Jurnal Komtika (Komputasi dan Informatika),
Vol. 9 No. 2 | Nov 2025

Implementasi Support Vector Machine dan Resampling dalam Analisis Ulasan Pengguna Google Maps

Umi Khultsum[1*], Eka Rahmawati[2], Annida Purnamawati[3], Bara Rifki Annajib[4], Christina Yuli Anggita[5]
[1,2,3,4,5]Sistem Informasi / Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
*email: umikhultsum.ukm@bsi.ac.id

DOI: https://doi.org/10.31603/komtika.v9i2.14813

Received: 16-09-2025, Revised: 31-10- 2025, Accepted: 04-11-2025

ABSTRACT

The development of information technology has driven the increasing use of digital services such as Google Maps, which functions not only as a navigation tool but also as a platform for users to provide reviews. These reviews serve as an important data source for sentiment analysis; however, they are often unstructured and contain noise. This study aims to conduct sentiment analysis using the Support Vector Machine (SVM) model with the application of resampling techniques to address data imbalance issues in user reviews of the Google Maps application. A total of 1,000 recent reviews were collected through a scraping process, followed by data cleaning (lowercasing, stopwords removal, stemming, and lemmatization) and data preprocessing. The SVM model combined with resampling techniques was then implemented and evaluated using accuracy, precision, and recall metrics. The results indicate that the SVM model achieved an accuracy of 81%, with a weighted average precision of 0.79, recall of 0.81, and F1-score of 0.76. These findings demonstrate that applying resampling techniques to SVM yields good performance in sentiment classification. The study is expected to contribute to the development of sentiment analysis methods using the SVM model with resampling in the context of Google Maps reviews.
Keywords: *Sentiment Analysis, Google Maps Reviews, Support Vector Machine, Resampling*

ABSTRAK

Perkembangan teknologi informasi mendorong meningkatnya penggunaan layanan digital seperti Google Maps, yang bukan sekadar berperan sebagai sarana penunjuk arah, melainkan juga menjadi wadah bagi pengguna dalam menyampaikan ulasan. Ulasan tersebut menjadi sumber data penting dalam analisis sentimen, namun sering kali tidak terstruktur dan mengandung *noise*. Penelitian ini mempunyai tujuan untuk melakukan analisis menggunakan model Support Vector Machine (SVM) dengan penerapan teknik *resampling* untuk mengatasi masalah ketidakseimbangan data terhadap sentiment ulasan terhadap applikasi Google Maps. Data penelitian sebanyak 1.000 ulasan yang paling baru didapatkan dengan teknik *scraping*, yang kemudian melalui tahapan *data cleaning* (lowercase, *stopwords removal*, stemming, dan lemmatization) serta *data pre-processing*. Model SVM dan teknik *resampling* kemudian diimplementasikan serta dievaluasi dengan metrik akurasi, presisi, dan recall. Hasil penelitian menunjukkan model SVM mampu mencapai akurasi 81%, dengan nilai presisi rata-rata tertimbang 0,79, recall 0,81, dan F1-score 0,76. Hasil penelitian menunjukkan penerapan teknik *resampling* pada SVM memiliki performa yang baik untuk klasifikasi sentimen. Penelitian yang dihasilkan diharapkan memberikan kontribusi pada pengembangan metode analisis sentimen menggunakan model SVM dengan penerapan resampling dalam konteks ulasan applikasi Google Maps.
Kata Kunci : Analisis Sentimen, Google Maps Reviews, Support Vector Machine, Resampling

PENDAHULUAN

Perubahan signifikan dalam teknologi informasi dan komunikasi telah memengaruhi bagaimana masyarakat menggunakan layanan digital dan mengakses berbagai informasi. Salah satu platform digital yang banyak digunakan adalah Google Maps, yang tidak hanya berfungsi sebagai alat navigasi, tetapi juga sebagai media bagi pengguna untuk memberikan

ulasan tentang tempat dan layanan. Ulasan-ulasan ini menjadi sumber data yang sangat berharga, baik bagi penyedia layanan untuk meningkatkan kualitas, maupun bagi pengguna lain untuk membuat keputusan yang lebih tepat. Banyaknya pengguna dari aplikasi tersebut tentu memberikan *feedback* untuk pengembangan aplikasi. Ulasan pengguna ini menjadi sumber data penting bagi bisnis untuk meningkatkan kualitas layanan dan bagi pengguna lain untuk membuat keputusan yang lebih baik[1]. Namun, ulasan yang sering kali tidak terstruktur dan mengandung noise ini dapat menjadi tantangan dalam analisis yang efektif[2]. *Support Vector Machine* (SVM) merupakan algoritma klasifikasi yang banyak digunakan dalam analisis teks, khususnya pada analisis sentimen ulasan pengguna[3]. SVM dikenal unggul dalam menangani data berukuran tinggi dan tetap memberikan performa yang stabil di berbagai bidang penerapan[4].

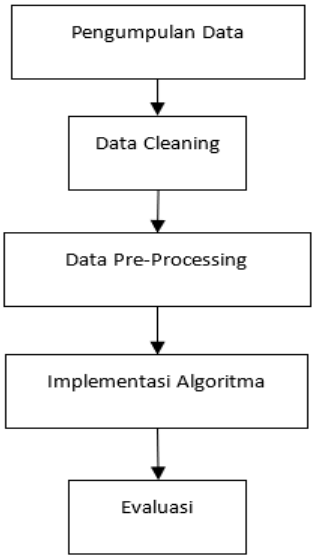
Ketidakseimbangan jumlah ulasan positif dan negatif menjadi tantangan yang sering dihadapi oleh model SVM. Ketidakseimbangan ini dapat menyebabkan bias dalam model, di mana prediksi lebih condong ke kelas yang dominan[5]. Salah satu cara mengatasi ketidakseimbangan data adalah dengan menerapkan teknik *resampling*. Metode ini meliputi oversampling pada kelas minoritas dengan tujuan menyeimbangkan distribusi data agar model klasifikasi seperti SVM dapat belajar lebih optimal serta menghasilkan prediksi yang akurat[6]. Implementasi teknik resampling diharapkan dapat meningkatkan kinerja SVM dalam analisis ulasan pengguna Google Maps dengan mengurangi bias yang disebabkan oleh ketidakseimbangan data. Pendekatan yang diusulkan dalam penelitian ini adalah penerapan teknik *resampling* guna mencapai distribusi data yang seimbang sebelum tahap pelatihan model SVM dimulai. Pendekatan tersebut memungkinkan algoritma untuk belajar lebih merata pada seluruh kelas sehingga mampu menghasilkan klasifikasi dengan tingkat akurasi yang lebih tinggi. Meskipun demikian, studi mengenai analisis sentimen pada ulasan pengguna Google Maps dengan sifat uniknya masih relatif jarang ditemukan. Penelitian ini mengisi celah tersebut dengan menerapkan berbagai teknik *resampling* dan mengevaluasi dampaknya terhadap kinerja model SVM secara menyeluruh melalui metrik akurasi, presisi, dan recall. Oleh karena itu, penelitian ini menawarkan kebaruan dengan fokus pada penerapan teknik resampling dalam konteks spesifik ini, serta evaluasi dampaknya terhadap kinerja algoritma SVM.

Berbagai penelitian telah mengkaji metode untuk mengatasi ketidakseimbangan data, termasuk penggunaan model klasifikasi yang lebih kompleks atau penerapan teknik resampling. Sebagai contoh, pengembangan *Synthetic Minority Over-sampling Technique* (SMOTE) yang telah menunjukkan peningkatan kinerja dalam berbagai domain[7]. Selain itu, SVM dan SMOTE juga sudah diimplementasikan dalam sentimen analisis kenaikan harga BBM[8]. Perbandingan teknik resampling juga telah dilakukan untuk memprediksi kinerja siswa dengan menggunakan model machine learning[9]. *State of the art* menunjukkan bahwa teknik resampling seperti SMOTE (*Synthetic Minority Over-sampling Technique*) telah diimplementasikan oleh sejumlah studi guna mengatasi masalah ketidakseimbangan data pada berbagai domain, seperti prediksi kinerja siswa, klasifikasi penyakit, serta analisis sentimen umum. Kombinasi SVM dan SMOTE juga telah diterapkan dalam kasus spesifik seperti analisis sentimen kenaikan harga BBM. Meskipun demikian, konteks analisis sentimen pada ulasan pengguna Google Maps yang memiliki karakteristik unik belum banyak diteliti. Penelitian ini mengisi celah tersebut dengan menerapkan berbagai teknik resampling dan

mengevaluasi dampaknya terhadap kinerja model SVM secara menyeluruh melalui metrik akurasi, presisi, dan recall. Oleh karena itu, penelitian ini menawarkan kebaruan dengan fokus pada penerapan teknik resampling dalam konteks spesifik ini, serta evaluasi dampaknya terhadap kinerja model SVM.

METODE

Bagian ini menunjukkan metodologi yang digunakan pada penelitian. Adapun tahapannya terdiri dari pengumpulan data, pembersihan, praproses, implementasi algoritma, dan evaluasi. Berikut alur metode penelitian dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

1. Pengumpulan Data
   Data ulasan pengguna Google Maps akan dikumpulkan menggunakan teknik web scraping. Pengumpulan data meliputi teks ulasan, rating bintang, tanggal ulasan, dan metadata lainnya.
2. Data *Cleaning*
   Pembersihan teks (*lowering case folding*, menghilangkan *stopwords*, *lemmatization* dan *stemming*). *Lowering case folding* merupakan tahapan mengubah seluruh karakter dalam teks menjadi huruf kecil[10]. Proses ini bertujuan agar kata dengan bentuk huruf berbeda, seperti ‘Data’ dan ‘data’, diperlakukan sama dalam analisis teks. Menghilangkan *stopwords* adalah proses mengeliminasi kata-kata umum yang kerap muncul dalam teks namun tidak memberikan banyak informasi bagi analisis[11]. roses penghapusan stopwords dilakukan untuk mengurangi gangguan (noise) dan memusatkan perhatian pada kata-kata penting dalam analisis teks. Adapun lemmatization adalah tahap normalisasi kata dengan mengembalikannya ke bentuk dasar atau lema.. Proses ini mempertimbangkan konteks morfologi dan gramatikal sehingga kata kerja seperti "running", "ran", dan "runs" diubah menjadi bentuk dasar "run". *Lemmatization* membantu dalam mengurangi variasi kata dan mengkonsolidasikan informasi untuk analisis lebih lanjut[12]. *Stemming* menjadi langkah terakhir pada proses pembersihan data, yang dilakukan dengan mengubah kata menjadi bentuk akar (stem) melalui

penghilangan awalan dan akhiran[13]. Berbeda dengan lemmatization, stemming tidak mempertimbangkan konteks morfologi dan gramatikal, sehingga sering menghasilkan bentuk kata yang tidak lazim dalam bahasa alami.

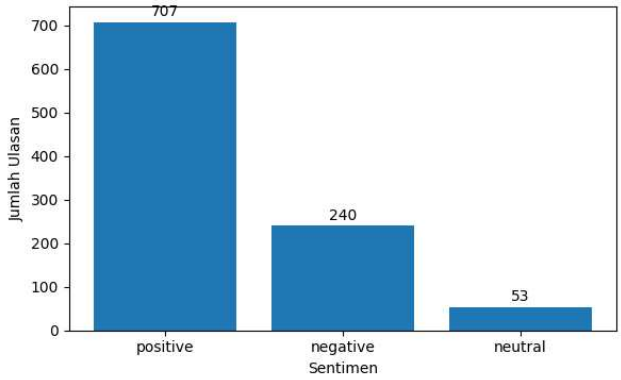
3. Data *Pre-Processing*
   Pada tahap ini diterapkan metode resampling dengan menggunakan teknik *Oversampling* yang merupakan metode penyeimbangan data yang bertujuan mengatasi ketidakseimbangan kelas dalam suatu dataset[14].
4. Implementasi Algoritma
   Tahap untuk mengimplementasikan model SVM dengan teknik *resampling*. *Resampling* dilakukan menggunakan metode *oversampling* untuk menyeimbangkan jumlah data antar kelas. Proses ini dilakukan sebelum tahap pelatihan SVM, di mana data dari kelas minoritas (ulasan negatif dan netral) direplikasi secara sintetik menggunakan pendekatan acak terkontrol agar distribusi data lebih proporsional. Langkah ini bertujuan agar model SVM tidak bias terhadap kelas dominan (positif).
5. Evaluasi
   Evaluasi dilakukan dengan beberapa metrik yaitu[15]:
   a. Akurasi: Persentase prediksi yang sesuai dengan hasil sebenarnya dibandingkan dengan total prediksi yang dilakukan.
   b. Presisi: Persentase prediksi positif yang benar dari keseluruhan prediksi yang dikategorikan positif.
   c. *Recall*: Persentase data positif yang berhasil diprediksi dengan benar dari seluruh data positif yang ada.

HASIL DAN PEMBAHASAN

Scrapping data menjadi tahap awal yang dilakukan sebelum penelitian dilaksanakan. Perolehan data berjumlah 1000 ulasan terbaru terhadap aplikasi Google Maps. Setelah tahap pembersihan dan tokenisasi, dataset dibagi menjadi data latih dan uji (80:20). Pada data latih dilakukan *oversampling* dengan fungsi *RandomOverSampler()* dari pustaka imblearn, kemudian model SVM dilatih menggunakan kernel linear. Proses ini dibandingkan dengan model SVM tanpa resampling untuk mengukur peningkatan performa. Labeling dilakukan berdasarkan rating yang diberikan pengguna. Pedoman klasifikasi terdapat pada Tabel 1. Hasil pelabelan terdapat pada Gambar 2.

Tabel 1. Klasifikasi Sentimen

Rating	Label
1-2	Negative
3	Neutral
4-5	Positive



Gambar 2. Hasil Distribusi Sentimen

Gambar 2 menunjukan hasil dari klasifikasi sentimen. Komentar positif berjumlah 707, negative 240 dan 53 netral. Pada tahapan data cleaning, dilakukan beberapa prosedur yaitu lowercase, stopwords removal, stemming dan lemmatization.

0	tingkatkan lagi, supaya pengguna tidak malah n...
1	Masi lelet
2	okay
3	bagus
4	maps tidak akurat

	cleaned_content
0	tingkatkan lagi supaya pengguna tidak malah na...
1	masi lelet
2	okay
3	bagus
4	maps tidak akurat

Gambar 3. Cleaned Content

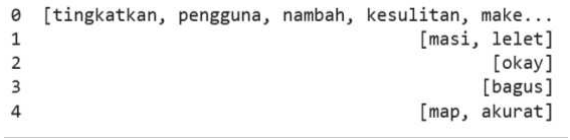
Gambar 3 menunjukkan perbandingan antara teks ulasan asli dengan hasil pembersihan awal menggunakan tahap *data cleaning*. Pada kolom teks asli, ulasan pengguna masih ditampilkan dalam bentuk mentah, misalnya “tingkatkan lagi, supaya pengguna tidak malah n...”, “Masi lelet”, “okay”, “bagus”, dan “maps tidak akurat”. Teks tersebut masih mengandung huruf kapital, tanda baca, serta bentuk penulisan yang tidak seragam. Setelah melalui proses *data cleaning*, hasil pada kolom cleaned_content tampak lebih rapi dan konsisten. Tahap pembersihan data ini meliputi sejumlah langkah, seperti mengonversi semua huruf ke bentuk kecil (*lowercasing*), menghapus tanda baca maupun karakter yang tidak diperlukan, serta melakukan normalisasi sederhana agar teks lebih konsisten.. Sebagai contoh, ulasan “Masi lelet” berubah menjadi “masi lelet”, sementara tanda baca koma pada kalimat panjang dihilangkan agar kalimat menjadi lebih sederhana

0	tingkatkan lagi supaya pengguna tidak malah na...
1	masi lelet
2	okay
3	bagus
4	maps tidak akurat

	tokens
0	[tingkatkan, pengguna, nambah, kesulitan, make...
1	[masi, lelet]
2	[okay]
3	[bagus]
4	[map, akurat]

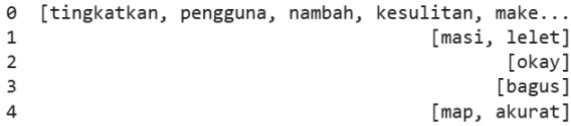
Gambar 4. Tokenization

Gambar 4 memperlihatkan hasil pemecahan teks ulasan menjadi unit kata atau token. Pada kolom ulasan yang telah dibersihkan sebelumnya, setiap kalimat masih berbentuk teks utuh, seperti “tingkatkan lagi supaya pengguna tidak malah na...”, “masi lelet”, “okay”, “bagus”, dan “maps tidak akurat”. Setelah melalui proses tokenisasi, teks tersebut diubah menjadi daftar kata terpisah. Misalnya, kalimat panjang dipecah menjadi token seperti [“tingkatkan”, “pengguna”, “nambah”, “kesulitan”, “make...”], sedangkan ulasan singkat seperti “masi lelet” diubah menjadi [“masi”, “lelet”], dan “maps tidak akurat” menjadi [“map”, “akurat”]. Proses tokenisasi ini sangat penting karena memecah teks menjadi bagian terkecil yang dapat dianalisis oleh algoritma.



Gambar 5. Stopwords Removal

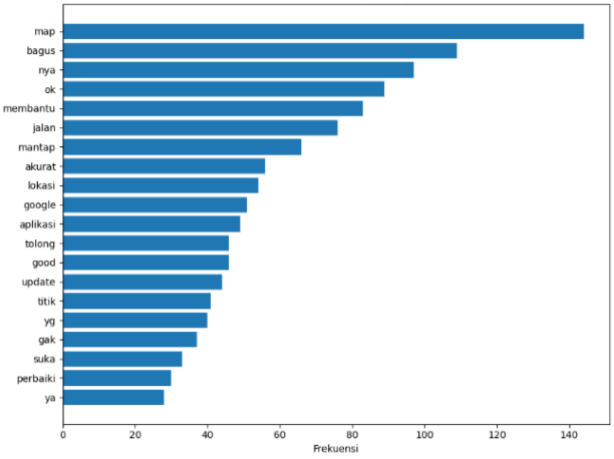
Gambar 5 menampilkan hasil pembersihan kata-kata yang dianggap tidak memiliki makna penting dalam analisis, atau disebut *stopwords*. Tahap tokenisasi sebelumnya menghasilkan teks yang masih mengandung kata sambung atau kata bantu, seperti ‘yang’, ‘dan’, ‘tidak’, yang tidak terlalu berpengaruh dalam analisis sentimen. Setelah melalui proses *stopwords removal*, daftar token menjadi lebih ringkas dan berfokus pada kata-kata bermakna utama. Sebagai contoh, kalimat panjang yang semula menghasilkan banyak token kini disaring sehingga hanya menyisakan kata penting seperti “tingkatkan”, “pengguna”, “nambah”, “kesulitan”, dan “make...”. Ulasan singkat seperti “masi lelet”, “okay”, “bagus”, dan “map akurat” juga tetap dipertahankan karena kata-kata tersebut mengandung informasi sentimen yang relevan. Tahap ini penting untuk mengurangi *noise* dalam data.



Gambar 6. Lemitization

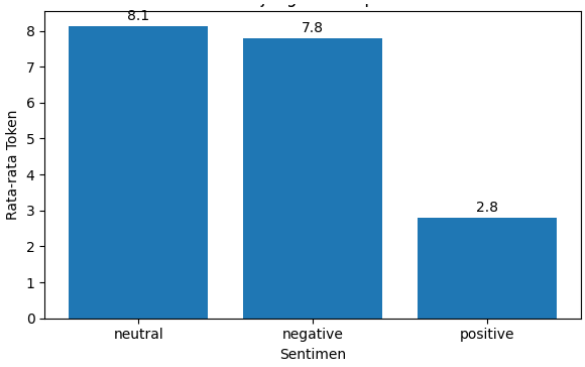
Gambar 6 memperlihatkan hasil pengolahan kata setelah melalui tahap normalisasi ke bentuk dasar atau lema (*lemma*). Pada tahap sebelumnya, data sudah melewati proses tokenisasi dan *stopwords removal*, namun masih terdapat variasi kata yang berbeda bentuk meskipun memiliki makna yang sama. Melalui lemmatization, kata-kata diubah ke bentuk dasarnya dengan mempertimbangkan struktur morfologi dan makna semantik. Misalnya, kata “nambah” akan dikembalikan ke bentuk dasar “tambah”, sedangkan kata lain seperti “masi lelet”, “okay”, “bagus”, dan “map akurat” tetap dipertahankan karena sudah berada dalam bentuk yang sesuai. Proses ini berbeda dengan *stemming* yang cenderung memangkas imbuhan secara mekanis, sehingga lemmatization menghasilkan bentuk kata yang lebih tepat dan alami dalam bahasa. Tahap ini penting agar distribusi kata menjadi lebih konsisten,

memudahkan algoritma pembelajaran mesin dalam mengenali pola, serta meningkatkan akurasi pada analisis sentimen.



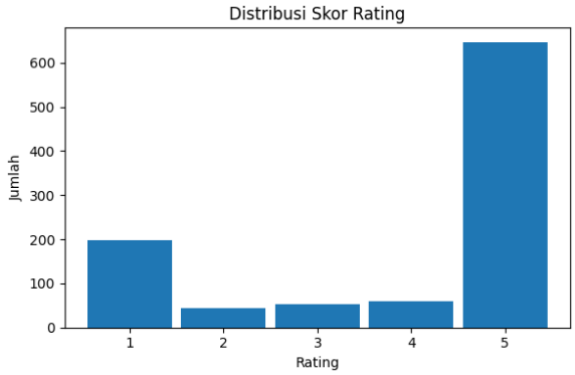
Gambar 7. Kata Paling Sering

Gambar 7 menunjukkan kata yang paling sering digunakan dalam pemberian ulasan Google Map. Kata pertama adalah map yang merujuk pada aplikasi Google Map dan kata setelahnya adalah bagus yang merujuk pada komentar positif. Dari 20 kata yang paling sering digunakan dalam ulasan, hanya 1 kata yang berkonotasi negatif yaitu kata ‘gak’ yang berarti tidak.



Gambar 8. Rata-rata panjang ulasan per sentimen

Gambar 8 menunjukkan perbandingan rata-rata panjang ulasan berdasarkan kategori sentimen. Ulasan dengan sentimen netral memiliki rata-rata panjang paling tinggi yaitu 8,1 token, diikuti oleh ulasan negatif dengan rata-rata 7,8 token. Sementara itu, ulasan dengan sentimen positif cenderung jauh lebih singkat dengan rata-rata hanya 2,8 token. Hal ini mengindikasikan bahwa pengguna lebih banyak menuliskan penjelasan panjang ketika menyampaikan opini netral atau keluhan (negatif), sedangkan ulasan positif biasanya disampaikan secara singkat dan padat.



Gambar 9. Distribusi Skor Rating

Gambar 9 menunjukkan bahwa sebagian besar pengguna memberikan penilaian dengan skor 5, yaitu lebih dari 600 ulasan, sehingga menjadi kategori yang paling dominan. Sementara itu, skor 1 juga cukup banyak diberikan, sekitar 200 ulasan. Adapun skor 2, 3, dan 4 memiliki jumlah ulasan yang relatif rendah dan hampir seimbang, yaitu di bawah 100 ulasan masing-masing. Hal ini mengindikasikan bahwa mayoritas pengguna merasa sangat puas (skor 5), meskipun masih terdapat sejumlah pengguna yang sangat tidak puas (skor 1).

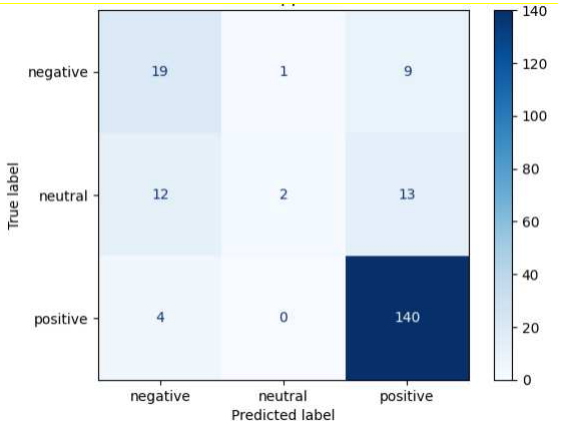
Setelah melalui prosedur data *cleaning*, tahap selanjutnya adalah implementasi algoritma. Pada tahap ini metode *resampling* diimplementasikan dengan algoritma *machine learning*.

Tabel 2. Hasil Evaluasi Model

Kelas Sentimen	Metode	Precision	Recall	F1-Score
Negative	SVM	0.69	0.55	0.61
	SVM+Resampling	0.62	0.57	0.60
Neutral	SVM	0.40	0.17	0.24
	SVM+Resampling	0.29	0.42	0.34
Positive	SVM	0.87	0.95	0.91
	SVM+Resampling	0.91	0.90	0.90
Accuracy	SVM	—	—	0.825
	SVM+Resampling	—	—	0.805
Macro Average	SVM	0.65	0.56	0.58
	SVM+Resampling	0.61	0.63	0.62
Weighted Average	SVM	0.80	0.82	0.81
	SVM+Resampling	0.82	0.81	0.81

Tabel 2 menampilkan hasil evaluasi performa model Support Vector Machine (SVM) dan SVM dengan Resampling (Oversampling) berdasarkan nilai precision, recall, dan F1-score pada masing-masing kelas sentimen. Secara umum, model SVM tanpa resampling menunjukkan performa yang lebih baik pada kelas positive, dengan nilai F1-score tertinggi sebesar 0.91, sedangkan kelas neutral memiliki kinerja terendah dengan F1-score 0.24. Hal ini menunjukkan bahwa model lebih mudah mengenali ulasan positif yang jumlahnya dominan dibandingkan dengan ulasan netral yang relatif sedikit. Setelah penerapan resampling, terjadi peningkatan pada recall kelas neutral (dari 0.17 menjadi 0.42), yang menunjukkan bahwa teknik oversampling mampu meningkatkan sensitivitas model terhadap kelas minoritas. Namun demikian, peningkatan tersebut tidak diikuti oleh peningkatan signifikan pada akurasi

keseluruhan. Nilai akurasi model SVM dengan resampling (0.805) sedikit lebih rendah dibandingkan model tanpa resampling (0.825). Perbedaan ini menunjukkan bahwa teknik resampling tidak selalu memberikan peningkatan performa, terutama pada dataset dengan distribusi alami yang sangat didominasi oleh satu kelas (positif). Model SVM tanpa resampling lebih stabil dalam mengklasifikasikan data sesuai pola dominan, sementara resampling berpotensi memperkenalkan duplikasi data yang tidak menambah informasi baru dan dapat menyebabkan overfitting pada kelas minoritas.



Gambar 11. Confusion Matrix

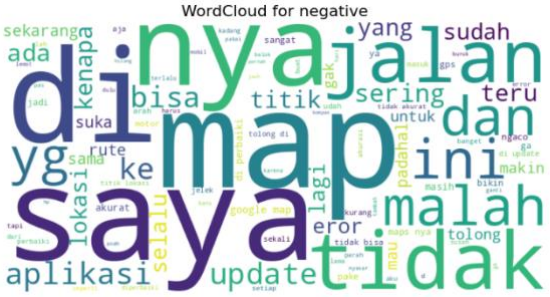
Gambar 11 memperlihatkan distribusi hasil prediksi model pada data uji untuk tiga kelas sentimen, yaitu negatif, netral, dan positif. Pada kelas negatif, dari total 29 data, sebanyak 19 ulasan berhasil terklasifikasi dengan benar, sedangkan 9 ulasan salah diprediksi sebagai positif dan 1 ulasan sebagai netral, terlihat kelemahan model yang signifikan: dari 27 data, hanya 2 ulasan yang diklasifikasikan dengan benar, sedangkan 12 salah dikategorikan sebagai negatif dan 13 sebagai positif. Hal ini mengonfirmasi bahwa model kesulitan dalam mengenali sentimen netral akibat jumlah data yang relatif sedikit. Untuk kelas positif, model justru menunjukkan akurasi tinggi, di mana 140 dari 144 ulasan berhasil dikategorikan dengan tepat, sedangkan 4 ulasan lainnya salah diklasifikasikan sebagai negative. Hasil confusion matrix ini selaras dengan laporan klasifikasi sebelumnya, di mana kelas positif mendominasi performa model, sedangkan kelas netral menjadi titik lemah. Secara keseluruhan, meskipun SVM efektif dalam mengenali ulasan positif.

Distribusi kata yang paling dominan pada tiap label dapat dilihat pada Gambar 12, Gambar 13, dan Gambar 14.

Gambar 12. Kata Sering Muncul pada Label Positif

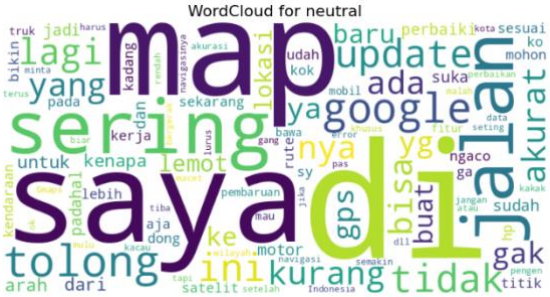
Gambar 12 menunjukkan *WordCloud* untuk ulasan berlabel positif. Kata yang mendominasi antara lain bagus, mantap, ok, baik, dan membantu. Hal ini mencerminkan

bahwa pengguna cenderung mengekspresikan kepuasan dan pengalaman yang menyenangkan dengan menggunakan kata-kata bernuansa apresiasi.



Gambar 13. Kata Sering Muncul pada Label Negatif

WordCloud pada gambar 13 menggambarkan ulasan berlabel negatif. Beberapa kata yang sering muncul antara lain tidak, map, saya, aplikasi, dan update. Temuan ini mengindikasikan bahwa sebagian besar keluhan pengguna berkaitan dengan fungsi aplikasi, proses pembaruan, serta pengalaman penggunaan yang tidak sesuai dengan ekspektasi.



Gambar 14. Kata Sering Muncul pada Label Netral

Gambar 14 memperlihatkan WordCloud untuk ulasan netral. Kata-kata yang dominan adalah map, saya, sering, tolong, akurasi, dan update. Ulasan netral umumnya berisi deskripsi atau permintaan pengguna tanpa muatan emosional yang kuat, baik positif maupun negatif, sehingga kata-katanya lebih informatif.

Berdasarkan *WordCloud*, ulasan berlabel positif didominasi kata bernuansa apresiasi seperti bagus dan mantap, ulasan negatif banyak memuat keluhan seperti tidak dan update, sedangkan ulasan netral lebih menekankan deskripsi atau permintaan seperti map dan tolong. Hal ini menunjukkan bahwa ulasan positif berfokus pada kepuasan layanan, negatif menyoroti masalah teknis, sementara netral cenderung informatif tanpa emosi yang kuat.

KESIMPULAN

Penelitian ini membuktikan bahwa model SVM dengan penerapan teknik resampling mampu memberikan kinerja yang cukup baik dalam analisis sentimen ulasan pengguna Google Maps. Model yang dibangun berhasil mencapai akurasi sebesar 81% dengan nilai presisi rata-rata tertimbang 0,79, recall 0,81, dan F1-score 0,76. Hasil ini menunjukkan bahwa SVM sangat andal dalam mengenali sentimen positif, cukup moderat dalam mengenali sentimen negatif, namun masih lemah dalam mengklasifikasikan sentimen netral akibat keterbatasan jumlah data. Analisis WordCloud turut menguatkan temuan tersebut, di mana

ulasan positif didominasi oleh kata-kata bernuansa apresiasi, ulasan negatif banyak memuat keluhan terkait fungsi aplikasi maupun pembaruan, sedangkan ulasan netral lebih menekankan deskripsi atau permintaan yang informatif.

Berdasarkan temuan ini, penelitian selanjutnya disarankan untuk memperluas jumlah dataset khususnya pada kelas netral agar distribusi data lebih seimbang, menerapkan teknik resampling lanjutan seperti SMOTE atau ADASYN untuk meningkatkan performa pada kelas minoritas, serta mengeksplorasi model lain seperti Random Forest, XGBoost, atau Deep Learning (CNN/LSTM) sebagai pembanding kinerja dengan SVM. Selain itu, pendekatan semantik berbasis word embeddings seperti Word2Vec, GloVe, atau BERT dapat dipertimbangkan untuk meningkatkan pemahaman konteks kata dalam analisis, sehingga hasil klasifikasi menjadi lebih akurat. Ke depan, penelitian ini juga dapat dikembangkan dalam bentuk aplikasi praktis berbasis web atau mobile yang mampu melakukan analisis sentimen secara real-time, sehingga hasilnya dapat dimanfaatkan secara langsung untuk mendukung pengambilan keputusan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Bina Sarana Informatika atas dukungan pendanaan yang diberikan sehingga penelitian ini dapat terlaksana.

DAFTAR PUSTAKA

[1] L. Nita, K. Pasi, And B. Sudaryanto, “Analisis Pengaruh Online Customer Reviews Dan Kualitas Pelayanan Terhadap Keputusan Pembelian Dengan Kepercayaan Sebagai Variabel Intervening (Studi Pada Konsumen Shopee di Kota Semarang),” *Diponegoro Journal of Management*, Vol. 10, No. 3, 2021, [Online]. Available: Http://Ejournal-S1.Undip.Ac.Id/Index.Php/Dbr

[2] M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad, And G. S. Choi, “Gbsvm: Sentiment Classification from Unstructured Reviews Using Ensemble Classifier,” *Applied Sciences (Switzerland)*, Vol. 10, No. 8, Apr. 2020, Doi: 10.3390/App10082788.

[3] A. Purnamawati, W. Nugroho, D. Putri, And W. F. Hidayat, “Infotekjar: Jurnal Nasional Informatika Dan Deteksi Penyakit Daun Pada Tanaman Padi Menggunakan Algoritma Decision Tree , Random Forest , Naïve Bayes , Svm Dan Knn,” Vol. 1, 2020.

[4] I. S. Aisah, B. Irawan, And T. Suprapti, “Algoritma Support Vector Machine (Svm) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur’an Digital,” 2023.

[5] I. W. Dharmana, I. G. A. Gunadi, And L. J. E. Dewi, “Deteksi Transaksi *Fraud* Kartu Kredit Menggunankan *Oversampling* Adasyn Dan Seleksi Fitur Svm-Rfecv,” *Jurnal Teknologi Informasi Dan Ilmu Komputer*, Vol. 11, No. 1, Pp. 125–134, Feb. 2024, Doi: 10.25126/Jtiik.20241117640.

[6] G. N. A. Kusuma, I. M. Pradipta, I. M. A. Santosa, And I. K. Dharmendra, “Penanganan Ketidakseimbangan Data Pada Klasifikasi Pengaduan Masyarakat,” *Jurnal Teknologi Informasi Dan Komputer*, Vol. 9, No. 5, Pp. 489–497, 2023, [Online]. Available: https://pengaduan.denpasarkota.go.id[6].

[7] M. Ibnu and C. Rachmatullah, “Penerapan Smote Untuk Meningkatan Kinerja Klasifikasi Penilaian Kredit,” *Jurnal Riset Komputer)*, Vol. 10, No. 1, Pp. 2407–389, 2023, Doi: 10.30865/Jurikom.V10i1.5612.

[8] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, And M. K. Anam, “Perbandingan Evaluasi Kernel Svm Untuk Klasifikasi Sentimen Dalam Analisis Kenaikan Harga

Bbm," *Malcom: Indonesian Journal Of Machine Learning And Computer Science*, Vol. 3, No. 2, Pp. 153–160, Oct. 2023, Doi: 10.57152/Malcom.V3i2.897.

[9] R. Ghorbani and R. Ghousi, "Comparing Different Resampling Methods In Predicting Students' Performance Using Machine Learning Techniques," *Ieee Access*, Vol. 8, Pp. 67899–67911, 2020, Doi: 10.1109/Access.2020.2986809.

[10] M. Hafizh Mahendra, D. Triantoro Murdiansyah, And K. Muslim Lhaksmana, "Dike: Jurnal Ilmu Multidisiplin Analisis Sentimen Tweet Covid-19 Menggunakan Metode K-Nearest Neighbors Dengan Ekstraksi Fitur Tf-Idf Dan Countvectorizer," 2023.

[11] D. Ananda and R. R. Suryono, "Jurnal Media Informatika Budidarma Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia Dengan Metode Support Vector Machine Dan Naïve Bayes," 2024, Doi: 10.30865/Mib.V8i2.7517.

[12] A. Iffan Alfanzar and I. Sudanawati Rozas, "Topic Modelling Skripsi Menggunakan Metode Latent Diriclhet Allocation," *Sistem Informasi* |, Vol. 7, No. 1, Pp. 7–13.

[13] S. Firman, W. Desena, A. Wibowo, M. I. Komputer, And U. B. Luhur, "Penerapan Algoritma Stemming Nazief & Adriani Pada Proses Klasterisasi Berita Berdasarkan Tematik Pada Laman (Web) Direktorat Jenderal Ham Menggunakan Rapidminer," 2022.

[14] N. Lau Kheng Seng, G. Wei Wei, And T. Ee Xion, "Oversampling Social Media-Sourced Image Datasets for Better Deep Learning Classification of Natural Disaster Damage Levels." [Online]. Available: www.Ijacsa.Thesai.Org

[15] R. Maulana, R. Narasati, R. Herdiana, R. Hamonangan, And S. Anwar, "Komparasi Algoritma Decision Tree dan Naive Bayes dalam Klasifikasi Penyakit Diabetes," 2023.