

OPTIMALISASI MODEL KLASIFIKASI MENGGUNAKAN ALGORITMA SVM DENGAN RESAMPLING TERHADAP DATA IMBALANCE ULASAN MYSUZUKI DI PLAYSTORE

Ahmad Satrio¹, Rudi Kurniawan², Bani Nurhakim³, Aris Pratama Putra⁴.

Program Studi Teknik Informatika¹²
Program Studi Manajemen Informatika³⁴

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
satrioahmad883@gmail.com

(*) Corresponding Author : satrioahmad883@gmail.com
Published : 20 Juli 2026

Abstract—This study aims to implement sentiment analysis on user reviews of the My Suzuki application on Google Play Store using the Support Vector Machine (SVM) algorithm, analyze the effect of resampling techniques on classification performance, and compare the performance of SVM models with and without resampling based on accuracy, precision, recall, and F1-score metrics. The main issue addressed in this study is the imbalance in sentiment data distribution, which may cause the classification model to become biased toward the majority class. The dataset used in this research consisted of 1,227 user reviews of the My Suzuki application, categorized into three sentiment classes: positive, negative, and neutral. The research methodology consisted of several stages, including data scraping, text preprocessing (lowercase conversion, cleaning, stopword removal, tokenizing, and stemming), manual labeling by a language expert, imbalance data handling using oversampling techniques, feature extraction using TF-IDF, and sentiment classification using the Support Vector Machine (SVM) algorithm. Model evaluation was conducted using a confusion matrix with accuracy, precision, recall, and F1-score metrics. The results showed that the use of resampling techniques significantly improved sentiment classification performance. The SVM model without resampling achieved an accuracy of 73%, while the SVM model using oversampling achieved an accuracy of 90%. In addition, the precision, recall, and F1-score values for all sentiment classes became more balanced after the resampling process. The oversampling technique proved effective in helping the model recognize patterns in minority classes, especially the neutral sentiment class, which was previously difficult to classify accurately. Furthermore, the WordCloud visualization revealed that negative sentiments were dominated by complaints related to service quality, spare part delivery, and application performance, while positive sentiments were mainly associated with ease of use and vehicle service accessibility. Based on the findings, it can be concluded that the combination of the Support Vector Machine (SVM) algorithm and resampling techniques provides an effective and optimal approach for sentiment analysis of My Suzuki application reviews on Google Play Store. This study is expected to assist application developers in understanding user perceptions and improving service quality based on user review data.

Keywords: Sentiment Analysis, Support Vector Machine, Resampling, Oversampling, TF-IDF, My Suzuki

Abstrak—Penelitian ini bertujuan untuk mengimplementasikan analisis sentimen terhadap ulasan pengguna aplikasi My Suzuki di Google Play Store menggunakan algoritma Support Vector Machine (SVM), menganalisis pengaruh teknik resampling terhadap performa klasifikasi, serta membandingkan performa model SVM dengan dan tanpa resampling berdasarkan nilai accuracy, precision, recall, dan F1-score. Permasalahan utama dalam penelitian ini adalah ketidakseimbangan distribusi data sentimen yang berpotensi menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas. Dataset yang digunakan terdiri dari 1.227 ulasan pengguna aplikasi My Suzuki yang diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Metode penelitian meliputi tahapan scraping data, preprocessing teks yang terdiri dari lowercase, cleaning, stopword removal, tokenizing, dan stemming, pelabelan data secara manual oleh ahli bahasa, penanganan imbalance data menggunakan teknik oversampling, ekstraksi fitur menggunakan TF-IDF, serta proses klasifikasi menggunakan algoritma Support Vector Machine (SVM). Evaluasi model dilakukan menggunakan confusion matrix dengan metrik accuracy, precision, recall, dan F1-score.

Hasil penelitian menunjukkan bahwa penggunaan teknik resampling mampu meningkatkan performa klasifikasi sentimen secara signifikan. Model SVM tanpa resampling memperoleh accuracy sebesar 73%, sedangkan model SVM menggunakan oversampling berhasil mencapai accuracy sebesar 90%. Selain itu, nilai precision, recall, dan F1-score pada seluruh kelas sentimen mengalami peningkatan yang lebih seimbang setelah dilakukan resampling. Teknik oversampling terbukti membantu model dalam mengenali pola data pada kelas minoritas, khususnya sentimen netral yang sebelumnya sulit diklasifikasikan dengan baik. Hasil visualisasi WordCloud juga menunjukkan bahwa sentimen negatif pengguna didominasi oleh keluhan terkait layanan, pengiriman suku cadang, dan performa aplikasi, sedangkan sentimen positif didominasi oleh kemudahan penggunaan aplikasi dan layanan servis kendaraan. Berdasarkan hasil penelitian, dapat disimpulkan bahwa kombinasi algoritma Support Vector Machine (SVM) dan teknik resampling merupakan pendekatan yang efektif dan optimal dalam analisis sentimen ulasan aplikasi My Suzuki di Google Play Store. Penelitian ini diharapkan dapat membantu pengembang aplikasi dalam memahami persepsi pengguna serta meningkatkan kualitas layanan aplikasi berbasis data ulasan pengguna.

Kata Kunci: Analisis Sentimen, Support Vector Machine, Resampling, Oversampling, TF-IDF, My Suzuki

INTRODUCTION

Perkembangan teknologi digital dalam satu dekade terakhir telah membawa perubahan signifikan terhadap pola konsumsi informasi dan interaksi masyarakat dengan layanan berbasis aplikasi mobile. Berbagai sektor industri mulai memanfaatkan teknologi digital untuk meningkatkan kualitas layanan kepada pelanggan, termasuk sektor otomotif. Salah satu fenomena yang sangat menonjol adalah meningkatnya ketergantungan masyarakat pada aplikasi mobile sebagai sarana untuk mengakses layanan, seperti pemesanan servis kendaraan, pembelian suku cadang, memperoleh informasi kendaraan, hingga layanan bantuan pelanggan secara langsung. Sebagai bagian dari transformasi digital tersebut, berbagai produsen kendaraan menghadirkan aplikasi pendukung layanan purnajual untuk memberikan kemudahan dan meningkatkan pengalaman pengguna. My Suzuki merupakan salah satu aplikasi yang dikembangkan oleh PT Suzuki Indomobil Sales sebagai bentuk peningkatan kualitas layanan berbasis digital yang ditujukan untuk mendukung pengalaman pelanggan secara lebih efisien dan responsif.

Seiring meningkatnya penggunaan aplikasi mobile, kualitas layanan digital sangat dipengaruhi oleh persepsi dan pengalaman pengguna. Dalam konteks ini, ulasan pengguna pada platform seperti Google Play Store menjadi sumber data penting untuk menilai tingkat kepuasan, mengidentifikasi keluhan, serta memahami harapan pengguna terhadap layanan yang diberikan. Menurut Baihaqi et al. (2025), analisis ulasan pada platform digital mampu memberikan insight yang bernilai karena sifat datanya yang organik, jujur, dan mencerminkan pengalaman langsung pengguna. Namun, jumlah

ulasan yang sangat banyak dan berbentuk data teks tidak terstruktur menyebabkan proses analisis secara manual menjadi sulit dilakukan dan membutuhkan waktu yang cukup lama.

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan analisis sentimen yang mampu mengelompokkan opini pengguna ke dalam kategori sentimen positif, negatif, maupun netral secara otomatis. Analisis sentimen merupakan salah satu cabang dari Natural Language Processing (NLP) dan text mining yang digunakan untuk memahami kecenderungan emosi atau opini dalam suatu teks. Dengan adanya analisis sentimen, pengembang aplikasi dapat mengetahui tingkat kepuasan pengguna secara lebih efektif serta memperoleh informasi penting untuk meningkatkan kualitas layanan aplikasi. Safrudin et al. (2024) menjelaskan bahwa analisis sentimen berbasis machine learning menjadi pendekatan populer karena kemampuannya dalam memodelkan pola linguistik yang kompleks dari data teks.

Dalam implementasinya, analisis sentimen banyak menggunakan metode machine learning karena mampu melakukan klasifikasi teks dengan tingkat akurasi yang baik. Salah satu algoritma machine learning yang sering digunakan dalam analisis sentimen adalah Support Vector Machine (SVM). Algoritma SVM memiliki kemampuan yang baik dalam mengklasifikasikan data berdimensi tinggi, termasuk data teks hasil transformasi TF-IDF. Selain itu, SVM dikenal mampu menghasilkan performa klasifikasi yang stabil dan akurat dalam berbagai penelitian analisis sentimen. Lamonda et al. (2025) menyatakan bahwa SVM secara konsisten menunjukkan performa tinggi pada tugas klasifikasi teks karena kemampuannya memisahkan kelas menggunakan hyperplane optimal. Penelitian lain yang dilakukan oleh Java et al. (2024) juga menunjukkan bahwa algoritma SVM menghasilkan

tingkat akurasi yang lebih tinggi pada klasifikasi ulasan aplikasi dibandingkan beberapa metode klasifikasi lainnya.

Meskipun sejumlah penelitian telah membahas penggunaan algoritma SVM untuk analisis sentimen pada berbagai aplikasi digital, terdapat variasi performa yang dipengaruhi oleh karakteristik dataset, teknik preprocessing, metode pelabelan, serta distribusi data pada masing-masing kelas sentimen. Salah satu permasalahan yang sering muncul dalam analisis sentimen adalah ketidakseimbangan distribusi data atau *imbalanced data*, yaitu kondisi ketika jumlah data pada suatu kelas jauh lebih banyak dibandingkan kelas lainnya. Ketidakseimbangan data dapat menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas sehingga kemampuan model dalam mengenali kelas minoritas menjadi menurun.

Dalam konteks aplikasi My Suzuki, ulasan pengguna sangat penting untuk menilai kualitas layanan digital seperti pemesanan servis, ketersediaan informasi, pengiriman suku cadang, maupun responsivitas sistem aplikasi. Namun, hingga saat ini penelitian yang secara khusus mengkaji analisis sentimen pada aplikasi My Suzuki masih sangat terbatas. Padahal, ulasan pada aplikasi otomotif memiliki karakteristik yang berbeda dibandingkan aplikasi pada domain hiburan, e-commerce, atau media sosial karena lebih banyak membahas layanan purnajual, performa sistem, serta pengalaman pengguna dalam melakukan servis kendaraan. Oleh karena itu, penelitian ini memiliki urgensi yang tinggi, baik dari aspek akademik maupun praktis.

Berdasarkan hasil pengumpulan data ulasan aplikasi My Suzuki, ditemukan bahwa distribusi sentimen masih belum seimbang, di mana jumlah data pada kelas tertentu lebih dominan dibandingkan kelas lainnya. Kondisi tersebut berpotensi mempengaruhi performa algoritma SVM dalam melakukan klasifikasi sentimen. Oleh karena itu, diperlukan teknik penanganan *imbalance data* agar model dapat mempelajari pola dari setiap kelas secara lebih seimbang.

Salah satu metode yang dapat digunakan untuk mengatasi ketidakseimbangan data adalah teknik *resampling*. Teknik *resampling* dilakukan dengan menyeimbangkan jumlah data pada masing-masing kelas sehingga distribusi data menjadi lebih proporsional. Dalam penelitian ini, metode yang digunakan adalah *oversampling*, yaitu teknik penambahan data pada kelas minoritas hingga jumlahnya setara dengan kelas mayoritas. Dengan distribusi data

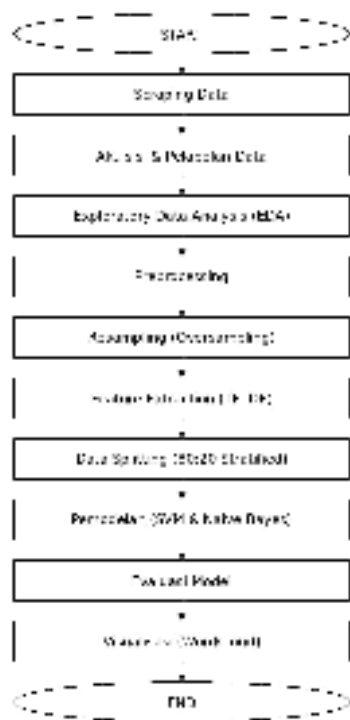
yang lebih seimbang, diharapkan performa algoritma Support Vector Machine (SVM) dapat meningkat dan menghasilkan klasifikasi sentimen yang lebih optimal.

Selain itu, berdasarkan hasil tinjauan penelitian terdahulu, masih terdapat kesenjangan metodologis pada proses penanganan ketidakseimbangan data (*imbalanced data*). Sebagian besar penelitian analisis sentimen hanya berfokus pada penggunaan algoritma klasifikasi tanpa memperhatikan distribusi data pada masing-masing kelas sentimen. Padahal, distribusi data yang tidak seimbang dapat menyebabkan model cenderung bias terhadap kelas mayoritas sehingga menurunkan performa klasifikasi pada kelas minoritas. Beberapa penelitian juga belum menerapkan teknik *resampling* secara optimal untuk mengatasi permasalahan tersebut, khususnya pada data ulasan aplikasi berbasis layanan otomotif.

Berdasarkan latar belakang tersebut, penelitian ini dilakukan dengan judul "Optimalisasi Algoritma Support Vector Machine (SVM) Menggunakan Resampling pada Analisis Sentimen Ulasan My Suzuki di Google Play Store". Penelitian ini diharapkan mampu memberikan kontribusi dalam pengembangan analisis sentimen berbasis machine learning, khususnya dalam penanganan data tidak seimbang pada domain aplikasi layanan otomotif, serta memberikan rekomendasi berbasis data bagi pengembang aplikasi dalam meningkatkan kualitas layanan digital berdasarkan opini pengguna secara objektif.

MATERIALS AND METHODS

Proses penelitian dalam bidang data mining dan machine learning memerlukan tahapan yang sistematis agar hasil analisis dapat dilakukan secara optimal dan menghasilkan model yang akurat. Setiap tahapan memiliki fungsi penting yang saling berhubungan mulai dari pengumpulan data hingga evaluasi model. Alur penelitian yang terstruktur membantu peneliti dalam memahami proses pengolahan data secara bertahap sehingga dapat meminimalkan kesalahan dalam proses analisis. Selain itu, penggunaan metode yang tepat pada setiap tahapan juga berpengaruh terhadap kualitas hasil klasifikasi yang diperoleh. Gambar berikut menunjukkan tahapan penelitian yang digunakan dalam proses analisis data teks menggunakan metode machine learning mulai dari scraping data hingga visualisasi hasil menggunakan WordCloud secara sistematis dan terintegrasi.



Gambar 1. Alur Penelitian

Tahap berikutnya adalah preprocessing yang bertujuan untuk membersihkan data dari noise, simbol, kata tidak penting, maupun data duplikat agar data menjadi lebih terstruktur dan siap digunakan dalam proses pemodelan. Setelah preprocessing selesai, dilakukan proses resampling menggunakan teknik oversampling untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas sehingga model dapat belajar secara lebih optimal. Selanjutnya dilakukan feature extraction menggunakan metode TF-IDF untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma machine learning.

Tahapan berikutnya adalah data splitting dengan rasio 80:20 secara stratified, yaitu pembagian data menjadi data latih dan data uji dengan mempertahankan proporsi kelas pada dataset. Setelah itu dilakukan proses pemodelan menggunakan algoritma Support Vector Machine (SVM) dan Naive Bayes untuk melakukan klasifikasi data. Hasil model kemudian dievaluasi menggunakan metrik evaluasi tertentu guna mengetahui tingkat performa model yang dihasilkan. Tahap terakhir adalah visualisasi menggunakan WordCloud untuk menampilkan kata-kata yang paling sering muncul dalam dataset sehingga memudahkan interpretasi hasil analisis secara visual.

RESULTS AND DISCUSSION

Pada tahap ini dilakukan proses pemodelan menggunakan algoritma Support Vector Machine (SVM) untuk menganalisis sentimen pada data ulasan pengguna aplikasi My Suzuki. Algoritma SVM dipilih karena memiliki kemampuan yang baik dalam melakukan klasifikasi data teks berdimensi tinggi, khususnya pada data hasil ekstraksi fitur menggunakan TF-IDF. Selain itu, SVM dikenal mampu menghasilkan performa klasifikasi yang stabil dan akurat dalam berbagai penelitian analisis sentimen.

Pada penelitian ini digunakan kernel linear dengan parameter $C=1C = 1C=1$. Kernel linear dipilih karena data teks hasil TF-IDF memiliki karakteristik sparse dan berdimensi tinggi sehingga lebih efektif diproses menggunakan hyperplane linear. Sementara itu, parameter $C=1C = 1C=1$ digunakan untuk mengontrol keseimbangan antara margin hyperplane dan tingkat kesalahan klasifikasi agar model tidak mengalami overfitting maupun underfitting.

Proses pemodelan dilakukan menggunakan dua skenario pengujian, yaitu: Model SVM tanpa resampling dan Model SVM menggunakan resampling

Kedua model dilatih menggunakan data latih (training data) dan diuji menggunakan data uji (testing data) dengan pembagian data sebesar 80:20 menggunakan teknik stratified split. Evaluasi model dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score.

Berikut merupakan hasil evaluasi model SVM menggunakan dataset asli tanpa proses penyeimbangan data (resampling):

Tabel 1.

Sentimen	Precision	Recall	F1-Score	Support
Negatif	0.71	0.73	0.72	128
Netral	0.08	0.02	0.04	43
Positif	0.77	0.87	0.82	121
Accuracy			0.73	199
Macro Avg	0.52	0.54	0.69	362
Weighted Avg	0.67	0.73	0.69	362

Berdasarkan Tabel 1 model SVM tanpa resampling memperoleh nilai accuracy sebesar 0.73. Model mampu mengklasifikasikan sentimen positif dan negatif dengan cukup baik, terlihat dari

nilai recall sentimen positif sebesar 0.87 dan sentimen negatif sebesar 0.73. Namun, performa model pada kelas netral masih sangat rendah dengan nilai recall sebesar 0.02 dan F1-score sebesar 0.04.

Rendahnya performa pada kelas netral menunjukkan bahwa model mengalami kesulitan dalam mengenali data minoritas. Kondisi tersebut terjadi karena distribusi dataset yang tidak seimbang (imbalanced data), di mana jumlah data netral jauh lebih sedikit dibandingkan kelas positif dan negatif. Akibatnya, model cenderung bias terhadap kelas mayoritas sehingga kemampuan klasifikasi pada kelas minoritas menjadi menurun.

Pada skenario kedua dilakukan proses resampling menggunakan teknik oversampling untuk menyeimbangkan jumlah data pada setiap kelas sentimen sebelum dilakukan pelatihan model:

Tabel 2

Sentimen	Precisi on	Recall	F1-Score	Support
Negatif	0.88	0.85	0.86	120
Netral	0.85	0.86	0.86	121
Positif	0.95	0.98	0.96	121
Accuracy			0.9	362
Macro Avg	0.89	0.89	0.89	362
Weighted Avg	0.89	0.9	0.89	362

Berdasarkan Tabel 2, model SVM menggunakan resampling memperoleh nilai accuracy sebesar 0.90, lebih tinggi dibandingkan model tanpa resampling. Selain peningkatan akurasi, seluruh nilai precision, recall, dan F1-score pada setiap kelas juga mengalami peningkatan yang signifikan.

Performa model pada kelas netral meningkat secara drastis dengan nilai recall sebesar 0.86 dan F1-score sebesar 0.87. Hal ini menunjukkan bahwa teknik resampling berhasil membantu model dalam mempelajari pola pada kelas minoritas secara lebih baik.

Peningkatan performa tersebut menunjukkan bahwa penanganan imbalance data memiliki pengaruh yang sangat penting dalam proses klasifikasi sentimen. Dengan distribusi data yang lebih seimbang, model SVM

mampu menghasilkan performa klasifikasi yang lebih stabil dan optimal pada seluruh kategori sentimen.

Tahap evaluasi dilakukan untuk mengukur kinerja model klasifikasi yang telah dibangun menggunakan algoritma Support Vector Machine (SVM) pada dua skenario pengujian, yaitu tanpa resampling dan menggunakan resampling. Evaluasi dilakukan untuk mengetahui pengaruh teknik resampling terhadap kemampuan model dalam mengklasifikasikan sentimen positif, negatif, dan netral secara lebih akurat dan seimbang.

Metrik evaluasi yang digunakan dalam penelitian ini meliputi accuracy, precision, recall, dan F1-score. Accuracy digunakan untuk mengukur tingkat ketepatan model secara keseluruhan, sedangkan precision, recall, dan F1-score digunakan untuk mengevaluasi performa model pada masing-masing kelas sentimen. Selain itu, evaluasi juga dilakukan menggunakan confusion matrix untuk mengetahui jumlah prediksi benar dan salah pada setiap kategori sentimen.

Tabel 3 Perbandingan Hasil Evaluasi Model

Model	Accuracy	Precision	Recall	F1-Score
SVM tanpa resampling	0.73	0.67\	0.73	0.69
SVM + Resampling	0.9	0.89	0.89	0.89

Berdasarkan Tabel 3 dapat diketahui bahwa model Support Vector Machine (SVM) menggunakan resampling memiliki performa yang lebih baik dibandingkan model SVM tanpa resampling. Hal ini terlihat dari peningkatan nilai accuracy dari 0.73 menjadi 0.90 setelah diterapkan teknik resampling.

Selain itu, nilai precision, recall, dan F1-score pada model SVM menggunakan resampling juga mengalami peningkatan yang signifikan dan lebih seimbang pada seluruh kelas sentimen. Kondisi ini menunjukkan bahwa teknik resampling berhasil membantu model dalam mengenali pola data pada kelas minoritas yang sebelumnya sulit diklasifikasikan pada model tanpa resampling.

Perbedaan performa paling terlihat pada kelas netral. Pada model tanpa resampling, nilai recall kelas netral sangat rendah karena model cenderung bias terhadap kelas mayoritas. Setelah dilakukan proses oversampling, distribusi data

Wordcloud Secara Keseluruhan Visualisasi ini menampilkan kumpulan kata dari seluruh dataset tanpa memandang label sentimen. Tujuannya adalah untuk melihat topik utama atau entitas yang paling sering dibicarakan oleh pengguna secara umum dalam konteks ulasan yang diteliti.



Visualisasi ini hanya mencakup kata-kata dari ulasan yang berlabel sentimen positif. Fokus utamanya adalah mengidentifikasi kata sifat atau kata kerja yang merepresentasikan kepuasan pengguna, kelebihan produk/layanan, atau apresiasi yang diberikan.



Visualisasi ini menampilkan kata-kata yang berasal dari ulasan berlabel sentimen negatif. Hal ini sangat berguna bagi peneliti untuk memetakan poin-poin keluhan, kekurangan, atau aspek yang menyebabkan ketidakpuasan pengguna.



Visualisasi ini merangkum kata-kata dari ulasan yang bersifat objektif atau tidak memiliki kecenderungan emosional yang kuat. Biasanya, kata-kata yang muncul berkaitan dengan fakta penggunaan atau informasi umum tanpa unsur pujian maupun kritikan.



Gambar 5 Wordcloud Netral

CONCLUSION

Berdasarkan hasil penelitian mengenai optimalisasi algoritma Support Vector Machine (SVM) menggunakan teknik resampling pada analisis sentimen ulasan aplikasi My Suzuki di Google Play Store, maka diperoleh beberapa kesimpulan sebagai berikut:

Penelitian ini berhasil mengimplementasikan analisis sentimen terhadap ulasan pengguna aplikasi My Suzuki menggunakan algoritma Support Vector Machine (SVM). Proses analisis dilakukan melalui beberapa tahapan, yaitu scraping data ulasan dari Google Play Store, preprocessing teks, pelabelan manual oleh ahli bahasa, penanganan imbalance data menggunakan teknik oversampling, transformasi fitur menggunakan TF-IDF, serta proses klasifikasi menggunakan algoritma SVM. Dataset yang digunakan terdiri dari 1.227 ulasan pengguna yang diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

Penggunaan teknik resampling terbukti memberikan pengaruh yang signifikan terhadap peningkatan performa klasifikasi sentimen menggunakan algoritma SVM. Pada model tanpa resampling, algoritma SVM hanya memperoleh accuracy sebesar 73% dan mengalami kesulitan dalam mengenali kelas minoritas, khususnya sentimen netral yang memiliki nilai recall sangat rendah. Setelah dilakukan proses oversampling untuk menyeimbangkan distribusi data, performa model meningkat secara signifikan dengan accuracy mencapai 90%. Selain itu, nilai precision, recall, dan F1-score pada seluruh kelas juga mengalami peningkatan yang lebih seimbang. Hasil tersebut menunjukkan bahwa penanganan imbalance data sangat penting dalam meningkatkan kemampuan model machine learning pada klasifikasi teks.

Berdasarkan hasil evaluasi menggunakan accuracy, precision, recall, dan F1-score, model SVM menggunakan resampling memiliki performa yang lebih baik dibandingkan model SVM tanpa resampling. Teknik oversampling berhasil membantu model dalam mempelajari

pola data pada kelas minoritas sehingga proses klasifikasi menjadi lebih optimal dan tidak bias terhadap kelas mayoritas. Dengan demikian, kombinasi algoritma Support Vector Machine (SVM) dan teknik resampling dapat dinyatakan sebagai pendekatan yang efektif dan optimal dalam analisis sentimen ulasan aplikasi My Suzuki di Google Play Store. Selain itu, hasil visualisasi WordCloud menunjukkan bahwa sebagian besar sentimen negatif pengguna berkaitan dengan kendala layanan, pengiriman suku cadang, dan performa aplikasi, sedangkan sentimen positif didominasi oleh kemudahan penggunaan aplikasi dan kemudahan akses layanan servis kendaraan.

REFERENCE

- [1] M. Baihaqi, L. Magdalena, and R. Fahrudin, "Analisis Sentimen Aplikasi Deepseek Menggunakan Metode Naive Bayes dan Support Vector Machine," *Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 3, 2025, doi: 10.31004/riggs.v4i3.2511.
- [2] M. Safrudin, M. Martanto, and U. Hayati, "Perbandingan kinerja naïve bayes dan support vector machine untuk klasifikasi sentimen ulasan game genshin impact," *JATI (Jurnal Mahasiswa Teknik Informatika)*, 2024, doi: 10.36040/jati.v8i3.8415.
- [3] M. Lamonda, E. Budiarto, and H. Rahendra Herlianto, "Perbandingan metode support vector machine (svm) dan multinomial naive bayes (mnb) pada analisis sentimen pengguna aplikasi shopee di google play store terkait fitur cash on delivery (cod)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 5, 2025, doi: 10.36040/jati.v9i5.14822.
- [4] M. A. Java, M. Syafrullah, and W. Windarto, "Analisis Sentimen Ulasan Pengguna Aplikasi Threads pada Google Play Store Menggunakan Multinomial Naive Bayes dan Support Vector Machine," *TICOM*, vol. 12, no. 2, 2024, doi: 10.70309/ticom.v12i2.112.