

Analisis Sentimen Pengguna GoPay pada Layanan Keuangan Digital dengan Perbandingan Naïve Bayes dan SVM

Dian Ardiansyah^{1*}, Riska Aryanti², Eka Fitriani³, Royadi Royadi⁴

^{1,3} Program Studi Sistem Informasi Kampus Kabupaten Karawang, Fakultas Teknik & Informatika, Universitas Bina Sarana Informatika, Indonesia

² Program Studi Informatika, Fakultas Teknik & Informatika, Universitas Bina Sarana Informatika, Indonesia

⁴ Program Studi Teknologi Informasi, Fakultas Teknik & Informatika, Universitas Bina Sarana Informatika, Indonesia

dian.did@bsi.ac.id^{1*}, riska.rts@bsi.ac.id², eka.ean@bsi.ac.id³, royadi.roo@bsi.ac.id⁴

Artikel Info : Diterima : 10-10-2025 | Direvisi : 20-11-2025 | Disetujui : 01-12-2025

Abstrak

Pesatnya perkembangan layanan keuangan digital menyebabkan meningkatnya penggunaan dompet digital, salah satunya aplikasi GoPay, yang menghasilkan ulasan pengguna dalam jumlah besar. Ulasan tersebut mengandung informasi penting mengenai tingkat kepuasan dan permasalahan layanan, sehingga diperlukan metode otomatis untuk menganalisis sentimen pengguna secara akurat. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi GoPay serta membandingkan kinerja algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) dalam klasifikasi sentimen. Penelitian ini menggunakan dataset sebanyak 132.393 ulasan pengguna GoPay yang diperoleh dari platform Kaggle. Data dilabeli berdasarkan rating pengguna ke dalam tiga kelas sentimen, yaitu positif, netral, dan negatif. Tahapan penelitian meliputi *preprocessing* teks, transformasi fitur menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF), klasifikasi menggunakan algoritma *Naïve Bayes* dan SVM, serta evaluasi kinerja model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil analisis menunjukkan bahwa 79,2% ulasan termasuk sentimen positif, 17,1% sentimen negatif, dan 3,7% sentimen netral. Berdasarkan evaluasi kinerja, algoritma SVM menunjukkan performa yang lebih unggul dengan nilai akurasi 90,65%, *precision* 90,7%, *recall* 90,65%, dan *F1-score* 89,05%, dibandingkan *Naïve Bayes* dengan akurasi 87,89%, *precision* 89,1%, *recall* 87,89%, dan *F1-score* 88,42%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa SVM merupakan metode yang lebih optimal untuk analisis sentimen ulasan GoPay, sedangkan *Naïve Bayes* tetap menjadi alternatif yang efisien dan kompetitif dalam klasifikasi teks berskala besar.

Kata Kunci: Analisis Sentimen, GoPay, *Naïve Bayes*, *Support Vector Machine*, Text Mining

Abstract

The rapid development of digital financial services has led to increased use of digital wallets, one of which is the GoPay application, resulting in a large volume of user reviews. These reviews contain valuable information regarding user satisfaction and service-related issues, making automated methods necessary to accurately analyze user sentiment. This study aims to analyze sentiment in GoPay user reviews and compare the performance of the *Naïve Bayes* and *Support Vector Machine* (SVM) algorithms for sentiment classification. This research uses a dataset of 132,393 GoPay user reviews obtained from the Kaggle platform. The data are labeled based on user ratings into three sentiment classes: positive, neutral, and negative. The research stages include text preprocessing, feature transformation using the *Term Frequency–Inverse Document Frequency* (TF-IDF) method, sentiment classification using the *Naïve Bayes* and SVM algorithms, and model performance evaluation using *accuracy*, *precision*, *recall*, and *F1-score* metrics. The results show that 79.2% of the reviews are classified as positive, 17.1% as negative, and 3.7% as neutral. Based on performance evaluation, the SVM algorithm demonstrates superior results with an accuracy of 90.65%, precision of 90.7%, recall of 90.65%, and *F1-score* of 89.05%, compared to *Naïve Bayes*, which achieves an accuracy of 87.89%, precision of 89.1%, recall of 87.89%, and *F1-score* of 88.42%. These findings indicate that SVM is a more optimal method for sentiment analysis of GoPay user reviews, while *Naïve Bayes* remains an efficient and competitive alternative for large-scale text classification.

Keywords: Sentiment Analysis, GoPay, *Naïve Bayes*, *Support Vector Machine*, Text Mining

PENDAHULUAN

Perkembangan teknologi informasi di Indonesia telah mendorong transformasi masif di sektor ekonomi, khususnya dengan menjamurnya layanan teknologi finansial (*fintech*) yang mengubah perilaku transaksi masyarakat menuju *cashless society*. Akselerasi ini didukung oleh penetrasi internet yang merata serta infrastruktur pembayaran digital yang semakin matang. Dompet digital atau *e-wallet* kini telah berevolusi menjadi instrumen



vital dalam ekosistem pembayaran nasional karena menawarkan efisiensi waktu dan keamanan transaksi yang lebih terjamin dibandingkan uang tunai (Anjeli et al., 2025). Di tengah persaingan industri yang ketat, GoPay berhasil mempertahankan posisinya sebagai salah satu pemimpin pasar dompet digital di Indonesia. Berdasarkan studi komparatif terhadap persepsi pengguna, GoPay bersaing ketat dengan kompetitor utama seperti ShopeePay dan OVO dalam merebut pangsa pasar milenial (Widianto et al., 2024). Keunggulan kompetitif GoPay tidak lepas dari integrasinya dalam ekosistem *super-app* Gojek, yang memungkinkan pengguna mengakses berbagai layanan mulai dari transportasi hingga pembayaran tagihan dalam satu aplikasi terpadu (Indarwati, 2023).

Tingginya aktivitas transaksi harian tersebut secara alami menghasilkan jejak digital berupa ulasan pengguna (*user reviews*) dalam jumlah besar di platform distribusi aplikasi. Ribuan ulasan yang masuk ke Google Play Store setiap harinya menjadi indikator penting untuk mengukur respons publik terhadap pembaruan fitur maupun kendala teknis yang terjadi (Surya et al., 2024). Ulasan-ulasan ini mengandung sentimen positif dan negatif yang merefleksikan kepuasan pelanggan secara jujur dan *real-time* (Riskawati et al., 2024).

Namun, tantangan utama muncul ketika perusahaan harus mengelola lonjakan volume data ulasan tersebut. Metode analisis manual, di mana staf harus membaca dan melabeli setiap komentar satu per satu, menjadi sangat tidak efisien dan memakan waktu lama (Mola et al., 2025). Selain masalah efisiensi, pendekatan manual juga rentan terhadap subjektivitas manusia yang dapat menyebabkan bias dalam menentukan apakah sebuah ulasan bernada kritik atau saran (Nadira et al., 2023).

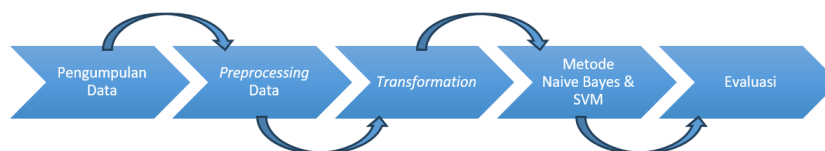
Oleh karena itu, penerapan teknologi *Natural Language Processing* (NLP) menjadi solusi krusial untuk mengotomatisasi proses ekstraksi informasi dari data teks yang tidak terstruktur. Teknik *text mining* memungkinkan pengolahan data ulasan dalam skala besar untuk diklasifikasikan ke dalam kategori sentimen secara cepat dan akurat (Khoiruddin et al., 2023). Salah satu tantangan teknis dalam proses ini adalah menangani ketidakseimbangan data (*imbalanced data*) yang sering ditemukan pada dataset ulasan aplikasi, di mana jumlah ulasan positif seringkali jauh lebih banyak daripada ulasan negatif (Yulianti et al., 2024).

Dalam ranah klasifikasi teks, algoritma *Naïve Bayes* sering dipilih sebagai metode utama karena karakteristiknya yang sederhana namun efektif. Algoritma ini bekerja dengan pendekatan probabilistik yang terbukti memiliki kecepatan komputasi tinggi, sehingga sangat cocok untuk memproses dataset ulasan produk atau layanan yang berjumlah ribuan (Dwilestari & Afifah, 2025). Penelitian terdahulu menunjukkan bahwa *Naïve Bayes* mampu menghasilkan akurasi yang kompetitif dalam mengidentifikasi opini pengguna pada layanan aplikasi berbasis public (Prasetyo et al., 2025).

Sebagai pembanding yang kuat, algoritma *Support Vector Machine* (SVM) menawarkan pendekatan berbasis ruang vektor untuk menemukan batas pemisah (*hyperplane*) terbaik antar kelas. SVM dikenal memiliki keunggulan dalam menangani data berdimensi tinggi dan fitur teks yang kompleks, seperti yang sering ditemukan pada ulasan dengan ragam bahasa campuran (Hidayati et al., 2023). Beberapa studi komparasi menunjukkan bahwa dalam kondisi dataset tertentu, SVM mampu memberikan performa generalisasi yang lebih stabil dibandingkan metode klasifikasi lainnya (Pulungan & Utami, 2025).

Penelitian ini bertujuan untuk melakukan analisis perbandingan kinerja (*head-to-head*) antara *Naïve Bayes* dan SVM secara spesifik pada studi kasus ulasan GoPay terbaru tahun 2025. Komparasi ini penting untuk memvalidasi algoritma mana yang paling optimal diterapkan pada data ulasan *fintech* terkini yang memiliki karakteristik bahasa yang dinamis (Fitri & Damayanti, 2024). Hasil evaluasi dari penelitian ini diharapkan dapat memberikan rekomendasi teknis bagi pengembangan sistem analisis sentimen yang lebih akurat guna mendukung pengambilan keputusan yang strategis.

METODE PENELITIAN



Gambar 1. Tahapan Penelitian

Berikut adalah penjelasan tahapan penemuan pengetahuan dalam basis data:

1. *Pengumpulan Data*: Data yang dipilih adalah ulasan teks pengguna aplikasi GoPay dari Kaggle.com
2. *Preprocessing Data*: Pemeriksaan dan pembersihan data teks dari noise yang tidak relevan.
3. *Transformation*: Mengubah data yang telah dipilih agar sesuai untuk proses data mining.
4. *SMetode*: Melakukan klasifikasi menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) diimplementasikan menggunakan library *Scikit-learn* pada bahasa pemrograman Python.
5. *Evaluasi*: Mengevaluasi hasil klasifikasi untuk menilai relevansi dan akurasi dari model yang dibangun.

Penelitian ini menggunakan dataset yang terdiri dari 132.393 ulasan pengguna aplikasi GoPay yang diambil

dari platform Kaggle. Dataset ini memiliki atribut utama berupa teks ulasan dan label sentimen yang ditentukan berdasarkan rating pengguna, dengan kategori sentimen positif untuk rating 4-5, sentimen netral untuk rating 3 dan sentimen negatif untuk rating 1-2.

Dalam penelitian ini, kami akan melakukan metode klasifikasi menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM).

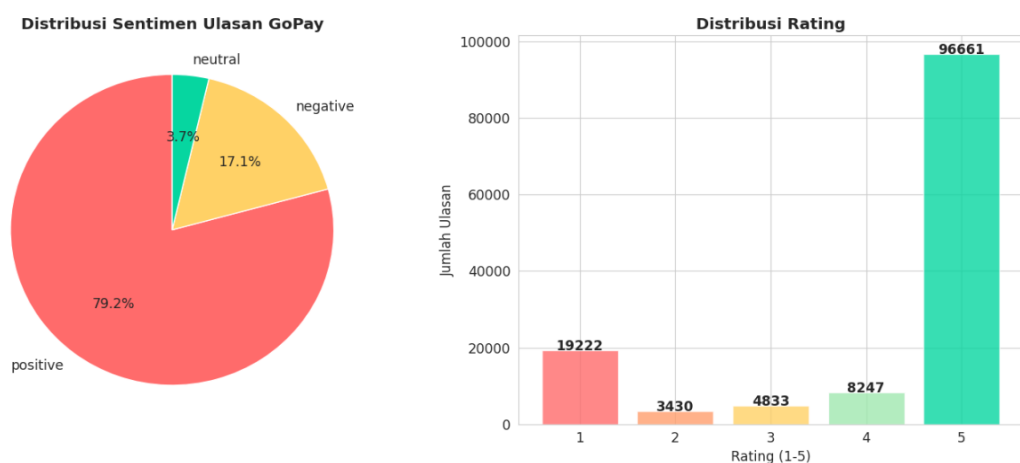
1. **Scikit-learn:** *Scikit-learn* merupakan salah satu library pembelajaran mesin pada Python yang banyak digunakan untuk implementasi algoritma klasifikasi dan evaluasi model, termasuk algoritma *Naïve Bayes* dan SVM (Ariboyo & Fairuz, 2024).
2. **Naïve Bayes:** Algoritma probabilistik yang menghitung probabilitas sebuah dokumen termasuk dalam kelas tertentu berdasarkan kata-kata yang ada di dalamnya (Hajaroh et al., 2024).
3. **Support Vector Machine (SVM):** Algoritma yang menemukan hyperplane pemisah optimal di antara kelas-kelas data untuk memaksimalkan margin (Nurqotimah et al., 2024)
4. **Confusion Matrix:** *Confusion Matrix* adalah alat untuk menilai kinerja model klasifikasi dalam machine learning. Matriks ini disajikan dalam bentuk tabel yang membandingkan prediksi model dengan data actual (Azmi & Voutama, 2024).
5. **TF-IDF (Term Frequency-Inverse Document Frequency):** TF-IDF adalah metode representasi numerik yang digunakan untuk mencerminkan seberapa penting sebuah kata dalam sebuah dokumen dalam kumpulan dokumen (korpus). Metode ini menggabungkan frekuensi kemunculan kata (TF) dengan kebalikan dari frekuensi dokumen yang mengandung kata tersebut (IDF) (Putra et al., 2024).

HASIL DAN PEMBAHASAN

reviewId	userName	userImage	content	score	thumbs	reviewCreated	at	replyContent	repliedAt	appVersion
a8653a31-d892	Pengguna Google	https://play-lh.google	Saya bisa bintang 5 klo iklan g	1	0		19/12/2024 13.26	Hai Kak Sumad	19/12/2024 13.10	
b2cd9f70-083c	Pengguna Google	https://play-lh.google	Good job	5	0	01.38.00	19/12/2024 13.26			01.38.00
b7d8ad28-59e2	Pengguna Google	https://play-lh.google	Mudah dimengeeti	5	0	01.38.00	19/12/2024 13.24			01.38.00
28bbfec0-fccc	Pengguna Google	https://play-lh.google	sangat bagus	5	0	01.39.00	19/12/2024 13.24			01.39.00
ce4aaaf8-b378	Pengguna Google	https://play-lh.google	Tnfo pinjaman	5	0	01.39.00	19/12/2024 13.23			01.39.00
0a168204-9fe6	Pengguna Google	https://play-lh.google	Good aja dah	4	0	01.38.00	19/12/2024 13.18			01.38.00
8e94a0d2-e274	Pengguna Google	https://play-lh.google	saya sudah lama pakai aplikas	5	0	01.38.00	19/12/2024 13.18			01.38.00
73e94cfe-31b3	Pengguna Google	https://play-lh.google	Menjengkelkan iklan promo d	1	0		19/12/2024 13.16	Hai Kak Budi, r	19/12/2024 13.54	
9bb88ddb-a7b0	Pengguna Google	https://play-lh.google	Baik	4	0	01.38.00	19/12/2024 13.12			01.38.00
907f60c4-ce61	Pengguna Google	https://play-lh.google	sangat puas	5	0	01.38.00	19/12/2024 13.11			01.38.00
cda81eba-01fd	Pengguna Google	https://play-lh.google	Karena taralu bayak iklan bari	1	0		19/12/2024 13.11	Hai Kak Fajar, n	19/12/2024 13.09	
f12fdab0-45eb	Pengguna Google	https://play-lh.google	mantap banyak promo nataru	5	0	01.38.00	19/12/2024 13.10			01.38.00
b11b7846-def4	Pengguna Google	https://play-lh.google	banyak promo y gopay... mant	5	0	01.38.00	19/12/2024 13.10			01.38.00
4aec94f4-3250	Pengguna Google	https://play-lh.google	Kenapa gopay pas mau engirir	5	0	01.38.00	19/12/2024 13.09			01.38.00
d88e535b-fb33	Pengguna Google	https://play-lh.google	Tolong jangan sampe ada kort	1	0	01.38.00	19/12/2024 13.09	Hai Kak Dadi, r	19/12/2024 13.52	01.38.00
734ce348-4423	Pengguna Google	https://play-lh.google	pake GoPay murah...diskon b	5	0	01.38.00	19/12/2024 13.08			01.38.00
ff4986f5-a1b9	Pengguna Google	https://play-lh.google	Sangat membatu dan mudah c	5	0	01.39.00	19/12/2024 13.07			01.39.00
b520c940-4385	Pengguna Google	https://play-lh.google	Boleh lah ðŸˆ¸ðŸˆ¸»	5	0		19/12/2024 13.04			
97df70d1-afab	Pengguna Google	https://play-lh.google	trima ksh gopay udh kasih byk	5	0	01.38.00	19/12/2024 13.01			01.38.00
4e9d0da8-14df	Pengguna Google	https://play-lh.google	Karna ada iklan terus kalau lag	1	0		19/12/2024 12.59	Hai Kak Pahri, r	19/12/2024 13.12	
30d408fc-2bf8	Pengguna Google	https://play-lh.google	Terbaik	3	0	01.37.01	19/12/2024 12.59			01.37.01

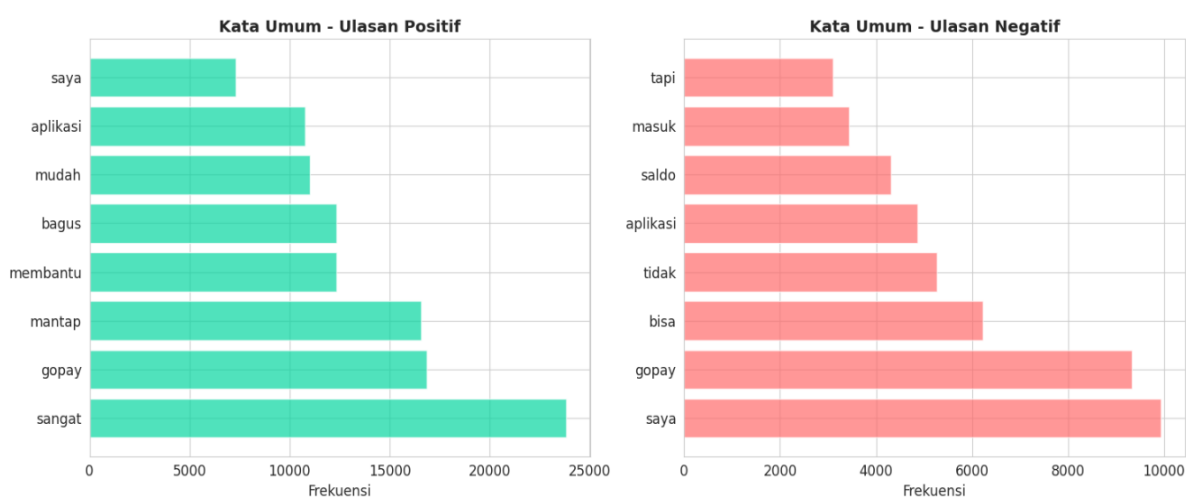
Gambar 2. Tampilan *Dataset*

Pada tahap awal penelitian, pemilihan data melibatkan proses pengambilan ulasan aplikasi GoPay dari platform Kaggle. Data yang diperoleh kemudian disimpan dalam format CSV untuk memudahkan proses analisis lebih lanjut.



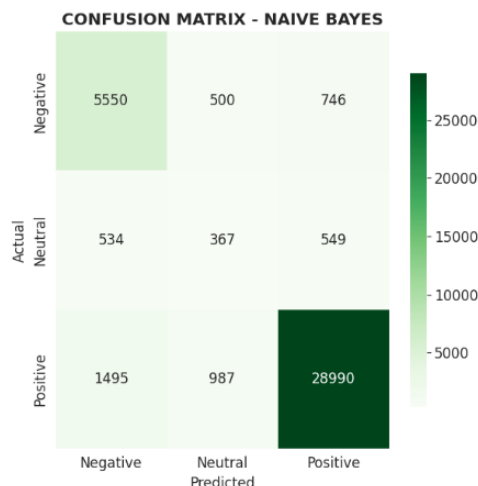
Gambar 3. Tampilan Hasil Distribusi Sentimen Ulasan dan Distribusi Rating

Pada gambar 3 setelah dilakukan *preprocessing* data teks yang meliputi *case folding*, *tokenizing*, *stopword removal*, dan *stemming* untuk membersihkan teks ulasan. Hasil analisis eksploratif menunjukkan distribusi sentimen yang didominasi oleh sentimen positif sebesar 79,2%, diikuti oleh sentimen negatif sebesar 17,1% dan netral sebesar 3,7%, yang diperoleh berdasarkan pelabelan dari rating pengguna. Grafik distribusi rating memperlihatkan bahwa rating 4 dan 5 memiliki jumlah ulasan sebanyak 104.908 data ini dikategorikan sebagai sentimen positif, sedangkan rating rendah (1 dan 2) memiliki jumlah ulasan sebanyak 22.652 data ini dikategorikan sebagai sentimen negatif dan untuk rating 3 dikategorikan sentimen netral memiliki 4.833 data. Hasil ini menunjukkan bahwa sebagian besar pengguna memberikan penilaian positif terhadap layanan GoPay.



Gambar 4. Tampilan Hasil Kata Umum Ulasan Positif dan Negatif

Pada gambar 4 menunjukkan hasil dari proses data transformation pada data teks ulasan pengguna aplikasi GoPay menggunakan pendekatan *Term Frequency-Inverse Document Frequency (TF-IDF)*. Pada tahap ini, teks ulasan yang telah melalui *preprocessing* diubah ke dalam bentuk fitur numerik dengan memberikan bobot pada setiap kata berdasarkan tingkat kemunculan dan tingkat kepentingannya dalam dokumen. Hasil transformasi menunjukkan bahwa pada ulasan positif, kata-kata seperti “sangat”, “gopay”, “mantap”, “membantu”, dan “mudah” memiliki bobot TF-IDF yang tinggi, yang merepresentasikan persepsi positif pengguna terhadap kemudahan dan manfaat layanan. Sebaliknya, pada ulasan negatif, kata-kata seperti “tidak”, “bisa”, “aplikasi”, dan “saldo” muncul dominan dengan bobot yang signifikan, yang mengindikasikan keluhan pengguna terkait kendala teknis dan transaksi. Perbedaan bobot kata antar kelas sentimen ini menunjukkan bahwa proses transformasi menggunakan TF-IDF berhasil menghasilkan fitur yang relevan dan diskriminatif, sehingga dapat digunakan secara efektif sebagai input pada metode klasifikasi *Naïve Bayes* dan *Support Vector Machine (SVM)*.



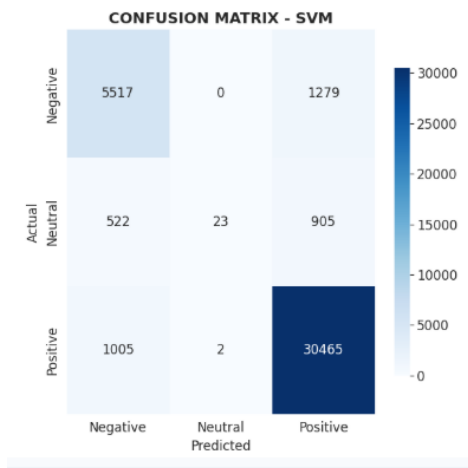
Gambar 5. Confusion Matrix Naïve Bayes

Confusion matrix model Naïve Bayes menunjukkan hasil prediksi klasifikasi sentimen ke dalam tiga kelas, yaitu Negatif, Netral, dan Positif. Matriks ini menggambarkan kinerja model berdasarkan jumlah prediksi yang benar dan salah pada masing-masing kelas, serta diperkuat dengan nilai precision dan recall sebagai indikator ketepatan dan kelengkapan prediksi.

Pada kelas Negatif, model berhasil mengklasifikasikan 5.550 data dengan benar. Namun, masih terdapat 500 data yang salah diprediksi sebagai Netral dan 746 data sebagai Positif. Nilai *precision* yang tinggi pada kelas Negatif menunjukkan bahwa sebagian besar prediksi Negatif yang dihasilkan model memang benar, sedangkan nilai *recall* yang cukup baik menandakan bahwa mayoritas data Negatif aktual berhasil dikenali oleh model. Untuk kelas Netral, Naïve Bayes mampu memprediksi 367 data dengan benar. Akan tetapi, terdapat 534 data Netral yang salah diklasifikasikan sebagai Negatif dan 549 data sebagai Positif. Nilai *recall* pada kelas Netral tergolong rendah, yang menunjukkan bahwa tidak semua data Netral dapat dikenali dengan baik. Meskipun demikian, dibandingkan dengan SVM, Naïve Bayes menunjukkan *precision* dan *recall* yang lebih seimbang pada kelas ini.

Pada kelas Positif, model menunjukkan kinerja yang sangat baik dengan 28.990 data berhasil diklasifikasikan secara benar. Kesalahan prediksi relatif lebih kecil, yaitu 1.495 data salah sebagai Negatif dan 987 data sebagai Netral. Nilai *recall* yang tinggi pada kelas Positif menunjukkan bahwa hampir seluruh data Positif aktual berhasil dikenali, sementara nilai *precision* yang juga tinggi menandakan bahwa prediksi Positif yang dihasilkan model sebagian besar akurat.

Secara keseluruhan, berdasarkan *confusion matrix* serta nilai *precision* dan *recall*, dapat disimpulkan bahwa model Naïve Bayes memiliki performa klasifikasi yang stabil dan relatif seimbang pada ketiga kelas, terutama pada kelas Positif dan Negatif. Meskipun kinerja pada kelas Netral masih perlu ditingkatkan, Naïve Bayes mampu memberikan hasil yang lebih konsisten dibandingkan model lain dalam mengenali distribusi sentimen secara menyeluruh.



Gambar 6. Confusion Matrix Support Vector Machine (SVM)

Confusion matrix model Support Vector Machine (SVM) menunjukkan hasil prediksi klasifikasi sentimen ke

dalam tiga kelas, yaitu Negatif, Netral, dan Positif. Matriks ini memperlihatkan distribusi prediksi benar dan salah yang menjadi dasar evaluasi kinerja model, didukung oleh nilai *precision* dan *recall*. Pada kelas Negatif, model SVM berhasil mengklasifikasikan 5.517 data dengan benar sebagai Negatif. Namun, masih terdapat 1.279 data Negatif yang salah diprediksi sebagai Positif, dan tidak ada data Negatif yang diklasifikasikan sebagai Netral. Nilai *recall* yang tinggi menunjukkan bahwa sebagian besar data Negatif aktual berhasil dikenali, sementara *precision* yang baik menandakan bahwa prediksi Negatif yang dihasilkan relatif akurat.

Untuk kelas Netral, kinerja SVM tergolong sangat rendah. Hanya 23 data Netral yang berhasil diprediksi dengan benar, sedangkan 522 data salah diklasifikasikan sebagai Negatif dan 905 data sebagai Positif. Kondisi ini menyebabkan nilai *recall* kelas Netral sangat rendah, yang menunjukkan bahwa sebagian besar data Netral gagal dikenali oleh model. *Precision* pada kelas ini juga rendah karena sedikitnya prediksi Netral yang benar.

Pada kelas Positif, model SVM menunjukkan performa yang sangat baik. Sebanyak 30.465 data berhasil diklasifikasikan dengan benar sebagai Positif. Kesalahan prediksi relatif kecil, yaitu 1.005 data salah sebagai Negatif dan hanya 2 data sebagai Netral. Nilai *recall* yang sangat tinggi pada kelas Positif menunjukkan bahwa hampir seluruh data Positif aktual berhasil dikenali, sedangkan *precision* yang tinggi menegaskan bahwa prediksi Positif yang dihasilkan model sangat akurat.

Secara keseluruhan, berdasarkan confusion matrix serta nilai *precision* sebesar 0,907 dan *recall* sebesar 0,9065, dapat disimpulkan bahwa model SVM sangat unggul dalam mengklasifikasikan sentimen Positif dan Negatif, namun kurang optimal dalam mengenali sentimen Netral. Ketimpangan ini menunjukkan bahwa meskipun performa global SVM sangat baik, peningkatan masih diperlukan pada kemampuan model dalam membedakan kelas Netral agar hasil klasifikasi menjadi lebih seimbang.

Model	Accuracy	Precision	Recall	F1-Score
SVM	0.9065	0.907	0.9065	0.8905
Naive Bayes	0.8789	0.891	0.8789	0.8842

Gambar 7. Hasil Perbandingan Model

Gambar 7 menunjukkan hasil evaluasi kinerja dua algoritma klasifikasi, yaitu *Support Vector Machine* (SVM) dan *Naive Bayes*, berdasarkan empat metrik utama, yakni *Accuracy*, *Precision*, *Recall*, dan F1-Score. Keempat metrik ini digunakan untuk memberikan gambaran menyeluruh mengenai tingkat akurasi serta keseimbangan performa masing-masing model dalam melakukan klasifikasi. Berdasarkan nilai *Accuracy*, model SVM memperoleh nilai sebesar 0,9065, lebih tinggi dibandingkan dengan model *Naive Bayes* yang mencapai 0,8789. Hasil ini menunjukkan bahwa secara keseluruhan SVM memiliki tingkat ketepatan prediksi yang lebih baik terhadap data uji. Pada metrik *Precision*, SVM kembali menunjukkan kinerja yang unggul dengan nilai 0,907, sedangkan *Naive Bayes* memperoleh nilai 0,891. Hal ini mengindikasikan bahwa prediksi yang dihasilkan oleh SVM cenderung lebih tepat, terutama ketika model memutuskan suatu data termasuk ke dalam kelas tertentu.

Nilai *Recall* pada model SVM sebesar 0,9065, lebih tinggi dibandingkan *Naive Bayes* yang bernilai 0,8789. Ini menunjukkan bahwa SVM memiliki kemampuan yang lebih baik dalam mengenali seluruh data yang seharusnya termasuk ke dalam kelas target, sehingga risiko kehilangan data penting (*false negative*) relatif lebih rendah. Pada metrik F1-Score, yang merepresentasikan keseimbangan antara *precision* dan *recall*, SVM kembali unggul dengan nilai 0,8905, sementara *Naive Bayes* memperoleh 0,8842. Meskipun perbedaannya tidak terlalu besar, hasil ini menegaskan bahwa SVM memiliki performa yang lebih stabil dalam menangani klasifikasi data. Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa model SVM memiliki kinerja yang lebih baik dibandingkan *Naive Bayes* pada seluruh metrik evaluasi yang digunakan. Namun demikian, nilai yang diperoleh oleh *Naive Bayes* juga tergolong tinggi, sehingga model ini tetap dapat dianggap andal dan kompetitif, khususnya pada kasus klasifikasi dengan karakteristik data yang sesuai dengan asumsi *Naive Bayes*.

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi GoPay pada layanan keuangan digital serta membandingkan kinerja algoritma *Naive Bayes* dan *Support Vector Machine* (SVM). Berdasarkan dataset sebanyak 132.393 ulasan pengguna, hasil analisis distribusi sentimen menunjukkan bahwa 79,2% ulasan tergolong sentimen positif, 17,1% sentimen negatif, dan 3,7% sentimen netral. Temuan ini mengindikasikan bahwa mayoritas pengguna memiliki persepsi positif terhadap layanan GoPay.

Hasil evaluasi kinerja model menunjukkan bahwa algoritma SVM memberikan performa yang lebih unggul dibandingkan *Naive Bayes* pada seluruh metrik evaluasi. SVM memperoleh nilai akurasi sebesar 90,65%, *precision* 90,7%, *recall* 90,65%, dan F1-score 89,05%, sedangkan *Naive Bayes* menghasilkan akurasi 87,89%, *precision* 89,1%, *recall* 87,89%, dan F1-score 88,42%. Hal ini menunjukkan bahwa SVM memiliki kemampuan klasifikasi yang lebih akurat dan konsisten dalam mengidentifikasi sentimen pengguna.

Meskipun demikian, kedua algoritma masih menunjukkan keterbatasan dalam mengklasifikasikan sentimen netral akibat ketidakseimbangan distribusi data. Secara keseluruhan, penelitian ini menyimpulkan bahwa SVM merupakan metode yang lebih optimal untuk analisis sentimen ulasan GoPay, sementara *Naïve Bayes* tetap dapat digunakan sebagai alternatif yang efisien dan kompetitif, khususnya pada data teks berskala besar.

REFERENSI

- Anjeli, R., Putri, D. C. S., Perengki, M., & Soleh, E. (2025). *PENGGUNAAN CASHLESS DI LINGKUNGAN MAHASISWA*. Penerbit Widina.
- Aribowo, B., & Fairuz, S. (2024). *Panduan praktis Machine Learning klasifikasi menggunakan Python: Diandra Kreatif*. Diandra Kreatif.
- Azmi, A. F., & Voutama, A. (2024). Prediksi Churn Nasabah Bank Menggunakan Klasifikasi Random Forest Dan Decision Tree Dengan Evaluasi Confusion Matrix. *Komputa: Jurnal Ilmiah Komputer Dan Informatika*, 13(1), 111–119.
- Dwilestari, G., & Afifah, T. A. (2025). Perbandingan Kinerja Algoritma Naive Bayes Dan Decision Tree Dalam Klasifikasi Kanker Paru-Paru. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1), 801–807.
- Fitri, D. A., & Damayanti, D. (2024). Komparasi algoritma random forest classifier dan support vector machine untuk sentimen masyarakat terhadap pinjaman online di media sosial. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(4), 2018–2029.
- Hajarah, H., Suprapti, T., & Narasati, R. (2024). Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Ulasan Produk Makanan Dan Minuman Di Tokopedia. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 111–118.
- Hidayati, N., Hamami, F., & Fa'rifah, R. Y. (2023). Aspect-Based Sentiment Analysis On FLIP Application Reviews (Play Store) Using Support Vector Machine (SVM) Algorithm. *Journal of Informatics and Telecommunication Engineering*, 7(1), 183–197.
- Indarwati, K. D. (2023). Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 10(1).
- Khoiruddin, Y., Fauzi, A., & Siregar, A. M. (2023). Analisis Sentimen Gojek Indonesia Pada Twitter Menggunakan Algoritme Naïve Bayes Dan Support Vector Machine. *Progresif: Jurnal Ilmiah Komputer*, 19(1), 391–400.
- Mola, S. A. S., Djawa, S. N. R., & Mauko, A. Y. (2025). *Text Mining: Analisis Sentimen dengan Naïve Bayes*. Kaizen Media Publishing.
- Nadira, A., Setiawan, N. Y., & Purnomo, W. (2023). Analisis Sentimen Pada Ulasan Aplikasi Mobile Banking Menggunakan Metode Naïve Bayes Dengan Kamus Inset. *Indexia*, 5(01), 35–47.
- Nurqotimah, D. R., Khudori, A. N., & Pradini, R. S. (2024). Implementasi Algoritma Support Vector Machine (SVM) Untuk Klasifikasi Penyakit Stroke. *Journal of Applied Computer Science and Technology*, 5(2), 179–185.
- Prasetyo, Y. A., Nastiti, F. E., & Nurohman, N. (2025). PENERAPAN MULTINOMIAL NAIVE BAYES UNTUK ANALISIS SENTIMEN ASPEK PADA ULASAN MALL DI KOTA SOLO. *Jurnal Informatika Teknologi Dan Sains (Jinteks)*, 7(3), 1093–1102.
- Pulungan, L. N. T., & Utami, E. (2025). PERBANDINGAN ALGORITMA SVM DAN RANDOM FOREST PADA KLASIFIKASI KESEHATAH MENTAL BERDASARKAN SLEEP DISORDERS. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 10(3), 2827–2837.
- Putra, R. R., Putri, N. A., & Putra, A. D. (2024). *Teknik Cosine Similarity dan TF-IDF dalam analisis data*. Serasi Media Teknologi.
- Riskawati, R., Fatihanursari, F., Iin, I., & Rinaldi, A. R. (2024). Penerapan Metode Naïve Bayes Classifier Pada Analisis Sentimen Aplikasi Gopay. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 346–353.
- Surya, M. R. A., Martanto, M., & Hayati, U. (2024). Analisis Sentimen Ulasan Pengguna Ovo Menggunakan Algoritma Naive Bayes Pada Google Play Store. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 2780–2786.
- Widianto, I. S., Ramadhan, Y. R., & Komara, M. A. (2024). ANALISIS SENTIMEN E-WALLET GOPAY, SHOPEEPAY, DAN OVO MENGGUNAKAN ALGORITMA NAIVE BAYES. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3S1).
- Yulianti, I., Ramdhani, L. S., & Koeswara, T. S. N. (2024). Analisis Sentimen Ulasan Aplikasi Gopay Menggunakan Naive Bayes dengan Teknik Oversampling. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 9(2).