

# EVALUASI *HYPERPARAMETER TUNING* PADA *SUPPORT VECTOR MACHINE* (SVM) DALAM KLASIFIKASI ULASAN HOTEL DI TRIPADVISOR

Fiashintha Dewi<sup>1)</sup>, Nur Cahyo Hendro Wibowo<sup>2)</sup>, Maya Rini Handayani<sup>3)</sup>, Khothibul Umam<sup>\*4)</sup>

1. Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Walisongo Semarang, Indonesia
2. Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Walisongo Semarang, Indonesia
3. Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Walisongo Semarang, Indonesia
4. Fakultas Dakwah dan Komunikasi, Universitas Islam Negeri Walisongo Semarang, Indonesia

## Article Info

**Kata Kunci:** *Hyperparameter Tuning; SVM; Tripadvisor*

**Keywords:** *Hyperparameter Tuning; SVM; Tripadvisor*

## Article history:

Received 5 May 2025

Revised 2 June 2025

Accepted 16 June 2024

Available online 1 September 2025

## DOI :

<https://doi.org/10.29100/jipi.v10i3.7774>

\* Corresponding author.

Khothibul Umam

E-mail address:

[khothibul\\_umam@walisongo.ac.id](mailto:khothibul_umam@walisongo.ac.id)

## ABSTRAK

Dengan adanya perkembangan teknologi para wisatawan sangat memudahkan dalam mengakses informasi mengenai pemesanan kamar hotel. Dengan adanya hal tersebut, maka ulasan dari pengguna lain sangatlah penting untuk menemukan tempat yang mereka inginkan. Studi ini membahas tentang analisa ulasan para wisatawan mengenai hotel pada Tripadvisor. Tripadvisor adalah salah satu platform panduan wisata terbesar di dunia, yang menawarkan wisatawan untuk merencanakan serta memperoleh perjalanan memuaskan. Data diambil melalui *website* Hugging Face yang kemudian dilanjutkan dengan proses *pre-processing* data. Dataset yang digunakan berjumlah 20.491 ulasan, terdiri dari 15.093 ulasan positif dan 5.938 ulasan negatif. Tujuan dari penelitian ini untuk mengevaluasi performa model SVM dalam melakukan klasifikasi sentimen pada ulasan hotel di Tripadvisor. Untuk mengoptimalkan performa model, dilakukan *hyperparameter tuning* menggunakan metode *GridSearchCV*. Hasil menunjukkan bahwa model default SVM memiliki akurasi 91%, namun recall pada kelas negatif masih rendah (0,75). Setelah tuning, akurasi sedikit menurun menjadi 90%, tetapi recall kelas negatif meningkat menjadi 0,77. Model terbaik diperoleh pada kombinasi parameter  $C = 10$ ,  $\gamma = 0,01$ , dan kernel = linear, dengan precision 0,92, recall 0,94, dan f1-score 0,80. Tuning terbukti meningkatkan keseimbangan klasifikasi antar kelas dan sensitivitas terhadap ulasan negatif. Hasil ini menegaskan pentingnya *hyperparameter tuning* dalam mengoptimalkan performa dan generalisasi model SVM pada analisis sentimen dengan data yang tidak seimbang.

## ABSTRACT

With the development of technology, travelers are facilitated in accessing information about booking hotel rooms. With this, reviews from other users are very important to find the place they want. This study discusses the analysis of travelers' reviews of hotels on Tripadvisor. Tripadvisor is one of the largest travel guide platforms in the world, offering travelers to plan and obtain satisfying trips. The data is collected through the Hugging Face website which is then followed by data pre-processing. The dataset used amounted to 20.491 reviews, consisting of 15.093 positive reviews and 5.938 negative reviews. The purpose of this research is to evaluate the performance of the SVM model in performing sentiment classification on hotel reviews on Tripadvisor. To optimize the performance of the model, hyperparameter tuning is performed using the GridSearchCV method. The results show that the default SVM model has 91% accuracy, but the recall on the negative class is still low (0,75). After tuning, the accuracy decreased slightly to 90%, but the recall of the negative class increased to 0,77. The best model was obtained at a combination of parameters  $C = 10$ ,  $\gamma = 0,01$ , and kernel = linear, with precision 0,92, recall 0,94, and f1-score 0,80. Tuning was shown to improve classification balance between classes and sensitivity to negative reviews. These results confirm the importance of hyperparameter tuning in optimizing the performance and generalization of SVM models in sentiment analysis with imbalanced data.

## I. PENDAHULUAN

**P**ERKEMBANGAN teknologi telah mengalami perubahan yang sangat signifikan dari berbagai bidang, baik dari segi ekonomi, budaya, social, maupun Pendidikan. Dengan adanya media *online* Masyarakat sangat memudahkan dalam mengakses informasi [1]. Salah satunya yaitu di industry pariwisata, sebagai contoh para wisatawan dalam merencanakan perjalanannya tidak perlu pergi ke *travel agent* hanya untuk membeli tiket atau sekedar memesan kamar di hotel. Untuk memenuhi semua kebutuhan perjalanan mereka, parawisatawan hanya perlu mengunduh aplikasi agen perjalanan *online* ke *smartphone* untuk melakukan pemesanan tiket [2].

Dengan adanya perubahan tersebut, maka ulasan tentang sesuatu hal sangatlah penting untuk menemukan tempat yang mereka inginkan. Para wisatawan akan memudahkan dalam mendapatkan Gambaran mengenai hotel yang akan ditinggali dengan melalui ulasan yang banyak beredar di media sosial maupun melalui situs web salah satunya yaitu Tripadvisor [3]. Tripadvisor adalah salah satu platform panduan wisata terbesar di dunia, yang menawarkan wisatan untuk merencanakan serta memperoleh perjalanan wisata yang memuaskan. Pada tahun 2013, Tripadvisor dibentuk sebagai platform untuk pencarian tempat wisata dan hotel yang paling sering diakses [2].

Ulasan online menyajikan Gambaran tentang kesan, penilaian, serta pengalaman yang dirasakan oleh pelanggan hotel. Proses penilaian biasanya dalam bentuk rating dan juga komentar yang berisi pengalaman, saran, dan masukkan dari pelanggan. Meskipun system rating sering digunakan untuk mengukur kualitas layanan hotel, namun system ini tidak dapat sepenuhnya menangkap keluhan atau kekurangan konsumen [4]. Pada fitur ulasan di berbagai situs sering kali tidak tersedia pilihan untuk memilah ulasan dengan pandangan positif ataupun ulasan dengan pandangan negatif, sehingga diperlukan analisis sentimen untuk membantu mengklasifikasikan ulasan dari pengunjung sebelumnya [5]. Analisis sentimen ialah teknik untuk menginterpretasikan, mengekstrak, dan memproses data tekstual dan tampilan secara otomatis untuk menentukan sentimen yang diungkapkan di dalamnya [6]. Analisis sentimen para wisatawan ditentukan oleh beberapa faktor, termasuk volume data teks, sumber data, dan efisiensi algoritma yang digunakan. Wisatawan memiliki penilaian mengenai pelayanan yang diberikan oleh pihak hotel tujuan, sehingga data tersebut dapat diklasifikasikan untuk menentukan apakah penilaiannya positif atau negatif. Namun terdapat tantangan utama dalam analisis sentiment yang mempengaruhi akurasi model klasifikasi, yaitu keanekaragaman gaya bahasa. Di Tripadvisor sendiri terdapat banyak ulasan mengenai hotel yang beraneka ragam gaya dan struktur bahasa. Banyak ulasan yang mengandung bahasa informal, slang, hingga ekspresi emosional yang tidak baku, untuk mengatasi hal tersebut dibutuhkan model klasifikasi yang kuat dan adaptif [7].

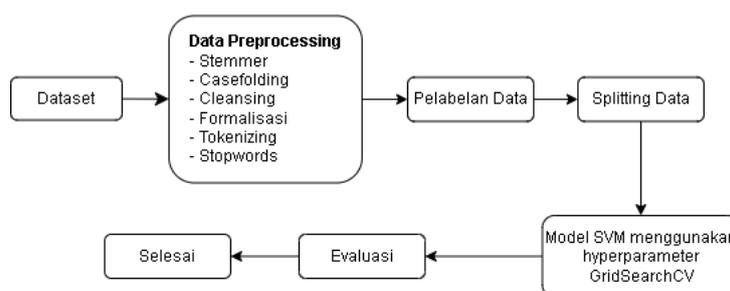
Klasifikasi teks dapat dilakukan melalui berbagai pendekatan, salah satunya metode *Support Vector Machine* (SVM). Dalam proses *machine learning*, terutama klasifikasi, umumnya melibatkan pengujian model yang berbeda pada dataset, yang kemudian diikuti dengan seleksi model yang memperlihatkan kinerja terbaik. Untuk memperoleh model yang mampu menghasilkan prediksi secara akurat, salah satu langkah yang dapat dilakukan yakni proses penyesuaian *hyperparameter* atau *hyperparameter tuning*. Penyesuaian *Hyperparameter tuning* ini memainkan peran penting dalam meningkatkan kinerja algoritma *machine learning* [8].

Pada penelitian yang dilakukan oleh Afis, Andi, dan Ferry dengan mengambil objek penelitian optimasi *hyperparameter* untuk klasifikasi penyakit tanaman padi. RMSProp, ukuran batch 32, learning rate 0,001, dan jumlah epoch 100 adalah hiperparameter yang memberikan hasil terbaik dalam penelitian ini, dengan nilai akurasi 97,56% [9]. Penelitian selanjutnya, mengenai *hyperparameter tuning* pada algoritma klasifikasi, untuk saat ini metodologi *hyperparameter* yang digunakan adalah *Grid Search*. Selain itu, hasilnya menunjukkan *Decision Tree* memiliki nilai terendah dan model *XGBoost* memiliki nilai tertinggi [8]. Selain itu, beberapa studi membahas penggunaan penyesuaian hiperparameter *RandomSearchCV* dalam Adaptive Boosting untuk meramalkan kelangsungan hidup pasien gagal jantung. Model ini mampu memperoleh akurasi klasifikasi 85%, spesifisitas 92,56%, sensitivitas 68,4%, dan AUC 89,7%, menurut data [10]. Penelitian yang dilakukan oleh Quratul, Ema, dan Asro dengan mengambil objek penelitian prediksi rating terhadap *reviews* wisatawan Tanjung Puting pada Tripadvisor, hasil evaluasi SVM memperoleh rata-rata 80% dari dataset sebanyak 390 data dengan pembagian dataset 30:70 [11]. Penelitian selanjutnya, mengenai analisis *hyperparameter tuning indobert* dalam pendeteksian berita palsu. Hasil penelitian menunjukkan bahwa metode *Bayesian optimization* merupakan pendekatan yang paling efektif, yang mencapai *precision* 88,79%, *recall* 94,5%, dan *F1-Score* 91,56% [12]. Penelitian berikutnya, mengenai ulasan aplikasi sayurbox di google play store, hasil akurasi SVM sebesar 89,29% dengan total 4201 data ulasan dengan perbandingan 80:20 [13]. Lalu ada penelitian dari Liana dan Gerry dengan mengambil objek penelitian optimasi *Hyperparameter GridSearchCV* untuk klasifikasi kualitas udara. Pada penelitian tersebut, menggunakan *Support Vector Machine* yang menghasilkan akurasi awal sebesar 96%, dimana nilai presisi 97%, *recall* 91%, dan *F1-score* 94%. Namun setelah melakukan proses optimasi, akurasi meningkat sebesar 2% yang menjadikan nilai akurasi akhir mencapai 98%. Nilai presisi tetap stabil di angka 97%, sedangkan *recall* meningkat menjadi 96%, dan *F1-score* juga naik menjadi 96% [14].

Metode *Support Vector Machine* (SVM) memiliki kelemahan yaitu pemilihan parameter yang sensitif, SVM sangat bergantung pada pemilihan parameter yang tepat (seperti C, kernel, dan gamma untuk RBF kernel). Pemilihan yang salah dapat menyebabkan model kurang efektif. Beberapa pendekatan populer yang sering digunakan yaitu *GridSearchCV*, *RandomSearchCV*, dan *Bayesian Optimization*. Seperti penelitian yang dilakukan oleh Muhamad Fajri dan Aji Primajaya [15] yang membandingkan dua metode optimasi hyperparameter, yaitu *Grid Search* dan *Random Search* pada algoritma SVM untuk permasalahan klasifikasi. Evaluasi dilakukan berdasarkan akurasi, penggunaan memori, dan waktu eksekusi. Hasil menunjukkan bahwa *Grid Search* memberikan akurasi yang lebih tinggi, namun dengan waktu komputasi yang lebih lama dibandingkan *Random Search*. Selain itu, pada penelitian yang dilakukan oleh Rini Muzayanah, dkk [16] yang membandingkan *GridSearchCV* dan *Bayesian Optimizatin* menggunakan algoritma Random Forest. Didapatkan hasil bahwa *GridSearchCV* menghasilkan akurasi sebesar 0,74 meskipun membutuhkan waktu lebih lama, sedangkan *Bayesian Optimization* memiliki akurasi lebih rendah dengan selisih 0,01 yaitu 0,73 dan membutuhkan waktu sedikit dibandingkan *GridSearchCV*. Oleh karena itu, dalam penelitian ini dipilih metode *GridSearchCV* karena metode ini terbukti efektif dalam berbagai studi terdahulu terkait klasifikasi teks menggunakan SVM. Menurut temuan penelitian sebelumnya, penelitian ini bertujuan untuk mengevaluasi kinerja algoritma *Support Vector Machine* (SVM) yang telah dioptimalkan melalui *hyperparameter tuning* menggunakan *GridSearchCV* dalam klasifikasi sentiment ulasan hotel di Tripadvisor. Selain itu, hasil penelitian ini diharapkan memberikan kontribusi nyata terhadap pengembangan model klasifikasi berbasis *Support Vector Machine* (SVM) dalam hal analisis sentiment yang lebih luas, termasuk dalam e-commerce, layanan pelanggan, dan industri lain yang mengandalkan ulasan pengguna sebagai umpan balik. Dengan optimalisasi parameter yang tepat, model tidak hanya akan lebih akurat dalam mengenali sentiment pengguna, tetapi juga lebih andal dalam menghadapi kompleksitas bahasa alami di berbagai sektor industri.

## II. METODE PENELITIAN

Bagian ini membahas tahapan-tahapan yang dilakukan peneliti dalam melakukan penelitian evaluasi analisis seperti, pengumpulan data ulasan yang akan digunakan sebagai *dataset*, *preprocessing*, *split data*, *training SVM* yang menggunakan *hyperparameter GridSearchCV*, dan evaluasi SVM. Gambar 1 menunjukkan langkah yang diambil penelitian ini. Menurut gambar ini, Langkah *preprocessing* yang meliputi *stemmer*, *casefolding*, *cleaning*, *normalisasi*, *tokenizing*, *stopwords*, dan *stemming* dilakukan setelah dataset berhasil diambil. Setelah langkah *preprocessing* selesai, data dipisahkan menjadi data latih dan data uji. Selanjutnya, dilakukan training model SVM menggunakan *hyperparameter tuning* menggunakan *GridSearchCV*. Setelah itu, model dievaluasi dengan menggunakan data uji hingga diperoleh hasil akhir berupa nilai akurasi model.



Gambar. 1. Tahapan Penelitian

### A. Pengumpulan Data

Dataset yang digunakan ialah dataset yang terkait dengan ulasan hotel. Data yang digunakan diperoleh dari situs *website* *Hungging Face* (ulasan hotel Tripadvisor). Penelitian ini memanfaatkan sebanyak 20.491 data yang diambil dari kolom review Tripadvisor. Selain mengambil data review, peneliti juga menyertakan beberapa informasi tambahan seperti rating.

### B. Preprocessing

Dalam proses ini, data terlebih dahulu diproses melalui *preprocessing* agar format ejaannya sesuai dan lebih mudah diproses oleh *text mining*, merupakan metode untuk menemukan informasi baru yang tidak diketahui sebelumnya oleh computer, dengan cara mengekstrak informasi secara otomatis dari berbagai sumber. Dalam *preprocessing*, akan diterapkan beberapa teknik seperti *casefolding*, *cleansing*, *tokenizing*, *stopwords*, dan *stemming* guna meningkatkan akurasi sert efisiensi hasil analisis.

- a. *Cleansing* merupakan tahapan untuk mengevaluasi dan meningkatkan kualitas data secara menyeluruh dengan cara pengubahan, penggantian, atau penghapusan elemen data yang tidak sesuai format yang biasanya ditemukan dalam dataset, guna memperoleh data yang lebih bersih dan berkualitas tinggi [17].
- b. *Case folding* merupakan proses mengubah seluruh karakter dalam teks menjadi huruf kecil. Langkah ini bertujuan untuk menyederhanakan proses pengolahan teks, sehingga perbedaan huruf kapital tidak memengaruhi analisis data [18].
- c. *Tokenizing* adalah proses memisahkan teks menjadi kata yang lebih kecil. Pada tahap ini, setelah teks dikonversi ke huruf kecil dan tanda baca dihapus, setiap elemen dalam kalimat diberikan *token* agar dapat dipecah menjadi kata-kata individual untuk memudahkan analisis lebih lanjut [11].
- d. *Stopwords* merupakan tahap eliminasi kata-kata yang dianggap tidak memiliki makna signifikan atau bobot informasi yang rendah dalam analisis teks. Tahapan ini bertujuan untuk menyederhanakan kalimat tanpa menghilangkan makna utama. Contoh kata yang termasuk *stopwords*, yaitu “i”, “me”, “my”, “the”, “and”, “is”, “of” dan lainnya [18].
- e. *Stemming* adalah tahapan dalam *preprocessing* teks yang bertujuan untuk mengembalikan kata ke bentuk dasarnya dengan menghapus imbuhan atau akhiran yang tidak diperlukan. Tahapan ini dilakukan untuk menyatukan berbagai bentuk kata yang memiliki arti serupa, sehingga analisis teks dapat dilakukan secara lebih efisien dan konsisten. [19].

### C. Pelabelan Data

Pelabelan data merupakan tahap pemberian label atau informasi tertentu pada data mentah sehingga data tersebut dapat dimanfaatkan dalam proses pelatihan model *machine learning*. Label yang diberikan umumnya berupa kategori, nilai, atau atribut yang mencerminkan target dari proses klasifikasi. Tujuan utama dari pelabelan adalah agar algoritma SVM mampu mempelajari pola yang terdapat dalam data berlabel dan menerapkannya untuk memprediksi label pada data baru yang belum diketahui sebelumnya.

Pada tahap ini, setiap data ulasan dan rating akan diberi label berdasarkan sentiment yang diberikan oleh pengguna Tripadvisor. Penelitian ini, menggunakan dua jenis label, yaitu positif dan negatif. Di mana label positif diberikan untuk rating antara 4 hingga 5, sedangkan label negatif untuk rating antara 1 hingga 3. Dari total dataset 20.491 data ulasan, terdapat 15.093 ulasan positif dan 5.398 ulasan negatif. Hal tersebut menunjukkan ketidakseimbangan data, dengan ulasan positif mendominasi lebih dari 70% total data. Ketidakseimbangan tersebut dapat mempengaruhi performa model, terutama dalam menurunkan kemampuan model untuk mendeteksi ulasan sentiment minoritas atau negatif. Model cenderung bias terhadap kelas mayoritas, yang dapat menyebabkan menurunnya nilai recall pada kelas dengan jumlah datanya yang lebih sedikit.

### D. Data Splitting

*Data Splitting* adalah suatu metode yang biasanya digunakan dalam proses data mining untuk membagi dataset ke dalam subset pelatihan, pengujian, dan validasi. Pendekatan tersebut berguna dalam proses penentuan *hyperparameter* dan dalam mengevaluasi kemampuan generalisasi model. Pada penelitian ini, data yang telah melalui tahap *preprocessing* kemudian dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan ketentuan 80:20. Pembagian dataset ini dilakukan secara otomatis dengan memanfaatkan *library scikit-learn*.

### E. Modelling

*Modelling* yaitu tahapan membangun model yang disesuaikan dengan permasalahan yang ingin diselesaikan, mempertimbangkan aspek kinerja serta kompleksitas dalam pemrosesan data. Tujuannya adalah untuk menghasilkan model yang mampu menyelesaikan permasalahan dengan tingkat efektivitas dan efisiensi yang tinggi. Pada penelitian ini, menggunakan model *Support Vector Machine* (SVM), dimana SVM menciptakan model yang dapat membedakan antar kategori dan data dengan menemukan *hyperplane* terbaik yang memisahkan setiap ulasan kedalam dua kelas yaitu positif dan negatif [20]. Data tersebut selanjutnya akan digunakan pada prediksi data testing.

### F. Hyperparameter Tuning dengan GridSearchCV

Dalam mengoptimalkan kinerja algoritma *machine learning*, yang memiliki peran yang sangat penting adalah *Hyperparameter tuning*. Untuk memperoleh hasil klasifikasi yang lebih optimal, pemilihan parameter yang tepat pada model yang digunakan sangatlah penting. *Hyperparameter* berperan dalam mengatur berbagai aspek dalam *machine learning* yang berdampak besar terhadap performa dan kualitas model [21]. Berbagai metode yang dapat digunakan untuk mengoptimalkan *hyperparameter*, seperti *GridSearchCV*, *RandomSearchCV*, *Bayesian Optimization*, dan *Evaluationary Optimization*.

Dalam penelitian ini, untuk menentukan nilai *hyperparameter* terbaik bagi model menggunakan metode *Grid Search*. *GridSearchCV* (*Grid Search Cross Validation*) merupakan metode yang mengevaluasi setiap kombinasi dari model dan *hyperparameter*, serta melakukan validasi terhadap masing-masing kombinasi tersebut [22]. Sebagai contoh, jika jumlah *n\_estimators*=[10, 20, 30, 40] dan *learning rate*=[0,01, 0,1, 1,0], maka *grid search* akan menguji total 12 kombinasi. Dari seluruh kemungkinan tersebut, akan dipilih kombinasi yang menghasilkan performa terbaik. Parameter utama yang akan dituning adalah C, Kernel, dan Gamma, yang masing-masing memengaruhi margin pemisah, jenis fungsi pemetaan data, dan pengaruh titik data terhadap model.

Sebelum pembuatan model dilakukan terlebih dahulu penentuan search space yaitu parameter-parameter yang akan digunakan pada proses *hyperparameter tuning*. Pemilihan parameter *search space* ini merujuk pada penelitian Fajri & Primajaya [15] serta Misriati & Aryanti [20], yang menunjukkan bahwa *search space* serupa dapat menghasilkan performa optimal pada klasifikasi berbasis SVM.

### G. Evaluasi Model

Pada proses evaluasi, tujuannya adalah mengevaluasi kinerja model SVM. Untuk mengevaluasi kinerja model, memanfaatkan *confusion matrix* guna menentukan nilai *accuracy*, *precision*, *recall*, dan *f1-score* dari model yang diterapkan. Hasil evaluasi digunakan sebagai dasar untuk menilai tingkat akurasi, sensitivitas, serta efektivitas kinerja sistem atau model dalam melakukan proses klasifikasi [23]. Selain itu, proses ini juga bertujuan menilai kinerja model dalam menangani kesalahan serta mengidentifikasi sumber yang menjadi penyebab terjadinya kesalahan tersebut. Evaluasi performa model dilakukan dengan menggunakan beberapa metrik klasifikasi berikut merupakan rumus perhitungannya (1), (2), (3), (4):

*Accuracy* merupakan persentase dari jumlah prediksi yang benar, dihitung memakai rumus:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

*Precision* merupakan nilai *true positive* dari sebuah nilai positif. Berikut adalah rumusnya:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

*Recall* merupakan persentase prediksi positif dengan *true positive*. Rumusnya adalah:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

*F1-Score* merupakan perbandingan rata-rata *precision* dan *recall*. Rumusnya sebagai berikut:

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

## III. HASIL DAN PEMBAHASAN

Tahapan ini membahas hasil dari penelitian yang berfokus pada hasil tingkat akurasi tertinggi dari model SVM yang telah dikembangkan melalui proses *hyperparameter tuning*. Nilai akurasi akhir diperoleh melalui perhitungan menggunakan rumus pada *confusion matrix*. Dengan demikian, setelah akurasi model diketahui, dapat diketahui kombinasi *hyperparameter* yang memberikan performa paling optimal.

### A. Pengumpulan Data

Dataset diperoleh dari *Hugging Face*, dataset ini memiliki bentuk format CSV (*a commaseparated values*). Dataset ini terdiri dari 20.491 data dengan beberapa variabel yaitu meliputi *text*, *inputs*, *prediction*, *prediction\_agent*, *annotation*, *annotation\_agent*, *multi\_label*, *explanation*, *id*, *metadata*, *status*, *event\_timestamps*, dan *metric*. Berikut adalah sampel data primer Tripadvisor review hotel.

	text	inputs	prediction	prediction_agent	annotation	annotation_agent	multi_label	explanation	id	metadata	status	event_timestamp	metrics
0	beautiful accommodations stayed hotel santa mar...	{'text': 'beautiful accomodations stayed hotel...	['label': '4', 'score': 1.0]]	Argilla	NaN	NaN	False	NaN	0001bee9-6e54-4f55-857a-70caa8e02a54	NaN	Default	2022-12-06 14:02:26.932135	{'text_length': 570}
1	great hotel great location hotel located la ra...	{'text': 'great hotel great location hotel loc...	['label': '5', 'score': 1.0]]	Argilla	NaN	NaN	False	NaN	0003ddb6-6d7d-43c7-b93e-322feacc1fb2	NaN	Default	2022-12-06 14:02:26.885039	{'text_length': 477}
2	beautiful hotel great location like beautiful...	{'text': 'beautiful hotel great location like ...	['label': '4', 'score': 1.0]]	Argilla	NaN	NaN	False	NaN	000630a6-da97-4cba-a9ef-964f9866ed2d	NaN	Default	2022-12-06 14:02:26.921625	{'text_length': 504}
3	great deal waikiki trip hawaii outrigger luana...	{'text': 'great deal waikiki trip hawaii outri...	['label': '4', 'score': 1.0]]	Argilla	NaN	NaN	False	NaN	00067c7e-2015-4bfb-8531-b6432f1b73e0	NaN	Default	2022-12-06 14:02:26.858496	{'text_length': 551}
4	choose airport hotel busy triparound australia...	{'text': 'choose airport hotel busy triparound...	['label': '1', 'score': 1.0]]	Argilla	NaN	NaN	False	NaN	000905b7-5e96-40a4-8a6c-f66ce0b465f9	NaN	Default	2022-12-06 14:02:26.914048	{'text_length': 307}

Gambar. 2. Sampel Hasil Pengambilan Data

Karena yang akan digunakan dalam penelitian ini hanya ulasan dan rating, maka pada variabel tersebut akan diambil kolom *text* dan kolom *prediction*, yang dimana pada kolom *prediction* akan dirubah menjadi kolom rating dengan mengambil labelnya saja. Pada Tabel I tertera data rating dari ulasan di Tripadvisor adalah rating bintang 1, 2, 3, 4, dan 5.

TABEL I  
 SAMPEL HASIL KOLOM PREDICTION MENJADI RATING

index	prediction	rating
0	[{'label': '4', 'score': 1.0}]	4
1	[{'label': '5', 'score': 1.0}]	5
2	[{'label': '4', 'score': 1.0}]	4
3	[{'label': '4', 'score': 1.0}]	4
4	[{'label': '1', 'score': 1.0}]	1

### B. Preprocessing Data

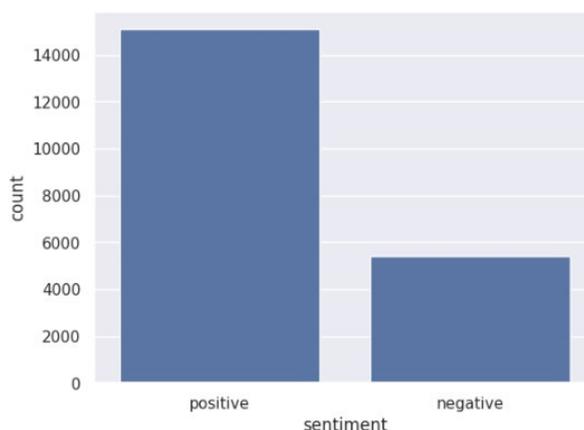
Dataset ulasan dari Tripadvisor yang berhasil dikumpulkan sebanyak 20.491 data, terdiri dari 15.093 ulasan dengan sentiment positif dan 5.398 ulasan dengan sentiment negatif. Data tersebut kemudian diproses melalui *preprocessing* karena teks yang digunakan dalam ulasan merupakan data teks tidak terstruktur yang memerlukan pemrosesan ulang. Proses *preprocessing* yang dilakukan mencakup beberapa tahapan seperti *casefolding*, *cleansing*, normalisasi, *tokenizing*, *stopwords*, dan *stemming*, dengan tujuan untuk meningkatkan akurasi dan efisiensi analisis. Dapat dilihat pada Gambar 3, disajikan hasil dari *preprocessing* data ulasan di Tripadvisor.

	text	rating	cleaned_text
0	beautiful accomodations stayed hotel santa mar...	4	0 beauti accomod stay hotel santa maria novella ...
1	great hotel great location hotel located la ra...	5	1 great hotel great locat hotel locat la rambla ...
2	beautiful hotel great location like beautifull...	4	2 beauti hotel great locat like beauti appoint r...
3	great deal waikiki trip hawaii outrigger luana...	4	3 great deal waikiki trip hawaii outrigg luana g...
4	choose airport hotel busy triparound australia...	1	4 choos airport hotel busi triparound australia ...

Gambar 3. Sampel Data Cleaned Text Ulasan Hotel Tripadvisor

### C. Pelabelan Data

Pelabelan data dilakukan dengan mempertimbangkan ulasan juga rating pengguna pada Tripadvisor. Proses ini mengkategorikan ulasan menjadi dua yaitu: positif untuk rating 4 hingga 5, dan negatif untuk rating antara 1 hingga 3.



Gambar 4. Hasil Pelabelan Data

Gambar 4. menunjukkan hasil pelabelan data yang telah melewati proses *cleaning* data dengan menghapus data kosong dari tabel sehingga data yang terdeteksi hanya 20.491, terdapat 15.093 data ulasan positif dan 5398 data ulasan negatif.

### D. Data Splitting

Dalam tahapan data *splitting*, data akan dibagi menjadi data training dan data testing, digunakan fungsi *train\_test\_split* dari *library sklearn* untuk memisahkan dataset ke dalam subset *train* dan *test*. Proses pembagian dilakukan dengan proporsi 80% untuk data pelatihan yaitu sebanyak 16.392 data dan 20% untuk data pengujian sebanyak 4.099 data. Setelah data terbagi, dilakukan evaluasi model melalui perbandingan nilai *accuracy*, *precision*, *recall*, dan *f1-score* dengan menggunakan *confusion matrix*.

### E. Modelling

Pada tahap pembuatan model, langkah pertama yang dilakukan adalah menentukan *search space*, yaitu kumpulan parameter yang akan dipakai dalam proses *hyperparameter tuning*. Daftar parameter tersebut tercantum pada Tabel II. Setelah menentukan *search space*, proses pembangunan model dilakukan. Dalam penelitian, model yang digunakan adalah *Support Vector classification (SVC)* dari *library scikit-learn*, yang merupakan implementasi dari SVM berbasis LibSVM. SVC merupakan jenis SVM yang dirancang untuk menyelesaikan masalah klasifikasi, baik untuk *binary class classification* maupun *multi-class classification* [15]. Proses *hyperparameter tuning* serta validasi model dilakukan dengan menggunakan metode *GridSearchCV*.

TABEL II. DAFTAR SEARCH SPACE

Parameter	Nilai
C	[0.1, 1, 10, 100]
Gamma	[scale, auto, 0.01, 0.1, 1]
Kernel	[linier, rbf, poly]

Proses uji validitas model dilakukan dengan menjalankan fungsi *fit* pada objek *GridSearchCV* yang telah dibuat. Dalam tahap ini, data akan dibagi menjadi lima bagian menggunakan teknik *5-fold cross-validation*, dimana masing-masing bagian secara bergantian digunakan sebagai data validasi sementara sisanya sebagai data pelatihan. Jadi setiap proses *GridSearchCV* akan mengevaluasi seluruh kombinasi parameter yang telah ditentukan sebelumnya dalam *search space*. Pada Gambar 5, disajikan hasil dari kombinasi parameter terbaik yang diperoleh dari proses ini yaitu  $C=10$ ,  $\gamma=0,01$ , dan  $\text{kernel}=\text{'linear'}$ .

 Best Parameters: {'svm\_c': 10, 'svm\_gamma': 0.01, 'svm\_kernel': 'linear'}

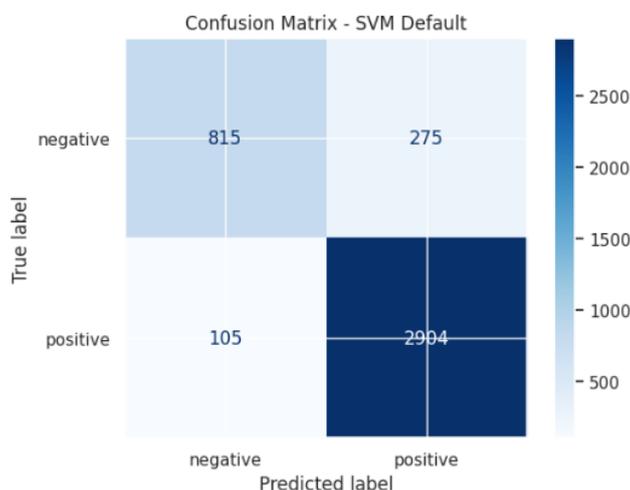
Gambar. 5. Hasil Parameter Terbaik

### F. Evaluasi Model

Setelah proses tuning selesai dilakukan, model terbaik diperoleh dengan kombinasi parameter yang optimal dan kemudian dievaluasi data pengujian yang sama. Evaluasi dilakukan melalui perhitungan metrik evaluasi dan penyajian diagram dari *confusion matrix*. Hasil evaluasi sebelum dilakukan *hyperparameter tuning*, model SVM menggunakan parameter default yaitu  $C=1$ ,  $\text{kernel}=\text{'rbf'}$  dan  $\gamma=\text{'scale'}$ .

TABEL III.  
HASILEVALUASI SEBELUM TUNING

	precision	recall	f1-score	support
negative	0,89	0,75	0,81	1090
positive	0,91	0,97	0,94	3009
accuracy			0,91	4099
macro avg	0,90	0,86	0,87	4099
weighted avg	0,91	0,91	0,90	4099



Gambar. 6. Hasil confusion Matrix SVM Default

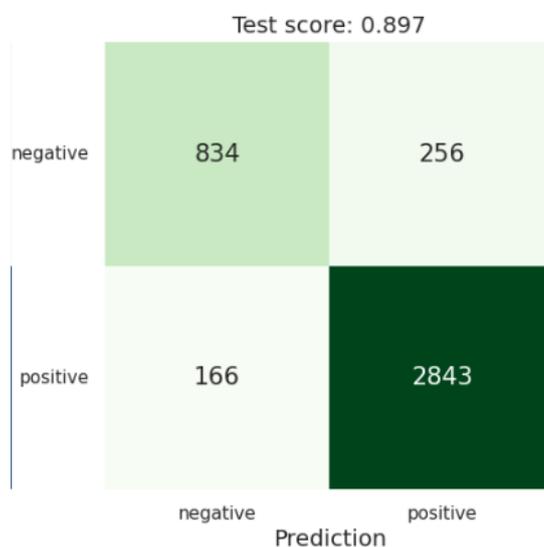
Dilihat dari Tabel III, akurasi yang diperoleh menggunakan nilai default model SVM yaitu 0,91 yang menunjukkan bahwa performa SVM cukup baik, meskipun nilai recall untuk kelas negative masih lebih rendah dibandingkan kelas positif, yang

mengindikasikan bahwa sebagian ulasan positif masih salah diklasifikasikan. Ketidakseimbangan jumlah data antara kelas positif dan negatif serta variasi gaya bahasa pada ulasan negative, dapat menjadi alasan menurunnya recall untuk kelas negatif. Dapat dilihat pada Gambar 6, model SVM default menunjukkan 815 ulasan negatif terdeteksi dengan benar atau true negative. 275 ulasan negatif yang terdeteksi positif atau false positif. Lalu terdapat 105 ulasan positif terdeteksi negatif atau false negative dan terdapat 2904 ulasan positif yang terdeteksi secara benar atau true positif.

Kemudian dilakukan pengujian menggunakan *hyperparameter tuning* menggunakan *GridSearchCV*. Hasil dari evaluasi *hyperparameter tuning* berupa nilai *accuracy*, *precision*, *recall*, dan *f1-score* yang dapat dilihat pada tabel III dan hasil *confusion matrix* terdapat pada Gambar 7.

TABEL IV.  
 HASIL EVALUASI MENGGUNAKAN HYPERPARAMETER TUNING

	precision	recall	f1-score	support
negative	0,84	0,77	0,80	1090
positive	0,92	0,94	0,93	3009
accuracy			0,90	4099
macro avg	0,88	0,85	0,86	4099
weighted avg	0,90	0,90	0,90	4099



Gambar. 7. Hasil *Confusion Matrix*

Diagram *Confusion Matrix* pada Gambar 7, menunjukkan bahwa model SVM yang telah melewati proses tuning mampu memberikan hasil prediksi yang cukup baik dengan tingkat akurasi sebesar 89,7%. Dari 4.099 data uji, model berhasil mengklasifikasikan 2.843 ulasan positif secara tepat dan salah mengklasifikasikan 166 ulasan positif sebagai negatif. Pada kelas negatif, model juga berhasil mengklasifikasi 834 ulasan negatif dengan benar, sementara 256 ulasan negatif masih salah diklasifikasikan sebagai ulasan positif.

Hasil perhitungan metrik evaluasi berdasarkan dari *confusion matrix* dapat dilihat pada Tabel IV, menunjukkan bahwa model meraih akurasi keseluruhan sebesar 0,90 dalam mengklasifikasikan sentimen ulasan hotel. Jika dibandingkan dengan model default, proses tuning menyebabkan penurunan kecil pada akurasi dari 0,91 menjadi 0,90. Namun demikian, pada kelas positif, model menunjukkan kinerja yang optimal dengan *precision* mencapai 0,92, *recall* sebesar 0,94, dan *f1-score* 0,93, hasil tersebut menunjukkan bahwa model sangat efektif dalam menganalisa ulasan dengan sentimen positif. Sementara itu, pada kelas negatif, nilai *precision* mencapai 0,84, *recall* sebesar 0,77, dengan *f1-score* 0,80. Meskipun akurasinya mengalami penurunan, tetapi terjadi peningkatan recall pada kelas negative dari 0,75 menjadi 0,77 dibandingkan dengan model default. Hal ini menunjukkan bahwa model hasil tuning lebih sensitive terhadap deteksi ulasan negatif. Hal ini penting dalam pengawasan kualitas layanan, karena ulasan negatif cenderung mewakili keluhan pelanggan.

Hasil tersebut merupakan hasil dari penerapan metode *GridSearchCV*, yang secara sistematis mengevaluasi berbagai kombinasi *hyperparameter* untuk menemukan parameter terbaik sesuai dengan metrik evaluasi yang telah ditetapkan. Dengan memanfaatkan *cross-validation*, *GridSearch* memastikan bahwa kinerja model diuji pada berbagai subset data, sehingga mampu meminimalkan kemungkinan overfitting.

Berdasarkan perbandingan hasil evaluasi sebelum dilakukan tuning dan sesudah dilakukan tuning, dapat dilihat bahwa tuning tidak hanya bertujuan untuk meningkatkan akurasi, melainkan juga menyeimbangkan kemampuan klasifikasi antar kelas khususnya pada data tidak seimbang. Hal tersebut sejalan dengan penelitian yang dilakukan Thomas dan Rumaisa [7], bahwa ulasan negatif dalam kasus perhotelan sering kali mengandung ekspresi tidak langsung, sehingga sulit diklasifikasikan tanpa proses tuning yang tepat. Selain itu, hasil kami juga mendukung penelitian dari Fajri dan Primajaya [15], menunjukkan bahwa

*hyperparameter tuning* pada SVM dapat meningkatkan sensitivitas model terhadap kelas minoritas, meskipun kadang disertai dengan penurunan kecil akurasi keseluruhan. Dalam penelitian ini, peningkatan nilai *recall* dan *F1-score* pada kelas negative setelah tuning menunjukkan dampak nyata dari optimasi parameter terhadap klasifikasi. Selain itu, Gonzalez-Castro et al [24], menyoroti bahwa tanpa optimasi *hyperparameter* yang sistematis, model cenderung gagal menangani pola non-linier yang kompleks, sehingga potensi penuh dari algoritma SVM tidak dapat tercapai. Secara keseluruhan, *hyperparameter tuning* dengan *GridSearchCV* dapat menjadi pendekatan yang efektif dalam meningkatkan keteepatan klasifikasi, khususnya pada data ulasan yang kompleks dan tidak seimbang. Dengan demikian penelitian ini memberikan kontribusi nyata dalam menunjukkan efektivitas tuning terhadap klasifikasi sentiment dalam hal ulasan hotel.

#### IV. KESIMPULAN

Penelitian ini membuktikan bahwa penerapan *hyperparameter tuning* menggunakan *GridSearchCV* pada algoritma SVM mampu meningkatkan keseimbangan performa klasifikasi sentimen ulasan hotel di Tripadvisor. Sebelum tuning, model menghasilkan akurasi sebesar 91% dengan *recall* kelas negatif sebesar 0,75. Setelah tuning, meskipun akurasi sedikit menurun menjadi 90%, *recall* pada kelas negatif meningkat menjadi 0,77, yang menunjukkan peningkatan sensitivitas terhadap data minoritas. Hal ini penting dalam konteks layanan pelanggan, di mana identifikasi ulasan negatif sangat krusial. Kombinasi *hyperparameter* terbaik diperoleh pada nilai  $C = 10$ ,  $\gamma = 0,01$ , dan kernel = linear. Hasil ini menunjukkan bahwa tuning tidak hanya berfokus pada akurasi, tetapi juga pada peningkatan generalisasi model terhadap data tidak seimbang. Untuk penelitian selanjutnya, disarankan dapat mengintegrasikan teknik penyeimbang data seperti SMOTE atau class weighting, serta mengeksplorasi model lain seperti Random Forest, Naive Bayes, LSTM, dan BERT untuk meningkatkan performa pada kelas minoritas.

#### DAFTAR PUSTAKA

- [1] A. M. Ndapamuri, D. Manongga, and A. Iriani, "Analisis Sentimen Ulasan Aplikasi Tripadvisor Dengan Metode Support Vector Machine, K-Nearest Neighbor, Dan Naive Bayes," *INOVTEK Polbeng - Seri Inform.*, vol. 8, no. 1, p. 127, 2023, doi: 10.35314/isi.v8i1.3260.
- [2] R. Sarudin, "Analisis Online Review Tripadvisor . Com Terhadap Minat Pembelian Produk Jasa Akomodasi Di Hotel Manhattan Tripadvisor . Com Review Online Analysis on the Interest of Buying Accommodation Services in Hotel," *J. Hosp. dan Pariwisata*, vol. 7, pp. 33–46, 2021.
- [3] S. Suparyati and A. Fathurrahman, "Analisis Sentimen Dengan Klasifikasi Naive Bayes pada Review Hotel Tripadvisor," *J. Ilm. Inform.*, vol. 10, no. 01, pp. 20–24, 2022, doi: 10.33884/jif.v10i01.4524.
- [4] B. Bayu Baskoro et al., "Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR)," *J. Informatics Inf. Syst. Softw. Eng. Appl. (INISTA)*, vol. Volume 3 N, no. 2, pp. 21–029, 2021, doi: 10.20895/INISTA.V3.
- [5] M. H. AUFAN, M. R. Handayani, A. B. Nurjanna, and N. C. Hendro, "THE PERCEPTIONS OF SEMARANG FIVE STAR HOTEL TOURISTS WITH SUPPORT VECTOR MACHINE ON GOOGLE REVIEWS PERSEPSI WISATAWAN HOTEL BINTANG LIMA SEMARANG DENGAN," *J. Tek. Inform.*, vol. 5, no. 5, pp. 1241–1247, 2025.
- [6] A. Saputra, R. C. Sigitta Hariyono, and N. M. Saraswati, "Analisis Sentimen Pengguna Aplikasi MyPertamina Menggunakan Algoritma Bidirectional Long Short Term Memory," *J. Eksplor Inform.*, vol. 13, no. 2, pp. 156–163, 2024, doi: 10.30864/eksplor.v13i2.973.
- [7] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1767, 2022, doi: 10.30865/mib.v6i3.4218.
- [8] W. Nugraha and A. Sasongko, "Hyperparameter Tuning on Classification Algorithm with Grid Search," *Sistemasi*, vol. 11, no. 2, p. 391, 2022, doi: 10.32520/stmsi.v11i2.1750.
- [9] Afis Julianto, Andi Sunyoto, and Ferry Wahyu Wibowo, "Optimasi Hyperparameter Convolutional Neural Network Untuk Klasifikasi Penyakit Tanaman Padi," *Tek. Teknol. Inf. dan Multimed.*, vol. 3, no. 2, pp. 98–105, 2022, doi: 10.46764/teknimedia.v3i2.77.
- [10] T. A. E. Putri, T. Widiarihari, and R. Santoso, "Penerapan Tuning Hyperparameter Randomsearchcv Pada Adaptive Boosting Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung," *J. Gauss*, vol. 11, no. 3, pp. 397–406, 2023, doi: 10.14710/j.gauss.11.3.397-406.
- [11] Q. Ain, E. Utami, and A. Nasiri, "Analisis sentimen: prediksi," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 3, pp. 1586–1595, 2024.
- [12] Anugerah Simanjuntak et al., "Research and Analysis of IndoBERT Hyperparameter Tuning in Fake News Detection," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 13, no. 1, pp. 60–67, 2024, doi: 10.22146/jnteti.v13i1.8532.
- [13] A. M. Yolanda and R. T. Mulya, "Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store," *VARIANSI J. Statistics Its Appl. Teach. Res.*, vol. 6, no. 2, pp. 76–83, 2024, doi: 10.35580/variansiunm258.
- [14] L. Trihardianingsih and G. S. Lasatira, "Optimasi Hyperparameter GridSearchCV pada Klasifikasi Kualitas Udara menggunakan Support Vector Machine," *JITU (Jurnal Inf. dan Teknol.)*, vol. 1, no. 2, pp. 40–47, 2024.
- [15] M. Fajri and A. Primajaya, "Komparasi Teknik Hyperparameter Optimization pada SVM untuk Permasalahan Klasifikasi dengan Menggunakan Grid Search dan Random Search," *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 14–19, 2023, doi: 10.30871/jaic.v7i1.5004.
- [16] R. Muzayanah, D. A. A. Pertiwi, M. Ali, and M. A. Muslim, "Comparison of gridsearchcv and bayesian hyperparameter optimization in random forest algorithm for diabetes prediction," *J. Soft Comput. Explor.*, vol. 5, no. 1, pp. 86–91, 2024, doi: 10.52465/jsosex.v5i1.308.
- [17] S. Nauli, S. S. Berutu, H. Budiati, and F. Maedjaja, "Klasifikasi kalimat perundungan pada twitter menggunakan algoritma support vector machine," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 1, pp. 107–122, 2025.
- [18] J. P. Hidayat and I. Nurhaida, "Analisis sentimen pada ulasan lms pembelajaran menggunakan metode natural language processing," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 1, pp. 565–575, 2025.
- [19] N. Hadi and D. Sugiarto, "Analisis Sentimen Pembangunan IKN pada Media Sosial X Menggunakan Algoritma SVM , Logistic Regression dan Naive Bayes," *J. Inform. J. Pengemb. IT*, vol. 10, no. 1, pp. 37–49, 2025, doi: 10.30591/jpit.v10i1.7106.
- [20] R. Titik Misriati; Aryanti, "Optimasi Random Forest dan Support Vector Machine dengan Hyperparameter GridSearchCV untuk Analisis Sentimen Ulasan PrimaKu," *J. Inf. Syst. Res.*, vol. 5, no. 4, pp. 1333–1341, 2024, doi: 10.47065/josh.v5i4.5347.
- [21] I. Muhamad Malik Matin, "Hyperparameter Tuning Menggunakan GridsearchCV pada Random Forest untuk Deteksi Malware," *Multinetics*, vol. 9, no. 1, pp. 43–50, 2023, doi: 10.32722/multinetics.v9i1.5578.
- [22] Z. Maisat, E. Darmawan, and A. Fauzan, "Implementasi Optimasi Hyperparameter GridSearchCV Pada Sistem Prediksi Serangan Jantung Menggunakan SVM Implementation of GridSearchCV Hyperparameter Optimization in Heart Attack Prediction System Using SVM," *Unipdu*, vol. 13, no. 1, pp. 8–15, 2023.
- [23] E. I. Saputra et al., "OPTIMALISASI ALGORITMA SUPPORT VECTOR MACHINE PADA ASPECT-BASED SENTIMENT ANALYSIS,"

*Zo. J. Sist. Inf.*, vol. 7, no. 1, pp. 271–278, 2025.

- [24] L. González-Castro, M. Chávez, P. Duflot, V. Bleret, G. Del Fiol, and M. López-Nores, “Impact of Hyperparameter Optimization to Enhance Machine Learning Performance: A Case Study on Breast Cancer Recurrence Prediction,” *Appl. Sci.*, vol. 14, no. 13, 2024, doi: 10.3390/app14135909.