

ANALISIS SENTIMEN TWITTER TERHADAP KONFLIK DI PAPUA MENGGUNAKAN PERBANDINGAN NAIVE BAYES DAN SVM

M.Rafli Saputra*¹⁾, Parjito²⁾

1. Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia
2. Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia

Article Info

Kata Kunci: Analisis Sentimen; Naïve Bayes; SVM; Konflik Papua; Papua

Keywords: *Sentiment Analysis; Naïve Bayes; SVM; Papua conflict; Papua*

Article history:

Received 11 November 2024

Revised 15 Desember 2024

Accepted 14 Januari 2025

Available online 15 Maret 2025

DOI :

<https://doi.org/10.29100/jupi.v10i3.6180>

* Corresponding author.

M.Rafli Saputra

E-mail address:

raflisaputra2370@gmail.com

ABSTRAK

Konflik di Papua merupakan isu yang kompleks dan telah berlangsung selama beberapa dekade, melibatkan berbagai faktor seperti politik, ekonomi, sosial, dan budaya. Ketegangan antara pemerintah Indonesia dan kelompok separatis Papua sering kali memicu konflik bersenjata, pelanggaran hak asasi manusia, dan ketidakstabilan regional. Konflik ini juga menarik perhatian berbagai pihak, termasuk masyarakat internasional dan pengguna media sosial, khususnya di platform Twitter.

Data yang dikumpulkan dari Twitter menggunakan kata kunci terkait konflik Papua. Proses analisis meliputi tahap pengumpulan data, pra-pemrosesan teks, dan pelabelan sentimen (positif, negatif, netral). Penelitian ini menggunakan 5723 data yang diperoleh melalui teknik web scraping terkait topik tersebut. Tujuan penelitian ini adalah membandingkan performa dua algoritma klasifikasi yang populer dalam analisis sentimen, yaitu Naïve Bayes dan Support Vector Machine (SVM). Sebelum perbandingan dilakukan, optimasi SMOTE (Synthetic Minority Over-sampling Technique) digunakan untuk menyeimbangkan jumlah data minoritas dan mayoritas, sehingga kedua algoritma dapat belajar secara lebih efektif dari setiap kelas sentimen. Hasil perbandingan menunjukkan bahwa algoritma Naïve Bayes memiliki akurasi sebesar 95%, sedangkan SVM mencapai akurasi 99%, dengan presisi 99%, recall 98%, dan F1-Score 99%. Evaluasi performa dilakukan dengan menganalisis confusion matrix dari setiap algoritma. Kesimpulannya, SVM dapat menjadi pilihan yang lebih baik untuk analisis sentimen mengenai konflik Papua. Penelitian ini memberikan kontribusi penting dalam memahami opini masyarakat terkait konflik di Papua.

ABSTRACT

The conflict in Papua is a complex issue and has been going on for several decades, involving various factors such as political, economic, social and cultural. Tensions between the Indonesian government and Papuan separatist groups often trigger armed conflict, human rights violations and regional instability. This conflict has also attracted the attention of various parties, including the international community and social media users, especially on the Twitter platform.

Data collected from Twitter uses keywords related to the Papua conflict. The analysis process includes data collection stages, text pre-processing, and sentiment labeling (positive, negative, neutral). This research uses 5723 data obtained through web scraping techniques related to this topic. The aim of this research is to compare the performance of two classification algorithms that are popular in sentiment analysis, namely Naïve Bayes and Support Vector Machine (SVM). Before the comparison is carried out, SMOTE (Synthetic Minority Over-sampling Technique) optimization is used to balance the amount of minority and majority data, so that both algorithms can learn more effectively from each sentiment class. The comparison results show that the Naïve Bayes algorithm has an accuracy of 95%, while the SVM achieves an accuracy of 99%, with a precision of 99%, a recall of 98%, and an F1-Score of 99%. Performance evaluation is carried out by analyzing the confusion matrix of each algorithm. In conclusion, SVM can be a better choice for sentiment analysis regarding the Papua conflict. This research makes an important contribution to understanding public opinion regarding the conflict in Papua.

I. PENDAHULUAN

Negara Indonesia memproklamasikan kemerdekaannya pada 17 Agustus 1945. Proklamasi tersebut menjadi titik awal perlawanan diplomatik dan bersenjata dalam Revolusi Nasional Indonesia melawan pasukan dan pendukung Belanda. Indonesia mengklaim seluruh wilayah Hindia Belanda, termasuk Papua Barat, sebagai bagian dari negara baru tersebut. Namun, Belanda menolak klaim ini karena menganggap Papua Barat masih menjadi provinsi Kerajaan Belanda. Dalam Konferensi Meja Bundar tahun 1949, Belanda dan Indonesia membahas status Papua Barat tetapi tidak mencapai kesepakatan akhir, meski setuju untuk membahasnya lagi dalam setahun[1]. Aksi kriminal dan separatis oleh Kelompok Kriminal Bersenjata di Papua telah menyebabkan korban sejak tahun 1965[2]. Siklus kekerasan di Papua tidak terbatas pada konflik bersenjata saja. Contohnya, insiden rasisme terhadap mahasiswa Papua di Surabaya pada Agustus 2019 memicu kerusuhan massal di Paniai, Fak-Fak, Sorong, dan Wamena. Dalam kerusuhan tersebut, dilaporkan 6 warga sipil dan 1 tentara menjadi korban, meskipun Polri membantah laporan ini. Konfirmasi jumlah korban sulit dilakukan karena pemerintah Indonesia menutup akses jurnalis dan internet saat kerusuhan berlangsung[3].

Penyebab konflik Papua meliputi sengketa historis terkait integrasi Irian Barat ke Indonesia, kasus pelanggaran hak asasi manusia yang belum diselesaikan, serta meningkatnya marginalisasi dan diskriminasi terhadap orang Papua, yang hingga kini masih belum terselesaikan[4]. Organisasi Papua Merdeka (OPM) adalah istilah yang digunakan untuk menggambarkan gerakan pro-kemerdekaan Papua yang merasa tidak puas dengan kebijakan pemerintah Indonesia sejak Papua bergabung dengan Indonesia[1]. Dalam wawancara dengan BBC News Indonesia, juru bicara TPNPB-OPM, Sebby Sambom, menyatakan bahwa selama lima tahun terakhir mereka telah meminta warga non-Papua untuk meninggalkan apa yang mereka sebut sebagai "wilayah perang" di Papua. "Wilayah perang tersebut mencakup Ndugama, Intan Jaya, Puncak Jaya, Puncak Papua, Pegunungan Bintang, dan Sorong-Maybrat," ujar Sebby kepada BBC Indonesia[5]. Kebijakan pemerintah yang tidak adil menyebabkan kekecewaan di kalangan orang Papua, sehingga gerakan separatis seperti Organisasi Papua Merdeka (OPM) tetap bertahan dan berkembang. OPM bertujuan meningkatkan kesejahteraan rakyat Papua, memberikan rasa aman dari militer Indonesia, dan mencapai kemerdekaan[6].

Analisis sentimen adalah bidang penelitian yang cukup populer dan dianggap memberikan manfaat dalam berbagai aspek. Biasanya, analisis sentimen digunakan dalam prediksi harga saham, isu politik, kepuasan terhadap produk atau layanan, analisis reputasi, dan lainnya. Salah satu penerapan analisis sentimen adalah terhadap isu-isu yang muncul di Twitter[7]. Metode yang digunakan untuk klasifikasi pengujian ini adalah Naïve Bayes dan Support Vector Machine. Naïve Bayes merupakan metode yang sangat populer karena kemudahannya dan kemampuannya untuk memenuhi persyaratan di berbagai domain yang berbeda. Keunggulan dari metode Support Vector Machine adalah kemampuannya mengidentifikasi hyperplane yang berbeda, sehingga meningkatkan margin antara dua kelas yang berbeda. Namun, metode Support Vector Machine juga memiliki kelemahan dalam pemilihan fitur yang tepat[8].

Naïve Bayes adalah metode klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dengan asumsi independensi yang kuat antara fitur-fitur. Dalam konteks analisis sentimen, fitur-fitur ini biasanya berupa kata-kata atau frasa dalam teks. Keunggulan Naïve Bayes adalah kemudahannya dalam implementasi, efisiensi komputasi yang tinggi, dan performa yang baik dalam berbagai domain. Namun, kelemahannya adalah asumsi independensinya yang jarang terpenuhi dalam data nyata, yang dapat mengurangi akurasi. Support Vector Machine (SVM) adalah metode klasifikasi yang mencari hyperplane optimal untuk memisahkan data dalam ruang fitur. Keunggulan SVM adalah kemampuannya untuk menangani data yang tidak linear melalui kernel trick dan meningkatkan margin antara dua kelas yang berbeda, yang dapat menghasilkan performa yang lebih baik pada data yang kompleks. Namun, SVM memiliki kelemahan dalam pemilihan fitur yang tepat dan kebutuhan komputasi yang lebih tinggi. Dalam konteks penelitian konflik Papua, Naïve Bayes dapat diterapkan untuk menganalisis tweet yang berkaitan dengan konflik dengan cepat dan efisien, meskipun dengan asumsi independensi yang mungkin tidak realistis. SVM dapat memberikan hasil yang lebih akurat dalam memisahkan sentimen positif dan negatif terkait konflik, tetapi membutuhkan lebih banyak waktu dan sumber daya untuk pelatihan dan pemilihan fitur. Kelebihan Naïve Bayes dalam kesederhanaan dan kecepatan membuatnya ideal untuk analisis skala besar, sementara SVM lebih cocok untuk analisis mendalam dengan data yang kompleks, meskipun dengan beban komputasi yang lebih besar.

Dalam penelitian sebelumnya menganalisis pengguna twitter terhadap konflik antara Palestina dan Israel, metode klasifikasi yang digunakan dalam penelitian ini adalah Naïve Bayes dan Support Vector Machine. Setelah klasifikasi dilakukan, evaluasi terhadap kedua model tersebut dilakukan. Hasil penelitian menunjukkan bahwa metode Naïve Bayes menghasilkan akurasi sebesar 74%, precision 74%, dan recall 56%. Sementara itu, metode Support Vector Machine menghasilkan akurasi sebesar 80%, precision 79%, dan recall 72%. Dari hasil ini, dapat disimpulkan bahwa metode Support Vector Machine dapat mengklasifikasikan dataset lebih baik dibandingkan dengan metode Naïve Bayes[9].

Penelitian selanjutnya menganalisis terkait konflik Palestina-Israel Pada Media Sosial X Menggunakan Algoritma Naïve Bayes Classifier, Data yang diambil dari media sosial X berfokus pada komentar dan postingan yang berada di Indonesia, dengan total sebanyak 1.715 data. Penelitian ini menggunakan algoritma Naïve Bayes Classifier dengan pembagian 80% untuk data latih dan 20% untuk data uji, setelah melalui proses pre-processing. Hasil dari penelitian ini menunjukkan akurasi sebesar 94%, precision 91%, recall 100%, serta F1-Score sebesar 95%. Hasil analisis dalam penelitian ini menunjukkan sentimen positif, sehingga dapat disimpulkan bahwa respon masyarakat Indonesia di media sosial X menunjukkan dukungan positif terhadap Palestina[10].

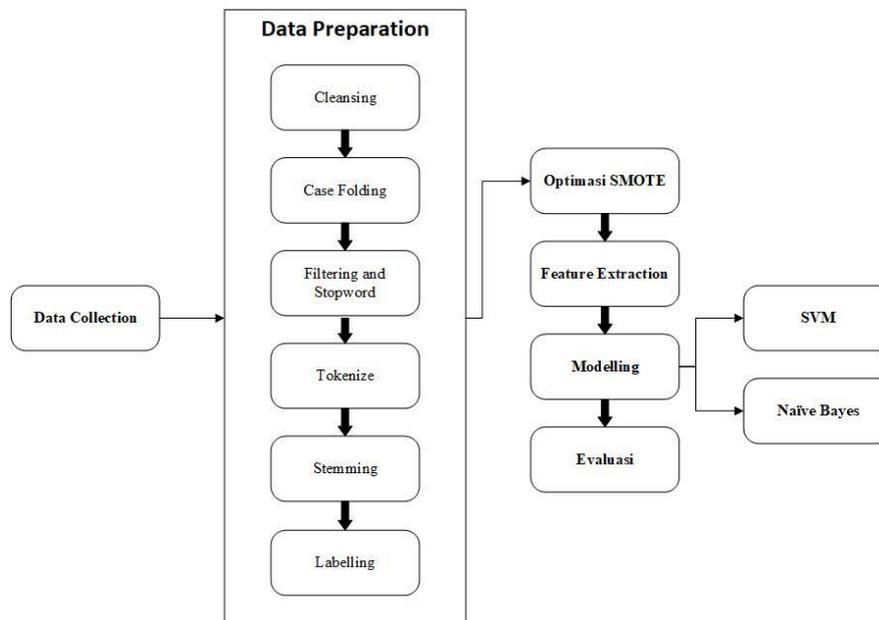
Penelitian lainnya menganalisis opini publik terkait kondisi bahasa daerah Papua Menggunakan Pendekatan Data Science, Dari 778 opini yang dianalisis, 92% menunjukkan sentimen positif. Analisis sentimen opini publik dilakukan menggunakan algoritma Naïve Bayes, yang memiliki akurasi model sebesar 78% dan precision sebesar 88%. Model machine learning yang dibangun, beserta analisis visualisasi word cloud, berhasil memberikan wawasan baru terkait kondisi bahasa Papua[11].

Berdasarkan masalah yang telah diidentifikasi dan berbagai penelitian terkait, penelitian ini bertujuan untuk membandingkan algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam analisis sentimen terhadap konflik di Papua. Penelitian ini berbeda signifikan dari penelitian sebelumnya dengan fokus yang lebih mendalam pada topik konflik Papua. Sentimen publik terhadap konflik Papua memberikan wawasan yang lebih akurat tentang opini masyarakat.

Penelitian ini menunjukkan bahwa metode Support Vector Machine (SVM) lebih unggul dalam analisis sentimen konflik Papua dibandingkan Naïve Bayes, serupa dengan penelitian terkait konflik Palestina-Israel yang juga menunjukkan keunggulan SVM. Namun, penelitian lain terkait konflik Palestina-Israel di media sosial X dengan Naïve Bayes Classifier mencapai akurasi tinggi 94%, kemungkinan karena pre-processing yang lebih baik atau dataset yang lebih spesifik dan bersih. Sebaliknya, analisis opini publik terkait kondisi bahasa daerah Papua dengan Naïve Bayes menunjukkan akurasi lebih rendah 78%, menyoroti pengaruh kualitas data dan proses pre-processing. Hasil penelitian ini dan studi terkait menunjukkan bahwa SVM cenderung memberikan performa lebih baik dalam analisis sentimen karena kemampuannya menangani data yang tidak linear dan kompleks, sementara variasi dalam hasil penelitian mungkin disebabkan oleh kualitas data, pre-processing, dan konteks spesifik dataset yang digunakan. Implikasinya, dalam analisis sentimen, sangat penting untuk mempertimbangkan karakteristik spesifik dataset dan tujuan analisis saat memilih algoritma, di mana SVM sering lebih unggul dalam data kompleks, namun Naïve Bayes bisa memberikan hasil baik dengan pre-processing yang tepat dan sesuai dengan asumsi independensi fitur.

II. METODOLOGI PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan, sehingga memiliki alur kerja yang terstruktur dari awal hingga akhir. Tahapan-tahapan penelitian ini dapat dilihat pada Gambar 1 berikut:



Gambar 1. Tahapan Penelitian

A. Data Collection

Penelitian ini dimulai dengan mengumpulkan data tweet dari Twitter menggunakan Application Programming Interface (API). Dalam proses crawling data di Twitter, API memungkinkan akses langsung dan real-time ke berbagai tweet, profil pengguna, serta informasi lainnya. API memfasilitasi pencarian dan filtrasi data berdasarkan kriteria tertentu, otomatisasi pengambilan data dalam skala besar, dan integrasi dengan alat analitik lainnya. Selain itu, API menyediakan kontrol dan keamanan melalui mekanisme otentikasi dan otorisasi bagi pengguna. Selama proses crawling data di Twitter, pengguna memerlukan token otentikasi dari media sosial Twitter yang akan digunakan bersama library Harvest dalam bahasa pemrograman Python. Untuk mendapatkan data tweet dalam skala besar, kata kunci yang digunakan dalam penelitian ini adalah "Konflik Papua". Peneliti berhasil mengumpulkan kurang lebih 5000 tweet melalui proses crawling pada periode waktu yang berbeda.

B. Data Understanding

Data understanding adalah tahap dalam analisis sentimen di mana peneliti harus terlebih dahulu memahami data yang ada. Melalui pemahaman data ini, peneliti dapat memecahkan masalah sesuai dengan tujuan yang telah ditentukan. Pada tahap ini, peneliti menganalisis data yang telah dikumpulkan. Dengan pemahaman yang mendalam tentang data tersebut, peneliti dapat menentukan langkah-langkah yang diperlukan untuk memecahkan masalah dalam penelitian[12].

C. Preprocessing Data

Tahapan preprocessing merupakan langkah awal dalam proses pengolahan data untuk mendapatkan kata-kata yang relevan untuk diklasifikasikan. Langkah ini mencakup semua rutinitas dan proses untuk mempersiapkan data yang akan digunakan dalam operasi knowledge discovery dalam sistem text mining. Proses yang dilakukan pada tahap ini meliputi case folding, tokenizing, stemming, labeling, dan filtering[13].

D. Cleaning dan Case Folding

Cleaning Data bertujuan untuk menghapus kata yang tidak penting sementara Case folding merupakan proses untuk mengubah kalimat atau menjadi huruf kecil[14]. Proses ini mencakup penghapusan URL, mention, hashtag, stopwords, emotikon, tanda baca, dan angka, serta mengubah semua huruf menjadi huruf kecil. Dalam penelitian ini, pembersihan stopwords dilakukan menggunakan library Python pypSastrawi.

TABEL I.
CLEANING DAN CASE FOLDING

Tweet text	Cleansing	Case Folding
Ancaman di wilayah Papua yang datang dari kelompok kriminal separatis opm yang menunjukkan eksistensinya terus diantisipasi melalui sinergitas aparat gabungan tni-polri. https://t.co/jtq3uuslua	Ancaman di wilayah Papua yang datang dari kelompok kriminal separatis opm yang menunjukkan eksistensinya terus diantisipasi melalui sinergitas aparat gabungan tnipolri	ancaman di wilayah papua yang datang dari kelompok kriminal separatis opm yang menunjukkan eksistensinya terus diantisipasi melalui sinergitas aparat gabungan tnipolri

E. Tokenizing dan Removal Stopword atau filtering

Proses tokenisasi digunakan untuk mengidentifikasi kata-kata dan memecah kalimat menjadi istilah berdasarkan spasi dan tanda baca, sementara *Removal stopwords* dilakukan untuk menghilangkan kata-kata penghubung atau kata berimbuhan seperti 'dan', 'yang', atau 'ke', sehingga hanya menyimpan kata-kata penting yang memiliki makna[15].

TABEL II.
TOKENIZING DAN FILTERING

Tokenizing	Filtering
['of', 'in', 'dob', 'papua', 'improve', 'generasi', 'muda', 'papua', 'menempuh', 'ancaman', 'di', 'wilayah', 'papua', 'yang', 'di']	['dob', 'papua', 'improve', 'public', 'generasi', 'muda', 'papua', 'menempuh', 'ancaman', 'wilayah', 'papua', 'yang']

F. Stemming dan Labeling

Stemming adalah proses seleksi kata melibatkan perubahan kata sambung, kata berimbuhan, kata ganti, dan kata kerja menjadi kata dasar dengan menghapus awalan atau akhiran[16], *Labeling* bertujuan untuk mengategorikan data ke dalam dua kelas, yakni kelas dengan label *positif* dan label *negatif*[12].

TABEL III.
STEMMING DAN LABELING

Stemming	Labeling
ancaman di wilayah papua yang datang dari kelompok kriminal separatis opm yang menunjukkan eksistensinya terus diantisipasi melalui sinergitas aparat gabungan tnipolri	Negatif

G. Optimasi Smote

Menurut Siringoringo, SMOTE adalah metode yang digunakan untuk mengatasi masalah data yang tidak seimbang, di mana satu kelompok data memiliki jumlah yang jauh lebih besar dibandingkan dengan kelompok lainnya. Ketidakseimbangan ini dapat menyebabkan model yang dibangun menjadi terlalu terfokus pada kelompok mayoritas, sehingga performa prediksi untuk kelompok minoritas menjadi buruk. SMOTE mengatasi masalah ini dengan menghasilkan data tambahan untuk kelompok minoritas, sehingga jumlah data di kedua kelompok menjadi seimbang. Dengan cara ini, model yang dikembangkan dapat belajar dengan lebih adil dan menghindari overfitting[17].

H. Feature Extraction

Feature Extraction merupakan teknik yang sering digunakan untuk mereduksi kompleksitas data, di mana data yang besar dan rumit diubah menjadi fitur-fitur yang lebih sederhana namun tetap informatif. Proses ini dapat secara signifikan meningkatkan performa model. Salah satu metode yang digunakan dalam ekstraksi fitur adalah Term Frequency (TF) dan Inverse Document Frequency (IDF), yang berfungsi untuk menentukan frekuensi kemunculan suatu kata dalam dokumen dibandingkan dengan jumlah total kata dalam dokumen tersebut[18]. Dalam proses pembobotan kata menggunakan TF dan IDF memiliki rumus dan mekanisme yang berbeda, Terlihat pada persamaan dibawah ini.

- Term Frequency (TF) : TF mengukur seberapa sering sebuah term muncul dalam sebuah dokumen

$$TF(t,d) = (\text{jumlah kemunculan term } t \text{ dalam dokumen } d) / (\text{Total jumlah term dalam dokumen } d)$$

- Inverse Document Frequency (IDF) : IDF mengukur seberapa penting sebuah term, dengan mengurangi bobot term yang sering muncul dalam banyak dokumen

$$IDF(t, D) = \log (N / (|\{ d \in D : t \in d \}|))$$

Penjelasan : N adalah total jumlah dokumen

$|\{ d \in D : t \in d \}|$ adalah jumlah dokumen yang mengandung term t.

- TF-IDF: TF-IDF menggabungkan kedua ukuran ini untuk memberikan bobot yang lebih akurat bagi sebuah term dalam konteks dokumen dan keseluruhan kumpulan dokumen

$$TF-IDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

I. Modeling

Setelah data diberi pembobotan kata menggunakan TF-IDF, langkah selanjutnya adalah membangun model klasifikasi teks dengan metode Naïve Bayes, dan Support Vector Machine (SVM).

1) Naïve Bayes

Naive Bayes Classifier adalah metode klasifikasi berbasis probabilitas sederhana. Metode ini terdiri dari dua tahap dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan, data sampel diolah agar dapat mewakili keseluruhan data, dengan menentukan probabilitas prior untuk setiap kategori berdasarkan sampel tersebut. Pada tahap klasifikasi, nilai kategori suatu data ditentukan berdasarkan istilah yang muncul dalam data yang diklasifikasikan. Persamaan teorema Bayes dapat dinyatakan sebagai berikut:[19]

$$P(H|X) = \frac{P(X|H)P(H)A}{P(H)}$$

Keterangan:

- X = Data dengan class yang belum diketahui
- H = Hipotesis data menggunakan suatu class yang spesifik
- P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X (Parteori Probabilitas)
- P(H) = Probabilitas hipotesis H (Prior Probabilitas)
- P(X|H) = Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) = Probabilitas H

Dalam Naïve Bayes, parameter utama adalah alpha, yang digunakan untuk smoothing dan menghindari probabilitas nol, dengan nilai yang lebih kecil (misalnya, 0.1) berisiko overfitting. Gaussian Naïve Bayes menggunakan var_smoothing untuk menambah stabilitas dengan menghindari divisi nol. Bernoulli Naïve Bayes memiliki alpha dan binarize untuk membinarisasi fitur. Pengaturan parameter ini mempengaruhi keseimbangan antara overfitting dan underfitting, stabilitas model, dan akurasi.

2) Support Vector Machine

Support Vector Machine (SVM) adalah algoritma machine learning yang menggunakan hyperplane untuk memisahkan data ke dalam wilayah yang mewakili setiap kelas. Hyperplane berfungsi sebagai garis pemisah antar kelas yang digunakan dalam prediksi kelas data, di mana SVM menetapkan label berdasarkan wilayah kelas tempat data berada. Algoritma ini sering digunakan pada dataset besar dari sumber online dan dikenal karena keefektifannya dalam klasifikasi teks. Prinsip dasar SVM adalah membangun hyperplane dengan margin yang merata di antara kelas-kelasnya, tanpa bias terhadap salah satu kelas. Hal ini dicapai dengan mengukur margin dan mencari titik maksimalnya. Fokus utama metode SVM adalah mencari hyperplane terbaik untuk memisahkan kelas[20].

Parameter penting dalam SVM termasuk C, yang mengatur trade-off antara margin maksimal dan kesalahan klasifikasi, serta kernel (linear, poly, rbf, sigmoid) yang menentukan transformasi fitur. Parameter gamma untuk kernel RBF dan sigmoid mengatur kompleksitas model, sementara degree dan coef0 mempengaruhi kernel polinomial dan sigmoid. Pengaturan parameter ini mempengaruhi waktu komputasi, akurasi, kemampuan generalisasi, dan stabilitas model terhadap outlier.

J. Evaluasi

Evaluasi dilakukan untuk menilai sejauh mana model dapat memprediksi sentimen dengan akurat. Kemampuan prediksi ini diukur menggunakan Confusion Matrix, akurasi, precision, recall, dan F1-Score. Confusion Matrix memberikan gambaran yang jelas mengenai kinerja sistem dalam mengklasifikasikan data. Dalam proses evaluasi ini, empat istilah utama digunakan untuk menilai kinerja klasifikasi: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Dengan menganalisis Confusion Matrix, kita dapat menghitung berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1-Score untuk menilai performa model klasifikasi secara keseluruhan.

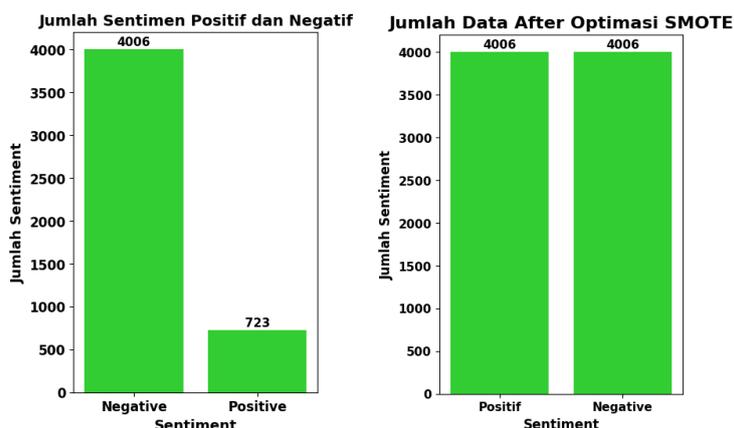
Evaluasi model menggunakan Confusion Matrix menghasilkan metrik seperti akurasi, precision, recall, dan F1-Score untuk menilai kinerja klasifikasi. Confusion Matrix mengidentifikasi True Positives (TP), True Negatives (TN), False Positives (FP), dan False Negatives (FN). Akurasi $(TP+TN)/(TP+TN+FP+FN)$, Precision $TP/(TP+FP)$, Recall $TP/(TP+FN)$, dan F1-Score $2 * TP / (2 * TP + FP + FN)$ memberikan gambaran

umum prediksi benar, tetapi bisa menyesatkan pada dataset tidak seimbang. Precision $(\frac{TP}{TP+FP})$ mengukur akurasi prediksi positif, sedangkan recall $(\frac{TP}{TP+FN})$ mengukur kemampuan deteksi positif. F1-Score $2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$ menggabungkan precision dan recall untuk keseimbangan. Akurasi mudah dihitung dan dipahami tetapi kurang efektif untuk dataset yang tidak seimbang. Precision berguna ketika biaya kesalahan positif tinggi, namun tidak mempertimbangkan contoh negatif yang diklasifikasikan dengan benar. Recall penting untuk meminimalkan false negatives, tetapi bisa tinggi jika banyak false positives. F1-Score berguna untuk mengatasi trade-off antara precision dan recall, tetapi tidak memberikan gambaran lengkap jika dataset sangat tidak seimbang. Memilih metrik tergantung pada konteks dan tujuan model.

III. HASIL DAN PEMBAHASAN

A. Perbandingan Akurasi

Setelah melalui tahapan preprocessing, data yang dihasilkan ada 4006 data dengan sentimen negatif dan 723 data dengan sentimen positif. Karena tidak seimbang antara data sentimen negatif dan positif menyebabkan model algoritma Naïve Bayes dan Support Vector Machine lebih cenderung belajar dari kata-kata negatif, sehingga akurasi dan performa klasifikasi untuk kelas negatif menjadi tinggi. Karena itu, hasil eksperimen tidak hanya berfokus pada nilai akurasi untuk mengevaluasi model algoritma. Untuk menentukan apakah Naïve Bayes atau Support Vector Machine memiliki performa terbaik, diperlukan evaluasi melalui classification report yang mencakup nilai precision, recall, dan F1-Score. Untuk mengatasi ketidakseimbangan data ini, penelitian ini menerapkan optimasi SMOTE (Synthetic Minority Over-sampling Technique). Dengan menggunakan SMOTE, jumlah data minoritas akan disamakan dengan data mayoritas, sehingga model algoritma tidak hanya belajar dari satu jenis sentimen saja.



Gambar 2. Optimasi Smote

Pada Gambar 2 diperlihatkan perbandingan jumlah sentimen positif dan negatif sebelum dan sesudah penerapan optimasi SMOTE. Setelah proses optimasi SMOTE, jumlah data sentimen positif dan negatif menjadi seimbang. Dengan menggunakan SMOTE, model algoritma dapat mempelajari setiap sentimen dengan lebih merata karena jumlah data antara sentimen positif dan negatif telah diseimbangkan. SMOTE berfungsi dengan menambahkan contoh-contoh baru dari kelas minoritas melalui sintesis data, sehingga tidak ada lagi ketidakseimbangan antara jumlah data minoritas dan mayoritas. Setelah proses optimasi, langkah berikutnya adalah melakukan pelatihan (training) dan pengujian (testing) pada model algoritma. Hal ini penting untuk memastikan bahwa model dapat belajar dan menggeneralisasi dengan baik terhadap data baru yang telah dioptimasi dengan SMOTE, sehingga meningkatkan performa dan akurasi dalam klasifikasi sentimen secara keseluruhan.

Penerapan SMOTE (Synthetic Minority Over-sampling Technique) bertujuan untuk menyeimbangkan data dengan menambahkan contoh-contoh baru dari kelas minoritas melalui sintesis data. Dampak SMOTE pada kedua model menunjukkan hasil yang berbeda. Untuk Naïve Bayes, akurasi meningkat, namun tidak sebesar peningkatan pada SVM. Precision pada Naïve Bayes menjadi lebih seimbang untuk kelas positif dan negatif setelah SMOTE, namun peningkatannya tidak setinggi SVM karena asumsi independensi antar fitur. Recall untuk kelas positif pada Naïve Bayes meningkat signifikan, sementara recall untuk kelas negatif sedikit menurun. F1-Score pada Naïve Bayes menunjukkan peningkatan, namun tidak sebesar SVM. SVM, di sisi lain, menunjukkan peningkatan akurasi

yang sangat tinggi setelah SMOTE, dengan precision dan recall yang juga sangat tinggi untuk kedua kelas. SVM mampu memanfaatkan data tambahan dari kelas minoritas untuk meningkatkan generalisasi dan mengurangi jumlah false positives secara signifikan, berkat kemampuan untuk memisahkan data dengan margin maksimal. Secara keseluruhan, SMOTE memberikan dampak yang lebih besar pada peningkatan performa SVM dibandingkan Naïve Bayes.

B. Tahap Pengujian

Dalam penelitian ini, untuk menentukan performa terbaik dari masing-masing model algoritma, peneliti menggunakan 70% data untuk pelatihan dan 30% untuk pengujian pada model Naïve Bayes dan Support Vector Machine. Data yang digunakan untuk menerapkan model algoritma ini telah melalui proses optimasi menggunakan teknik SMOTE, sehingga data tersebut lebih seimbang dan dapat meningkatkan akurasi model

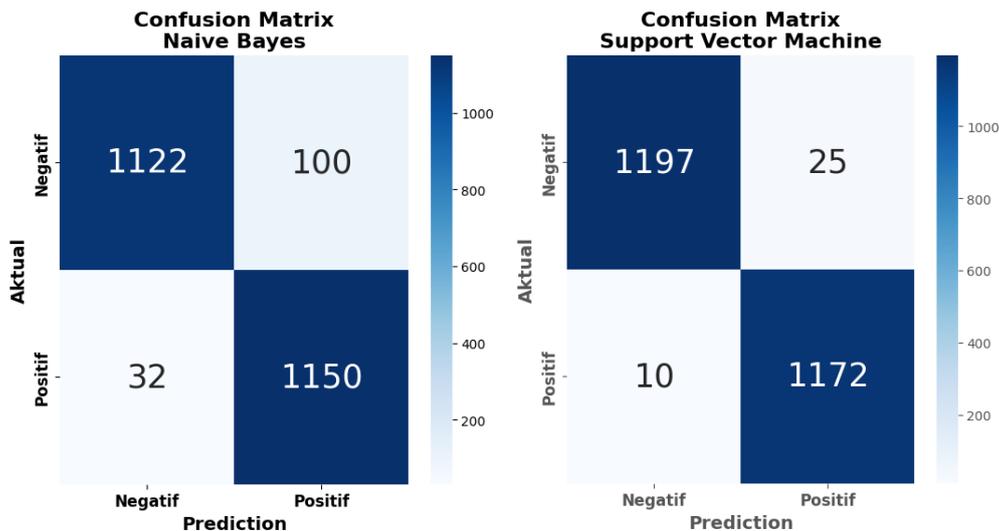
TABEL IV.
 MODELLING NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

Modeling		Accuracy	Precision	recall	F1-Score
Naïve Bayes	Negatif	95%	97%	94%	94%
	Positif		92%	97%	95%
Support Vector Machine	Negatif	99%	99%	98%	99%
	Positif		98%	99%	99%

Pada tabel 4, hasil eksperimen menggunakan 70% data untuk pelatihan dan 30% data untuk pengujian setelah optimasi SMOTE menunjukkan bahwa model Support Vector Machine (SVM) memiliki performa dengan akurasi, presisi, recall, dan F1-Score semuanya mencapai 99% dan 98% untuk kedua kelas (negatif dan positif). Di sisi lain, model Naïve Bayes juga menunjukkan performa yang sangat baik dengan akurasi 95% dan keseimbangan yang sangat baik antara presisi dan recall untuk kedua kelas. SVM menunjukkan keunggulan dalam konsistensi dan keseimbangan performa, menjadikannya pilihan yang lebih baik untuk mengklasifikasikan sentimen negatif dan positif dalam penelitian ini. Namun, sebelum benar-benar menentukan model mana yang terbaik, kita harus mempertimbangkan aspek lain seperti presisi, recall, dan F1-score. Pada algoritma Naïve Bayes, akurasi mencapai 95% sedangkan untuk Support Vector Machine (SVM) mencapai 99%. Untuk sentimen negatif, presisi Naïve Bayes berada di angka 97% dan SVM di 99%, recall Naïve Bayes berada di 94% dan SVM di 98%, serta F1-Score dari Naïve Bayes 94% dan SVM 99%. Untuk sentimen positif, presisi Naïve Bayes adalah 92% dan SVM 98%, recall Naïve Bayes adalah 97% dan SVM 99%, serta F1-Score dari Naïve Bayes 95% dan SVM 99%.

Support Vector Machine (SVM) menunjukkan performa yang lebih baik dibandingkan Naïve Bayes dalam klasifikasi sentimen setelah penerapan SMOTE. Alasan utama SVM lebih unggul adalah kemampuannya dalam memodelkan garis batas yang lebih jelas melalui margin maksimal, yang memaksimalkan jarak antara kelas-kelas yang berbeda dalam dataset. SVM juga menggunakan kernel trick untuk memetakan data ke dimensi yang lebih tinggi, memungkinkan pemisahan non-linear yang sangat penting dalam analisis teks. Selain itu, SVM memiliki mekanisme regularisasi yang kuat untuk menghindari overfitting, yang sangat berguna dalam dataset dengan dimensi tinggi seperti data teks. Di sisi lain, Naïve Bayes mengasumsikan independensi antar fitur yang seringkali tidak realistis dalam data teks, mengurangi akurasinya. Meskipun SMOTE telah menyeimbangkan data, SVM tetap lebih efektif dalam menangani data yang sebelumnya tidak seimbang, karena berfokus pada margin maksimal yang membantu mengatasi bias dari ketidakseimbangan data awal. Faktor-faktor ini, termasuk distribusi data dan dimensi fitur, memberikan keunggulan signifikan bagi SVM dalam menangani kompleksitas dan ketidakseimbangan dalam dataset teks dibandingkan dengan Naïve Bayes.

Untuk menentukan model algoritma terbaik, peneliti melakukan eksperimen perbandingan pada kedua algoritma yang digunakan dengan menggunakan confusion matrix. Hasil perbandingan kedua model algoritma dapat dilihat secara visual pada gambar 3 di bawah ini.



Gambar 3. Confusion Matrix

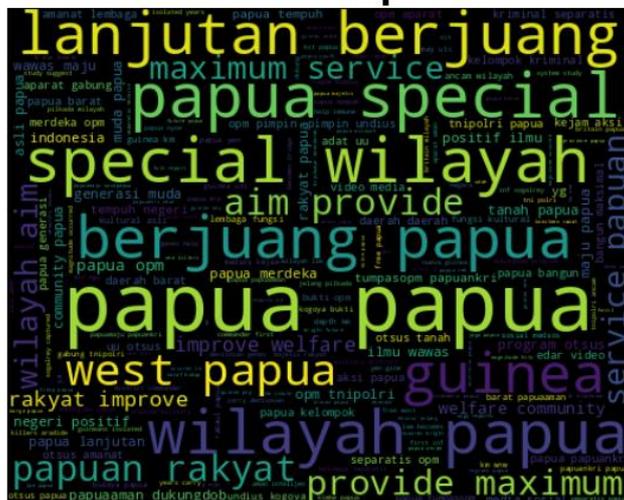
Setelah proses optimasi, setiap algoritma dievaluasi menggunakan confusion matrix untuk menilai kinerja model klasifikasi dengan membandingkan prediksi model terhadap nilai sebenarnya. Berdasarkan gambar di atas, algoritma Support Vector Machine mengklasifikasikan 1197 data sebagai True Negatif dan 1172 sebagai True Positif. Sementara itu, algoritma Naïve Bayes mengklasifikasikan 1122 data sebagai True Negatif dan 1150 sebagai True Positif. Jika dibandingkan, algoritma SVM menunjukkan akurasi yang lebih tinggi dibandingkan dengan Naïve Bayes, dengan akurasi mencapai 99%.

Selanjutnya, kedua algoritma dibandingkan berdasarkan akurasi (precision) dan kelengkapan (recall) dalam klasifikasi sentimen, baik untuk kelas negatif maupun positif. Untuk mengevaluasi algoritma terbaik, F1-score digunakan sebagai indikator kinerja. F1-score dihitung berdasarkan precision dan recall. Algoritma SVM memiliki nilai F1-score sebesar 99%, menjadikannya pilihan terbaik dalam analisis ini. Performanya yang optimal menunjukkan kemampuan algoritma SVM dalam menangani data dengan presisi dan recall yang tinggi, sehingga dapat memberikan hasil yang lebih akurat dibandingkan dengan algoritma Naïve Bayes.

C. Visualisasi Data

Dalam penelitian ini, untuk memudahkan identifikasi kata-kata yang paling sering muncul dalam dataset dan memahami topik utama atau sentimen dominan, peneliti menggunakan visualisasi wordcloud. Setiap kata ditampilkan dengan ukuran font yang bervariasi, di mana kata yang lebih besar menunjukkan frekuensi kemunculannya yang lebih tinggi dibandingkan dengan kata yang lebih kecil.

WordCloud Analisis Sentiment Konflik Papua



Gambar 4. Wordcloud Konflik Papua

Pada hasil wordcloud di atas menunjukkan bahwa kata Papua adalah salah satu kata yg paling banyak disebut

diskusi sering kali berkisar pada aspek geografis dan administratif dari konflik, menunjukkan fokus pada daerah tertentu yang menjadi pusat perhatian. Kata “opm” mengacu pada Organisasi Papua Merdeka, yang merupakan salah satu aktor utama dalam konflik, sering disebut dalam konteks aktivitas mereka, perjuangan kemerdekaan, atau konflik dengan aparat keamanan Indonesia. Kata “special” sering muncul dalam konteks pembicaraan mengenai status atau perhatian khusus yang diberikan kepada Papua. Hal ini bisa merujuk pada Otonomi Khusus Papua atau dukungan khusus yang diberikan untuk menangani isu-isu di wilayah tersebut. Kata “lanjutan” menunjukkan keberlanjutan dari upaya atau konflik yang sedang berlangsung di Papua. Ini bisa merujuk pada tindakan atau kebijakan yang terus berlanjut dari pihak pemerintah atau aktivitas yang berkelanjutan dari kelompok separatistis. Kata “berjuang” mengindikasikan adanya perjuangan yang diakui dalam narasi publik, baik itu perjuangan dari kelompok separatistis untuk kemerdekaan, maupun perjuangan masyarakat Papua dalam menghadapi berbagai tantangan di wilayah mereka. Kata “rakyat” menekankan aspek humanis dan sosial dalam diskusi tentang Papua. Ini menunjukkan perhatian terhadap kondisi dan kesejahteraan masyarakat Papua, yang sering menjadi pusat perhatian dalam berbagai diskusi. Kata “km” bisa jadi merujuk pada satuan kilometer, yang mungkin digunakan dalam konteks deskripsi geografis atau laporan kejadian yang terjadi di wilayah Papua. Ini menunjukkan bagaimana detail geografis atau lokasi spesifik sering dibahas dalam narasi mengenai Papua.

Secara keseluruhan, analisis frekuensi kata ini memberikan wawasan tentang bagaimana masyarakat membicarakan Konflik Papua di media sosial. Ini menunjukkan bahwa selain aspek konflik dan keamanan, terdapat juga perhatian khusus pada status khusus Papua, keberlanjutan upaya dan perjuangan, serta kesejahteraan rakyat. Pemahaman ini bisa menjadi dasar penting untuk merancang strategi komunikasi dan kebijakan yang lebih responsif terhadap isu-isu yang dibahas oleh masyarakat.

Penelitian ini membandingkan hasil analisis sentimen terhadap konflik Papua dengan studi sebelumnya yang menganalisis pengguna Twitter terhadap konflik Palestina-Israel menggunakan metode Naïve Bayes dan Support Vector Machine (SVM). Kedua penelitian menggunakan algoritma yang sama dan menunjukkan bahwa SVM menghasilkan performa yang lebih baik dibandingkan Naïve Bayes, dengan penelitian sebelumnya mencapai akurasi 80%, precision 79%, dan recall 72% untuk SVM, sementara Naïve Bayes mencapai akurasi 74%, precision 74%, dan recall 56%. Penelitian ini menunjukkan bahwa setelah penerapan SMOTE, SVM mencapai akurasi hampir 99%, menunjukkan peningkatan signifikan dibandingkan hasil sebelumnya, yang disebabkan oleh penerapan SMOTE yang menyeimbangkan data dan meningkatkan performa model. Hasil penelitian ini menegaskan pentingnya teknik penyeimbangan data seperti SMOTE dalam analisis sentimen konflik, memberikan kontribusi dengan menekankan efektivitas SMOTE dalam meningkatkan performa model klasifikasi sentimen, serta validasi ulang bahwa SVM tetap unggul dibandingkan Naïve Bayes dalam skenario ini. Penelitian ini memperkuat temuan sebelumnya bahwa SVM lebih baik daripada Naïve Bayes, dan menunjukkan bagaimana teknik penyeimbangan data dapat mengubah hasil secara signifikan, memberikan wawasan tambahan bagi peneliti yang ingin menerapkan analisis sentimen pada dataset yang tidak seimbang, terutama dalam konteks konflik atau isu serupa.

IV. KESIMPULAN

Berdasarkan hasil penelitian ini, telah dilakukan analisis sentimen terhadap opini masyarakat Indonesia mengenai konflik Papua di media sosial Twitter. Penelitian ini menggunakan 5723 data yang berfokus pada topik Luhut Binsar Pandjaitan. Dalam penelitian ini, dibandingkan performa algoritma Naïve Bayes dan Support Vector Machine (SVM). Sebelum melakukan perbandingan, dilakukan optimasi SMOTE untuk menyeimbangkan jumlah data minoritas dan mayoritas. Hasil komparasi menunjukkan bahwa algoritma Naïve Bayes mencapai akurasi 95%, sementara SVM mencapai akurasi 99%. Kesimpulan dari penelitian ini menunjukkan bahwa SVM memiliki performa terbaik dengan akurasi 99%, presisi 99%, recall 98%, dan F1-Score 99%.

Dari hasil tersebut, dapat disimpulkan bahwa algoritma SVM lebih unggul dibandingkan dengan Naïve Bayes dalam analisis sentimen terhadap opini masyarakat Indonesia mengenai konflik Papua di Twitter. Meskipun kedua algoritma memiliki akurasi yang cukup tinggi, SVM menunjukkan performa yang lebih baik dengan akurasi 99%, presisi 99%, recall 98%, dan F1-Score 99%. Ini menunjukkan bahwa SVM mampu mengklasifikasikan sentimen positif dan negatif dengan lebih akurat dibandingkan Naïve Bayes. Oleh karena itu, dalam analisis sentimen terhadap konflik Papua, algoritma SVM dapat menjadi pilihan yang lebih baik dibandingkan Naïve Bayes..

DAFTAR PUSTAKA

- [1] D. A. Kaisupy and S. G. Maing, “Proses negosiasi konflik papua: dialog Jakarta-Papua,” *Jurnal Ilmu Sosial Dan Humaniora*, vol. 10, no. 1, pp. 82–98, 2021, Accessed: Jul. 02, 2024. [Online]. Available: <https://ejournal.undiksha.ac.id/index.php/JISH/article/view/27056/17936>

- [2] A. G. Syahviar, S. J. I. Ismail, and M. R. Alfari, "Analisis Parsing Data Sosial Media Terhadap Konflik Kelompok Kriminal Bersenjata Di Papua Dengan Pendekatan Lexicon Based," *eProceedings of Applied Science*, vol. 10, no. 4, 2023, Accessed: Jul. 12, 2024. [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/appliedscience/article/view/20868>
- [3] M. A. Ramdhan, "Analisis Dimensi Internasional Konflik Papua dalam Model Counterinsurgency (COIN)," *Jurnal Ilmiah Hubungan Internasional*, vol. 17, no. 1, pp. 139–152, 2021, Accessed: Jul. 02, 2024. [Online]. Available: <https://journal.unpar.ac.id/index.php/JurnalIlmiahHubunganInternasional/article/view/3532/3324>
- [4] W. A. S. Tangguh Chairil, "Konflik Papua: Pemerintah Perlu Mengubah Pendekatan Keamanan dengan Pendekatan Humanis," <https://ir.binus.ac.id/>. Accessed: Jul. 03, 2024. [Online]. Available: <https://ir.binus.ac.id/2020/10/05/konflik-papua-pemerintah-perlu-mengubah-pendekatan-keamanan-dengan-pendekatan-humanis/>
- [5] BBC News Indonesia, "Konflik Papua: Serangan OPM di Pegunungan Bintang, warga pendatang 'pasrah saja, tunggu dievakuasi,'" *Bbc.com*. Accessed: Jul. 02, 2024. [Online]. Available: <https://www.bbc.com/indonesia/indonesia-64226326>
- [6] J. Hasibuan, "Distingsi Pengelolaan Konflik Papua Yang Tidak Pernah Selesai," *Asketik: Jurnal Agama dan Perubahan Sosial*, vol. 6, no. 1, pp. 69–84, 2022, Accessed: Jul. 02, 2024. [Online]. Available: <https://jurnalfuda.iainkediri.ac.id/index.php/asketik/article/view/189/103>
- [7] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan metode naïve bayes dan support vector machine pada analisis sentimen twitter," *SMATIKA Jurnal: STIKI Informatika Jurnal*, vol. 10, no. 02, pp. 71–76, 2020, Accessed: Jul. 02, 2024. [Online]. Available: <https://file.stiki.ac.id/SMATIKA/article/view/455>
- [8] M. I. Petiwi, A. Triyudi, and I. D. Sholihati, "Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, pp. 542–550, 2022, Accessed: Jul. 03, 2024. [Online]. Available: <https://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/3530/2462>
- [9] R. B. Sinaga, H. R. Al Fajri, H. Mubarak, A. D. Pangestu, and D. S. Prasvita, "Analisis sentimen pengguna Twitter terhadap konflik antara Palestina dan Israel menggunakan metode Naïve Bayesian classification dan Support Vector Machine," in *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, 2021, pp. 166–175. Accessed: Jul. 03, 2024. [Online]. Available: <https://conference.upnvj.ac.id/index.php/senamika/article/view/1632>
- [10] S. D. Simamora, F. Irwiensyah, and F. N. Hasan, "Analisis Sentimen Terkait Konflik Palestina-Israel Pada Media Sosial X Menggunakan Algoritma Naïve Bayes Classifier," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 1, pp. 146–157, 2024, Accessed: Jul. 03, 2024. [Online]. Available: <https://ejurnal.seminar-id.com/index.php/bits/article/view/5344>
- [11] N. F. Hasan, A. Aisyah, R. Rahman, and H. Wonda, "Sentiment Analysis of Public Opinion Regarding Papuan Local Languages Condition Using Data Science Approach," *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 2, pp. 125–139, 2022, Accessed: Jul. 03, 2024. [Online]. Available: <https://pustaka-psm.unilak.ac.id/index.php/dz/article/view/11545>
- [12] B. Ramadhani and R. R. Suryono, "Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 2, pp. 714–725, 2024, Accessed: Jul. 12, 2024. [Online]. Available: <https://www.stmik-budidarma.ac.id/ejurnal/index.php/mib/article/view/7458/3831>
- [13] D. Gunawan, D. Riana, D. Ardiansyah, F. Akbar, and S. Alfari, "Komparasi Algoritma Support Vector Machine Dan Naïve Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Calon Gubernur Jabar 2018-2023," 2020. Accessed: Jul. 12, 2024. [Online]. Available: <https://repository.bsi.ac.id/index.php/unduh/item/234011/Jurnal-Komparasi-Algoritma-Support-Vector-Machine-.pdf>
- [14] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naïve Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *Jurnal Tekno Kompak*, vol. 15, no. 1, pp. 131–145, 2021, Accessed: Jul. 12, 2024. [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/view/744>
- [15] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi, and R. Watrionthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 5, no. 1, pp. 157–163, 2021, Accessed: Jul. 12, 2024. [Online]. Available: <https://stmik-budidarma.ac.id/ejurnal/index.php/mib/article/view/2580>
- [16] N. S. Marga, "Sentimen Analisis Tentang Kebijakan Pemerintah Terhadap Kasus Corona Menggunakan Metode Naïve Bayes," *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, vol. 2, no. 4, pp. 453–463, 2021, Accessed: Jul. 12, 2024. [Online]. Available: <https://jim.teknokrat.ac.id/index.php/informatika/article/view/1602/500>
- [17] H. Hidayatullah, P. Purwanto, and Y. Umaidah, "Penerapan Naïve Bayes Dengan Optimasi Information Gain Dan Smote Untuk Analisis Sentimen Pengguna Aplikasi Chatgpt," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, pp. 1546–1553, 2023, Accessed: Jul. 13, 2024. [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/article/view/6887/4111>
- [18] R. Yunita and M. Kamayani, "Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Penghapusan Kewajiban Skripsi," *Indonesian Journal of Computer Science*, vol. 12, no. 5, 2023, Accessed: Jul. 13, 2024. [Online]. Available: <http://3.8.6.95/ijcs/index.php/ijcs/article/view/3415>
- [19] R. Noviana and I. Rasal, "Penerapan Algoritma Naïve Bayes Dan Svm Untuk Analisis Sentimen Boy Band Bts Pada Media Sosial Twitter," *Jurnal Teknik dan Science*, vol. 2, no. 2, pp. 51–60, 2023, Accessed: Jul. 13, 2024. [Online]. Available: <https://journal.admi.or.id/index.php/JTS/article/view/791/959>
- [20] R. S. Arischo and D. Damayanti, "Analisis Sentimen Pinjaman Online di Twitter dengan Metode Naïve Bayes Classifier dan SVM," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 2, pp. 1120–1130, 2024, Accessed: Jul. 13, 2024. [Online]. Available: <https://stmik-budidarma.ac.id/ejurnal/index.php/mib/article/view/7406/3874>