

**ANALISIS SENTIMEN PADA REVIEW TOKOPEDIA
MENGUNAKAN ALGORITMA SVM DENGAN
HYPERPARAMETER TUNING DAN METODE GRID
SEARCH**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

ERLAM MUHAMMAD FAJAR

20.83.0479

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS SENTIMEN PADA REVIEW TOKOPEDIA
MENGUNAKAN ALGORITMA SVM DENGAN
HYPERPARAMETER TUNING DAN METODE GRID
SEARCH**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

ERLAM MUHAMMAD FAJAR

20.83.0479

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PADA REVIEW TOKOPEDIA MENGGUNAKAN
ALGORITMA SVM DENGAN HYPERPARAMETER TUNING DAN
METODE GRID SEARCH**

yang disusun dan diajukan oleh

Erlam Muhammad Fajar

20.83.0479

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 September 2024

Dosen Pembimbing,



Robert Marco, S.T., M.T., Ph.D.

NIK. 190302228

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PADA REVIEW TOKOPEDIA MENGGUNAKAN
ALGORITMA SVM DENGAN HYPERPARAMETER TUNING DAN
METODE GRID SEARCH**

yang disusun dan diajukan oleh

Erlam Muhammad Fajar

20.83.0479

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Desember 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Senie Destya, S.T., M.Kom.
NIK. 190302312



Muhammad Kopravi, S.Kom., M.Eng.
NIK. 190302454



Robert Marco, S.T., M.T., Ph.D.
NIK. 190302228

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Desember 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Erlam Muhammad Fajar
NIM : 20.83.0479

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN PADA REVIEW TOKOPEDIA MENGGUNAKAN ALGORITMA SVM DENGAN HYPERPARAMETER TUNING DAN METODE GRID SEARCH

Dosen Pembimbing : Robert Marco, S.T., M.T., Ph.D.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 17 Desember 2024

Yang Menyatakan,



The image shows a handwritten signature in black ink over a 1000 Rupiah postage stamp. The stamp features the Garuda Pancasila emblem and the text 'SPULUH RIBU RUPAH', '1000', 'METERAI TEMPEL', and the serial number 'BE7EAMX073190427'.

Erlam Muhammad Fajar

HALAMAN PERSEMBAHAN

Saya mengucapkan rasa terima kasih dan penghormatan yang setinggi-tingginya kepada kedua orang tua saya atas dukungan moral maupun finansial yang tak henti-hentinya diberikan selama menempuh pendidikan ini. Tanpa doa, kasih sayang, dan pengorbanan mereka, pencapaian ini tidak akan mungkin terwujud.

Rasa terima kasih juga saya tujukan kepada seluruh dosen dan staf di Universitas Amikom Yogyakarta, yang telah memberikan ilmu, pengalaman, serta fasilitas yang mendukung studi saya. Terima kasih atas ilmu yang berharga serta kesempatan untuk mengembangkan diri di lingkungan akademik yang penuh dukungan.

Ucapan terima kasih ini juga saya sampaikan kepada teman-teman seperjuangan, yang selalu siap membantu, berbagi suka dan duka, serta saling mendukung di setiap tahap perkuliahan hingga proses penyusunan tugas akhir ini. Kehadiran kalian semua membuat perjalanan ini menjadi lebih bermakna dan berkesan.

Terakhir, penghargaan dan terima kasih kepada semua pihak yang tidak dapat saya sebutkan satu per satu, yang telah memberikan bantuan dan dukungannya, baik secara langsung maupun tidak langsung. Semoga kebaikan dan dukungan kalian mendapat balasan yang berlipat ganda.

Dengan rasa syukur dan rendah hati, karya ini saya persembahkan.

KATA PENGANTAR

Segala puji dan syukur penulis haturkan kepada Tuhan Yang Maha Esa atas karunia dan anugerah-Nya, sehingga penulisan skripsi dengan judul “Analisis Sentimen pada Review Tokopedia Menggunakan Algoritma SVM dengan Hyperparameter Tuning dan Metode Grid Search” dapat diselesaikan dengan baik. Penulis berharap hasil penelitian ini dapat memberikan manfaat bagi banyak pihak, khususnya dalam pengembangan ilmu di bidang analisis sentimen dan pemrosesan bahasa alami (NLP).

Penelitian ini disusun sebagai salah satu syarat untuk menyelesaikan program studi di Teknik Komputer pada Universitas Amikom Yogyakarta. Dalam proses penyusunan skripsi ini, penulis menyadari bahwa skripsi ini tidak akan dapat terselesaikan tanpa bantuan, dukungan, dan bimbingan dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat, penulis mengucapkan terima kasih kepada:

1. Bapak, Ibu, Mama, Kakak serta Adik saya atas dukungan moral dan kasih sayang yang terus berbalas tiada hentinya
2. Bapak Robert Marco, M.T., Ph.D. yang dengan sabar memberikan arahan, bimbingan, dan motivasi selama proses penyusunan penelitian ini.
3. Rekan TK01 & TK02, Sadder Saddest dan Keluarga The Kontrakans yang telah memberikan semangat, bantuan, dan kebersamaan yang berarti selama proses penyusunan skripsi ini.
4. *Last but not least, I wanna thank for me, I wanna thank me for believing in me, I wanna thank me for doing all this hard works, I wanna thank me for having no days off, I wanna thank me for never quitting, I wanna thank me for always being a giver and tryna give more than I receive, I wanna thank me for tryna do more right than wrong and I wanna thank me for just being me at all times.*

Demikian kata pengantar ini Terima Kasih untuk semuanya.

Yogyakarta, 12 November 2024

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	13

2.2.1 Jupyter Notebook	13
2.2.2 Python	13
2.2.3 Tokopedia	13
2.2.4 Analisis Sentimen	14
2.2.5 Support Vector Machine (SVM).....	14
2.2.6 Hyperparameter Tuning.....	14
2.2.7 Metode Grid Search	15
2.2.8 Term Frequency-Inverse Document Frequency (TF-IDF)	15
BAB III METODE PENELITIAN	16
3.1 Objek Penelitian	16
3.2 Alur Penelitian.....	16
3.3 Data dan Pustaka	17
3.3.1 Dataset.....	17
3.3.2 Pustaka dan Perangkat Lunak	18
3.4 Exploratory Data Analysis (EDA)	19
3.4.1 Pengecekan Struktur Data.....	19
3.4.2 Pengecekan Nilai Kosong.....	19
3.4.3 Visualisasi WordCloud	19
3.5 Pre-Processing Data	20
3.5.1 Case Folding dan Penghapusan Karakter Khusus.....	21
3.5.2 Tokenization	21
3.5.3 Stopword Removal dan Lemmatizing.....	21
3.5.4 Stopword Removal.....	22
3.5.5 Lemmatizing	22
3.5.6 Koreksi Penulisan	22

3.5.7	Hapus teks “NaN”	23
3.5.8	Stemming	23
3.6	Model dan Teknik Evaluasi.....	24
3.6.1	Teknik Evaluasi Model	25
3.6.2	Teknik Validasi Model	26
3.7	Eksperimen dan Pengaturan Hyperparameter	27
3.7.1	Evaluasi Test Size	27
3.7.2	Hyperparameter Tuning	27
BAB IV HASIL DAN PEMBAHASAN		29
4.1	Hasil Exploratory Data Analyst (EDA)	29
4.1.1	Karakteristik Data	29
4.1.2	Frekuensi Kata dalam Teks.....	29
4.2	Hasil Modeling.....	31
4.2.1	Hasil Hyperparameter Tuning.....	31
4.2.2	Kombinasi Hyperparameter Terbaik.....	31
4.2.3	Hasil Evaluasi Model	32
4.2.4	Hasil Evaluasi Performa Model	33
4.3	Visualisasi Confusion Matrix.....	34
4.3.1	Penjelasan Confusion Matrix	34
4.3.2	Analisis Hasil Confusion Matrix	35
BAB V PENUTUP		36
5.1	Kesimpulan.....	36
5.2	Saran.....	37
REFERENSI		38
LAMPIRAN.....		40

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	8
Tabel 3.1 Aspek dan Deskripsi Data penelitian	17
Tabel 3.2 Pustaka dan Perangkat Lunak penelitian	18
Tabel 3.3 Hyperparameter Tuning	28
Tabel 4.1 Jumlah Sentimen	29
Tabel 4.2 Hyperparameter Tuning	32
Tabel 4.3 Hasil Evaluasi Model	32

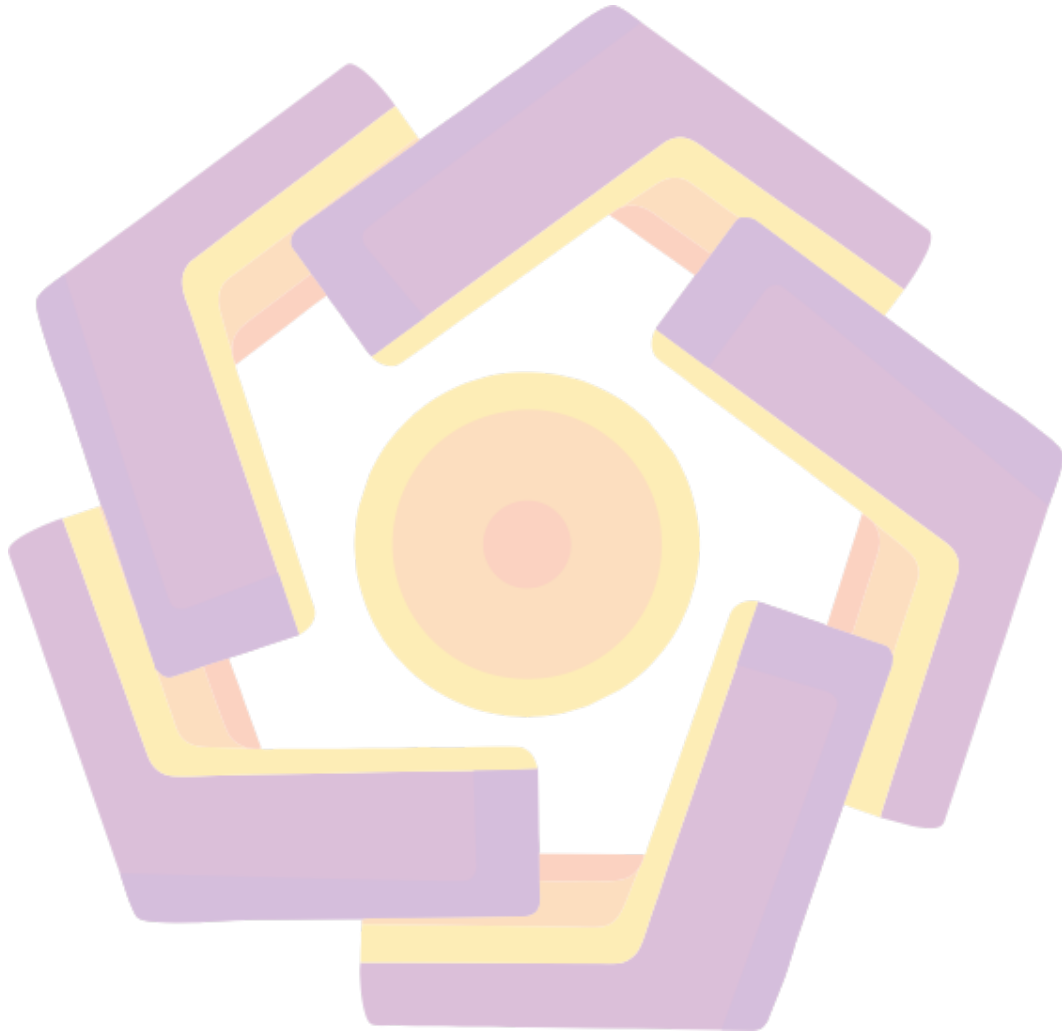


DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	17
Gambar 3.2 Data Hasil Scraping	18
Gambar 3.3 Pengecekan Struktur Data	19
Gambar 3.4 Pengecekan Nilai Kosong dan Diagram Data	19
Gambar 3.5 Positive Word Cloud	20
Gambar 3.6 Negative Word Cloud	20
Gambar 3.7 Case Folding dan Penghapusan Karakter Khusus	21
Gambar 3.8 Tokenization	21
Gambar 3.9 Stopword Removal dan Lemmatizing	22
Gambar 3.10 Koreksi Penulisan	23
Gambar 3.11 Hapus Teks “NaN”	23
Gambar 3.12 Stemming	24
Gambar 3.13 SVM Classifier Performance with Different Test Sizes	27
Gambar 4.1 Frekuensi Kata dalam Dataset	30
Gambar 4.1 Frekuensi Kata pada Review Positif	30
Gambar 4.2 Frekuensi Kata pada Review Negatif	31
Gambar 4.3 Hasil Performa Model	33
Gambar 4.4 Confusion Matrix	34

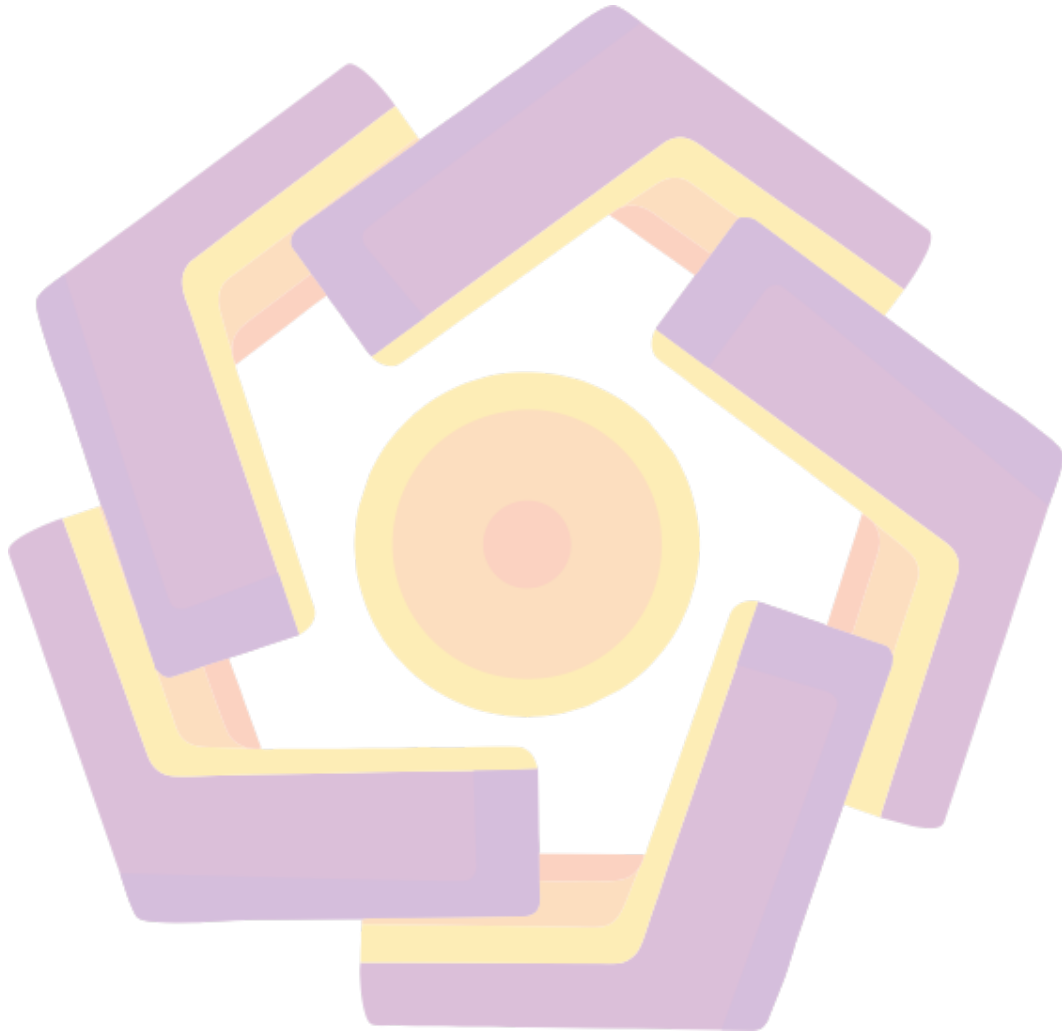
DAFTAR LAMPIRAN

Lampiran 1: File Scraping Data	40
Lampiran 2: File Preprocessing Data	41
Lampiran 3: File Hyperparameter Tuning	47



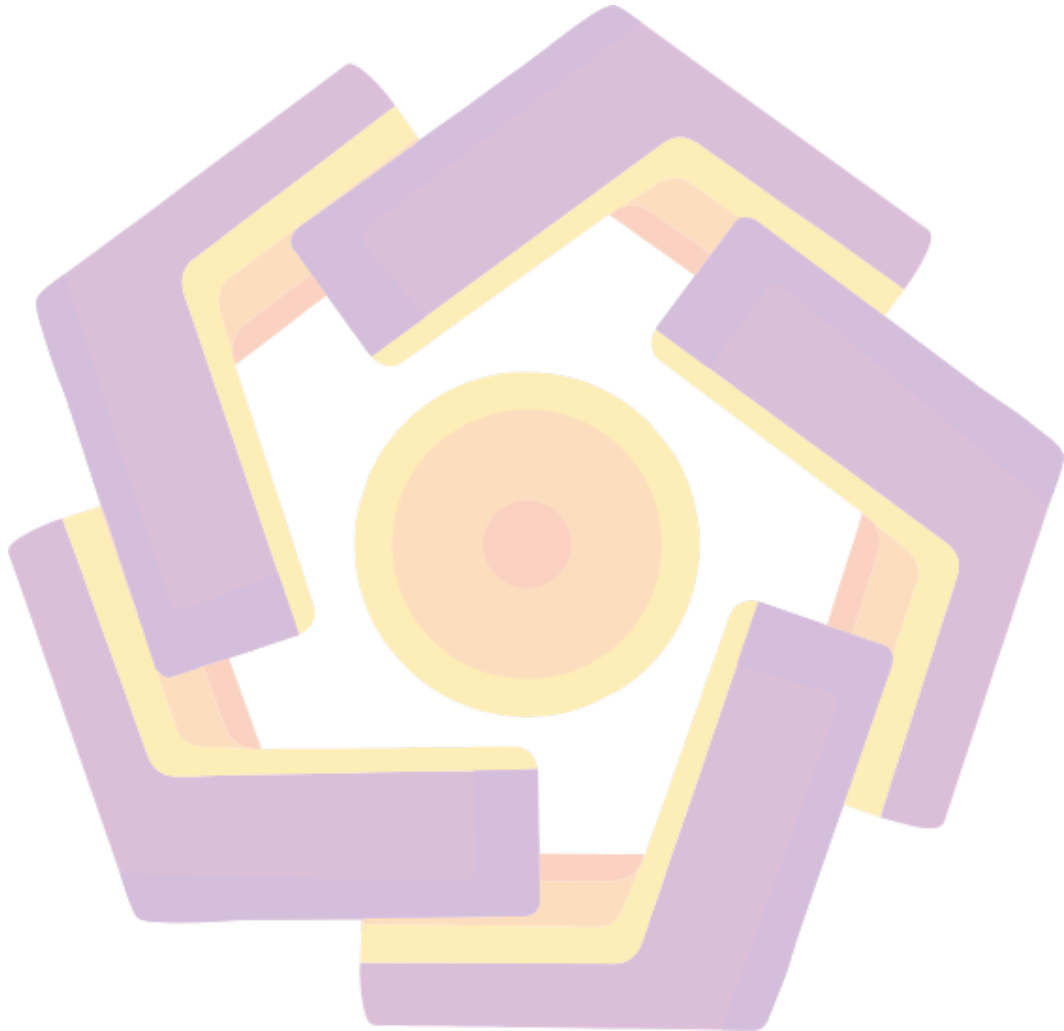
DAFTAR LAMBANG DAN SINGKATAN

TF-IDF	Term Frequency-Inverse Document Frequency
EDA	Exploratory Data Analyst
SVM	Support Vector Machines



DAFTAR ISTILAH

Case Folding	Menyamarkan penggunaan huruf kapital dalam data teks
Dataset	Sekumpulan data
Stemming	Pemotongan imbuhan



INTISARI

Analisis sentimen pada ulasan pengguna merupakan salah satu cara efektif untuk memahami persepsi pelanggan terhadap layanan yang diberikan oleh suatu platform, termasuk e-commerce seperti Tokopedia. Penelitian ini bertujuan untuk menganalisis sentimen pada ulasan pengguna Tokopedia dengan menerapkan algoritma Support Vector Machine (SVM) sebagai model klasifikasi. Algoritma SVM dipilih karena kemampuannya dalam menangani data dengan dimensi tinggi dan memberikan hasil yang baik pada klasifikasi teks.

Untuk meningkatkan akurasi dan performa model, penelitian ini menggunakan teknik hyperparameter tuning dengan metode Grid Search. Grid Search memungkinkan pemilihan kombinasi hyperparameter terbaik dengan cara mengevaluasi performa model pada berbagai konfigurasi parameter. Evaluasi model dilakukan dengan menggunakan metrik akurasi, precision, recall, dan F1-score untuk mengukur efektivitas dari model yang dibangun.

Hasil penelitian menunjukkan bahwa model SVM yang telah dioptimalkan menggunakan Grid Search mampu mencapai akurasi yang tinggi dalam klasifikasi sentimen pada ulasan Tokopedia. Nilai akurasi yang diperoleh sebesar **91,28%**, dengan **precision 91,40%**, **recall 91,28%**, dan **F1-score 91,30%**. Hasil ini menunjukkan bahwa model yang dibangun memiliki keseimbangan yang baik antara ketepatan dan sensitivitas dalam mendeteksi sentimen positif dan negatif.

Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan sistem analisis sentimen di industri e-commerce, serta menjadi acuan bagi penelitian selanjutnya dalam pengembangan model klasifikasi teks dengan performa yang lebih baik.

Kata kunci: Analisis Sentimen, Tokopedia, Support Vector Machine (SVM), Hyperparameter Tuning, Grid Search.

ABSTRACT

Sentiment Analysis on User Reviews as an Effective Way to Understand Customer Perceptions of Services Provided by Platforms, Including E-commerce such as Tokopedia

This study aims to analyze sentiment in Tokopedia user reviews by applying the Support Vector Machine (SVM) algorithm as a classification model. SVM is chosen for its capability to handle high-dimensional data and its effectiveness in text classification.

To improve the model's accuracy and performance, this research uses hyperparameter tuning with the Grid Search method. Grid Search enables the selection of the best hyperparameter combination by evaluating the model's performance across various parameter configurations. The model evaluation is conducted using accuracy, precision, recall, and F1-score metrics to measure the effectiveness of the developed model.

*The study results indicate that the SVM model optimized with Grid Search achieved high accuracy in sentiment classification of Tokopedia reviews. The accuracy achieved is **91.28%**, with a precision of **91.40%**, recall of **91.28%**, and **F1-score of 91.30%**. These results demonstrate that the developed model maintains a good balance between accuracy and sensitivity in detecting positive and negative sentiment.*

This research is expected to contribute to the development of sentiment analysis systems in the e-commerce industry and serve as a reference for future research in developing text classification models with better performance.

Keyword: *Sentiment Analysis, Tokopedia, Support Vector Machine (SVM), Hyperparameter Tuning, Grid Search*