkan dapat berkontribusi dalam literatur analisis opini publik di media sosial, khususnya dalam domain politik, yang hingga kini masih kurang dieksplorasi. Dengan adanya kajian ini, diharapkan diperoleh pemahaman yang lebih komprehensif mengenai bagaimana opini politik masyarakat Indonesia terbentuk dan terdistribusi di ruang digital, sekaligus memberikan rekomendasi algoritma yang paling tepat digunakan untuk konteks serupa pada penelitian berikutnya.

2. METODOLOGI PENELITIAN

Penelitian ini dilakukan untuk membandingkan performa algoritma Naïve Bayes, Support Vector Machine (SVM), dan Random Forest dalam menganalisis sentimen komentar YouTube pada kanal Indonesia Lawyers Club (ILC) dan Rakyat Bersuara. Dataset yang diperoleh sebanyak 15.257 komentar kemudian melalui tahapan penelitian. Berikut alur penelitian yang akan dilakukan yang ada pada Gambar 1, sebagai berikut:



Gambar 1. Alur Penelitian

Alur penelitian yang dilakukan dalam studi ini ditunjukkan pada Gambar 1, yang menggambarkan tahapan penelitian mulai dari pengumpulan data, preprocessing, pelabelan data, pembagian data, pelatihan model, hingga evaluasi dan kesimpulan. Setiap tahapan memiliki peran penting agar hasil analisis sentimen dapat lebih optimal.



2.1 Pengumpulan Data

Data dikumpulkan dengan metode web scraping, yaitu teknik pengambilan data otomatis menggunakan bahasa pemrograman Python. Library seperti Selenium atau BeautifulSoup digunakan untuk mengekstrak komentar, kemudian data disimpan dalam format terstruktur (CSV/Excel). Proses ini sejalan dengan penelitian terdahulu yang juga menggunakan teknik scraping untuk mengambil data teks dari media sosial seperti Twitter (Qi & Shabrina, 2023) maupun ulasan hotel di TripAdvisor (Utama et al., 2021). Jumlah data yang berhasil dikumpulkan dalam penelitian ini adalah 15.257 komentar. Komentar yang diperoleh masih dalam bentuk mentah (raw data), mengandung berbagai *noise* seperti tanda baca, hyperlink, emotikon, serta kata tidak baku. Oleh karena itu, diperlukan tahap preprocessing agar data siap digunakan pada proses pelabelan dan pemodelan selanjutnya.

2.2 Preprocessing

Tahap preprocessing merupakan langkah penting dalam mempersiapkan data teks agar siap diolah pada tahap analisis sentimen. Data komentar YouTube yang telah dikumpulkan masih berupa data mentah dengan berbagai *noise* seperti emotikon, tanda baca, singkatan, maupun kata tidak baku. Oleh karena itu, dilakukan beberapa tahapan preprocessing sebagai berikut:

1. *Cleaning Data*

Proses ini bertujuan untuk menghapus elemen-elemen yang tidak relevan seperti tanda baca, angka, URL, emotikon, dan karakter khusus lainnya. Cleaning dilakukan agar teks menjadi lebih bersih dan fokus pada konten utama (Ardinata et al., 2024).

2. *Case Folding*

Semua huruf dalam teks diubah menjadi huruf kecil (lowercase). Langkah ini penting untuk menyeragamkan teks sehingga kata dengan huruf besar maupun kecil dapat dianggap sama. Misalnya kata *Presiden* dan *presiden* akan diperlakukan identik.

3. *Tokenizing*

Teks kemudian dipecah menjadi unit kata atau token. Tokenisasi memungkinkan kata-kata dianalisis secara terpisah sehingga dapat dihitung frekuensi kemunculannya maupun digunakan dalam pembobotan fitur (Alita & Shodiqin, 2023).

4. *Normalisasi*

Normalisasi dilakukan untuk mengubah kata tidak baku atau slang menjadi bentuk standar. Dalam penelitian ini digunakan kamus slang yang memuat daftar kata gaul atau singkatan yang sering muncul pada komentar YouTube, misalnya kata *gk* → *tidak*, *bgt* → *banget*. Tahap ini penting untuk meningkatkan akurasi model karena komentar di media sosial sering kali menggunakan bahasa informal (Nurdiansyah & Dewi, 2023) (Ardinata et al., 2024).

5. *Stopword Removal*

Kata-kata umum yang sering muncul namun tidak memberikan kontribusi signifikan terhadap sentimen dihapus dari teks. Contohnya adalah kata “dan”, “atau”, “yang”, “di”. Penghapusan *stopword* membantu mengurangi *noise* dalam data sehingga model fokus pada kata yang lebih bermakna (Ardinata et al., 2024).

6. *Stemming*

Kata berimbuhan dikembalikan ke bentuk dasarnya menggunakan stemmer Bahasa Indonesia, misalnya *berjalan*, *berjalannya* → *jalan*. Dengan *stemming*, variasi kata dapat disatukan menjadi bentuk akar sehingga representasi fitur lebih ringkas dan akurat.

2.3 Pelabelan Data

Pelabelan data dilakukan dengan pendekatan Lexicon-based, yaitu metode analisis sentimen yang memanfaatkan kamus Lexicon. Kamus ini berisi daftar kata yang telah diberi bobot sentimen yang digunakan sebagai acuan dalam menghitung nilai sentimen dari setiap kata (Pratama et al., 2023). Ketika dilakukan pembobotan kata, setiap kata dalam kalimat dikelompokkan menggunakan metode Lexicon. Kata-kata dengan sentimen positif akan diberi skor +1, sementara kata-kata dengan sentimen negatif akan diberi skor -1, dan kata dengan sentiment netral diberi skor 0 (Arif Widiasan Subagio et al., 2024). Untuk memperjelas proses ini, Tabel 1 menampilkan contoh kata dalam kamus InSet beserta nilai lexicon yang telah ditentukan. Nilai positif menunjukkan kecenderungan sentimen positif, sedangkan nilai negatif menunjukkan kecenderungan sentimen negatif.

Tabel 1. Contoh Kata Pada Kamus InSet

No	Kata Pada InSet	Nilai Lexicon
1	Sayang	3
2	Kualitas	4
3	Mati	-3
4	Lelah	-5

Setiap teks akan dicari kata yang terdapat pada kamus lexicon lalu diberi nilai untuk setiap katanya, selanjutnya seluruh nilai kata pada satu teks akan dijumlahkan sehingga mendapatkan skor yang dapat menentukan teks tersebut termasuk sentimen positif, negatif, atau netral.



2.4 Split Data

Splitting data dilakukan untuk membagi dataset yang telah diberi label menjadi data latih (training data) dan data uji (testing data). Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur performa model terhadap data yang belum pernah dilihat sebelumnya (Anggina et al., 2022).

2.5 Pelatihan Model

Tahap pelatihan model dilakukan setelah data melalui proses preprocessing, pelabelan, dan pembagian menjadi data latih dan data uji. Pada penelitian ini digunakan tiga algoritma machine learning, yaitu Naïve Bayes, Support Vector Machine (SVM), dan Random Forest, yang dipilih karena masing-masing memiliki karakteristik berbeda dalam klasifikasi teks. Naïve Bayes dikenal sederhana dan cepat, SVM unggul dalam pemisahan kelas dengan margin optimal, sedangkan Random Forest mampu menangani data berdimensi tinggi dan mengurangi overfitting. Ketiga algoritma ini dipilih karena merupakan metode klasifikasi teks yang paling umum digunakan dalam penelitian analisis sentimen, serta sering dijadikan benchmark dalam membandingkan performa model pada berbagai domain. Algoritma lain seperti Logistic Regression atau neural network tidak digunakan karena penelitian ini difokuskan pada algoritma klasik yang terbukti stabil, efisien, dan memiliki interpretabilitas lebih baik untuk dataset berbahasa Indonesia.

2.6 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma dalam mengklasifikasikan komentar YouTube yang telah diberi label. Penelitian ini menggunakan ukuran akurasi, presisi, recall, dan F1-score. Akurasi digunakan untuk melihat ketepatan keseluruhan prediksi, presisi menilai ketepatan prediksi pada suatu kelas, recall mengukur sejauh mana model dapat menemukan data yang benar dalam suatu kelas, sedangkan F1-score digunakan untuk menilai keseimbangan antara presisi dan recall. Selain itu, digunakan Confusion matrix untuk mengetahui distribusi hasil klasifikasi pada setiap kelas sentimen. Agar hasil evaluasi lebih reliabel, penelitian ini juga menerapkan k-fold cross validation sehingga setiap data mendapat kesempatan untuk menjadi data latih maupun data uji (Hassan et al., 2024).

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Dataset

Data penelitian diperoleh dari komentar pada kanal YouTube Indonesia Lawyers Club (ILC) dan Rakyat Bersuara yang membahas isu pemakzulan wakil presiden. Dari proses *web scraping* menggunakan Python dengan library Selenium dan BeautifulSoup, berhasil dikumpulkan sebanyak 15.257 komentar. Jumlah ini cukup besar dan menggambarkan bahwa isu politik tersebut mendapatkan perhatian luas dari masyarakat. Karakteristik data sesuai dengan penelitian (Qi & Shabrina, 2023) dan (Utama et al., 2021) yang menegaskan bahwa data media sosial cenderung masif, tidak terstruktur, banyak mengandung *noise* seperti singkatan, *slang*, emotikon, dan kata tidak baku, sehingga memerlukan preprocessing yang intensif.

3.2 Preprocessing dan Pelabelan Data

3.2.1 Preprocessing

Tahap *preprocessing* menjadi krusial untuk mengubah data mentah menjadi bentuk yang siap dianalisis. Langkah-langkah yang dilakukan adalah *cleaning*, *case folding*, *tokenizing*, normalisasi, *stopword removal*, dan *stemming*. Proses ini sejalan dengan penelitian (Alfandi Safira & Hasan, 2023) yang menekankan bahwa pembersihan teks mampu meningkatkan kualitas fitur sehingga model lebih efektif.

Tabel 2 memperlihatkan contoh transformasi teks dari bentuk mentah hingga menjadi token akhir. Dari tabel tersebut dapat diamati bagaimana teks komentar politik yang awalnya penuh dengan singkatan dan kata tidak baku diubah menjadi representasi kata dasar yang lebih sederhana dan bermakna. Hal ini penting karena menurut (Anggina et al., 2022), representasi teks yang ringkas dan baku dapat meningkatkan performa model klasifikasi sentimen secara signifikan.

Tabel 2. *Preprocessing Data*

Text Komentar	clean_content	tokens	normalized_tokens	final_tokens	stemmed_tokens
Klu mau debat seruh undang pak luhut vs sunarko,brani ngk kong karni elias	klu mau debat seruh undang pak luhut vs sunarkobrani ngk kong karni elias	['klu', 'mau', 'debat', 'seruh', 'undang', 'pak', 'luhut', 'vs', 'sunarkobrani', 'ngk', 'kong', 'karni', 'elias']	['klu', 'mau', 'debat', 'seruh', 'undang', 'bapak', 'luhut', 'vs', 'sunarkobrani', 'tidak', 'kong', 'karni', 'elias']	['klu', 'debat', 'seruh', 'undang', 'luhut', 'vs', 'sunarkobrani', 'kong', 'karni', 'elias']	['klu', 'debat', 'seruh', 'undang', 'luhut', 'vs', 'sunarkobrani', 'kong', 'karni', 'elias']
Klo byk yg terlibat dan byk	klo byk yg terlibat dan byk	['klo', 'byk', 'yg', 'terlibat', 'dan']	['kalau', 'banyak', 'yang', 'terlibat']	['terlibat', 'penjilatnya']	['libat', 'jilat', 'ya', 'gin']



Text Komentar	clean_content	tokens	normalized_tokens	final_tokens	stemmed_tokens
penjilatnya ya gini, semua kasus ga beres2.....!!! Krn mrk jg takut terbongkar. δÿ~ ,δÿ⊘£δÿ~ ,δÿ°	penjilatnya ya gini semua kasus ga beres krn mrk jg takut terbongkar	'byk', 'penjilatnya', 'ya', 'gini', 'semua', 'kasus', 'ga', 'beres', 'krn', 'mrk', 'jg', 'takut', 'terbongkar']	'dan', 'banyak', 'penjilatnya', 'ya', 'gini', 'semua', 'kasus', 'tidak', 'selesai', 'karena', 'mrk', 'juga', 'takut', 'terbongkar']	'ya', 'gini', 'selesai', 'mrk', 'takut', 'terbongkar']	'selesai', 'mrk', 'takut', 'bongkar']
Hahaha kumis tebal jgn cuma ngomong ingatkan pernah melakukan penghianatan melakukan makar	hahaha kumis tebal jgn cuma ngomong ingatkan pernah melakukan penghianatan melakukan makar	'kumis', 'tebal', 'jgn', 'cuma', 'ngomong', 'ingatkan', 'pernah', 'melakukan', 'penghianatan', 'melakukan', 'makar']	['hahaha', 'hanya', 'ngomong', 'ingatkan', 'pernah', 'melakukan', 'penghianatan', 'melakukan', 'makar']	['hahaha', 'kumis', 'tebal', 'ngomong', 'ingatkan', 'penghianatan', 'makar']	['hahaha', 'kumis', 'tebal', 'ngomong', 'ingat', 'penghianatan', 'makar']
Angkat lagi tema, Pemakzulan Presiden Ke-4 RI. GUSDUR, atas tuduhan" nya yang sampe sekarang tidak terbukti.	angkat lagi tema pemakzulan presiden ke ri gusdur atas tuduhan nya yang sampe sekarang tidak terbukti	['angkat', 'lagi', 'tema', 'pemakzulan', 'presiden', 'ke', 'ri', 'gusdur', 'atas', 'tuduhan', 'nya', 'yang', 'sampe', 'sekarang', 'tidak', 'terbukti']	['angkat', 'lagi', 'tema', 'pemakzulan', 'presiden', 'ke', 'ri', 'gusdur', 'atas', 'tuduhan', 'nya', 'yang', 'sampai', 'sekarang', 'tidak', 'terbukti']	['angkat', 'tema', 'pemakzulan', 'presiden', 'ri', 'gusdur', 'tuduhan', 'nya', 'terbukti']	['angkat', 'tema', 'makzul', 'presiden', 'ri', 'gusdur', 'tuduh', 'nya', 'bukti']
Negarawan Rela di Maksulkan demi negaranya.	negarawan rela di maksulkan demi negaranya	['negarawan', 'rela', 'di', 'maksulkan', 'demi', 'negaranya']	['negarawan', 'rela', 'di', 'maksulkan', 'demi', 'negaranya']	['negarawan', 'rela', 'maksulkan', 'negaranya']	['negarawan', 'rela', 'maksulkan', 'negara']
.....
Yang liat karna ada rocky gerung ada ?δÿ~ ,	yang liat karna ada rocky gerung ada	['yang', 'liat', 'karna', 'ada', 'rocky', 'gerung', 'ada']	['yang', 'lihat', 'karena', 'ada', 'rocky', 'gerung', 'ada']	['lihat', 'rocky', 'gerung']	['lihat', 'rocky', 'gerung']
malu punya wapres seperti itu.δÿ⦶²	malu punya wapres seperti itu	['malu', 'punya', 'wapres', 'seperti', 'itu']	['malu', 'punya', 'wapres', 'seperti', 'itu']	['malu', 'wapres']	['malu', 'wapres']
Klu mau debat seruh undang pak luhut vs sunarko,brani ngk kong karni elias	klu mau debat seruh undang pak luhut vs sunarkobrani ngk kong karni elias	['klu', 'mau', 'debat', 'seruh', 'undang', 'pak', 'luhut', 'vs', 'sunarkobrani', 'ngk', 'kong', 'karni', 'elias']	['klu', 'mau', 'debat', 'seruh', 'undang', 'bapak', 'luhut', 'vs', 'sunarkobrani', 'tidak', 'kong', 'karni', 'elias']	['klu', 'debat', 'seruh', 'undang', 'luhut', 'vs', 'sunarkobrani', 'kong', 'karni', 'elias']	['klu', 'debat', 'seruh', 'undang', 'luhut', 'vs', 'sunarkobrani', 'kong', 'karni', 'elias']
Klo byk yg terlibat dan byk penjilatnya ya gini, semua kasus ga beres2.....!!! Krn mrk jg takut terbongkar. δÿ~ ,δÿ⊘£δÿ~ ,δÿ°	klo byk yg terlibat dan byk penjilatnya ya gini semua kasus ga beres krn mrk jg takut terbongkar	['klo', 'byk', 'yg', 'terlibat', 'dan', 'byk', 'penjilatnya', 'ya', 'gini', 'semua', 'kasus', 'ga', 'beres', 'krn', 'mrk', 'jg', 'takut', 'terbongkar']	['kalau', 'banyak', 'yang', 'terlibat', 'dan', 'banyak', 'penjilatnya', 'ya', 'gini', 'semua', 'kasus', 'tidak', 'selesai', 'karena', 'mrk', 'juga', 'takut', 'terbongkar']	['terlibat', 'penjilatnya', 'ya', 'gini', 'selesai', 'mrk', 'takut', 'terbongkar']	['libat', 'jilat', 'ya', 'gin', 'selesai', 'mrk', 'takut', 'bongkar']

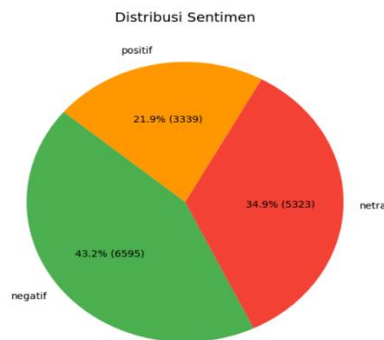
3.2.2 Pelabelan Data

Pelabelan dilakukan menggunakan pendekatan *Lexicon-based* dengan kamus **InSet** yang berisi 3.609 kata positif dan 6.609 kata negatif dengan bobot nilai antara -5 hingga +5 (Arif Widiyan Subagio et al., 2024). Dengan metode ini, setiap kata diberi skor polaritas dan kemudian dijumlahkan untuk menentukan kelas sentimen.

Tabel 3. Pelabelan sentiment menggunakan pendekatan Lexicon-based

stemmed_tokens	sentiment	polarity_score	pos_count	neg_count
['klu', 'debat', 'seruh', 'undang', 'luhut', 'vs', 'sunarkobrani', 'kong', 'karni', 'elias']	positif	2	3	1
['libat', 'jilat', 'ya', 'gin', 'selesai', 'mrk', 'takut', 'bongkar']	negatif	-1	3	4
['hahaha', 'kumis', 'tebal', 'ngomong', 'ingat', 'penghianatan', 'makar']	netral	0	1	1
['angkat', 'tema', 'makzul', 'presiden', 'ri', 'gusdur', 'tuduh', 'nya', 'bukti']	positif	1	2	1
['negarawan', 'rela', 'maksulkan', 'negara']	netral	0	1	1
.....
['lihat', 'rocky', 'gerung']	positif	1	1	0
['malu', 'wapres']	netral	0	1	1

Hasil pelabelan data menggunakan pendekatan Lexicon-based menghasilkan tiga kategori sentimen, yaitu positif, netral, dan negatif. Dari total 15.257 komentar YouTube, diperoleh distribusi sebagai berikut : kategori negatif berjumlah 6.595 komentar (43,2%), kategori netral sebanyak 5.323 komentar (34,9%), dan kategori positif sejumlah 3.339 komentar (21,9%). Distribusi ini divisualisasikan pada Gambar 2, yang memperlihatkan dominasi opini negatif. Temuan ini konsisten dengan penelitian (Busrayan & Andrianingsih, 2025) yang menemukan bahwa isu politik di media sosial cenderung memicu sentimen negatif lebih dominan dibanding positif.



Gambar 2. Distribusi Sentimen

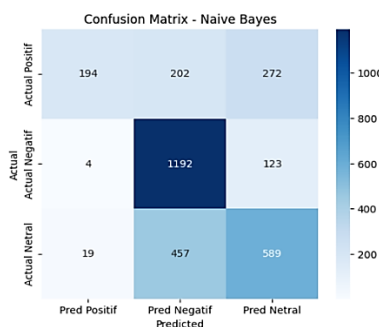
3.3 Pembagian Data Training dan Testing

Dataset dibagi menjadi 80% data latih (12.205 komentar) dan 20% data uji (3.052 komentar). Proporsi ini sejalan dengan praktik umum pada penelitian NLP (Ridwansyah, 2022) dan (Fathoni et al., 2024) yang merekomendasikan porsi 80/20 karena mampu menjaga keseimbangan antara data latih yang memadai dan data uji yang representatif.

3.4 Hasil Eksperimen Model

3.4.1 Algoritma Naïve Bayes

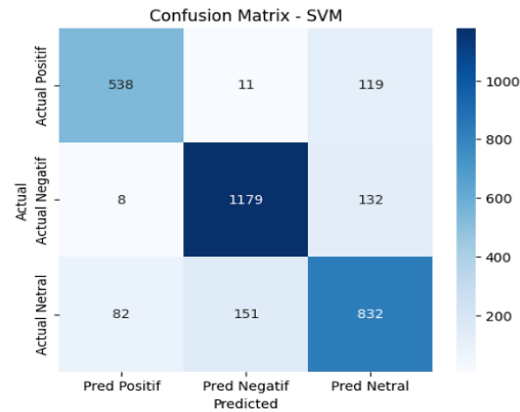
Setelah dilakukan pelatihan dan pengujian terhadap algoritma Naïve Bayes, Gambar 3. Merupakan visualisasi Confusion matrix setelah dilakukan pelatihan model:



Gambar 3. Confusion matrix Naïve Bayes

Confusion matrix algoritma Naïve Bayes menunjukkan bahwa model cukup baik dalam mengenali komentar negatif dengan 1.192 prediksi benar, meskipun masih terdapat kesalahan klasifikasi ke kelas netral sebanyak 123 dan ke kelas positif sebanyak 4 komentar. Untuk kelas positif, hanya 194 komentar yang terdeteksi dengan benar, sementara sisanya banyak salah diprediksi sebagai negatif (202) maupun netral (272), sehingga akurasi pada kelas ini relatif rendah. Sementara itu, pada kelas netral, sebanyak 589 komentar berhasil diprediksi dengan benar, tetapi masih terdapat kesalahan prediksi cukup besar, yaitu 457 komentar netral diprediksi sebagai negatif dan 19 sebagai positif. Naïve Bayes menunjukkan akurasi rendah dengan kelemahan utama pada kelas positif. Hal ini disebabkan asumsi independensi antar fitur yang tidak sesuai dengan komentar politik, di mana kata-kata sering muncul dalam konteks tertentu dan saling terkait (Darwis et al., 2021). Akibatnya, banyak kesalahan klasifikasi terutama pada komentar bernuansa positif.

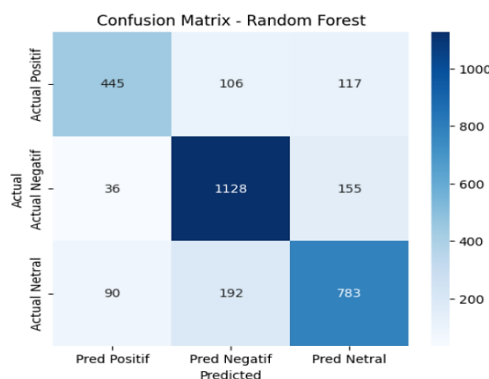
3.4.2 Algoritma SVM



Gambar 4. Confusion matrix SVM

Confusion matrix algoritma SVM menunjukkan kinerja yang lebih baik dibandingkan Naïve Bayes, terutama dalam mengenali semua kelas sentimen. Pada kelas positif, sebanyak 538 komentar berhasil diprediksi dengan benar, dengan sedikit kesalahan ke kelas negatif (11) dan netral (119). Pada kelas negatif, performa model sangat baik dengan 1.179 prediksi benar, sementara hanya 8 komentar salah diprediksi sebagai positif dan 132 sebagai netral. Untuk kelas netral, sebanyak 832 komentar berhasil diklasifikasikan dengan benar, meskipun masih ada 151 komentar diprediksi sebagai negatif dan 82 sebagai positif. Secara keseluruhan, SVM menghasilkan performa tertinggi, dengan akurasi 83,52%. Kelebihan SVM adalah kemampuannya memisahkan kelas dengan margin optimal dan ketahanannya terhadap *noise* (Fathoni et al., 2024). Hal ini sangat relevan untuk data komentar YouTube yang penuh variasi bahasa. Hasil ini juga konsisten dengan (Miftahusalam et al., 2022) yang membuktikan bahwa SVM unggul dalam analisis sentimen Twitter, serta (Azmi et al., 2023) yang menekankan bahwa penggunaan kernel SVM efektif untuk data dengan pola non-linear seperti komentar politik.

3.4.3 Algoritma Random Forest



Gambar 5. Confusion matrix Random Forest

Confusion matrix algoritma Random Forest memperlihatkan bahwa model ini cukup baik dalam mengklasifikasikan ketiga kelas sentimen, meskipun performanya tidak setinggi SVM. Pada kelas positif, sebanyak 445 komentar berhasil diprediksi dengan benar, sementara masih terdapat kesalahan prediksi ke kelas negatif sebanyak 106 dan ke kelas netral 117 komentar. Untuk kelas negatif, hasilnya cukup dominan dengan 1.128 prediksi benar, walaupun masih ada 36 komentar yang salah diprediksi sebagai positif dan 155 sebagai netral. Pada kelas netral, sebanyak 783 komentar berhasil diklasifikasikan dengan benar, namun masih ditemukan 192 komentar diprediksi sebagai negatif dan



90 sebagai positif. Random Forest menunjukkan performa menengah dengan akurasi 77,20%. Kelebihannya ada pada kestabilan prediksi, tetapi kelemahannya adalah bias terhadap kelas dominan (negatif). Menurut (Miftahusalam et al., 2022), kelemahan Random Forest pada data ketidakseimbangan kelas dapat membuat prediksi kelas minoritas (positif) kurang optimal.

3.5 Perbandingan dan Diskusi

Hasil evaluasi menunjukkan bahwa algoritma Naïve Bayes memiliki kinerja paling rendah dengan akurasi 64,71%, presisi 68,29%, recall 64,71%, dan F1-score 62,16%, yang menandakan model ini kurang optimal dalam mengklasifikasikan sentimen komentar. Algoritma Random Forest memberikan hasil lebih baik dengan akurasi 77,20%, presisi 77,14%, recall 77,20%, dan F1-score 77,02%, meskipun masih terdapat kelemahan dalam menjaga konsistensi prediksi pada semua kelas. Sementara itu, algoritma Support Vector Machine (SVM) terbukti menjadi yang paling unggul dengan akurasi 83,52%, presisi 83,55%, recall 83,52%, dan F1-score 83,52%, sehingga dapat disimpulkan bahwa SVM memiliki performa paling stabil dan akurat dalam melakukan analisis sentimen komentar YouTube dibandingkan dengan dua algoritma lainnya.

Tabel 4. Hasil Performa Algoritma

Confusion matrix	Algoritma		
	Naïve Bayes	SVM	Random Forest
Accuracy	64.71 %	83.52 %	77.20 %
Precision	68.29 %	83.55 %	77.14 %
Recall	64.71 %	83.52 %	77.20 %
F1-Score	62.16 %	83.52 %	77.02 %

Tabel 5. Hasil K-fold cross validation

Fold	Akurasi Algoritma		
	Naïve Bayes	SVM	Random Forest
Fold 2	63.24 %	84.14 %	75.88 %
Fold 3	63.96 %	82.90 %	76.21 %
Fold 4	65.07 %	81.98 %	76.67 %
Fold 5	64.74 %	83.75 %	76.67 %
Fold 6	64.61 %	83.09 %	77.33 %
Fold 7	62.91 %	83.75 %	76.28 %
Fold 8	65.31 %	83.67 %	77.84 %
Fold 9	65.38 %	83.21 %	77.84 %
Fold 10	65.31 %	83.08 %	78.16 %

Hasil penelitian ini memperkuat literatur bahwa SVM unggul dalam analisis sentimen politik. Misalnya, (Fathoni et al., 2024) melaporkan akurasi >85% pada data publik, sedangkan (Miftahusalam et al., 2022) juga menemukan SVM lebih baik dibanding Random Forest pada analisis sentimen Twitter. Sebaliknya, kelemahan Naïve Bayes pada data politik sejalan dengan (Ardinata et al., 2024) yang menemukan akurasi hanya 79% karena asumsi independensi fitur.

Secara teoritis, keunggulan SVM dapat dijelaskan karena data komentar politik memiliki fitur yang dapat dipisahkan secara non-linear. Kernel SVM memungkinkan pembentukan ruang vektor berdimensi tinggi untuk memisahkan kelas lebih optimal (Azmi et al., 2023). Random Forest tetap relevan untuk kestabilannya (Busrayan & Andrianingsih, 2025), tetapi sensitif pada *class imbalance*. Naïve Bayes justru kurang cocok karena komentar politik sering kali memiliki kata-kata yang saling berkaitan sehingga asumsi independensinya tidak terpenuhi (Ramadhani & Suryono, 2024).

4. KESIMPULAN

Penelitian ini menganalisis sentimen komentar politik pada YouTube dengan menggunakan tiga algoritma machine learning, yaitu Naïve Bayes, Random Forest, dan Support Vector Machine (SVM), serta pelabelan berbasis Lexicon (InSet). Dari total 15.257 komentar yang dikaji, distribusi sentimen menunjukkan dominasi komentar negatif (43,2%), disusul netral (34,9%), dan positif (21,9%). Hal ini mengindikasikan bahwa opini publik pada isu politik cenderung lebih kritis dibanding mendukung. Evaluasi model memperlihatkan bahwa Naïve Bayes memiliki performa paling rendah (akurasi 64,71%), yang menunjukkan keterbatasannya dalam menangani kompleksitas teks politik karena asumsi independensi antar fitur tidak terpenuhi. Random Forest memberikan performa menengah (akurasi 77,20%), cukup stabil namun masih rentan terhadap ketidakseimbangan kelas. Sebaliknya, SVM konsisten unggul dengan akurasi 83,52% dan hasil K-Fold Cross Validation mencapai 84,14%, yang menegaskan bahwa algoritma ini paling efektif dalam mengklasifikasikan komentar politik dengan karakteristik bahasa yang beragam, informal, dan penuh noise. Temuan ini sejalan dengan berbagai penelitian terdahulu yang juga menempatkan SVM sebagai algoritma unggulan



dalam analisis sentimen, khususnya pada domain yang kompleks seperti politik. Namun, penelitian ini memiliki keterbatasan, yaitu: (1) ketergantungan pada metode pelabelan Lexicon-based yang bisa kurang akurat dalam menangkap konteks ironi atau sarkasme, (2) distribusi kelas yang tidak seimbang sehingga memengaruhi performa algoritma tertentu, serta (3) ruang lingkup yang terbatas pada dua kanal YouTube sehingga generalisasi hasil masih terbatas. Untuk penelitian mendatang, disarankan penggunaan metode pelabelan hibrida yang menggabungkan Lexicon-based dan supervised learning agar klasifikasi lebih akurat. Selain itu, teknik penyeimbangan data seperti SMOTE dapat diterapkan guna mengatasi class imbalance. Perbandingan dengan pendekatan berbasis deep learning seperti LSTM atau BERT juga perlu dilakukan untuk menguji apakah model modern dapat memberikan peningkatan signifikan dibanding algoritma klasik. Penelitian juga dapat diperluas pada isu politik lain maupun platform media sosial berbeda untuk memperoleh gambaran opini publik yang lebih komprehensif dan representatif.

REFERENCES

- Alfandi Safira, & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier. *ZONAsi: Jurnal Sistem Informasi*, 5(1), 59–70. <https://doi.org/10.31849/zn.v5i1.12856>
- Alita, D., & Shodiqin, R. A. (2023). Sentimen Analisis Vaksin Covid-19 Menggunakan Naive Bayes Dan Support Vector Machine. *Journal of Artificial Intelligence and Technology Information (JAITI)*, 1(1), 1–12. <https://doi.org/10.58602/jaiti.v1i1.20>
- Anggina, S., Setiawan, N. Y., & Bachtiar, F. A. (2022). Analisis Ulasan Pelanggan Menggunakan Multinomial Naive Bayes Classifier dengan Lexicon-Based dan TF-IDF Pada Formaggio Coffee and Resto. *Is The Best Accounting Information Systems and Information Technology Business Enterprise This Is Link for OJS Us*, 7(1), 76–90. <https://doi.org/10.34010/aisthebest.v7i1.7072>
- Ardinata, P. M. S., Permana, A. A. J., & Wijaya, I. N. S. W. (2024). Identifikasi Dan Normalisasi Teks Slang Dengan Fasttext Pada Twitter Dalam Bahasa Indonesia. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 21(1), 34–44.
- Arif Widiyan Subagio, Anggraini Puspita Sari, & Andreas Nugroho Sihananto. (2024). Klasifikasi Lexicon-Based Sentiment Analisis Tragedi Kanjuruhan pada Twitter Menggunakan Algoritma Convolutional Neural Network. *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer*, 4(1), 166–177. <https://doi.org/10.55606/juisik.v4i1.759>
- Azmi, T. A. U., Hakim, L., Novitasari, D. C. R., & Utami, W. D. U. D. (2023). Application Random Forest Method for Sentiment Analysis in Jamsostek Mobile Review. *Telematika*, 20(1), 117. <https://doi.org/10.31315/telematika.v20i1.8868>
- Busrayan, I., & Andrianingsih. (2025). Analisis Sentimen Pelanggan Terhadap Aplikasi Wondr By Bni Menggunakan Naive Bayes, Support Vector Machine (Svm), Dan K-Nearest Neighbor (Knn). *Journal of Computer Science and Information Technology*, 2(2), 263–274.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Duei Putri, D., Nama, G. F., & Sulistiono, W. E. (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika Dan Teknik Elektro Terapan*, 10(1), 34–40. <https://doi.org/10.23960/jitet.v10i1.2262>
- Fathoni, M. F. N., Puspaningrum, E. Y., & Sihananto, A. N. (2024). Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM. *Modem: Jurnal Informatika Dan Sains Teknologi*, 2(3), 62–76.
- Fauzi, F., Setiayani, W., Utami, T. W., Yuliyanto, E., & Harmoko, I. W. (2023). Comparison of Random Forest and Naive Bayes Classifier Methods in Sentiment Analysis on Climate Change Issue. *Barekeng*, 17(3), 1439–1448. <https://doi.org/10.30598/barekengvol17iss3pp1439-1448>
- Hassan, F. M., Yacob, A., & Ghazali, N. E. (2024). Sentiment Analysis of ChatGPT Using the KNN Algorithm and K-Fold Cross-Validation Optimization of the K Value. *International Journal of Informatics and Computing*, 1(2), 49–55. <https://www.researchgate.net/publication/389515999>
- Hudha, M., Supriyati, E., & Listyorini, T. (2022). Analisis Sentimen Pengguna Youtube Terhadap Tayangan #Matanajwamenantiterawan Dengan Metode Naive Bayes Classifier. *JIKO (Jurnal Informatika Dan Komputer)*, 5(1), 1–6. <https://doi.org/10.33387/jiko.v5i1.3376>
- Marga, N. S. (2022). Sentimen Analisis Tentang Kebijakan Pemerintah Terhadap Kasus Corona Menggunakan Metode Naive Bayes. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 2(4), 453–463. <https://doi.org/10.33365/jatika.v2i4.1602>
- Miftahusalam, A., Nuraini, A. F., Khoirunisa, A. A., & Pratiwi, H. (2022). Perbandingan Algoritma Random Forest, Naive Bayes, dan Support Vector Machine Pada Analisis Sentimen Twitter Mengenai Opini Masyarakat Terhadap Penghapusan Tenaga Honorer. *Seminar Nasional Official Statistics*, 2022(1), 563–572. <https://doi.org/10.34123/semnasoffstat.v2022i1.1410>
- Munandar, A. A., Farikhin, F., & Widodo, C. E. (2023). Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM. *JOINTECS (Journal of Information Technology and Computer Science)*, 8(2), 77. <https://doi.org/10.31328/jointecs.v8i2.4747>
- Nurdiansyah, R. L., & Dewi, K. E. (2023). Pengaruh Information Gain Dan Normalisasi Kata Pada Analisis Sentimen Berbasis Aspek. *KOMPUTA : Jurnal Ilmiah Komputer Dan Informatika*, 12(2), 80–90.



- Prasetya, Y. N., Winarso, D., & Syahril. (2021). Penerapan Lexicon Based Untuk Analisis Sentimen Pada Twitter Terhadap Isu Covid-19. *Jurnal Fasilkom*, 11(2), 97–103. <https://ejournal.umri.ac.id/index.php/JIK/article/view/2772/1566>
- Pratama, M. R., Ramadha, Y. R., & Komara, M. A. (2023). Analisis Sentimen BRI Mo dan BCA Mobile Menggunakan Support Vector Machine dan Lexicon Based. *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 12(3), 1439. <https://doi.org/10.35889/jutisi.v12i3.1431>
- Qi, Y., & Shabrina, Z. (2023). Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach. *Social Network Analysis and Mining*, 13(1), 1–14. <https://doi.org/10.1007/s13278-023-01030-x>
- Ramadhani, B., & Suryono, R. R. (2024). Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse. *Jurnal Media Informatika Budidarma*, 8(2), 714. <https://doi.org/10.30865/mib.v8i2.7458>
- Ridwansyah, T. (2022). Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naïve Bayes Classifier. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 2(5), 178–185. <https://doi.org/10.30865/klik.v2i5.362>
- Utama, I. P. A. M., Prasetyowati, S. S., & Sibaroni, Y. (2021). Multi-Aspect Sentiment Analysis Hotel Review Using RF, SVM, and Naïve Bayes based Hybrid Classifier. *Jurnal Media Informatika Budidarma*, 5(2), 630. <https://doi.org/10.30865/mib.v5i2.2959>