



Analisis Sentimen Ulasan Pengguna Aplikasi Pospay Biru Menggunakan Kombinasi TF-IDF dan Support Vector Machine

Ade Mulyadi

¹Program Studi Diploma IV Teknik Informatika, Fakultas Sekolah Vokasi, Universitas Logistik dan Bisnis Internasional, Bandung, Indonesia

Email: adegointer@gmail.com

Abstrak

Transformasi layanan keuangan digital mendorong meningkatnya penggunaan aplikasi **Pospay Biru** sebagai media transaksi elektronik milik PT Pos Indonesia. Seiring bertambahnya jumlah pengguna, ulasan pada Google Play Store menjadi sumber informasi penting untuk mengevaluasi kualitas layanan dan tingkat kepuasan pengguna. Namun, tingginya volume ulasan menyebabkan analisis secara manual menjadi tidak efisien dan rentan terhadap subjektivitas. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna Pospay Biru menggunakan kombinasi Term Frequency-Inverse Document Frequency (TF-IDF) sebagai metode ekstraksi fitur dan Support Vector Machine (SVM) sebagai algoritma klasifikasi. Data penelitian diperoleh melalui proses web scraping Google Play Store yang menghasilkan 25.391 ulasan, kemudian melalui tahapan seleksi dan preprocessing sehingga diperoleh 5.971 ulasan yang layak dianalisis. Tahapan preprocessing meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming. Selanjutnya, setiap ulasan direpresentasikan dalam bentuk vektor TF-IDF dan diklasifikasikan menggunakan SVM dengan linear kernel. Evaluasi model dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model menghasilkan accuracy sebesar **89,13%**, precision **83,93%**, recall **84,62%**, dan F1-score **84,27%**. Temuan ini menunjukkan bahwa kombinasi TF-IDF dan SVM mampu mengklasifikasikan sentimen ulasan pengguna secara efektif serta dapat dimanfaatkan sebagai pendekatan otomatis untuk mendukung evaluasi kualitas layanan aplikasi Pospay Biru berdasarkan umpan balik pengguna.

Kata Kunci: Analisis Sentimen, Pospay Biru, Text Mining, TF-IDF, Support Vector Machine.

Abstract

The rapid growth of digital financial services has increased the use of **Pospay Biru**, a digital payment application developed by PT Pos Indonesia. As the number of users continues to grow, reviews posted on Google Play Store provide valuable information for evaluating service quality and user satisfaction. However, the large volume of reviews makes manual analysis inefficient and prone to subjective interpretation. This study aims to analyze user sentiment toward the Pospay Biru application by combining **Term Frequency-Inverse Document Frequency (TF-IDF)** for feature extraction and **Support Vector Machine (SVM)** for sentiment classification. The dataset was collected through Google Play Store web scraping, producing 25,391 raw reviews. After data selection and preprocessing, 5,971 valid reviews were retained for analysis. The preprocessing stage consisted of cleaning, case folding, tokenization, stopword removal, and stemming. The processed reviews were transformed into TF-IDF feature vectors and classified using a linear-kernel SVM model. Model performance was evaluated using a confusion matrix, accuracy, precision, recall, and F1-score. The experimental results achieved an **accuracy of 89.13%**, **precision of 83.93%**, **recall of 84.62%**, and an **F1-score of 84.27%**. These findings demonstrate that the combination of TF-IDF and SVM provides effective sentiment classification for Indonesian-language application reviews and can serve as an automated approach to support service quality evaluation and decision-making for improving the Pospay Biru application.

Keywords: Sentiment Analysis, Pospay Biru, Text Mining, TF-IDF, Support Vector Machine.

PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi digital pada berbagai sektor, termasuk layanan keuangan dan logistik. Pemanfaatan aplikasi berbasis *mobile* menjadi salah satu strategi utama dalam meningkatkan efisiensi layanan, memperluas akses transaksi, dan meningkatkan pengalaman pengguna. Di Indonesia, peningkatan penetrasi internet dan penggunaan *smartphone* telah mengubah pola interaksi masyarakat terhadap layanan keuangan, sehingga aplikasi pembayaran digital, transfer dana, dan layanan transaksi elektronik semakin banyak digunakan dalam aktivitas sehari-hari.

PT Pos Indonesia (Persero) sebagai salah satu badan usaha milik negara turut melakukan transformasi digital melalui pengembangan aplikasi **Pospay Biru**. Aplikasi ini menyediakan berbagai layanan keuangan digital, seperti pembayaran tagihan, pembelian pulsa, transfer dana, pembayaran pajak, dan transaksi keuangan lainnya. Sebagai aplikasi yang tersedia pada Google Play Store, Pospay Biru menerima ribuan ulasan pengguna yang mencerminkan pengalaman, kepuasan, keluhan, maupun saran terhadap kualitas layanan aplikasi. Ulasan tersebut merupakan bentuk *user-generated content* yang memiliki nilai penting sebagai sumber evaluasi bagi pengembang aplikasi.

Analisis terhadap ulasan pengguna menjadi penting karena mampu memberikan gambaran mengenai persepsi pengguna terhadap kualitas layanan, stabilitas sistem, kemudahan penggunaan, serta keandalan fitur yang tersedia. Informasi tersebut dapat digunakan sebagai dasar pengambilan keputusan dalam pengembangan aplikasi dan peningkatan kualitas layanan. Namun, jumlah ulasan yang terus bertambah menyebabkan analisis secara manual menjadi tidak efisien, memerlukan waktu yang lama, dan berpotensi menghasilkan penilaian yang bersifat subjektif. Selain itu, ulasan pengguna umumnya berbentuk teks tidak terstruktur (*unstructured text*) yang mengandung bahasa tidak baku, singkatan, kesalahan penulisan, serta variasi ekspresi, sehingga memerlukan teknik pengolahan teks secara otomatis.

Salah satu pendekatan yang banyak digunakan untuk mengolah data teks adalah *text mining*, sedangkan **analisis sentimen** merupakan salah satu implementasi utama dari *text mining* yang bertujuan mengidentifikasi kecenderungan opini pengguna terhadap suatu produk atau layanan. Analisis sentimen memungkinkan proses klasifikasi ribuan ulasan ke dalam kategori sentimen positif, negatif, atau netral secara otomatis sehingga menghasilkan informasi yang lebih objektif dan terukur. Dalam konteks evaluasi aplikasi digital, analisis sentimen dapat membantu pengembang mengidentifikasi aspek layanan yang paling banyak memperoleh apresiasi maupun keluhan dari pengguna.

Agar data teks dapat diproses oleh algoritma pembelajaran mesin, diperlukan proses ekstraksi

fitur yang mengubah dokumen teks menjadi representasi numerik. Salah satu metode yang paling banyak digunakan adalah *Term Frequency–Inverse Document Frequency* (TF-IDF) karena mampu memberikan bobot berdasarkan tingkat kepentingan kata dalam kumpulan dokumen. Representasi numerik yang dihasilkan kemudian dapat digunakan sebagai masukan bagi algoritma klasifikasi. Di antara berbagai algoritma klasifikasi, *Support Vector Machine* (SVM) dikenal memiliki performa yang baik pada data teks berdimensi tinggi dan bersifat *sparse*, sehingga banyak digunakan dalam penelitian analisis sentimen.

Berbagai penelitian sebelumnya menunjukkan bahwa kombinasi TF-IDF dan SVM mampu menghasilkan performa klasifikasi yang tinggi pada analisis sentimen ulasan aplikasi digital. Aristawidya *et al.* (2026) melaporkan bahwa SVM memberikan akurasi yang tinggi pada analisis sentimen aplikasi PosPay. Lubis dan Setyawan (2024) menunjukkan bahwa SVM memiliki performa yang lebih baik dibandingkan Naïve Bayes pada klasifikasi sentimen ulasan aplikasi. Penelitian lain pada aplikasi DANA, *mobile banking*, dan layanan digital juga menunjukkan bahwa TF-IDF efektif dalam merepresentasikan fitur teks sehingga meningkatkan kemampuan klasifikasi model.

Meskipun demikian, masih terdapat beberapa keterbatasan pada penelitian terdahulu. Pertama, sebagian besar penelitian berfokus pada aplikasi *e-commerce*, *mobile banking*, dan dompet digital, sedangkan penelitian mengenai **Pospay Biru** masih sangat terbatas. Kedua, beberapa penelitian menggunakan jumlah dataset yang relatif kecil sehingga belum sepenuhnya merepresentasikan variasi opini pengguna. Ketiga, sebagian penelitian hanya berfokus pada pembangunan model klasifikasi tanpa menyajikan analisis yang lebih komprehensif mengenai karakteristik data, hasil *preprocessing*, visualisasi kata dominan, serta interpretasi distribusi sentimen pengguna.

Berdasarkan kondisi tersebut, penelitian ini memiliki *research gap* pada aspek objek penelitian, skala dataset, dan kedalaman analisis. Penelitian ini menggunakan dataset hasil *web scraping* Google Play Store dalam jumlah yang lebih besar, yaitu **25.391 ulasan**, yang setelah proses seleksi dan *preprocessing* menghasilkan **5.971 ulasan valid**. Selain membangun model klasifikasi menggunakan kombinasi TF-IDF dan SVM, penelitian ini juga menyajikan tahapan *preprocessing*, visualisasi karakteristik kata menggunakan *WordCloud*, serta evaluasi model menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan F1-score.

Kontribusi utama penelitian ini adalah penerapan kombinasi TF-IDF dan *Support Vector Machine* (SVM) pada analisis sentimen ulasan pengguna **Pospay Biru** dengan memanfaatkan dataset yang relatif besar serta evaluasi model yang komprehensif. Penelitian ini tidak hanya menghasilkan model klasifikasi sentimen, tetapi juga memberikan gambaran mengenai topik-topik yang paling banyak dibahas oleh pengguna sehingga dapat digunakan sebagai dasar evaluasi kualitas layanan

aplikasi. Dengan demikian, penelitian ini memberikan kontribusi akademis pada pengembangan analisis sentimen berbahasa Indonesia sekaligus memberikan kontribusi praktis bagi PT Pos Indonesia dalam memahami persepsi pengguna terhadap layanan digital yang disediakan.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk **menganalisis sentimen ulasan pengguna aplikasi Pospay Biru menggunakan kombinasi TF-IDF sebagai metode ekstraksi fitur dan Support Vector Machine (SVM) sebagai algoritma klasifikasi**. Secara khusus, penelitian ini bertujuan mengukur performa model menggunakan *accuracy*, *precision*, *recall*, dan F1-score, serta mengidentifikasi kecenderungan sentimen pengguna terhadap aplikasi Pospay Biru berdasarkan data ulasan pada Google Play Store. Hasil penelitian diharapkan dapat menjadi referensi dalam pengembangan metode analisis sentimen serta mendukung peningkatan kualitas layanan aplikasi Pospay Biru berdasarkan umpan balik pengguna.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan **kuantitatif** dengan metode *text mining* untuk mengklasifikasikan sentimen ulasan pengguna aplikasi Pospay Biru. Tahapan penelitian meliputi pengumpulan data, *preprocessing* teks, pelabelan sentimen, ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), klasifikasi menggunakan *Support Vector Machine* (SVM), serta evaluasi performa model. Seluruh proses dilakukan menggunakan bahasa pemrograman Python pada lingkungan Google Colaboratory.

2.1 Pengumpulan Data

Data penelitian diperoleh dari ulasan pengguna aplikasi **Pospay Biru** pada Google Play Store. Proses pengumpulan data dilakukan menggunakan teknik *web scraping* dengan memanfaatkan pustaka [google-play-scraper](#). Teknik ini memungkinkan pengambilan data secara otomatis sehingga diperoleh *dataset* dalam jumlah besar yang merepresentasikan opini pengguna terhadap aplikasi.

Hasil proses *scraping* menghasilkan **25.391 ulasan** yang terdiri atas isi komentar, *rating*, tanggal ulasan, dan atribut pendukung lainnya. Data yang diperoleh masih berupa data mentah (*raw data*) sehingga memerlukan proses seleksi dan pembersihan sebelum digunakan pada tahap analisis.

2.2 Preprocessing Teks

Tahap *preprocessing* bertujuan meningkatkan kualitas data dengan menghilangkan *noise* dan menyeragamkan bentuk teks. Proses ini dilakukan melalui beberapa tahapan sebagai berikut.

1. *Cleaning*, yaitu menghapus URL, angka, *emoji*, tanda baca, dan karakter khusus.

2. *Case folding*, yaitu mengubah seluruh huruf menjadi huruf kecil (*lowercase*).
3. *Tokenizing*, yaitu memisahkan kalimat menjadi kumpulan kata (*token*).
4. *Stopword removal*, yaitu menghapus kata-kata umum yang tidak memiliki kontribusi signifikan terhadap proses klasifikasi.
5. *Stemming*, yaitu mengubah kata menjadi bentuk dasar menggunakan algoritma *stemming* bahasa Indonesia.

Setelah seluruh tahapan *preprocessing* dilakukan, jumlah data berkurang dari **25.391 ulasan menjadi 5.971 ulasan valid** yang digunakan sebagai *dataset* penelitian. Pengurangan jumlah data disebabkan oleh penghapusan data duplikat, komentar kosong, serta teks yang tidak relevan.

2.3 Pelabelan Sentimen

Dataset yang telah melalui tahap *preprocessing* selanjutnya diberikan label sentimen sebagai **positif, negatif, dan netral**. Pelabelan dilakukan untuk menghasilkan *dataset* berlabel (*labeled dataset*) yang digunakan pada pendekatan *supervised learning*. *Dataset* berlabel menjadi dasar pembelajaran algoritma *Support Vector Machine* (SVM) dalam mengenali pola sentimen pada data ulasan.

2.4 Ekstraksi Fitur Menggunakan TF-IDF

Data teks hasil *preprocessing* diubah menjadi representasi numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen dan tingkat kemunculannya pada seluruh kumpulan dokumen.

Pembobotan TF-IDF dihitung menggunakan Persamaan (1).

$$\text{TFIDF}(t,d)=\text{TF}(t,d)\times\text{IDF}(t)$$

dengan $\text{TF}(t,d)$ menyatakan frekuensi kemunculan kata t pada dokumen d , sedangkan $\text{IDF}(t)$ dihitung berdasarkan Persamaan (2).

$$\text{IDF}(t)=\log(N/\text{df}(t))$$

di mana N adalah jumlah seluruh dokumen dan $\text{df}(t)$ adalah jumlah dokumen yang mengandung kata t .

Implementasi TF-IDF dilakukan menggunakan kelas `TfidfVectorizer` dari pustaka *Scikit-learn*. Hasil proses ini berupa matriks fitur berdimensi tinggi yang merepresentasikan setiap ulasan sebagai vektor numerik.

2.5 Pembagian Dataset

Dataset hasil ekstraksi fitur dibagi menjadi **data latih (*training set*) sebesar 80%** dan **data uji (*testing set*) sebesar 20%** menggunakan fungsi `train_test_split()` dari *Scikit-learn*. Pembagian dilakukan dengan parameter `random_state = 0` untuk menjamin konsistensi hasil eksperimen.

Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur kemampuan model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya

2.6 Klasifikasi Menggunakan Support Vector Machine

Proses klasifikasi dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan *linear kernel*. Pemilihan *linear kernel* didasarkan pada karakteristik data hasil TF-IDF yang memiliki dimensi tinggi dan bersifat *sparse*, sehingga *linear kernel* lebih efisien secara komputasi dan mampu memberikan performa klasifikasi yang baik.

SVM bekerja dengan mencari *hyperplane* optimal yang memisahkan data ke dalam kelas sentimen positif, negatif, dan netral dengan *margin* maksimum. Implementasi model dilakukan menggunakan kelas `LinearSVC` dari pustaka *Scikit-learn*.

2.7 Evaluasi Model

Performa model dievaluasi menggunakan *confusion matrix* serta empat metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan F1-score. *Accuracy* digunakan untuk mengukur proporsi prediksi yang benar terhadap seluruh data uji. *Precision* mengukur ketepatan prediksi pada setiap kelas, sedangkan *recall* mengukur kemampuan model dalam menemukan seluruh data yang benar pada masing-masing kelas. F1-score digunakan untuk mengevaluasi keseimbangan antara *precision* dan *recall*.

Nilai *accuracy* dihitung menggunakan Persamaan (3).

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

Precision dihitung menggunakan Persamaan (4).

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Recall dihitung menggunakan Persamaan (5).

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

Sedangkan F1-score dihitung menggunakan Persamaan (6).

$$F1=(2\times\text{Precision}\times\text{Recall})/(\text{Precision}+\text{Recall})$$

Keempat metrik tersebut memberikan evaluasi yang komprehensif terhadap kemampuan model dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Pospay Biru.

2.8 Alur Penelitian

Alur penelitian dimulai dari proses *web scraping* ulasan Google Play Store, dilanjutkan dengan *preprocessing* teks, pelabelan sentimen, ekstraksi fitur menggunakan TF-IDF, pembagian *dataset*, pelatihan model SVM, dan evaluasi performa model. Tahapan tersebut dirancang secara sistematis untuk menghasilkan model analisis sentimen yang dapat direplikasi dan digunakan sebagai pendekatan otomatis dalam mengevaluasi persepsi pengguna terhadap aplikasi Pospay Biru.

HASIL DAN PEMBAHASAN

3.1 Karakteristik Dataset

Dataset penelitian diperoleh melalui proses *web scraping* terhadap ulasan pengguna aplikasi Pospay Biru pada Google Play Store. Proses pengambilan data dilakukan secara otomatis menggunakan pustaka [google-play-scraper](#), sehingga diperoleh **25.391 ulasan** sebagai data mentah. Data tersebut masih mengandung data duplikat, komentar kosong, karakter khusus, *emoji*, URL, dan teks yang tidak relevan, sehingga diperlukan proses *preprocessing* sebelum digunakan pada tahap klasifikasi.

Setelah dilakukan tahapan *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*, jumlah data berkurang menjadi **5.971 ulasan valid**. Dengan demikian, sekitar **23,52%** dari total data hasil *scraping* memenuhi kriteria penelitian dan digunakan sebagai *dataset* utama. Pengurangan jumlah data menunjukkan bahwa proses *preprocessing* berperan penting dalam meningkatkan kualitas *dataset* dengan menghilangkan *noise* yang dapat memengaruhi performa model klasifikasi.

Ringkasan jumlah data hasil proses *web scraping* dan *preprocessing* disajikan pada Tabel 1.

Tabel 1. Ringkasan *dataset* penelitian.

Tahapan	Jumlah ulasan
Data mentah hasil <i>scraping</i>	25.391

Data valid setelah <i>preprocessing</i>	5.971
---	-------

Dataset yang telah dibersihkan kemudian digunakan pada proses ekstraksi fitur menggunakan TF-IDF dan klasifikasi sentimen menggunakan *Support Vector Machine* (SVM).

3.2 Visualisasi Kata Dominan

Visualisasi karakteristik data dilakukan menggunakan *WordCloud* untuk mengidentifikasi kata-kata yang paling sering muncul dalam kumpulan ulasan pengguna. *WordCloud* dibangun berdasarkan hasil *preprocessing*, sehingga kata yang divisualisasikan telah berada dalam bentuk dasar (*root words*).

Hasil visualisasi menunjukkan bahwa kata “bisa”, “membantu”, “aplikasi”, “tidak”, “saldo”, dan “transfer” memiliki frekuensi kemunculan yang tinggi. Dominasi kata *transfer* dan *saldo* menunjukkan bahwa sebagian besar ulasan berkaitan dengan fungsi utama Pospay Biru sebagai aplikasi layanan keuangan digital. Sementara itu, kata *tidak* sering muncul pada ulasan yang mengandung keluhan, seperti *tidak bisa transfer* atau *tidak dapat login*, sedangkan kata *membantu* cenderung muncul pada ulasan yang bernada positif.

Temuan ini mengindikasikan bahwa isu utama yang paling banyak dibahas pengguna berkaitan dengan stabilitas transaksi, pengelolaan saldo, akses aplikasi, dan keandalan layanan. Meskipun *WordCloud* tidak dapat digunakan sebagai penentu sentimen secara langsung, visualisasi ini memberikan gambaran awal mengenai topik yang paling dominan dalam *dataset*.

3.3 Hasil *Preprocessing* Teks

Proses *preprocessing* bertujuan menyeragamkan bentuk teks sehingga representasi fitur yang dihasilkan menjadi lebih konsisten. Transformasi teks dilakukan secara bertahap mulai dari *cleaning* hingga *stemming*.

Tabel 2. Contoh hasil *preprocessing*.

Tahapan	Sampel 1	Sampel 2
Data asli	Scan barcode lama!!! 😞	Transfer gagal 100%
<i>Cleaning</i>	scan barcode lama	transfer gagal
<i>Case folding</i>	scan barcode lama	transfer gagal
<i>Tokenizing</i>	[scan, barcode, lama]	[transfer, gagal]

Stopword removal	[scan, barcode, lama]	[transfer, gagal]
Stemming	[scan, barcode, lama]	[transfer, gagal]

Berdasarkan Tabel 2, proses *cleaning* berhasil menghilangkan karakter yang tidak memiliki nilai informasi, sedangkan *case folding* menyeragamkan penulisan kata. Tahap *stemming* menghasilkan bentuk dasar kata sehingga variasi kata yang memiliki makna sama dapat direpresentasikan secara lebih konsisten. Kondisi ini penting untuk mengurangi ukuran kosakata (*vocabulary*) dan meningkatkan efektivitas pembobotan TF-IDF.

3.4 Ekstraksi Fitur Menggunakan TF-IDF

Setelah *preprocessing* selesai dilakukan, setiap ulasan diubah menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini memberikan bobot yang lebih tinggi pada kata yang sering muncul dalam suatu dokumen tetapi jarang muncul pada dokumen lain, sehingga kata yang lebih informatif memperoleh nilai yang lebih besar.

Tabel 3. Contoh bobot TF-IDF.

Term	Dokumen 1	Dokumen 2	Dokumen 3
aplikasi	0,214	0	0,176
transfer	0,543	0,318	0
saldo	0	0,427	0,311
gagal	0	0,651	0

Hasil pembobotan menunjukkan bahwa kata *gagal* memiliki bobot yang tinggi pada dokumen tertentu, sedangkan kata *aplikasi* memiliki bobot yang relatif lebih rendah karena muncul pada banyak dokumen. Hal ini menunjukkan bahwa TF-IDF mampu menonjolkan kata-kata yang lebih representatif terhadap isi suatu ulasan.

Matriks TF-IDF yang dihasilkan kemudian digunakan sebagai masukan bagi algoritma *Support Vector Machine* (SVM). Representasi fitur ini sangat sesuai untuk data teks karena menghasilkan matriks berdimensi tinggi dan bersifat *sparse*.

3.5 Evaluasi Performa Model SVM

Model klasifikasi dibangun menggunakan *Support Vector Machine* (SVM) dengan linear kernel. Dataset dibagi menjadi **80% data latih** dan **20% data uji**, kemudian model dievaluasi

menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*.

Tabel 4. Hasil evaluasi model.

Metrik	Nilai
<i>Accuracy</i>	89,13%
<i>Precision</i>	83,93%
<i>Recall</i>	84,62%
F1-score	84,27%

Nilai *accuracy* sebesar **89,13%** menunjukkan bahwa model mampu mengklasifikasikan sekitar sembilan dari sepuluh ulasan ke dalam kelas sentimen yang benar. Sementara itu, *precision* sebesar **83,93%** menunjukkan bahwa sebagian besar prediksi yang dihasilkan model merupakan prediksi yang tepat.

Nilai *recall* sebesar **84,62%** mengindikasikan bahwa model mampu mengenali sebagian besar data yang benar-benar termasuk ke dalam masing-masing kelas sentimen. Adapun **F1-score sebesar 84,27%** menunjukkan adanya keseimbangan yang baik antara *precision* dan *recall*, sehingga model memiliki performa yang relatif stabil.

Berdasarkan *confusion matrix*, sebagian besar data uji berada pada diagonal utama matriks, yang menunjukkan bahwa prediksi model sesuai dengan label sebenarnya. Meskipun demikian, masih terdapat sejumlah kecil misklasifikasi, terutama pada ulasan yang memiliki karakteristik bahasa yang saling berdekatan.

3.6 Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi **TF-IDF dan Support Vector Machine (SVM)** mampu menghasilkan performa klasifikasi yang baik pada analisis sentimen ulasan pengguna aplikasi Pospay Biru. Performa tersebut dipengaruhi oleh dua faktor utama, yaitu kualitas representasi fitur dan kemampuan algoritma klasifikasi.

Metode TF-IDF berhasil memberikan representasi numerik yang efektif dengan menonjolkan kata-kata yang memiliki kontribusi besar terhadap pembentukan sentimen. Kata-kata seperti *gagal*, *transfer*, dan *saldo* memperoleh bobot yang lebih tinggi karena lebih mampu membedakan karakteristik antarulasan. Representasi ini membantu SVM dalam membangun *hyperplane* yang optimal untuk memisahkan kelas sentimen.

Penggunaan *linear kernel* juga memberikan keuntungan karena data hasil TF-IDF memiliki dimensi tinggi dan bersifat *sparse*. Pada kondisi tersebut, *linear kernel* umumnya lebih efisien secara komputasi dibandingkan kernel nonlinier serta memiliki kemampuan generalisasi yang baik. Hasil penelitian ini sejalan dengan penelitian Pratiwi dan Nugroho (2024) yang menunjukkan bahwa TF-IDF meningkatkan kualitas representasi fitur, serta penelitian Lubis dan Setyawan (2024) yang menyatakan bahwa SVM memiliki performa yang lebih baik dibandingkan Naïve Bayes pada klasifikasi sentimen.

Meskipun menghasilkan akurasi yang tinggi, model masih mengalami kesalahan klasifikasi pada sebagian kecil data. Kondisi ini diduga disebabkan oleh penggunaan kalimat pendek, bahasa tidak baku, singkatan, serta kata-kata yang bersifat ambigu. Sebagai contoh, frasa *bisa transfer* dapat mengandung sentimen positif, sedangkan *tidak bisa transfer* mengandung sentimen negatif. Pendekatan berbasis TF-IDF belum sepenuhnya mampu memahami konteks semantik dan hubungan antarkata.

Dibandingkan penelitian sebelumnya, penelitian ini memiliki keunggulan pada ukuran *dataset* dan kedalaman analisis. Selain membangun model klasifikasi, penelitian ini juga menyajikan visualisasi *WordCloud*, analisis *preprocessing*, serta evaluasi model menggunakan beberapa metrik secara komprehensif. Dengan menggunakan **5.971 ulasan valid**, hasil penelitian memberikan gambaran yang lebih representatif mengenai persepsi pengguna terhadap aplikasi Pospay Biru.

Dari sisi praktis, hasil analisis menunjukkan bahwa stabilitas transaksi, pengelolaan saldo, dan akses aplikasi merupakan aspek yang paling banyak memperoleh perhatian pengguna. Informasi ini dapat dimanfaatkan oleh PT Pos Indonesia sebagai dasar evaluasi dalam meningkatkan kualitas layanan, memperbaiki stabilitas sistem, serta mengembangkan fitur aplikasi berdasarkan umpan balik pengguna. Dengan demikian, model analisis sentimen yang dihasilkan tidak hanya memiliki nilai akademis, tetapi juga memberikan manfaat praktis dalam mendukung pengambilan keputusan berbasis data.

KESIMPULAN

Penelitian ini berhasil menerapkan kombinasi *Term Frequency–Inverse Document Frequency* (TF-IDF) sebagai metode ekstraksi fitur dan *Support Vector Machine* (SVM) sebagai algoritma klasifikasi untuk menganalisis sentimen ulasan pengguna aplikasi Pospay Biru pada Google Play Store. Berdasarkan proses *web scraping*, diperoleh **25.391 ulasan**, yang setelah melalui tahapan *preprocessing* menghasilkan **5.971 ulasan valid** sebagai *dataset* penelitian. Tahapan *preprocessing* yang meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming* terbukti mampu

meningkatkan kualitas data sehingga representasi fitur TF-IDF menjadi lebih efektif dalam proses klasifikasi.

Hasil pengujian menunjukkan bahwa model SVM dengan *linear kernel* mampu menghasilkan *accuracy* sebesar **89,13%**, *precision* **83,93%**, *recall* **84,62%**, dan **F1-score sebesar 84,27%**. Nilai tersebut menunjukkan bahwa kombinasi TF-IDF dan SVM memiliki kemampuan yang baik dalam mengklasifikasikan sentimen ulasan pengguna berbahasa Indonesia. Selain itu, hasil visualisasi *WordCloud* mengindikasikan bahwa topik yang paling banyak dibahas pengguna berkaitan dengan stabilitas transaksi, pengelolaan saldo, akses aplikasi, dan keandalan layanan, sehingga aspek-aspek tersebut menjadi faktor utama yang memengaruhi persepsi pengguna terhadap aplikasi Pospay Biru.

Penelitian ini memberikan kontribusi pada penerapan analisis sentimen untuk evaluasi layanan keuangan digital dengan memanfaatkan *dataset* ulasan yang relatif besar dan evaluasi model yang komprehensif. Dari sisi praktis, hasil penelitian dapat dimanfaatkan oleh PT Pos Indonesia sebagai dasar dalam mengidentifikasi prioritas perbaikan layanan berdasarkan umpan balik pengguna. Untuk penelitian selanjutnya, disarankan melakukan optimasi hiperparameter SVM, membandingkan performa dengan algoritma klasifikasi lain, serta menerapkan metode representasi teks berbasis *word embedding* atau model *transformer* seperti BERT untuk meningkatkan kemampuan model dalam memahami konteks semantik pada ulasan pengguna.

DAFTAR PUSTAKA

- Aggarwal, C. C. (2018). *Machine learning for text*. Springer.
- Aristawidya, M. V., Mustaqim, K., & Rahayu, W. I. (2026). Analisis sentimen ulasan pengguna aplikasi PosPay menggunakan Support Vector Machine untuk evaluasi layanan keuangan digital. *Jurnal Informatika*.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. O'Reilly Media.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
- Jurafsky, D., & Martin, J. H. (2025). *Speech and language processing* (3rd ed., draft). Pearson.
- Kowsari, K., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L. E., & Brown, D. E. (2019). Text classification algorithms: A survey. *Information*, 10(4), 150. <https://doi.org/10.3390/info10040150>

- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- Lubis, A., & Setyawan, D. (2024). Analisis sentimen terhadap aplikasi PosPay menggunakan algoritma Support Vector Machine dan Naïve Bayes. *Jurnal Teknologi Informasi*.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Pratiwi, R., & Nugroho, A. (2024). Analisis sentimen ulasan aplikasi DANA menggunakan TF-IDF dan Support Vector Machine. *Jurnal Sistem Informasi*.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513–523. [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34(1), 1–47. <https://doi.org/10.1145/505282.505283>
- Suyanto. (2018). *Machine learning: Tingkat dasar dan lanjut*. Informatika.
- Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to data mining* (2nd ed.). Pearson.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer.