

Perbandingan Kinerja Support Vector Machine dan Naive Bayes dalam Klasifikasi Sentimen Twitter Terhadap Pelayanan BPJS

Feby Amandasari¹, Damayanti^{*2}

^{1,2}Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia
Email: feby_amandasari@teknokrat.ac.id, damayanti@teknokrat.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis dan membandingkan performa algoritma Support Vector Machine (SVM) dan Naive Bayes dalam mengklasifikasikan sentimen publik terhadap pelayanan BPJS menggunakan data dari Twitter. Data yang digunakan dalam penelitian ini dikumpulkan melalui API Twitter dengan kata kunci "BPJS", menghasilkan total 3.503 tweet. Proses preprocessing dilakukan untuk membersihkan dan mempersiapkan data, yang mencakup tahapan seperti cleaning, case folding, normalisasi kata, tokenization, stopword removal, dan stemming. Setelah preprocessing, data diberi label menggunakan metode lexicon-based, yang menghasilkan 1.572 tweet positif (44,88%), 1.247 tweet negatif (35,60%), dan 684 tweet netral (19,53%). Dataset kemudian dibagi dengan rasio 80:20 untuk pelatihan dan pengujian. Hasil menunjukkan bahwa SVM memiliki akurasi 79,32%, lebih tinggi dibandingkan Naive Bayes yang hanya mencapai 61,63%. Selain itu, SVM juga unggul dalam precision dan recall untuk semua kategori sentimen, menegaskan keandalannya dalam analisis sentimen berbasis media sosial. Hasil penelitian ini dapat membantu pengambil kebijakan dan penyedia layanan kesehatan dalam memahami opini masyarakat serta meningkatkan pelayanan berbasis data.

Kata kunci: analisis sentimen, BPJS, naive bayes, SVM, twitter.

Comparison of SVM and Naive Bayes Algorithm Performance in Classifying Twitter Sentiments on BPJS Services

Abstract

This study aims to analyze and compare the performance of the Support Vector Machine (SVM) and Naive Bayes algorithms in classifying public sentiment toward BPJS services using data from Twitter. The data used in this study was collected through the Twitter API using the keyword "BPJS," resulting in a total of 3,503 tweets. The preprocessing steps were carried out to clean and prepare the data, including cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. After preprocessing, the data was labeled using a lexicon-based method, resulting in 1,572 positive tweets (44.88%), 1,247 negative tweets (35.60%), and 684 neutral tweets (19.53%). The dataset was then split with an 80:20 ratio for training and testing. The results show that SVM achieved an accuracy of 79.32%, higher than Naive Bayes, which only reached 61.63%. Moreover, SVM also outperformed Naive Bayes in terms of precision and recall for all sentiment categories, reinforcing its reliability in social media-based sentiment analysis. The findings of this study can assist policymakers and healthcare service providers in understanding public opinion and improving data-driven services.

Keywords: BPJS, naive bayes, SVM, sentiment analysis, twitter.

1. PENDAHULUAN

Dalam era digital yang berkembang pesat saat ini, media sosial menjadi platform utama bagi masyarakat untuk menyampaikan opini, pengalaman, serta keluhan mereka, termasuk terkait pelayanan publik seperti BPJS Kesehatan. Pada 1 Agustus 2024, jumlah peserta BPJS Kesehatan di Indonesia mencapai 276.520.647 jiwa, atau sekitar 98,19% dari total penduduk Indonesia, yang mencerminkan komitmen pemerintah dalam menyediakan layanan kesehatan bagi hampir seluruh masyarakat Indonesia [1][2]. Namun, seiring dengan jumlah peserta yang sangat besar, tantangan terkait antrean panjang dan waktu tunggu yang lama di fasilitas kesehatan semakin meningkat[3]. Isu ini kerap menjadi sorotan di media sosial, yang memengaruhi persepsi publik terhadap kualitas layanan yang diberikan. Untuk mendapatkan pemahaman lebih mendalam tentang pandangan masyarakat terhadap BPJS Kesehatan, analisis sentimen di media sosial menggunakan algoritma seperti Support Vector Machine (SVM) atau Naive Bayes telah menjadi pendekatan yang efektif [4]. Analisis ini bertujuan untuk

mengklasifikasikan opini masyarakat mengenai layanan BPJS, serta mengidentifikasi area yang perlu mendapat perhatian untuk perbaikan, yang memungkinkan pengembangan kebijakan lebih responsif dalam meningkatkan kualitas pelayanan bagi peserta BPJS [1].

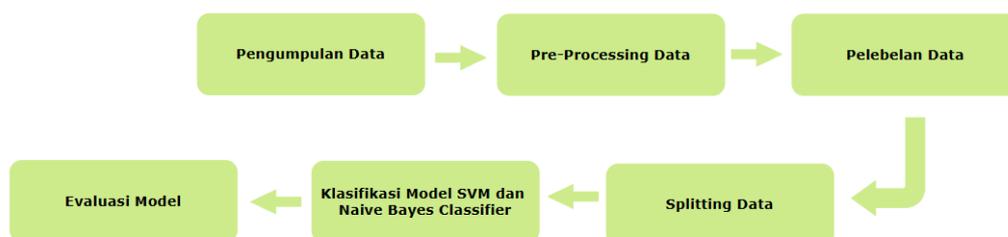
Analisis sentimen merupakan metode yang digunakan untuk mengekstrak dan mengklasifikasikan opini dari teks, yang dapat memberikan wawasan berharga tentang kepuasan dan ketidakpuasan masyarakat terhadap layanan yang diberikan. Dalam konteks BPJS, analisis ini dapat membantu pihak pengelola untuk memahami masalah yang dihadapi oleh peserta dan merespons dengan lebih efektif. Berbagai algoritma telah digunakan dalam analisis sentimen, di antaranya SVM dan Naive Bayes, yang masing-masing memiliki keunggulan dan kelemahan tersendiri. SVM dikenal efektif dalam menangani data dengan dimensi tinggi dan dapat memberikan hasil yang akurat, sementara Naive Bayes lebih sederhana dan cepat dalam proses klasifikasi. Penelitian oleh [5] menunjukkan bahwa SVM dapat mencapai akurasi yang lebih tinggi dalam klasifikasi teks dibandingkan dengan Naive Bayes, terutama pada dataset yang besar dan kompleks.

Beberapa penelitian sebelumnya telah menunjukkan bahwa SVM sering kali menghasilkan akurasi yang lebih tinggi dibandingkan dengan Naive Bayes dalam berbagai konteks analisis sentimen. Misalnya, penelitian oleh [6] menunjukkan bahwa SVM mampu mencapai akurasi hingga 92% dalam mengklasifikasikan sentimen terhadap produk tertentu, sementara Naive Bayes hanya mencapai 85%. Selain itu, penelitian oleh [7] menemukan bahwa SVM lebih unggul dalam mengklasifikasikan sentimen negatif, yang sering kali lebih sulit untuk diidentifikasi. Namun, Naive Bayes tetap menjadi pilihan yang populer karena kemudahan implementasi dan kecepatan dalam memproses data. Penelitian oleh [8] juga menunjukkan bahwa meskipun Naive Bayes memiliki akurasi yang sedikit lebih rendah, algoritma ini lebih efisien dalam hal waktu pemrosesan, yang menjadi pertimbangan penting dalam aplikasi real-time.

Dengan melakukan analisis sentimen terhadap tweet yang berkaitan dengan BPJS, penelitian ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai persepsi masyarakat terhadap layanan yang diberikan. Hasil dari penelitian ini tidak hanya akan memberikan kontribusi akademis, tetapi juga dapat menjadi masukan berharga bagi pengelola BPJS dalam meningkatkan kualitas layanan mereka. Melalui perbandingan kinerja SVM dan Naive Bayes, diharapkan dapat ditemukan algoritma yang lebih efektif dalam mengklasifikasikan sentimen masyarakat, sehingga dapat digunakan sebagai alat bantu dalam pengambilan keputusan yang lebih baik. Penelitian oleh [9] menunjukkan bahwa analisis sentimen dapat membantu organisasi dalam memahami kebutuhan dan harapan masyarakat, yang pada gilirannya dapat meningkatkan kepuasan pelanggan. Selain itu, penelitian oleh [10] menekankan pentingnya penggunaan teknik analisis sentimen dalam konteks pelayanan publik untuk meningkatkan transparansi dan akuntabilitas.

2. METODE PENELITIAN

Penelitian ini dimulai dengan pengumpulan data melalui proses crawling pada platform X (Twitter) untuk memperoleh opini masyarakat terkait layanan BPJS. Data yang diperoleh kemudian diproses melalui tahap preprocessing yang mencakup cleaning (menghapus karakter yang tidak relevan), *case folding* (mengubah teks menjadi huruf kecil), *normalisasi* (mengonversi kata tidak baku ke bentuk standar), *tokenization* (memecah teks menjadi kata-kata individu), *stopword removal* (menghapus kata umum yang kurang bermakna), dan *stemming* (mengubah kata ke bentuk dasar). Setelah preprocessing, data diberi label sentimen ke dalam tiga kategori: positif, negatif, dan netral. Selanjutnya, dataset dibagi menjadi dua bagian, yaitu data latih (80%) untuk melatih model dan data uji (20%) untuk mengukur kinerja model. Dua algoritma klasifikasi yang digunakan dalam penelitian ini adalah *Support Vector Machine (SVM)* dan *Naive Bayes*, yang bertujuan untuk memprediksi sentimen dari opini masyarakat. Evaluasi model dilakukan menggunakan Confusion Matrix dengan menghitung metrik akurasi, presisi, recall, dan F1-score guna menilai kinerja masing-masing model dalam mengklasifikasikan sentimen secara akurat



Gambar 1. Diagram Alir Penelitian

2.1. Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan autentikasi Twitter untuk mengakses tweet yang relevan dengan topik pelayanan BPJS dari media sosial twitter. Proses ini melibatkan pencarian kata kunci "BPJS", "pelayanan BPJS", dan "kesehatan". Data yang diambil mencakup tweet yang dipublikasikan dalam rentang waktu 2022-01-01 sampai 2024-12-30 untuk memastikan relevansi dan aktualitas informasi. Crawling data dari media sosial twitter berhasil di scraper sebanyak 3.656 twets. Penelitian oleh [11] menunjukkan bahwa penggunaan autentikasi Twitter efektif dalam mengumpulkan data untuk analisis sentimen, memungkinkan peneliti untuk mendapatkan data yang besar dan beragam.

2.2. Preprocessing Data

Preprocessing data adalah langkah penting dalam analisis teks untuk membersihkan dan mempersiapkan data sebelum analisis lebih lanjut [12]. Proses ini mencakup beberapa tahapan, seperti *cleaning* untuk menghapus karakter khusus, tautan, dan simbol yang tidak diperlukan. *Case folding* untuk mengubah semua teks menjadi huruf kecil guna menghindari duplikasi kata. *Normalisasi kata* untuk mengubah variasi kata menjadi bentuk baku, misalnya "BPJS" dan "bpjs" menjadi "bpjs". Selain itu, dilakukan *tokenization* untuk memecah teks menjadi unit-unit kata atau token, *stopword removal* untuk menghapus kata-kata umum yang tidak memberikan makna signifikan seperti "dan", "atau", dan "yang", serta *stemming* untuk mengubah kata ke bentuk dasarnya guna mengurangi variasi kata. Penelitian menunjukkan bahwa preprocessing yang baik dapat meningkatkan kualitas data dan hasil analisis, terutama dalam analisis sentimen, seperti yang diuraikan dalam Tabel 1.

Tabel 1. Hasil Preprocessing

Hasil Preprocessing	
Full_Text	BPJS Kesehatan Jangan Memberi Stigma Orang Miskin Tidak Boleh Sakit Bedakan Peserta yang Enggan dan Tidak Mampu Bayar @PKSejahtera @FPKSDPRRI @NettyPrasetya https://t.co/HNG7JCwCY1
Cleaning	BPJS Kesehatan Jangan Memberi Stigma Orang Miskin Tidak Boleh Sakit Bedakan Peserta yang Enggan dan Tidak Mampu Bayar
Case Folding	bpjs kesehatan jangan memberi stigma orang miskin tidak boleh sakit bedakan peserta yang enggan dan tidak mampu bayar
Normalisasi Kata	bpjs kesehatan jangan memberi stigma orang miskin tidak boleh sakit bedakan peserta yang enggan dan tidak mampu bayar
Tokenization	['bpjs', 'kesehatan', 'jangan', 'memberi', 'stigma', 'orang', 'miskin', 'tidak', 'boleh', 'sakit', 'bedakan', 'peserta', 'yang', 'enggan', 'dan', 'tidak', 'mampu', 'bayar']
Stopword Removal	['bpjs', 'kesehatan', 'stigma', 'orang', 'miskin', 'sakit', 'bedakan', 'peserta', 'enggan', 'bayar']
Steaming	bpjs sehat stigma orang miskin sakit beda serta enggan bayar

2.3. Pelabelan Data

Pelabelan data dilakukan dengan menggunakan metode lexicon-based, di mana setiap kata dalam dataset diberi label berdasarkan nilai kata yang dimiliki oleh kamus laxicon based. Metode Laxicon akan diberi label jika nilai score ≥ 0 sentiment label Positif, Jika nilai score ≤ 0 nilai sentiment label Negatif dan jika nilai score $= 0$ maka nilai label sentiment Netral. Penelitian oleh [13] menunjukkan bahwa pendekatan lexicon-based efektif dalam mengklasifikasikan sentimen, terutama dalam konteks media sosial.

Tabel 2. Hasil Labelling

Hasil Labelling		
bpjs sehat stigma orang miskin sakit beda serta enggan bayar	-3	Negatif
salam sila bayar sesuai tanggal tentu status aktif ya sahabat terima kasih starla	1	Positif
salam sahabat kartu aktif layan sehat sila periksa faskes daftar butuh tindak khusus tindak rs rujuk sesuai indikasi medis terima kasih rilla	0	Netral

2.4. Splitting Data

Dataset yang telah melalui tahap pebelan data kemudian dibagi menjadi dua bagian, yaitu data latih dan data uji. Pembagian dilakukan dengan rasio 80:20, di mana 80% data digunakan untuk melatih model, sementara 20% sisanya digunakan untuk menguji kinerjanya. Dalam penelitian ini, data latih terdiri dari 2.801 sampel, sedangkan data uji berjumlah 701 sampel. Pembagian ini bertujuan untuk memastikan bahwa model mampu melakukan generalisasi dengan baik terhadap data baru yang belum pernah diolah sebelumnya. Menurut penelitian oleh [14], pembagian dataset yang tepat memiliki pengaruh signifikan terhadap akurasi model, khususnya dalam analisis sentimen.

2.5. Klasifikasi

Penelitian ini membandingkan algoritma Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen tweet yang telah diberi label. Naive Bayes adalah algoritma probabilistik yang memanfaatkan teorema Bayes untuk menghitung kemungkinan suatu kelas berdasarkan fitur yang ada. Dengan mengasumsikan bahwa setiap fitur bersifat independen, algoritma ini menjadi cepat dan efisien, sehingga sering digunakan dalam analisis teks. Penelitian oleh [13] mengungkapkan bahwa Naive Bayes mampu memberikan hasil yang baik pada klasifikasi sentimen, terutama ketika diterapkan pada dataset berukuran besar. Meski sederhana, algoritma ini cocok untuk tugas-tugas yang memerlukan analisis probabilistik dasar. Di sisi lain, Support Vector Machine (SVM) bekerja dengan mencari hyperplane optimal yang memisahkan kelas-kelas dalam data secara maksimal. SVM sangat efektif dalam menangani data berdimensi tinggi dan sering digunakan untuk analisis data yang kompleks. SVM menggunakan kernel trick untuk memetakan data non-linear ke ruang yang lebih tinggi, sehingga kelas dapat dipisahkan dengan lebih baik. Menurut penelitian oleh [15], SVM sering kali unggul dibandingkan algoritma lain dalam analisis sentimen, terutama ketika data memiliki pola yang kompleks atau distribusi yang tidak seimbang. Algoritma ini dikenal karena keakuratannya yang tinggi dalam berbagai jenis dataset, termasuk yang memiliki dimensi tinggi atau distribusi kelas yang tidak merata.

2.6. Evaluasi Model

Dalam penelitian ini evaluasi model dilakukan menggunakan confusion matrix dan classification report. Confusion matrix adalah alat yang digunakan untuk menilai kinerja model klasifikasi dengan membandingkan prediksi model dengan label sebenarnya, dari confusion matrix seperti akurasi, presisi, recall, dan F1-score dapat dihitung untuk menilai efektivitas model. Penelitian oleh [1] menekankan pentingnya evaluasi yang tepat untuk memahami kekuatan dan kelemahan model dalam analisis sentimen. Selain itu, dalam tahap evaluasi kinerja sistem, classification report juga digunakan untuk menghitung metrik-metrik evaluasi seperti recall, precision, dan F1-score yang umum digunakan dalam analisis klasifikasi, termasuk dalam analisis sentimen. Recall mengukur kemampuan model untuk mengidentifikasi semua instance positif yang relevan, sementara precision mengukur kemampuan model untuk mengidentifikasi instance positif secara akurat. F1-score memberikan keseimbangan antara kedua metrik tersebut. Misalnya, recall yang tinggi menunjukkan kemampuan model dalam mengidentifikasi sebagian besar sentimen positif atau negatif secara keseluruhan, sementara precision yang tinggi menunjukkan bahwa model jarang memberikan prediksi yang salah. F1-score menggabungkan kedua aspek ini untuk memberikan gambaran yang lebih komprehensif tentang kinerja model secara keseluruhan. Berikut adalah rumus untuk menghitung accuracy, precision, recall, dan F1-score yang dapat pada persamaan di bawah ini:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Tabel 3. Confusion Matrix

Klasifikasi	Prediksi		
	Positif	Positif	Negatif
Aktual	Positif	True Positif (TP)	True Negatif (TN)
	Negatif	False Positif (FP)	Flase Negatif (FN)

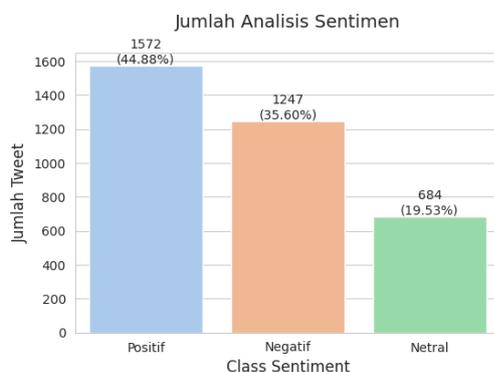
Berdasarkan tabel klasifikasi dan prediksi diatas *True Positif (TP)*, Jumlah data yang benar-benar positif dan diprediksi sebagai positif oleh model. *True Negatif (TN)*, Jumlah data yang benar-benar negatif dan diprediksi sebagai negatif oleh model. *False Positif (FP)*, Jumlah data yang sebenarnya negatif tetapi diprediksi sebagai positif oleh model. *False Negatif (FN)*, Jumlah data yang sebenarnya positif tetapi diprediksi sebagai negatif oleh model.

3. HASIL DAN PEMBAHASAN

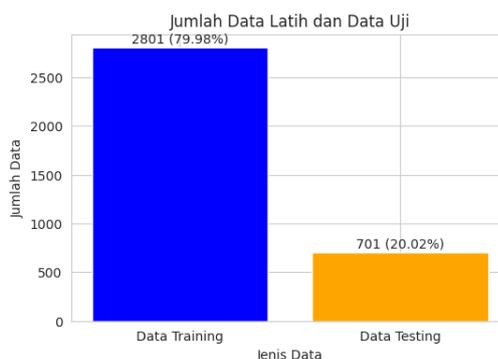
3.1. Tahap Pengujian

Penelitian ini mengumpulkan 3.656 tweet menggunakan kata kunci "BPJS" untuk mendapatkan opini masyarakat terkait layanan BPJS. Data yang terkumpul diproses melalui tahapan preprocessing, meliputi: cleaning (menghapus karakter khusus dan simbol), case folding (mengubah teks menjadi huruf kecil), normalisasi kata (mengubah variasi kata menjadi bentuk baku), tokenization (memecah teks menjadi token), stopword removal (menghapus kata-kata umum seperti "dan", "atau"), dan stemming (mengubah kata ke bentuk dasar) dari hasil preprocessing data ada sebanyak 3.502 data. Setelah preprocessing, data siap untuk klasifikasi.

Setelah dilakukan preprocessing, data dilabeli menggunakan metode *Lexicon-Based* Labeling dengan bantuan *InSet Lexicon*, menghasilkan tiga kategori sentimen. Sentimen positif mendominasi dengan 1.572 tweet (44.88%), menunjukkan pandangan positif masyarakat terhadap layanan BPJS. Sentimen negatif tercatat sebanyak 1.247 tweet (35.60%), mencerminkan kritik atau ketidakpuasan terhadap layanan tersebut. Sementara itu, sentimen netral berjumlah 684 tweet (19.53%), mengindikasikan konten yang tidak memiliki kecenderungan sentimen tertentu. Proporsi ini menggambarkan bahwa sentimen positif lebih menonjol dibandingkan sentimen negatif dan netral dalam pandangan masyarakat terhadap layanan BPJS. Hasil Labelling dapat di lihat pada gambar di bawah ini:



Gambar 2. Hasil Pelebelan Data



Gambar 3. Hasil Pembagian Data

Sebelum masuk pada proses klasifikasi, dataset dibagi menjadi dua bagian utama, yaitu data latih dan data uji atau rasion 80:20. Data latih terdiri dari 2.801 data atau 79.98% dari keseluruhan dataset, yang digunakan untuk melatih model agar dapat mengenali pola dan hubungan dalam data. Sementara itu, data uji berjumlah 701 data atau 20.02%, yang digunakan untuk menguji performa model terhadap data yang belum pernah dilihat sebelumnya. Pembagian ini penting untuk memastikan model tidak hanya bekerja baik pada data latih, tetapi juga mampu melakukan generalisasi pada data baru. Hasil pembagian dataset dapat di lihat pada gambar di bawah ini :

3.2. Klasifikasi Model

Setelah proses pembagian dataset selesai, langkah berikutnya adalah penerapan algoritma klasifikasi untuk mengevaluasi performa model. Dalam hal ini, algoritma yang digunakan adalah Naive Bayes (MultinomialNB) dan Support Vector Machine (SVM). Berikut adalah hasil evaluasi dari kedua algoritma tersebut:

Tabel 4. Evaluasi Kelas Sentimen Naïve Bayes

Sentimen	Precision		Recall		F1-Score		Support	
	NBC	SVM	NBC	SVM	NBC	SVM	NBC	SVM
Negatif	0.68	0.80	0.63	0.80	0.65	0.80	251	251
Netral	1.00	0.59	0.02	0.62	0.04	0.61	141	141
Positif	0.58	0.88	0.88	0.86	0.70	0.87	309	309

Tabel 4 menunjukkan evaluasi performa klasifikasi sentimen dengan algoritma *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* menggunakan metrik *Precision*, *Recall*, dan *F1-Score* dalam persentase. Pada sentimen Negatif, *precision*, *recall*, dan *F1-Score* untuk NBC masing-masing adalah 68%, 63%, dan 65%, sedangkan untuk SVM adalah 80%, 80%, dan 80% dengan jumlah data (*support*) sebanyak 251. Untuk sentimen Netral, NBC memiliki *precision* tinggi sebesar 100%, tetapi *recall* dan *F1-Score* sangat rendah, masing-masing hanya 2% dan 4%, sementara SVM mencatat *precision* 59%, *recall* 62%, dan *F1-Score* 61% dengan support 141 data. Pada sentimen Positif, NBC memiliki *precision* 58%, *recall* 88%, dan *F1-Score* 70%, sedangkan SVM menunjukkan performa lebih baik dengan *precision* 88%, *recall* 86%, dan *F1-Score* 87% pada support sebanyak 309 data. Secara keseluruhan, SVM menunjukkan performa lebih konsisten dan unggul pada sebagian besar metrik dibandingkan NBC.

Tabel 5. Hasil Accuracy

Model Algoritma	Accuracy
<i>Naïve Bayes Classifier</i>	0.6163
<i>Support Vector Machine</i>	0.7932

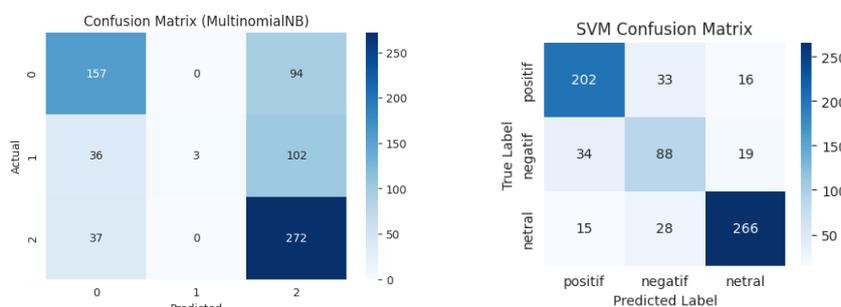
Dari tabel 5, terlihat bahwa algoritma *Support Vector Machine* memiliki akurasi yang lebih tinggi dibandingkan dengan *Naïve Bayes Classifier*. *Support Vector Machine* mencapai akurasi sebesar 79.32%, sedangkan *Naïve Bayes Classifier* hanya mencapai 61.63%. Akurasi yang lebih tinggi menunjukkan bahwa *Support Vector Machine* lebih mampu memprediksi label secara benar dalam dataset ini dibandingkan *Naïve Bayes Classifier*. Hasil evaluasi berdasarkan *precision*, *recall*, dan *f1-score* ditampilkan dalam tabel berikut:

Tabel 6. Hasil Evaluasi Precision, Recall, dan F1-Score

Model Algoritma	Precision (macro avg)	Recall (macro avg)	F1-Score (macro avg)	Precision (weighted avg)	Recall (weighted avg)	F1-Score (weighted avg)
<i>Naïve Bayes Classifier</i>	0.75	0.51	0.46	0.70	0.62	0.55
<i>Support Vector Machine</i>	0.76	0.76	0.76	0.80	0.7	0.79

Pada tabel 6, Algoritma SVM menunjukkan nilai rata-rata *precision*, *recall*, dan *f1-score* pada *macro average* sebesar 76%, lebih tinggi dibandingkan *Naïve Bayes Classifier* yang hanya mencapai 46% untuk *f1-score*. Hal ini menunjukkan bahwa SVM lebih konsisten dalam mendeteksi semua kelas dibandingkan dengan *Naïve Bayes Classifier*. *Naïve Bayes Classifier* memiliki kelemahan terutama pada kelas Netral, dengan *precision* dan *recall* yang rendah, sehingga memengaruhi hasil keseluruhan. Selain itu, SVM memiliki *weighted average* yang lebih

baik di semua metrik evaluasi (*precision*, *recall*, dan *f1-score*), masing-masing mencapai 80%, 79%, dan 79%. Hal ini menunjukkan performa SVM yang lebih baik dalam mempertimbangkan proporsi kelas yang berbeda. Sebaliknya, *Naive Bayes Classifier*, meskipun menunjukkan performa yang cukup baik pada kelas tertentu, memiliki *weighted average* yang lebih rendah dengan *f1-score* hanya 55%.



Gambar 4. Hasil Confution Matrix

Hasil confusion matrix di atas menunjukkan bahwa model Naive Bayes memiliki kelemahan dalam membedakan kelas, terutama pada kelas netral yang sering salah diklasifikasikan sebagai positif (94 kesalahan) dan kelas negatif yang banyak salah diprediksi sebagai netral (102 kesalahan), meskipun performanya cukup baik pada kelas netral dengan 272 prediksi benar. Sebaliknya, model SVM menunjukkan performa yang lebih baik secara keseluruhan, dengan prediksi yang lebih akurat di semua kelas, seperti pada kelas positif (202 prediksi benar), netral (266 prediksi benar), dan kesalahan yang relatif lebih sedikit dibandingkan Naive Bayes. SVM lebih unggul dalam mengenali pola antar kelas, sehingga dapat memberikan hasil prediksi yang lebih akurat.

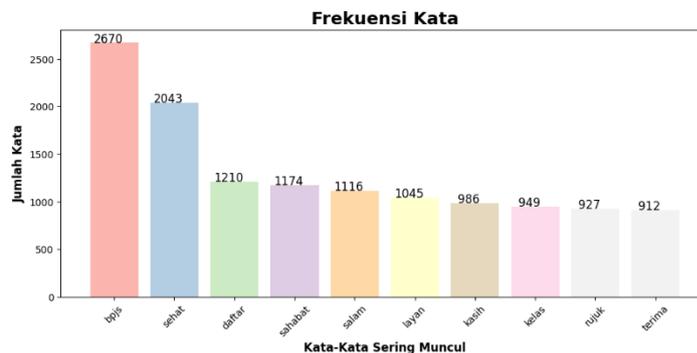
Hasil evaluasi kinerja dari algoritma Naive Bayes Classifier (NBC) dan Support Vector Machine (SVM) pada klasifikasi sentimen Twitter terhadap pelayanan BPJS menunjukkan bahwa SVM unggul dalam beberapa metrik utama. Pada Tabel 4, yang menunjukkan evaluasi kelas sentimen berdasarkan Precision, Recall, dan F1-Score, SVM secara konsisten mengungguli NBC. Untuk sentimen Negatif, SVM mencatatkan precision, recall, dan F1-Score masing-masing sebesar 80%, 80%, dan 80%, jauh lebih tinggi dibandingkan dengan NBC yang hanya mencapai 68%, 63%, dan 65%. Pada sentimen Netral, meskipun NBC memiliki precision yang sangat tinggi (100%), nilai recall dan F1-Score-nya sangat rendah, hanya 2% dan 4%. Sebaliknya, SVM menunjukkan hasil yang lebih seimbang dengan precision 59%, recall 62%, dan F1-Score 61%. Untuk sentimen Positif, SVM kembali unggul dengan precision 88%, recall 86%, dan F1-Score 87%, sementara NBC hanya mencapai precision 58%, recall 88%, dan F1-Score 70%.

Selain itu, pada Tabel 5 yang menunjukkan hasil akurasi, SVM mencapai akurasi sebesar 79.32%, jauh lebih tinggi dibandingkan NBC yang hanya mencapai 61.63%. Ini menunjukkan bahwa SVM lebih efektif dalam memprediksi label sentimen dengan benar. Tabel 6 lebih mendalam menggambarkan performa kedua algoritma dalam hal precision, recall, dan F1-Score. Rata-rata macro SVM mencapai 76%, sementara NBC hanya 46% pada F1-Score, yang menunjukkan ketidakseimbangan dalam kemampuan deteksi antar kelas pada NBC, terutama pada kelas Netral. Sebaliknya, SVM menunjukkan performa yang lebih stabil dengan *weighted average* yang lebih baik di semua metrik (*precision* 80%, *recall* 79%, dan *F1-Score* 79%), yang menunjukkan kemampuan model untuk mempertimbangkan proporsi kelas yang berbeda dengan lebih baik.

Hasil ini mendukung bahwa SVM lebih unggul dalam hal stabilitas dan ketepatan dalam mengklasifikasikan berbagai kelas sentimen. Dengan hasil evaluasi yang lebih konsisten dan lebih baik, penerapan SVM dapat membantu meningkatkan analisis sentimen terhadap pelayanan BPJS dengan memberikan hasil yang lebih akurat. Hal ini bisa digunakan oleh BPJS untuk memahami umpan balik publik dengan lebih baik, khususnya dalam merespons kritik atau keluhan yang mungkin kurang terdeteksi dengan metode lain, serta untuk merancang kebijakan yang lebih tepat guna dalam meningkatkan pelayanan berbasis sentimen masyarakat

3.3 Visualisasi Data

Gambar 5, grafik menunjukkan frekuensi kata yang paling sering muncul dalam data, dengan kata "bpjs" sebagai kata yang paling dominan muncul sebanyak 2670 kali, diikuti oleh "sehat" sebanyak 2043 kali. Kata-kata lainnya, seperti "daftar," "sahabat," "salam," dan "layan," juga sering muncul dengan jumlah antara 1210 hingga 1045 kali. Sementara itu, kata "kasih," "kelas," "rujuk," dan "terima" memiliki frekuensi kemunculan yang lebih rendah, berkisar antara 986 hingga 912 kali. Hal ini menunjukkan fokus data pada topik terkait layanan kesehatan, khususnya yang berhubungan dengan BPJS dan kesehatan.



Gambar 5. Hasil Frekuensi kata



Gambar 6. Wordcloud Kelas Sentimen

Gambar 6 menunjukkan word cloud untuk masing-masing kategori sentimen: negatif, positif, dan netral. Pada sentimen negatif, kata-kata seperti "bpjs," "sehat," "rawat inap," "kelas," dan "dokter" mendominasi, mencerminkan keluhan terkait layanan kesehatan, fasilitas rawat inap, serta standar kelas. Sentimen positif didominasi oleh kata-kata seperti "terima kasih," "bpjs," "sehat," dan "sahabat," yang menunjukkan apresiasi terhadap layanan kesehatan dan pengalaman positif pengguna. Sementara itu, pada sentimen netral, kata-kata seperti "bpjs," "sehat," "terima kasih," dan "indikasi medis" sering muncul, mengindikasikan komentar deskriptif atau informasi netral. Secara keseluruhan, kata "bpjs" dan "sehat" konsisten muncul di semua kategori, menunjukkan bahwa topik utama pembahasan adalah layanan kesehatan terkait BPJS.

4. KESIMPULAN

Penelitian ini membandingkan kinerja algoritma Support Vector Machine (SVM) dan Naive Bayes Classifier (NBC) dalam klasifikasi sentimen masyarakat terhadap layanan BPJS di Twitter. Dari 3.565 tweet yang dikumpulkan, data diproses melalui tahap preprocessing dan diberi label menggunakan metode lexicon-based, menghasilkan tiga kategori sentimen: positif (44,88%), negatif (35,60%), dan netral (19,52%). Hasil evaluasi menunjukkan bahwa SVM memiliki performa yang lebih unggul dibandingkan NBC, dengan akurasi sebesar 79,32% dibandingkan dengan 61,63% pada NBC. Selain itu, SVM menunjukkan nilai presisi, recall, dan f1-score yang lebih baik pada kategori sentimen negatif (0,80), netral (0,61), dan positif (0,87). Dengan akurasi yang lebih tinggi, SVM terbukti lebih efektif dalam memprediksi label sentimen yang benar pada dataset ini dibandingkan NBC, sehingga dapat disimpulkan bahwa SVM adalah algoritma yang lebih tepat untuk analisis sentimen. Penelitian ini menegaskan bahwa SVM lebih efektif untuk analisis sentimen Twitter terhadap layanan BPJS dibandingkan dengan Naive Bayes. Temuan ini dapat dimanfaatkan untuk meningkatkan pemahaman pemerintah terhadap opini masyarakat, serta mengembangkan sistem pemantauan opini publik berbasis kecerdasan buatan. Namun, penelitian ini memiliki beberapa keterbatasan. Pertama, penelitian ini hanya menggunakan Twitter sebagai sumber data, yang mungkin tidak mencerminkan seluruh opini masyarakat. Oleh karena itu, studi selanjutnya dapat mempertimbangkan platform lain seperti facebook, instagram dan tik tok untuk analisis yang lebih komprehensif. Kedua, penggunaan metode lexicon-based dalam pelabelan sentimen terbatas pada kata-kata yang ada dalam kamus sentimen dan dapat terpengaruh oleh konteks atau ironi. Untuk penelitian mendatang, disarankan untuk menggunakan pendekatan berbasis pembelajaran mendalam yang dapat menangani konteks dan ironi dalam teks lebih baik.

DAFTAR PUSTAKA

- [1] T. D. Dikiyanti, *Analisis Sentimen dan Pemodelan Topik Mengenai BPJS Kesehatan Berdasarkan Data Hasil Crawling Twitter Menggunakan Indonesian Sentiment* 2020.
- [2] A. Z. Najib, S. Achmadi, and K. Auliasari, "SISTEM KLASIFIKASI DATA PENDUDUK UNTUK MENENTUKAN TEMPAT PEMUNGUTAN SUARA (TPS) DENGAN METODE K-NEAREST NEIGHBOR (KNN) BERBASIS WEBSITE (STUDI KASUS : PEMILU DESA BULULAWANG)," vol. 8, no. 2, pp. 1323–1330, 2024.
- [3] M. Taufik Sugandi, Martanto, and U. Hayati, "Analisis Sentimen Komentar Pengguna Youtube terhadap Kebijakan Baru Badan Penyelenggara Jaminan Kesehatan Sosial Menggunakan Naive Bayes," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 218–227, 2024.
- [4] Z. Ardika and A. D. Wowor, "Analisis Sentimen Masyarakat Terhadap Program Badan Penyelenggara Jaminan Sosial (Bpjs) Menggunakan Data Twitter," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 1, pp. 90–99, 2024.
- [5] D. Atmajaya, A. Febrianti, and H. Darwis, "Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," *The Indonesian Journal of Computer Science*, vol. 12, no. 4. 2023.
- [6] B. A. Maulana and M. J. Fahmi, "Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," vol. 4, no. April, pp. 375–384, 2024.
- [7] Y. A. Singgalen, "Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine," vol. 4, no. 1, pp. 323–332, 2023.
- [8] A. Wandani, "Sentimen Analisis Pengguna Twitter pada Event Flash Sale Menggunakan Algoritma K-NN , Random Forest , dan Naive Bayes," vol. 5, no. September, pp. 651–665, 2021.
- [9] D. A. N. D. Sosial, "BUATAN DALAM SISTEM INFORMASI : TINJAUAN LITERATUR TENTANG APLIKASI , ETIKA ," vol. 6, pp. 1163–1169, 2023.
- [10] H. Oktavian *et al.*, "Akuntabilitas Birokrasi Dalam Pelayanan Publik Pada Dinas Ketenakerjaan Dan Transmigrasi Kabupaten Sikka , Provinsi Nusa Tenggara Timur Bureaucratic Accountability in Public Services at the Employment and Transmigration Office of Sikka Regency , East Nusa," vol. 2, no. 1, pp. 19–28, 2024.
- [11] D. Hernikawati, "Kecenderungan Tanggapan Masyarakat Terhadap Vaksin Sinovac Berdasarkan Lexicon Based Sentiment Analysis The Trend of Public Response to Sinovac Vaccine Based on Lexicon Based Sentiment Analysis," *Jurnal Ilmu Pengetahuan dan Teknologi Komunikasi*, vol. 23, no. 1. pp. 21–31, 2021.
- [12] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN," *J. KomtekInfo*, vol. 10, pp. 1–7, 2023.
- [13] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 4, pp. 411–423, 2022.
- [14] T. Sugihartono and R. R. C. Putra, "Penerapan Metode Support Vector Machine Dalam Classifikasi Ulasan Pengguna Aplikasi Mobile Jkn," *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 7, no. 2, pp. 144–153, 2024.
- [15] R. A. Husen, R. Astuti, L. Marlia, R. Rahmaddeni, and L. Efrizoni, "Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 211–218, 2023.