



Analisis Sentimen Pada Ulasan Aplikasi Bank Syariah Indonesia Mobile Menggunakan *Support Vector Machine* dan Naïve Bayes

Nabila Fadia Aqilla^{*}, Muhammad Jazman, Syaifullah, Medyantiwi Rahmawita

Fakultas Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ^{1,*}12050323291@students.uin-suska.ac.id, ²jazman@uin-suska.ac.id, ³syaifullah@uin-suska.ac.id,
⁴medyantiwi.rahmawita@uin-suska.ac.id

Email Penulis Korespondensi: 12050323291@students.uin-suska.ac.id

Submitted: **29/12/2024**; Accepted: **26/02/2025**; Published: **01/03/2025**

Abstrak—Internet memiliki peranan yang sangat penting dalam mempermudah berbagai kegiatan manusia, termasuk dalam sektor layanan perbankan elektronik, yang mencakup berbagai layanan keuangan seperti ATM, internet banking, SMS banking, dan mobile banking. Semua ini bertujuan untuk meningkatkan kualitas layanan dengan fokus pada keamanan, kenyamanan, dan efektivitas. BSI adalah salah satu bank yang menawarkan layanan perbankan mobile. Berdasarkan ulasan pengguna, aplikasi BSI Mobile sering mengalami masalah teknis seperti bug dan kegagalan saat melakukan transaksi. Untuk mengukur tingkat kepuasan terhadap aplikasi, peneliti menggunakan metode analisis sentimen. Metode ini juga membantu calon pelanggan untuk mengenali aspek-aspek yang perlu diperbaiki atau dikembangkan dalam produk dan layanan guna meningkatkan kualitasnya. Penelitian ini menggunakan algoritma Support Vector Machine (SVM) dan Naïve Bayes. Hasil pengujian menunjukkan bahwa algoritma Naïve Bayes menghasilkan akurasi sebesar 74,37%, recall 74,37%, precision 75,46%, dan F1-score 74,5%. Sementara itu, algoritma SVM memperoleh akurasi 77,39%, precision 77,8%, recall 77,39%, dan F1-score 77,38%. Temuan ini mengindikasikan bahwa SVM terbaik terhadap tugas klasifikasi sentimen dibandingkan Naïve Bayes. Dengan kinerjanya yang lebih baik, SVM menjadi pilihan algoritma yang lebih tepat untuk menganalisis persepsi pengguna terhadap aplikasi BSI Mobile. Oleh karena itu, hasil penelitian ini dapat berkontribusi pada pengembangan strategi layanan digital yang lebih inovatif dan dapat meningkatkan daya saing di era digital.

Kata Kunci: Analisis Sentimen; BSI Mobile; Naïve Bayes; *Support Vector Machine* (SVM); Ulasan

Abstract—The internet plays a crucial role in facilitating various human activities, including in the field of electronic banking services, which encompasses various financial services such as ATMs, internet banking, SMS banking, and mobile banking. All of these aim to enhance service quality with a focus on security, convenience, and effectiveness. BSI is one of the banks offering mobile banking services. Based on user reviews, the BSI Mobile app often experiences technical issues such as bugs and transaction failures. To assess the level of satisfaction with the app, the researcher uses sentiment analysis methods. This method also helps potential customers identify aspects that need improvement or development in the products and services to enhance their quality. The study employs Support Vector Machine (SVM) and Naïve Bayes algorithms. The test results show that the Naïve Bayes algorithm achieves an accuracy of 74.37%, recall of 74.37%, precision of 75.46%, and an F1-score of 74.5%. Meanwhile, the SVM algorithm achieves an accuracy of 77.39%, precision of 77.8%, recall of 77.39%, and an F1-score of 77.38%. These findings indicate that SVM performs better in sentiment classification tasks compared to Naïve Bayes. With its superior performance, SVM is the more suitable algorithm for analyzing user perceptions of the BSI Mobile app. Therefore, the findings of this study can contribute to the development of more innovative digital service strategies and enhance competitiveness in the digital era.

Keywords: BSI Mobile; Naïve Bayes; Review; Sentiment Analysis; *Support Vector Machine* (SVM)

1. PENDAHULUAN

Internet berperan penting dalam mempermudah berbagai aktivitas manusia, termasuk layanan perbankan. Pertumbuhan pengguna internet membuka peluang bagi perbankan untuk mengembangkan produk teknologi yang memungkinkan transaksi keuangan berjalan tanpa terbatas pada waktu atau tempat [1]. Perbankan elektronik merupakan sebuah inovasi yang mencakup berbagai layanan keuangan, seperti ATM, internet banking, SMS banking, dan mobile banking. Tujuan utama dari inovasi ini adalah untuk meningkatkan kualitas layanan dengan menekankan pada keamanan, kenyamanan, dan efektivitas [2].

Mobile banking dapat membantu pelanggan dalam melakukan berbagai transaksi keuangan langsung melalui ponsel [3]. Sebelum adanya layanan ini, pelanggan perlu memahami format pesan teks SMS serta nomor tujuan pengiriman tertentu untuk dapat berkomunikasi dan melakukan transaksi dengan pihak bank. Kehadiran *mobile banking* mempermudah akses layanan keuangan, karena pelanggan tidak lagi perlu mengunjungi ATM atau kantor bank secara langsung untuk melakukan transaksi [4]. Salah satu bank menggunakan layanan mobile banking yaitu Bank Syariah Indonesia (BSI).

Bank Syariah Indonesia (BSI) secara resmi berdiri pada 1 Februari 2021. oleh Presiden Republik Indonesia. Salah satu layanannya, BSI *Mobile* dirancang untuk mempermudah pengguna melaksanakan transaksi tanpa perlu mengunjungi bank secara langsung. Umpulan pengguna sangat diperlukan untuk mengevaluasi tingkat kepuasan dan meningkatkan kinerja aplikasi. Berdasarkan data dari *Google PlayStore*, BSI *Mobile* telah digunakan lebih dari 5 juta pengguna, dengan pertumbuhan tahunan mencapai 37%. Hingga Maret 2023, jumlah transaksi kumulatif melalui aplikasi ini mencapai 86,7 juta atau sekitar 57% dari total transaksi. Aplikasi ini juga telah menerima 163.332 ulasan pengguna baik positif maupun negatif.

Berdasarkan ulasan pengguna, aplikasi BSI *Mobile* sering menghadapi masalah seperti *bug* dan kegagalan pada saat melakukan transaksi. Pengguna juga mengeluhkan stabilitas aplikasi yang kurang baik, dengan seringnya muncul



informasi kegagalan transksi meskipun koneksi internet dalam kondisi stabil. Selain itu, masalah *timeout* dan permintaan input berulang kali menjadikan keluhan utama. Kondisi ini sangat mengganggu, terutama bagi pengguna yang mengandalkan aplikasi untuk transaksi sehari-hari. Kegagalan tersebut tidak hanya memakan waktu tetapi juga berpotensi menghambat aktivitas bisnis pengguna.

Untuk mengukur tingkat kepuasan terhadap aplikasi, peneliti menggunakan metode analisis sentimen [5]. Metode ini memungkinkan peneliti memahami persepsi atau penilaian individu terhadap sebuah produk, peristiwa, atau entitas tertentu. Analisis sentimen juga membantu calon klien atau pelanggan yang berminat pada produk maupun layanan untuk mengidentifikasi aspek yang perlu diperbaiki atau dikembangkan guna meningkatkan kualitas layanan [6].

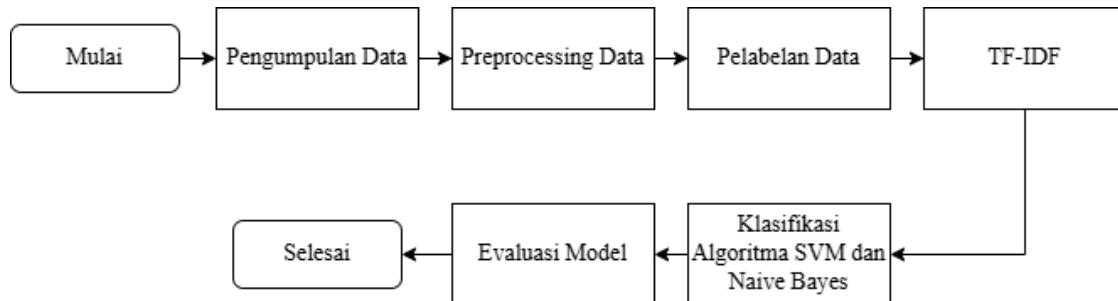
Penelitian ini menerapkan algoritma SVM dan *Naïve Bayes* untuk analisis sentimen. *Support Vector Machine* memanfaatkan fungsi kernel untuk menemukan *hyperplane* optimal dalam klasifikasi data [7], sementara *Naïve Bayes* mengklasifikasi data berdasarkan probabilitas dengan metode yang sederhana dan efektif [8].

Pada penelitian yang dilakukan Nuraeni Herlinawati dan kawan-kawan (2020) menyatakan bahwa SVM 6,85% lebih baik daripada dengan *Naïve Bayes* [9]. Pada penelitian R Kusumawati dan kawan-kawan (2019) menyatakan bahwa SVM lebih unggul dengan akurasi 83,34% [10]. Namun, penelitian yang dilakukan Sri Lestari dan kawan-kawan (2024) menyatakan bahwa *Naïve Bayes* lebih unggul dengan akurasi 96,14% [11]. Selanjutnya, penelitian oleh Safitri Linawati dan kawan-kawan (2020) didapatkan algoritma *Naïve Bayes* memiliki akurasi yang lebih tinggi yaitu 90,90% dibandingkan dengan metode SVM yaitu 89,25% [12].

Berdasarkan uraian diatas maka penelitian ini memanfaatkan algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes* untuk menganalisis sentimen terhadap aplikasi BSI Mobile di *Google PlayStore*. Tujuannya adalah mengelompokkan ulasan pengguna ke dalam kategori netral, negatif, dan positif, sehingga dapat menjadi acuan bagi calon pengguna dan membantu perusahaan mengenali kekurangan aplikasi. Harapannya, penelitian ini dapat memberikan pemahaman lebih baik kepada masyarakat tentang analisis sentimen aplikasi BSI Mobile di *Google PlayStore*.

2. METODOLOGI PENELITIAN

Penelitian ini dilakukan dalam beberapa tahap secara terstruktur untuk mencapai tujuan yang telah ditentukan. Tahapan penelitian ini disajikan dalam bentuk diagram alir agar dapat berjalan sesuai tujuan yang akan dicapai. Metodologi yang digunakan dalam penelitian ini disajikan pada Gambar 1.



Gambar 1. Metodologi Penelitian

Penelitian ini bertujuan untuk mengidentifikasi sentimen dalam ulasan pengguna aplikasi BSI *Mobile* yang tersedia di *Google Playstore*. Data ulasan dikumpulkan melalui teknik *crawling*, kemudian diolah melalui beberapa tahap *preprocessing*. Selanjutnya, fitur diekstraksi menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) dan data tersebut diberi label. Tahap berikutnya melibatkan proses pengklasifikasian data menggunakan algoritma SVM dan *Naïve Bayes*. Kemudian akan diperoleh akurasi masing-masing algoritma dan selanjutnya masuk ke tahap evaluasi model.

2.1 Analisis Sentimen

Analisis sentimen adalah bidang ilmu yang mempelajari opini, penilaian, sentimen, sikap, dan emosi individu terhadap suatu topik, peristiwa, isu, atau atribut tertentu. Dalam bidang ini, terdapat beberapa istilah yang saling berkaitan, seperti analisis sentimen, opinion mining, dan sentiment mining, yang semuanya termasuk dalam cakupan analisis sentimen [13].

Analisis sentimen digunakan oleh pelaku bisnis, pemerintah, dan peneliti sebagai alat untuk memahami opini publik, suasana hati masyarakat, mendapatkan wawasan strategis, dan mendukung pengambilan keputusan yang lebih baik [14]. Tugas utama dalam analisis ini adalah mengkategorikan sentimen pengguna menjadi tiga kelompok [15].

- a. Sentimen positif: sentimen positif merupakan sebuah perilaku atau anggapan seseorang yang meningkatkan nilai terhadap sesuatu.



- b. Sentimen negatif: sentimen negatif adalah sikap atau pandangan yang dapat mengurangi nilai suatu hal dan berpotensi menurunkan tren. Kalimat dengan sentimen negatif biasanya ditandai oleh penggunaan kata-kata negasi atau bentuk penyangkalan.
- c. Sentimen netral: sentimen netral adalah respons atau pandangan yang bersifat netral dan tidak memihak. Kalimat dengan sentimen netral biasanya tidak mengandung ekspresi yang menunjukkan sifat positif maupun negatif.

2.2 Pengumpulan Data

Pengumpulan data merupakan langkah untuk mengumpulkan data mengenai aplikasi BSI *Mobile* dari *Google Playstore*. Proses pengumpulan data ini dilakukan menggunakan teknik *scrapping* untuk mengambil ulasan yang terdapat pada aplikasi BSI *Mobile* di *google playstore*.

2.3 Preprocessing Data

Preprocessing data merupakan tahap yang dilakukan guna mengatasi data yang tidak sempurna, mengurangi gangguan data, serta membersihkan data yang tidak seragam atau masih berisi informasi yang tidak penting [16]. Terdapat beberapa tahapan *preprocessing* data untuk penelitian ini, yaitu [17]:

- a. *Case folding*, tahapan untuk mengubah teks atau kalimat yang berhuruf kapital menjadi huruf kecil, menghapus karakter, *username*, dan lain-lain.
- b. *Tokenizing*, tujuan dari langkah ini adalah untuk memecah kalimat menjadi kata-kata individual.
- c. *Stopword Removal*, langkah ini bertujuan untuk menghilangkan kata-kata yang tidak berkaitan dan tidak memiliki makna penting.
- d. *Stemming*, tujuan tahap ini adalah untuk mengubah kata yang memiliki imbuhan menjadi kata dasar, dengan menggunakan library Sastrawi sebagai bantuan.

2.4 Pelabelan

Proses pelabelan pada data komentar dilakukan dengan mengklasifikasikan komentar ke dalam tiga kategori berbeda: positif (misalnya, puji atau kepuasan pelanggan), negatif (kritik atau keluhan), dan netral (pertanyaan tentang produk, promo, atau informasi lainnya) [18]. Penilaian ini bertujuan menentukan sentimen masing-masing komentar.

2.5 TF-IDF

Metode Term Frequency-Inverse Document Frequency (TF-IDF) diterapkan untuk mengonversi data teks ke dalam format numerik dengan cara menghitung bobot masing-masing kata atau fitur yang terdapat dalam sebuah dokumen [19]. *Term Frequency* (TF) mengukur relevansi suatu istilah dalam dokumen dengan menghitung frekuensi kemunculannya. Sementara itu, *Inverse Document Frequency* (IDF) menyesuaikan bobot kata dengan mengurangi nilai jika kata tersebut sering muncul di banyak dokumen, sehingga memberikan fokus pada kata-kata yang lebih unik [20]. Adapun rumus dari pembobotan kata TF-IDF yaitu :

$$W(t, d) = tf(d, t) \times idf \quad (1)$$

Dalam penelitian ini, bobot suatu dokumen terhadap kata tertentu dihitung menggunakan rumus $W(t, d)$, yang mempresentasikan bobot kata ke- t dalam dokumen ke- d . Perhitungan bobot ini melibatkan dua komponen utama, yaitu frekuensi kata ($tf(d, t)$) dan *Inversed Document Frequency* (idf).

Frekuensi kata ($tf(d, t)$) menggambarkan seberapa sering kata ke- t muncul dalam dokumen ke- d . Semakin sering kata tersebut muncul, semakin besar nilai tf -nya. Sementara itu, idf adalah ukuran yang menunjukkan seberapa unik atau jarangnya kata tersebut dalam seluruh koleksi dokumen. Kata yang muncul dalam banyak dokumen akan memiliki nilai idf yang lebih rendah, sedangkan kata yang jarang muncul akan memiliki nilai idf yang lebih tinggi.

Penggabungan antara tf dan idf pada perhitungan $W(t, d)$ memungkinkan perhitungan bobot kata dalam dokumen menjadi relevan, dengan memberikan penekanan pada kata-kata yang tidak hanya sering muncul dalam dokumen tertentu, tetapi juga jarang ditemukan dalam dokumen lain. Pendekatan ini meningkatkan akurasi analisis teks, terutama dalam proses klasifikasi dan pengelompokan dokumen.

2.6 Support Vector Machine (SVM)

Algoritma Support Vector Machine (SVM) pertama kali dikembangkan oleh Vapnik dan timnya pada akhir tahun 1970-an. Algoritma ini merupakan metode pembelajaran berbasis kernel yang banyak digunakan dalam berbagai aplikasi *machine learning*, khususnya untuk *image classification* [7]. SVM bekerja dengan memisahkan kelas secara linear dengan memetakan fitur ke dalam ruang kernel. Secara umum, algoritma ini menambahkan dimensi fitur untuk menemukan *hyperplane* optimal yang memisahkan data. Pembagian data dilakukan dengan mencari vektor terdekat dari *hyperplane* pada setiap kelompok data [21]. Prinsip utama SVM adalah klasifikasi linear, di mana tingkat akurasi dipengaruhi oleh parameter yang digunakan serta fungsi kernel yang diterapkan [22]. Terdapat beberapa kernel pada algoritma ini diantaranya [23] :

- a. *Linear*

$$K(x, y) = x \cdot y \quad (2)$$

b. *Polynomial*

$$K(xi, xj) = ((xi \cdot xj) + c)^d \quad (3)$$

c. *Gaussian*

$$K(xi, xj) = \exp\left(-\frac{|xi - xj|^2}{2\sigma^2}\right) \quad (4)$$

d. *Sigmoid*

$$K(xi, xj) = \tanh(\alpha xi \cdot xj + \beta) \quad (5)$$

Fungsi kernel $K(xi, xj)$, yang berfungsi untuk mengukur tingkat kesamaan atau hubungan antara dua data, yaitu data ke-i (xi) dan data ke-j (xj). Fungsi ini memanfaatkan operasi eksponen (exp) untuk meningkatkan kemampuan dalam mengenali pola yang lebih kompleks pada data.

Fungsi kernel memiliki dua parameter utama, yaitu d dan σ . Parameter d mengatur sensitivitas terhadap perubahan data, sedangkan σ menentukan skala penyebaran data dalam ruang fitur. Hasil pemrosesan kernel ini kemudian dikombinasikan dengan matriks output *weight* (β), yang digunakan untuk mengoptimalkan kinerja model dalam proses prediksi atau klasifikasi. Pendekatan kernel ini memungkinkan data untuk dipetakan ke ruang fitur berdimensi tinggi, sehingga pola non-linear dalam data dapat diidentifikasi dan dianalisis dengan lebih baik.

2.7 Naïve Bayes

Naïve Bayes adalah metode klasifikasi yang mudah diterapkan, yang didasarkan pada teori statistik Bayesian atau teorema Bayes. Pendekatan ini memanfaatkan data historis serta analisis statistik dan probabilitas untuk memprediksi kemungkinan di masa depan. Algoritma ini bekerja dengan menentukan kelompok yang paling relevan melalui perhitungan probabilitas tertinggi dari hasil tes [8]. Selain itu, *Naïve Bayes* banyak diterapkan dalam penambangan teks, khususnya pada analisis sentimen, karena kemampuannya yang unggul dalam klasifikasi data, efisiensi komputasi, dan tingkat akurasi yang tinggi [24]. Berikut ini adalah rumus yang digunakan dalam metode *Naïve Bayes* [25]:

$$P(C|X) = \frac{P(X|C)P(c)}{P(x)} \quad (6)$$

Dalam konteks klasifikasi, terdapat beberapa elemen yang menggambarkan hubungan probabilistik antara input dan class. Misalnya, X merupakan vektor input yang digunakan untuk menentukan kelas spesifik, yang disimbolkan dengan c . $P(C|X)$ merujuk pada probabilitas dari suatu kelas berdasarkan vektor input yang diketahui, yang dikenal sebagai posterior probability. Sementara itu, $P(c)$ adalah probabilitas dari kelas yang dicari (prior probability) di seluruh data yang tersedia. $P(X|C)$ menggambarkan probabilitas dari setiap input yang diberikan, berdasarkan kondisi kelas yang ditetapkan, dan $P(x)$ adalah probabilitas dari suatu input dalam keseluruhan data yang ada.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Tahapan pengumpulan data merupakan langkah penting dalam penelitian ini, yang berfokus pada pengambilan data dari *Google PlayStore* terkait ulasan pengguna terhadap aplikasi *BSI Mobile*. Proses pengumpulan data dilakukan melalui penerapan teknik web scraping, yang bertujuan untuk mengekstraksi komentar atau ulasan yang tersedia pada halaman aplikasi *BSI Mobile* di platform tersebut. Dalam proses ini, digunakan bahasa pemrograman *python* sebagai alat utama untuk mengimplementasikan teknik scraping. Hasil dari tahap ini berhasil mengumpulkan sebanyak 1.000 data ulasan dari aplikasi *BSI Mobile* di *Google PlayStore*. Hasil dari proses pengumpulan data dapat dilihat pada Tabel 1.

Tabel 1. Hasil Scrapping data

No.	Ulasan
1	Terlalu sering pembaruan aplikasi tanpa manfaat yg jelas! Sangat mengganggu!
2	Mksh....Apk sangat bagus....
3	Kenapa ya saya tidak bisa aktivasi di BSI Mobile, selalu gagal katanya data tidak sesuai padahal semua data no hp, no ktp dan no rekening sudah sesuai...
...	...
1000	Sebentar bentar harus di perbarui kalo gga di perbarui gga bisa buka geraaaam

Setelah data diperoleh melalui proses scrapping seperti yang tercantum pada Tabel 1, tahap berikutnya adalah melakukan pembersihan teks atau *preprocessing* data teks. Proses ini melibatkan beberapa tahap, seperti *case folding*,



tokenizing, stopword removal, dan stemming. Tujuan dari *preprocessing* ini adalah untuk mempersiapkan data agar lebih siap dianalisis dan meningkatkan efisiensi serta akurasi model.

3.2 Hasil Preprocessing Data

Tahap *preprocessing* data adalah langkah awal yang bertujuan untuk mempersiapkan data mentah sebelum memasuki proses analisis lebih lanjut. Dalam penelitian ini, proses *preprocessing* dilakukan menggunakan bahasa pemrograman *python* dengan bantuan platform *google colab*. Tahapan *preprocessing* meliputi *case folding* (mengubah teks menjadi huruf kecil), *tokenization* (memecah teks menjadi kata-kata), *stopword removal* (menghapus kata-kata umum yang tidak relevan), dan *stemming* (mengubah kata ke bentuk dasarnya). Hasil dari *preprocessing* data ini disajikan pada Tabel 2.

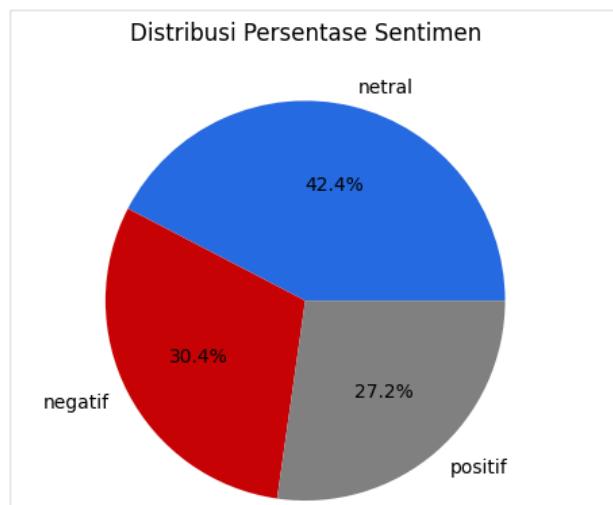
Tabel 2. Hasil Preprocessing Data

Tahapan	Ulasan
Data Awal	Kenapa ya saya tidak bisa aktivasi di BSI Mobile, selalu gagal katanya data tidak sesuai padahal semua data no hp, no ktp dan no rekening sudah sesuai_
Cleaning	kenapa ya saya tidak bisa aktivasi di bsi mobile, selalu gagal katanya data tidak sesuai padahal semua data no hp, no ktp dan no rekening sudah sesuai_
Case Folding	kenapa ya saya tidak bisa aktivasi di bsi mobile selalu gagal katanya data tidak sesuai padahal semua data no hp no ktp dan no rekening sudah sesuai
Tokenizing	[‘kenapa’, ‘ya’, ‘saya’, ‘tidak’, ‘bisa’, ‘aktivasi’, ‘di’, ‘bsi’, ‘mobile’, ‘selalu’, ‘gagal’, ‘katanya’, ‘data’, ‘tidak’, ‘sesuai’, ‘padahal’, ‘semua’, ‘data’, ‘no’, ‘hp’, ‘no’, ‘ktp’, ‘dan’, ‘no’, ‘rekening’, ‘sudah’, ‘sesuai’]
Stopword Removal	[‘ya’, ‘aktivasi’, ‘bsi’, ‘mobile’, ‘gagal’, ‘data’, ‘sesuai’, ‘data’, ‘no’, ‘hp’, ‘no’, ‘ktp’, ‘no’, ‘rekening’]
Stemming	[‘ya’, ‘aktivasi’, ‘bsi’, ‘mobile’, ‘gagal’, ‘data’, ‘sesuai’, ‘data’, ‘no’, ‘hp’, ‘no’, ‘ktp’, ‘no’, ‘rekening’]

Berdasarkan Tabel 2, langkah selanjutnya adalah pelabelan komentar untuk mengidentifikasi sentimen yang terkandung didalamnya, yaitu negatif, netral, dan positif. Pelabelan ini penting sebagai dasar untuk analisis selanjutnya, seperti pemodelan dan evaluasi sentimen.

3.3 Pelabelan Data

Proses pelabelan data ulasan dilakukan dengan menilai setiap komentar untuk menentukan kategori sentimen yang sesuai. Komentar yang memiliki sentimen positif mencakup ulasan yang berisi pujian, rasa puas, atau apresiasi terhadap produk atau layanan. Komentar dengan sentimen negatif mencakup kritik, keluhan, atau ketidakpuasan pengguna. Sementara itu, komentar dengan sentimen netral biasanya berupa pertanyaan, informasi, atau diskusi terkait produk, promosi, dan topik lainnya yang tidak menunjukkan emosi tertentu. Visualisasi hasil pengolahan data dapat dilihat pada Gambar 2.



Gambar 2. Hasil Pelabelan Data

Berdasarkan analisis visualisasi data, distribusi sentimen ulasan menunjukkan bahwa sentimen netral mendominasi dengan jumlah sekitar 400 ulasan atau 42,4% diikuti oleh sentimen negatif sebanyak 300 ulasan atau 30,4%, dan sentimen positif dengan jumlah paling sedikit, yaitu 250 ulasan atau 27,2%. Hal ini mengindikasikan bahwa mayoritas ulasan tidak secara tegas menilai kualitas produk atau layanan, meskipun jumlah ulasan negatif lebih banyak dibandingkan dengan ulasan positif.



3.4 TF-IDF

Setelah melakukan tahap *text preprocessing*, langkah berikutnya adalah menerapkan pembobotan menggunakan metode TF-IDF. Pendekatan ini digunakan untuk menentukan nilai bobot masing-masing kata berdasarkan tingkat pengaruh atau frekuensi kemunculannya dalam teks. Bobot yang dihasilkan memberikan gambaran seberapa penting suatu kata dalam dokumen tertentu. Hasil transformasi pembobotan TF-IDF terdapat pada Tabel 3.

Tabel 3. Hasil TF-IDF

No	aan	abal	abdet	abis	account	ad	ada	...	yg
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.247980
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.280767
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0

Tabel 3 menunjukkan hasil TF-IDF yang menggambarkan bobot setiap kata dalam dokumen berdasarkan frekuensi lokal (TF) dan tingkat kepentingan global (IDF). Kata dengan bobot 0 menandakan bahwa kata tersebut tidak muncul atau tidak memiliki signifikansi dalam dokumen. Di sisi lain, nilai positif pada kolom "yg" menunjukkan bahwa kata tersebut memiliki tingkat relevansi yang lebih besar dalam dokumen tersebut dibandingkan dengan kata-kata lainnya.

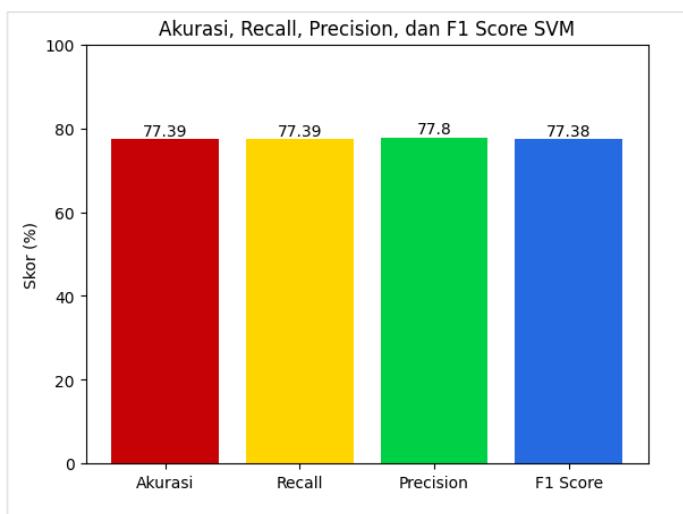
3.5 Hasil Klasifikasi Support Vector Machine

Klasifikasi sentimen dengan algoritma SVM menggunakan pendekatan kernel linear dilakukan dengan membagi data menjadi dua bagian, yaitu data latih dan data uji, dengan proporsi 80:20. Data latih (80%) digunakan untuk melatih model, sementara data uji (20%) digunakan untuk menguji kemampuan model dalam mengklasifikasikan sentimen. Kinerja model dievaluasi dengan mengukur metrik seperti *accuracy*, *recall*, *precision*, dan *f1-score*. Metrik-metrik ini memberikan gambaran seberapa baik model dalam mengklasifikasikan sentimen secara akurat dan efektif. Hasil evaluasi yang tercantum dalam Tabel 4 digunakan untuk menilai efektivitas algoritma SVM dalam menganalisis sentimen pada ulasan aplikasi BSI Mobile.

Tabel 4. Hasil Uji SVM

Hasil Uji Support Vector Machine			
Accuracy	Precision	Recall	F1-Score
77.39%	77.8%	77.39%	77.38%

Berdasarkan Tabel 3, hasil pengujian algoritma SVM pada data komentar menunjukkan akurasi sebesar 77,39%. Selain itu, model ini juga menghasilkan *precision* 77,8%, *recall* 77,39%, dan *f1-score* 77,38%. *Precision* menggambarkan persentase prediksi positif yang benar, sedangkan *recall* mengukur kemampuan model untuk mendeteksi data positif secara keseluruhan. *F1-score* memberikan keseimbangan antara *precision* dan *recall*. Hasil evaluasi ini kemudian divisualisasikan dalam bentuk diagram batang (*bar chart*) yang terdapat pada Gambar 3, yang memberikan gambaran visual mengenai kinerja model SVM dalam mengklasifikasikan data komentar.

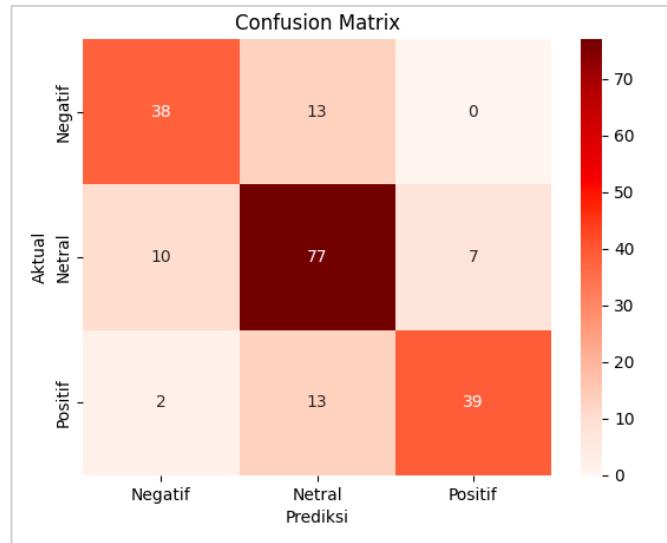


Gambar 3. Hasil Klasifikasi SVM

Setelah menghitung akurasi, *recall*, *precision*, dan *f1-score*, langkah berikutnya adalah menganalisis *confusion matrix*. *confusion matrix* digunakan untuk membandingkan hasil prediksi model dengan nilai asli dari dataset, dengan



menampilkan empat komponen utama, yaitu *true positive*, *false positive*, *true negatives*, dan *false negatives*. Analisis model ini memberikan wawasan tentang kesalahan prediksi model dan mendukung perhitungan metrik evaluasi lainnya. Hasil *confusion matrix* divisualisasikan dalam bentuk *heatmap* yang terdapat pada Gambar 4.



Gambar 4. Confussion Matrix Support Vector Machine

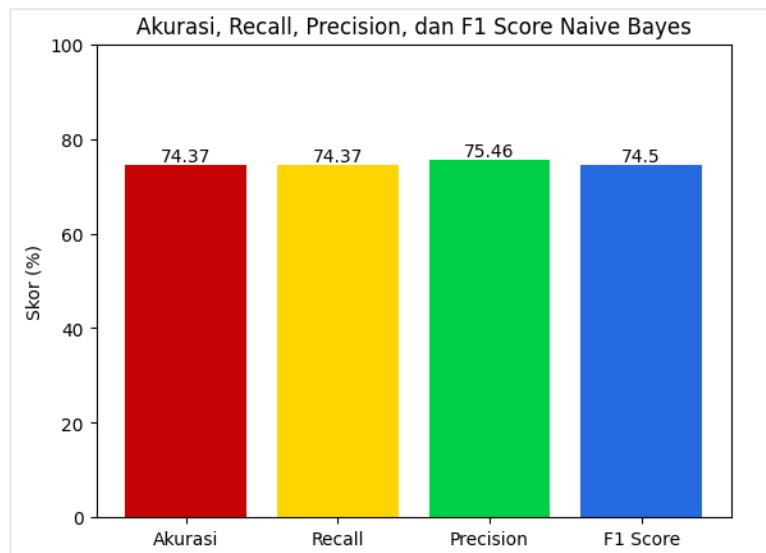
3.6 Hasil Klasifikasi Naïve Bayes

Klasifikasi sentimen menggunakan algoritma Naïve Bayes dievaluasi dengan metrik *accuracy*, *recall*, *precision*, dan *f1-score*. Metrik-metrik ini menilai seberapa efektif model dalam mengklasifikasikan sentimen dengan benar, mendeteksi sentimen positif yang ada, serta menyeimbangkan antara presisi dan *recall*. Hasil evaluasi yang diperoleh, yang tercantum dalam Tabel 5, digunakan untuk mengukur sejauh mana algoritma Naïve Bayes efektif dalam menganalisis sentimen pada data yang digunakan dalam penelitian ini.

Tabel 5. Hasil Uji Naïve Bayes

Hasil Uji Naïve Bayes			
Accuracy	Precision	Recall	F1-Score
74.37%	75.46%	74.37%	74.5%

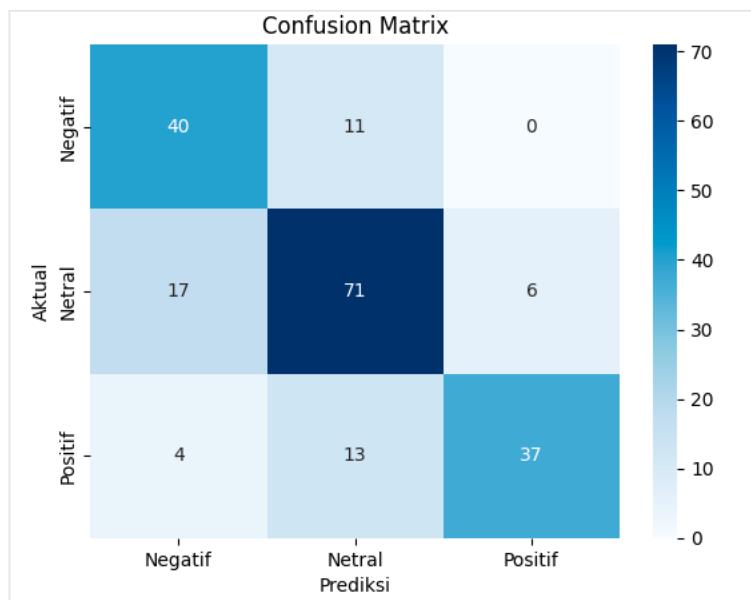
Berdasarkan Tabel 5, hasil pengujian algoritma Naïve Bayes pada data komentar menunjukkan akurasi sebesar 74,37%. Selain itu, model ini juga menghasilkan *precision* 75,46%, *recall* 74,37%, dan *f1-score* 74,5%. *Precision* menggambarkan persentase prediksi positif yang benar, sedangkan *recall* mengukur kemampuan model untuk mendeteksi data positif secara keseluruhan. *F1-score* memberikan keseimbangan antara *precision* dan *recall*. Hasil evaluasi ini kemudian divisualisasikan dalam bentuk diagram batang (*bar chart*) yang terdapat pada Gambar 5, yang memberikan gambaran visual mengenai kinerja model Naïve Bayes dalam mengklasifikasikan data komentar.



Gambar 5. Hasil Klasifikasi Naïve Bayes



Setelah menghitung akurasi, *recall*, *precision*, dan *f1-score*, langkah selanjutnya adalah menganalisis *confusion matrix* yang membandingkan nilai yang diprediksi oleh model dengan nilai sebenarnya dalam dataset. Matrix ini menunjukkan empat elemen utama, yaitu *true positives*, *false positives*, *true negatives*, dan *false negatives*. Untuk mempermudah interpretasi, *confusion matrix* divisualisasikan dalam bentuk *heatmap* yang terdapat pada Gambar 6. *Heatmap* ini memberikan gambaran visual tentang distribusi kesalahan prediksi model, dengan warna yang lebih gelap menunjukkan intensitas kesalahan yang lebih besar.



Gambar 6. Confussion Matrix Naïve Bayes

Pengujian dilakukan pada dua algoritma, yaitu SVM dan Naïve Bayes, dengan pembagian data sebesar 80% untuk data latih dan 20% untuk data uji. Algoritma SVM mencatat akurasi 77,39%, *precision* 77,8%, *recall* 77,39%, dan *f1-score* 77,38%. Sementara itu, algoritma Naïve Bayes menghasilkan akurasi sebesar 74,37%, *recall* 74,37%, *precision* 75,46%, dan *f1-score* 74,5%. Berdasarkan hasil pengujian, algoritma *Support Vector Machine* menunjukkan performa yang lebih baik dibandingkan Naïve Bayes dengan selisih akurasi sebesar 3,02%.

Dengan hasil tersebut, algoritma Support Vector Machine lebih efektif dalam menganalisis sentimen pengguna dan memahami pandangan terhadap aplikasi BSI Mobile. Hal ini memberikan kontribusi positif bagi Bank Syariah Indonesia dalam meningkatkan kualitas layanan aplikasi BSI Mobile, mendukung visi BSI untuk memberikan pelayanan terbaik kepada nasabahnya dan meningkatkan daya saing di era digital.

4. KESIMPULAN

Berdasarkan hasil penelitian, algoritma *Support Vector Machine* (SVM) mencapai tingkat akurasi 77,39% dalam klasifikasi sentimen aplikasi BSI Mobile, melebihi Naïve Bayes yang memiliki akurasi 74,37%. Selain itu, SVM juga unggul dalam nilai *precision*, *recall*, dan *f1-score*, menggambarkan kemampuannya yang lebih andal dalam mengklasifikasikan sentimen netral, negatif, maupun positif. Sementara Naive Bayes lebih sesuai untuk data sederhana dengan asumsi independensi fitur, SVM memiliki keunggulan dalam menangani pola data yang kompleks dan non-linear. Dengan demikian, penerapan SVM dapat membantu Bank Syariah Indonesia (BSI) memahami persepsi nasabah dengan lebih baik, meningkatkan kualitas layanan aplikasi BSI Mobile, dan mendukung inovasi digital untuk meningkatkan daya saing di era modern.

REFERENCES

- [1] I. Dewi Onantya and P. Pandu Adikara, “Analisis Sentimen Pada Ulasan Aplikasi BCA Mobile Menggunakan BM25 Dan Improved K-Nearest Neighbor”, vol.3, no. 3, hlm. 2575-2580, Maret 2019. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [2] F. Gunawan, M. A. Fauzi, and P. P. Adikara, “Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile),” *Syst. Inf. Syst. Inform. J.*, vol. 3, no. 2, pp. 1–6, Dec. 2017, doi: 10.29080/systemic.v3i2.234.
- [3] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes,” *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 4, no. 2, p. 113, Dec. 2018, doi: 10.26418/jp.v4i2.27526.
- [4] D. Pratmanto, R. Rousyati, F. F. Wati, A. E. Widodo, S. Suleman, and R. Wijianto, “App Review Sentiment Analysis Shopee Application In Google Play Store Using Naive Bayes Algorithm,” *J. Phys. Conf. Ser.*, vol. 1641, no. 1, p. 012043, Nov. 2020, doi: 10.1088/1742-6596/1641/1/012043.



- [5] L. B. Ilmawan and M. A. Mude, "Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 154–161, Aug. 2020, doi: 10.33096/ilkom.v12i2.597.154-161.
- [6] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online JD.ID Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, 2019.
- [7] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 6308–6325, 2020, doi: 10.1109/JSTARS.2020.3026724.
- [8] N. Hayatin, G. I. Marthasari, and L. Nuraini, "Optimization of Sentiment Analysis for Indonesian Presidential Election using Naïve Bayes and Particle Swarm Optimization," *J. Online Inform.*, vol. 5, no. 1, pp. 81-88, June 2020, doi: 10.15575/join.v5i1.558.
- [9] N. Herlinawati *et al.*, "Analisis Sentimen Zoom Cloud Meetings Di Play Store Menggunakan Naïve Bayes Dan Support Vector Machine," vol. 5, No.2, Juli 2020, doi: <https://doi.org/10.24114/cess.v5i2.18186>.
- [10] R. Kusumawati, A. D'arofah, and P. A. Pramana, "Comparison Performance of Naive Bayes Classifier and Support Vector Machine Algorithm for Twitter's Classification of Tokopedia Services," *J. Phys. Conf. Ser.*, vol. 1320, no. 1, p. 012016, Oct. 2019, doi: 10.1088/1742-6596/1320/1/012016.
- [11] S. Berliani and S. Lestari, "Analisis Sentimen Masyarakat Terhadap Isu Pecat Sri Mulyani Pada Twitter Menggunakan Metode Naive Bayes Dan Support Vector Machine," *J. Sains Dan Teknol.*, vol. 5, no. 3, pp. 951–960, Apr. 2024, doi: 10.55338/saintek.v5i3.2746.
- [12] S. Linawati, R. A. Safitri, A. R. Alfiyan, W. E. Pangesti, and M. N. Winnarto, "Perbandingan Algoritma Klasifikasi Naïve Bayes Dan Svm Pada Studi Kasus Pemberian Beasiswa PPA," *Swabumi*, vol. 8, no. 1, pp. 71–75, Mar. 2020, doi: 10.31294/swabumi.v8i1.7708.
- [13] S. Pandya and P. Mehta, "A Review On Sentiment Analysis Methodologies, Practices And Applications," *Int. J. Sci. Technol. Res.*, vol. 9, pp. 601-609, 2020.
- [14] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl.-Based Syst.*, vol. 226, p. 107134, Aug. 2021, doi: 10.1016/j.knosys.2021.107134.
- [15] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Soc. Netw. Anal. Min.*, vol. 11, no. 1, p. 81, Dec. 2021, doi: 10.1007/s13278-021-00776-6.
- [16] A. Muhamadin and I. A. Sobari, "Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma SVM Dan NBC," *Reputasi J. Rekayasa Perangkat Lunak*, vol. 2, no. 2, pp. 85–91, Dec. 2021, doi: 10.31294/reputasi.v2i2.785.
- [17] R. A. Raharjo, I. M. G. Sunarya, and D. G. H. Divayana, "Perbandingan Metode Naïve Bayes Classifier dan Support Vector Machine Pada Kasus Analisis Sentimen Terhadap Data Vaksin Covid-19 Di Twitter," *Elkom J. Elektron. Dan Komput.*, vol. 15, no. 2, pp. 456–464, Dec. 2022, doi: 10.51903/elkom.v15i2.918.
- [18] M. A. S. Putra, I. Permana, M. Mustakim, and M. Afdal, "Analisis Sentimen Masyarakat Mengenai Gerakan Childfree di Media Sosial X Menggunakan Algoritma NBC dan SVM: Sentiment Analysis of Childfree Campaign on X Social Media Using NBC and SVM Algorithms," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1189–1198, Jul. 2024, doi: 10.57152/malcom.v4i4.1356.
- [19] J. A. Septian, T. M. Fachrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," *J. Intell. Syst. Comput.*, vol. 1, no. 1, pp. 43–49, Aug. 2019, doi: 10.52985/insyst.v1i1.36.
- [20] O. I. Gifari, Muh. Adha, F. Freddy, and F. F. S. Durrand, "Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, Mar. 2022, doi: 10.46229/jifotech.v2i1.330.
- [21] M. Muhathir, M. H. Santoso, and D. A. Larasati, "Wayang Image Classification Using SVM Method and GLCM Feature Extraction," *J. Inform. Telecommun. Eng.*, vol. 4, no. 2, pp. 373–382, Jan. 2021, doi: 10.31289/jite.v4i2.4524.
- [22] U. Amelia, J. Indra, and A. F. N. Masruriyah, "Implementasi Algoritma Support Vector Machine (Svm) Untuk Prediksi Penyakit Stroke Dengan Atribut Berpengaruh," vol. 3, no. 2, pp. 254-259, Juli 2022.
- [23] A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm," *J. Data Min. Dan Sist. Inf.*, vol. 2, no. 1, p. 31, Feb. 2021, doi: 10.33365/jdmsi.v2i1.1021.
- [24] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 1, p. 157, Jan. 2021, doi: 10.30865/mib.v5i1.2580.
- [25] M. Al Khadafi, Kurnia Paraniha Kartika, and Filda Febrinita, "Penerapan Metode Naïve Bayes Classifier dan Lexicon Based Untuk Analisis Sentimen Cyberbullying Pada Bpjs," *JATI J. Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 725–733, Oct. 2022, doi: 10.36040/jati.v6i2.5633.