VOL. 03. NO. 05 SEPTEMBER 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

ANALISIS SENTIMEN ULASAN MYTELKOMSEL MENGGUNAKAN SVM DENGAN PEMBOBOTAN TF-IDF DAN COUNT VECTORIZER

Muhammad Bagus Mustofa Madinah[1], Martanto[2], Fatihanursari Dikananda[3], Rosidin[4].

Program Studi Teknik Informatika[13]
Program Studi Manajemen Informatika[2]
Program Studi Sistem Informasi[4]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
bagusmustofa18@gmail.com

(*) Corresponding Author : bagusmustofa18@gmail.com
Published : 17 Juli 2026

Abstract—*The rapid growth of MyTelkomsel application users on the Google Play Store has generated a large volume of reviews, reflecting public perception of service quality. However, these reviews are often unstructured and possess significant class imbalances, making manual analysis difficult. This study aims to perform sentiment analysis on MyTelkomsel user reviews using the Support Vector Machine (SVM) algorithm with Term Frequency-Inverse Document Frequency (TF-IDF) feature weighting and Synthetic Minority Over-sampling Technique (SMOTE) data balancing. A total of 8,934 reviews were collected through web scraping techniques. Initial analysis revealed a dominance of negative sentiment (6,691 data) compared to positive (1,315 data) and neutral (497 data) sentiments. Experiments were conducted by comparing the performance of Count Vectorizer and TF-IDF feature extraction combined with SMOTE. The results showed that the combination of SVM with TF-IDF and SMOTE provided the best performance with an accuracy rate of 81.42%, outperforming the Count Vectorizer method (68.08%) by 13.34%. This study concludes that TF-IDF weighting is more effective in capturing discriminative word features in ICT service reviews. Despite high overall accuracy, the neutral class remains a challenge with a low F1-score (0.08) due to feature ambiguity. These findings provide insights for MyTelkomsel developers to focus on improving network quality and pricing policies, which are the primary user complaints.*

Keywords: *Sentiment Analysis, MyTelkomsel, SVM, TF-IDF, SMOTE.*

Abstrak—Pertumbuhan pesat pengguna aplikasi MyTelkomsel di Google Play Store menghasilkan volume ulasan yang besar, yang mencerminkan persepsi publik terhadap kualitas layanan. Namun, ulasan tersebut seringkali tidak terstruktur dan memiliki ketidakseimbangan kelas yang signifikan, sehingga menyulitkan analisis manual. Penelitian ini bertujuan untuk melakukan analisis sentimen pada ulasan pengguna MyTelkomsel menggunakan algoritma *Support Vector Machine* (SVM) dengan pembobotan fitur *Term Frequency-Inverse Document Frequency* (TF-IDF) dan teknik penyeimbangan data *Synthetic Minority Over-sampling Technique* (SMOTE). Data penelitian sebanyak 8.934 ulasan berhasil dikumpulkan melalui teknik *web scraping*. Hasil analisis awal menunjukkan dominasi sentimen negatif (6.691 data) dibandingkan positif (1.315 data) dan netral (497 data). Eksperimen dilakukan dengan membandingkan performa ekstraksi fitur *Count Vectorizer* dan TF-IDF yang dikombinasikan dengan SMOTE. Hasil pengujian menunjukkan bahwa kombinasi SVM dengan TF-IDF dan SMOTE memberikan performa terbaik dengan tingkat akurasi sebesar 81,42%, unggul 13,34% dibandingkan metode *Count Vectorizer* (68,08%). Penelitian ini menyimpulkan bahwa pembobotan TF-IDF lebih efektif dalam menangkap fitur kata yang diskriminatif pada ulasan layanan TIK. Meskipun akurasi keseluruhan tinggi, kelas netral masih menjadi tantangan dengan *F1-score* rendah (0,08) akibat ambiguitas fitur. Temuan ini memberikan wawasan bagi pengembang MyTelkomsel untuk fokus pada perbaikan kualitas jaringan dan kebijakan harga yang menjadi keluhan utama pengguna.

Kata Kunci: Analisis Sentimen, MyTelkomsel, SVM, TF-IDF, SMOTE.

INTRODUCTION

Dalam era transformasi digital yang semakin pesat, aplikasi *mobile* menjadi salah satu sarana utama dalam penyediaan layanan digital kepada pengguna. Salah satu bentuk interaksi pengguna dengan aplikasi adalah melalui ulasan (*user*

reviews) pada platform distribusi aplikasi seperti Google Play Store. Ulasan tersebut mencerminkan pengalaman, kepuasan, serta permasalahan yang dihadapi pengguna terhadap suatu layanan digital secara langsung, termasuk pada aplikasi MyTelkomsel [1], [2].

Namun, jumlah ulasan yang sangat besar menyebabkan proses analisis secara manual menjadi tidak efektif dan memerlukan waktu yang lama. Oleh karena itu, diperlukan pendekatan otomatis untuk mengelompokkan opini pengguna ke dalam kategori sentimen tertentu agar dapat memberikan informasi yang bermanfaat bagi pengembang aplikasi dalam mengevaluasi distribusi kepuasan pelanggan [3].

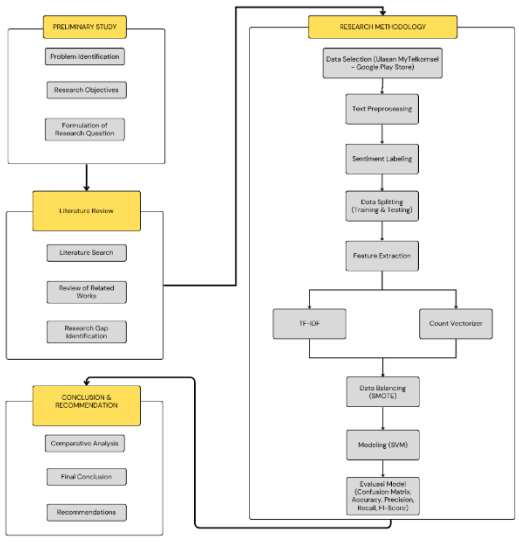
Analisis sentimen berbasis *Natural Language Processing* (NLP) telah banyak digunakan untuk mengatasi permasalahan tersebut. Salah satu algoritma yang sering digunakan adalah *Support Vector Machine* (SVM) karena memiliki kemampuan klasifikasi yang baik pada data teks serta efisiensi dalam proses pelatihan model [4]. Namun, dalam implementasinya, seringkali ditemukan masalah ketidakseimbangan kelas (*imbalanced data*) di mana jumlah ulasan positif dan negatif tidak proporsional, sehingga diperlukan teknik penyeimbangan data seperti SMOTE untuk menjaga objektivitas model.

Meskipun demikian, terdapat keterbatasan pada penelitian sebelumnya, di mana sebagian besar penelitian hanya menggunakan satu metode ekstraksi fitur tanpa melakukan perbandingan secara komprehensif. Padahal, pemilihan metode representasi fitur seperti TF-IDF dan *Count Vectorizer* memiliki pengaruh yang signifikan terhadap performa model klasifikasi sentimen, terutama pada karakteristik bahasa Indonesia yang unik [5].

Dengan demikian, perbandingan performa metode ekstraksi fitur TF-IDF dan *Count Vectorizer* yang diintegrasikan dengan teknik SMOTE pada algoritma SVM menjadi penting untuk diteliti lebih lanjut. Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menganalisis distribusi sentimen ulasan pengguna aplikasi MyTelkomsel serta mengevaluasi dan membandingkan performa kedua metode ekstraksi fitur tersebut guna menghasilkan model klasifikasi yang optimal

MATERIALS AND METHODS

Penelitian ini disusun dengan pendekatan metodologis yang terstruktur guna memastikan setiap tahapan berjalan secara sistematis dan menghasilkan temuan yang valid serta dapat dipertanggungjawabkan. Alur penelitian dimulai dari tahap identifikasi masalah hingga penarikan kesimpulan, dengan dukungan kajian literatur yang relevan dan proses pengolahan data yang komprehensif. Rangkaian tahapan ini dirancang untuk membangun model analisis sentimen yang akurat melalui pemilihan metode, teknik preprocessing, serta evaluasi yang tepat. Gambar berikut memperlihatkan keseluruhan kerangka kerja penelitian yang digunakan.



Gambar 1. Alur Penelitian

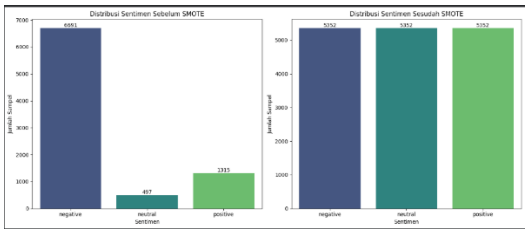
Berdasarkan gambar tersebut, alur penelitian dimulai dari tahap *preliminary study* yang mencakup identifikasi masalah, penentuan tujuan penelitian, serta perumusan pertanyaan penelitian sebagai dasar arah studi. Tahap ini dilanjutkan dengan *literature review* yang bertujuan untuk mengkaji penelitian terdahulu, melakukan pencarian literatur, serta mengidentifikasi celah penelitian (*research gap*) yang akan diisi oleh penelitian ini.

Selanjutnya, masuk ke tahap *research methodology* yang merupakan inti dari proses penelitian. Pada tahap ini dilakukan seleksi data berupa ulasan aplikasi MyTelkomsel dari Google Play Store, kemudian dilanjutkan dengan *text preprocessing* untuk membersihkan dan menormalisasi data teks. Setelah itu, dilakukan *sentiment labeling* untuk memberikan kategori sentimen pada data, serta *data splitting* menjadi data latih dan data uji. Tahap berikutnya adalah *feature extraction* menggunakan dua pendekatan, yaitu TF-IDF dan Count Vectorizer, yang bertujuan mengubah data teks menjadi representasi numerik.

Untuk mengatasi ketidakseimbangan data, diterapkan teknik *SMOTE* pada tahap *data balancing*. Selanjutnya, model dibangun menggunakan algoritma Support Vector Machine (SVM) yang dikenal efektif untuk klasifikasi teks. Kinerja model kemudian dievaluasi menggunakan beberapa metrik seperti confusion matrix, akurasi, presisi, recall, dan F1-score guna memastikan performa model secara menyeluruh. Tahap akhir adalah *conclusion and recommendation* yang meliputi analisis komparatif, penarikan kesimpulan akhir, serta pemberian rekomendasi untuk pengembangan penelitian selanjutnya.

RESULTS AND DISCUSSION

Untuk mengatasi ketidakseimbangan label, digunakan teknik *SMOTE* (*Synthetic Minority Oversampling Technique*) yang menghasilkan data sintetis untuk kelas minoritas. Setelah penerapan *SMOTE*, dataset menjadi seimbang dengan jumlah sekitar 5352 ulasan per kelas.



Gambar 2 Distribusi label setelah penerapan *SMOTE*

Sumber: Diolah Oleh Peneliti (2025)

Gambar 2 memperlihatkan perbandingan distribusi label sentimen sebelum dan sesudah penerapan metode *SMOTE* (*Synthetic Minority Over-sampling Technique*). Pada grafik sebelah kiri, terlihat bahwa data awal memiliki ketidakseimbangan yang signifikan, di mana sentimen negatif mendominasi dengan 6.691 data, sementara positif hanya 1.315 dan netral 497. Kondisi ini menunjukkan adanya *class imbalance* yang berpotensi menurunkan kinerja model klasifikasi. Setelah penerapan *SMOTE*, seperti terlihat pada grafik sebelah kanan, jumlah data pada ketiga kelas negatif, netral, dan positif menjadi seimbang dengan masing-masing sekitar 5.352 data. Penerapan *SMOTE* berhasil melakukan *oversampling* terhadap kelas minoritas (positif dan netral), sehingga distribusi data menjadi lebih proporsional. Hasil ini penting untuk meningkatkan performa model *machine learning*, karena memungkinkan algoritma pembelajaran mengenali pola dari setiap kelas secara lebih adil dan mengurangi bias terhadap kelas mayoritas.

Hasil Evaluasi Model

Model yang dievaluasi adalah *Support Vector Machine (SVM)* dengan dua konfigurasi berbeda: *SVM* + TF–IDF + *SMOTE* (kernel: polynomial) *SVM* + *Count Vectorizer* + *SMOTE* (kernel: linear) Data dibagi menjadi data latih (80%) dan data uji (20%). Evaluasi dilakukan menggunakan metrik: akurasi, presisi, recall, dan F1-score.

Hasil evaluasi menunjukkan bahwa konfigurasi *SVM* dengan TF–IDF dan *SMOTE* memberikan performa terbaik. Secara ringkas:

SVM (kernel='poly') + TF–IDF + *SMOTE* — akurasi *tertinggi: 81,42%.*

```
Evaluasi Model SVM dengan TF-IDF dan SMOTE:
              precision    recall  f1-score   support

    negative       0.85      0.95      0.89      1339
     neutral       0.15      0.05      0.08        99
    positive       0.65      0.42      0.51       263

    accuracy                           0.81      1701
   macro avg       0.55      0.47      0.49      1701
weighted avg       0.78      0.81      0.79      1701

Akurasi: 0.8142269253380364
```

Gambar 3 Hasil Evaluasi Model SVM

Sumber: Diolah Oleh Peneliti (2025)

Gambar 3 menunjukkan hasil evaluasi model *Support Vector Machine (SVM)* yang dikombinasikan dengan metode *Term Frequency–Inverse Document Frequency (TF-IDF)* dan teknik penyeimbangan data *Synthetic Minority Over-sampling Technique (SMOTE)*. Berdasarkan hasil pengujian, model mencapai tingkat akurasi sebesar 81,42%. Dari tabel evaluasi terlihat bahwa kelas "negative" memiliki performa paling tinggi dengan presisi 0,85, recall 0,95, dan F1-score 0,89. Sementara itu, kelas "neutral" menunjukkan performa terendah dengan presisi 0,15 dan recall 0,05, yang menandakan model masih sulit membedakan teks netral. Kelas "positive" memiliki presisi 0,65 dan recall 0,42 dengan F1-score 0,51, yang menandakan prediksi cukup moderat. Nilai *macro average* untuk F1-score sebesar 0,49 menunjukkan keseimbangan antar kelas masih belum optimal, namun nilai *weighted average* sebesar 0,79 memperlihatkan performa keseluruhan model cukup baik. Secara umum, kombinasi *TF-IDF* dan *SMOTE* berhasil meningkatkan akurasi model dengan memberikan distribusi data yang lebih seimbang pada tiap kelas sentimen.

SVM (kernel='linear') + *Count Vectorizer* + *SMOTE* — performa lebih rendah dibandingkan kombinasi TF–IDF.

Evaluasi Model SVM dengan Count Vectorizer dan SMOTE:

	precision	recall	f1-score	support
negative	0.88	0.76	0.81	1339
neutral	0.06	0.18	0.09	99
positive	0.53	0.49	0.51	263
accuracy			0.68	1701
macro avg	0.49	0.48	0.47	1701
weighted avg	0.78	0.68	0.72	1701

Akurasi: 0.6807760141093474

Gambar 4 Hasil Evaluasi Model SVM dengan *Count Vectorizer* dan *SMOTE*
Sumber: Diolah Oleh Peneliti (2025)

Gambar 4 menampilkan hasil evaluasi model *SVM* yang menggunakan *Count Vectorizer* dan teknik penyeimbangan data *SMOTE*. Berdasarkan hasil pengujian, model memperoleh tingkat akurasi sebesar 68,08%. Kelas negative menunjukkan performa terbaik dengan presisi sebesar 0,88, recall 0,76, dan F1-score 0,81. Hal ini menunjukkan bahwa model mampu mengenali data dengan sentimen negatif secara cukup baik. Sementara itu, kelas neutral masih memiliki performa rendah dengan presisi 0,06 dan recall 0,18, menandakan kesulitan model dalam mengidentifikasi teks yang bersifat netral. Untuk kelas positive, nilai presisi 0,53 dan recall 0,49 menghasilkan F1-score sebesar 0,51 yang menunjukkan kemampuan prediksi sedang. Nilai macro average F1-score sebesar 0,47 dan weighted average F1-score sebesar 0,72 mengindikasikan bahwa performa model secara keseluruhan masih dipengaruhi oleh dominasi data pada kelas negative. Dengan demikian, meskipun penggunaan *SMOTE* membantu menyeimbangkan data, representasi fitur menggunakan *Count Vectorizer* belum mampu memberikan hasil sebaik metode *TF-IDF* dalam meningkatkan kinerja klasifikasi sentimen.

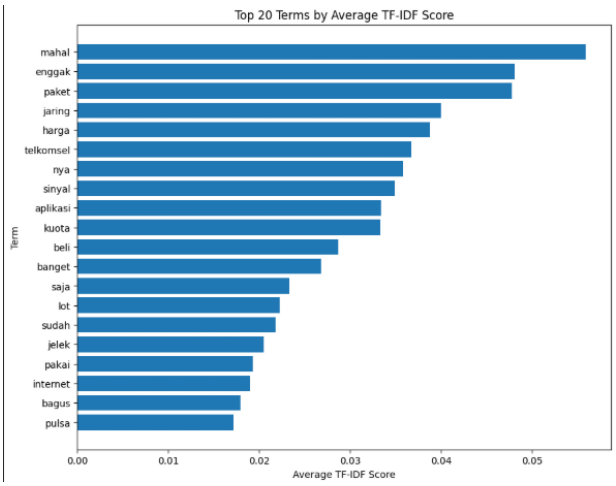
Evaluasi model dilakukan menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score yang diperoleh dari hasil *confusion matrix*. Untuk mengetahui metode yang memberikan performa terbaik, dilakukan perbandingan antara model SVM yang menggunakan *Count Vectorizer* dan model SVM yang menggunakan TF-IDF.

Membandingkan performa TF-IDF dan Count Vectorizer dalam klasifikasi sentimen berbasis SVM menggunakan metrik akurasi, presisi, recall, dan F1-score.

Sebagai tahap akhir eksperimen, peneliti melakukan perbandingan mendalam terhadap efektivitas metode TF-IDF dan *Count Vectorizer* dalam merepresentasikan teks ulasan. Perbandingan ini tidak hanya merujuk pada skor akurasi akhir, tetapi juga didasarkan pada analisis visual terhadap *confusion matrix* dari kedua skenario pemodelan untuk melihat seberapa akurat model dalam memprediksi setiap label sentimen.

Ekstraksi Fitur dengan TF–IDF dan *Count Vectorizer* Dua metode ekstraksi fitur diuji: *Count Vectorizer* dan TF–IDF. *Count Vectorizer*: Menghitung frekuensi kemunculan kata di setiap dokumen. TF–IDF: Memberi bobot lebih tinggi pada kata yang jarang muncul namun informatif untuk membedakan dokumen.

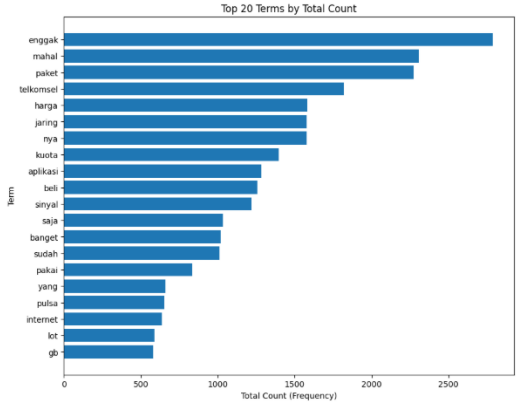
Analisis menunjukkan TF–IDF menghasilkan representasi yang lebih informatif untuk tugas klasifikasi sentimen: kata-kata umum pada semua ulasan (mis. "aplikasi", "pakai") memiliki bobot kecil, sementara kata khas yang mencerminkan ketidakpuasan atau masalah teknis (mis. "error", "lemot", "gagal") memperoleh bobot lebih tinggi.



Gambar 5 Hasil Ekstraksi Fitur Dengan TF-IDF
Sumber: Diolah Oleh Peneliti (2025)

Gambar 5 menunjukkan hasil ekstraksi fitur menggunakan metode *TF-IDF (Term Frequency–Inverse Document Frequency)* yang menampilkan 20 kata dengan skor tertinggi berdasarkan rata-rata nilai TF-IDF. Kata **"**mahal**"** menempati posisi teratas, diikuti oleh kata "enggak", "paket", "jaring", dan "harga". Hal ini menunjukkan bahwa kata-kata tersebut memiliki bobot paling penting dan paling sering muncul dalam dokumen ulasan yang dianalisis. Dominasi kata "mahal" dan "harga" mengindikasikan bahwa aspek biaya menjadi perhatian utama pengguna, sementara kata "jaring", "sinyal", dan "internet" menunjukkan fokus ulasan pada kualitas layanan jaringan. Selain itu, kata "jelek" dan "bagus" menggambarkan adanya ekspresi evaluatif dari pengguna, baik dalam konteks positif maupun negatif. Secara keseluruhan, hasil ekstraksi ini memberikan gambaran mengenai topik utama dan sentimen dominan dalam ulasan, serta menjadi dasar penting

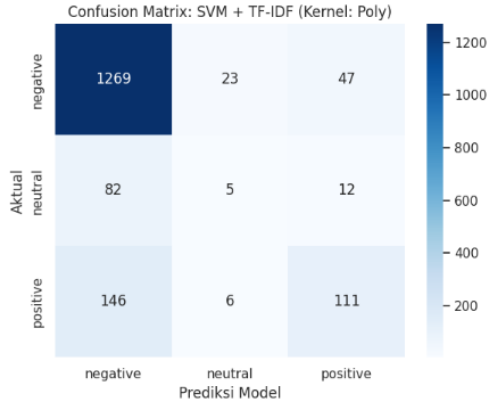
dalam analisis lanjutan seperti klasifikasi sentimen atau identifikasi aspek layanan yang paling banyak dibicarakan.



Gambar 6 Hasil Ekstraksi Fitur Dengan *Count Vectorizer*
Sumber: Diolah Oleh Peneliti (2025)

Gambar 6 menampilkan hasil ekstraksi fitur menggunakan metode *Count Vectorizer*, yang menunjukkan 20 kata dengan frekuensi kemunculan tertinggi dalam kumpulan data ulasan. Berdasarkan grafik tersebut, kata "enggak" menempati posisi teratas dengan jumlah kemunculan paling banyak, diikuti oleh kata "mahal", "paket", "telkomsel", dan "harga". Pola ini mengindikasikan bahwa sebagian besar pengguna mengekspresikan pendapat terkait harga dan kualitas layanan, khususnya yang berkaitan dengan paket data dan jaringan. Selain itu, munculnya kata-kata seperti "jaring", "sinyal", dan "internet" menunjukkan bahwa permasalahan teknis menjadi topik yang sering dibahas. Kata "jelek", "bagus", dan "banget" mencerminkan ekspresi emosional pengguna terhadap layanan yang diterima, baik dalam konteks negatif maupun positif. Hasil ekstraksi dengan *Count Vectorizer* ini menggambarkan fokus utama opini pelanggan secara umum, serta menjadi landasan awal dalam memahami persepsi publik terhadap layanan yang dianalisis sebelum dilakukan analisis sentimen lebih mendalam

Confusion Matrix



Gambar 7 Confusion Matrix TF-IDF
Sumber: Diolah Oleh Peneliti (2025)

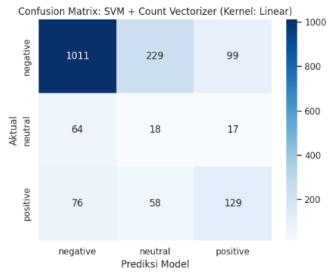
Gambar 7 tersebut, dapat dilakukan analisis mendalam mengenai performa klasifikasi model SVM dengan ekstraksi fitur TF-IDF sebagai berikut:

Dominasi Prediksi Benar pada Kelas Negatif: Model menunjukkan kemampuan yang sangat kuat dalam mengenali sentimen negatif, di mana sebanyak 1.269 ulasan berhasil diprediksi secara tepat sesuai dengan label aktualnya. Hal ini merepresentasikan nilai *recall* sebesar 0,95, yang berarti model mampu mendeteksi hampir seluruh keluhan pengguna secara akurat.

Kinerja Prediksi Kelas Positif: Pada kelas positif, model berhasil mengklasifikasikan secara benar sebanyak 111 ulasan. Meskipun demikian, terdapat sejumlah data positif yang salah diprediksi sebagai negatif, yang mengindikasikan adanya kemiripan pola kata atau karakteristik data yang cukup moderat pada kelas ini.

Tantangan pada Kelas Netral: Model mengalami kesulitan paling signifikan pada kelas netral, di mana hanya 5 ulasan yang berhasil diprediksi dengan benar. Rendahnya performa ini disebabkan oleh karakteristik data netral yang cenderung ambigu serta jumlah data yang lebih terbatas dibandingkan kelas lainnya.

Efektivitas Keseluruhan: Konsentrasi angka tertinggi yang berada pada diagonal utama matriks menunjukkan bahwa mayoritas data telah diklasifikasikan dengan benar, sehingga mendukung pencapaian akurasi total model sebesar 81,42%.



Gambar 8 Confusion Matrix Count Vectorizer
Sumber: Diolah Oleh Peneliti (2025)

Gambar 8 tersebut, dapat dilakukan analisis mendalam mengenai performa klasifikasi model SVM dengan metode ekstraksi fitur Count Vectorizer sebagai berikut:

Kemampuan Deteksi Kelas Negatif: Model memiliki kemampuan yang cukup baik dalam mengidentifikasi sentimen negatif, dengan jumlah prediksi benar sebanyak 1.011 ulasan. Meskipun demikian, terdapat kesalahan prediksi yang cukup signifikan di mana 229 ulasan negatif salah diklasifikasikan sebagai netral dan 99 ulasan lainnya salah diklasifikasikan sebagai positif.

Performa pada Kelas Positif: Pada kelas positif, model berhasil melakukan prediksi benar sebanyak 129 ulasan. Namun, model juga sering melakukan kesalahan klasifikasi dengan memprediksi 76 data positif sebagai negatif dan 58 data lainnya sebagai netral.

Tantangan Signifikan pada Kelas Netral: Model menunjukkan kesulitan yang besar dalam mengenali sentimen netral, dengan hanya 18 ulasan yang berhasil diprediksi secara tepat. Hal ini mengonfirmasi hasil evaluasi sebelumnya bahwa model masih sulit membedakan teks netral karena karakteristik data yang ambigu.

Interpretasi Keseluruhan: Persebaran angka di luar diagonal utama yang masih cukup tinggi menunjukkan bahwa representasi fitur menggunakan *Count Vectorizer* belum mampu memberikan pemisahan antar-kelas yang seoptimal metode TF-IDF, sehingga menghasilkan akurasi yang lebih rendah yaitu sebesar 68,08%

Analisis Komparasi Performa Akhir

Tabel 2 Perbandingan Performa Model Klasifikasi

Metode	akurasi
SVM + *Count Vectorizer*	68,08%
SVM + TF-IDF	81,42%

Sumber: Diolah Oleh Peneliti (2025)

Berdasarkan Tabel 2 dapat diketahui bahwa model SVM dengan metode ekstraksi fitur TF-IDF menghasilkan performa klasifikasi yang lebih baik dibandingkan model SVM dengan *Count Vectorizer*. Model SVM dengan *Count Vectorizer* memperoleh akurasi sebesar 68,08%, sedangkan model SVM dengan TF-IDF memperoleh akurasi sebesar 81,42%.

CONCLUSION

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan sebagai berikut:

1. Distribusi sentimen pada ulasan pengguna aplikasi MyTelkomsel yang diperoleh dari Google Play Store menunjukkan bahwa sentimen negatif mendominasi dibandingkan sentimen positif dan netral. Dari total data yang telah diolah, sebagian besar ulasan mengandung keluhan terkait harga layanan dan kualitas jaringan. Kondisi ini menunjukkan bahwa persepsi pengguna terhadap aplikasi cenderung negatif, serta mengindikasikan adanya aspek layanan yang perlu diperbaiki oleh pengembang.
2. Performa klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM) yang diintegrasikan dengan teknik SMOTE menunjukkan hasil yang cukup baik pada kedua metode ekstraksi fitur. Model dengan TF-IDF dan SMOTE memperoleh akurasi sebesar **81,42%** dengan nilai F1-score rata-rata sebesar **0,79**, sedangkan model dengan *Count Vectorizer* dan SMOTE memperoleh akurasi sebesar **68,08%**. Penerapan SMOTE terbukti mampu mengurangi bias terhadap kelas mayoritas, meskipun model masih mengalami keterbatasan dalam mengklasifikasikan kelas netral secara optimal.
3. Perbandingan performa antara metode ekstraksi fitur menunjukkan bahwa TF-IDF memiliki kinerja yang lebih unggul dibandingkan *Count Vectorizer* dalam klasifikasi sentimen menggunakan algoritma SVM. Hal ini ditunjukkan oleh selisih akurasi sebesar 13,34%, di mana TF-IDF mampu menghasilkan representasi fitur yang lebih diskriminatif karena mempertimbangkan bobot kepentingan kata dalam dokumen, bukan hanya frekuensi kemunculannya.

REFERENCE

[1] S. A. H. Bahtiar, C. K. Dewa, and A. Luthfi, "Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling," *Journal of Information Systems and Informatics*, vol. 5, no. 3, pp. 915–927, 2023, doi: 10.51519/journalisi.v5i3.539.

[2] K. D. Indarwati and H. Februariyanti, "Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 10, no. 1, 2023, doi: 10.35957/jatisi.v10i1.2643.

[3] A. Nurian, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.

[4] H. Firda, "Comparison of Rating-based and Inset Lexicon-based Labeling in Sentiment Analysis using SVM (Case Study: GoBiz Application Reviews on Google Play Store)," *Sistemasi: Jurnal Sistem Informasi*, vol. 14, no. 2, p. 516, 2025, doi: 10.32520/stmsi.v14i2.4795.

[5] R. Vincent, I. Maulana, and O. Komarudin, "Perbandingan Klasifikasi Naive Bayes dan Support Vector Machine Dalam Analisis Sentimen Dengan Multiclass Di Twitter," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 4, 2024, doi: 10.36040/jati.v7i4.7152.