

# Artikel Jurnal Computing

## ANALISIS SENTIMEN ULASAN PENGGUNA E-COMMERCE SHOPEE MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM)

Nama Penulis

Ardi Kamal Karima  
NIM 301230023

**ABSTRAK:** Perkembangan platform e-commerce mendorong meningkatnya jumlah ulasan pengguna yang berisi opini terhadap layanan yang diberikan. Ulasan tersebut dapat dimanfaatkan untuk mengetahui tingkat kepuasan pengguna, namun jumlah data yang besar menyebabkan analisis manual menjadi tidak efektif. Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis sentimen ulasan pengguna Shopee menggunakan metode Support Vector Machine (SVM). Dataset yang digunakan berasal dari ulasan pengguna Shopee yang diperoleh dari Kaggle dan berisi teks ulasan serta skor penilaian. Tahapan penelitian meliputi text preprocessing (case folding, cleaning, stopword removal), ekstraksi fitur menggunakan TF-IDF, pelatihan model SVM dengan kernel Linear dan RBF, serta evaluasi menggunakan confusion matrix dan metrik akurasi. Selain itu, sistem dikembangkan dalam bentuk aplikasi web berbasis Streamlit yang memungkinkan pengguna mengunggah dataset dan melakukan prediksi sentimen secara interaktif. Hasil penelitian menunjukkan bahwa kernel Linear memberikan performa yang lebih baik dibandingkan kernel RBF pada data teks berdimensi tinggi. Sistem yang dikembangkan mampu mengklasifikasikan sentimen positif dan negatif dengan baik serta dapat digunakan sebagai alat bantu analisis opini pengguna.

**Kata kunci:** Analisis Sentimen, Support Vector Machine, TF-IDF, Text Mining, Streamlit

**ABSTRACT:** *The rapid growth of e-commerce platforms has increased the number of user reviews containing opinions about services. These reviews can be utilized to measure user satisfaction; however, manual analysis becomes inefficient due to large data volumes. This study aims to perform sentiment analysis of Shopee user reviews using the Support Vector Machine (SVM) method. The dataset was obtained from Kaggle and consists of review texts and rating scores. The research stages include text preprocessing (case folding, cleaning, stopword removal), feature extraction using TF-IDF, training SVM models with Linear and RBF kernels, and evaluation using confusion matrix and accuracy metrics. Furthermore, a web-based application was developed using Streamlit to allow users to upload datasets and perform sentiment predictions interactively. The results show that the Linear kernel outperforms the RBF kernel for high-dimensional text data. The developed system is capable of accurately classifying positive and negative sentiments and can be used as a decision-support tool for opinion analysis.*

**Keywords:** Sentiment Analysis, Support Vector Machine, TF-IDF, Text Mining, Streamlit

## PENDAHULUAN

Perkembangan teknologi informasi telah mendorong pertumbuhan pesat platform e-commerce di Indonesia, salah satunya Shopee. Setiap transaksi yang dilakukan pengguna sering kali disertai dengan ulasan yang berisi opini dan penilaian terhadap kualitas layanan, pengiriman, serta fitur aplikasi. Ulasan tersebut menjadi sumber data yang sangat berharga bagi perusahaan untuk mengevaluasi kinerja layanan dan meningkatkan kepuasan pelanggan.

Namun, jumlah ulasan yang sangat besar menyebabkan analisis secara manual menjadi sulit dan tidak efisien. Oleh karena itu, diperlukan suatu metode otomatis yang mampu mengklasifikasikan opini pengguna ke dalam sentimen positif atau negatif. Analisis sentimen merupakan salah satu cabang dari text mining yang bertujuan untuk mengekstraksi informasi subjektif dari data teks.

Support Vector Machine (SVM) merupakan salah satu algoritma klasifikasi yang banyak digunakan dalam analisis sentimen karena kemampuannya dalam menangani data berdimensi tinggi seperti teks. Selain itu, penggunaan teknik ekstraksi fitur TF-IDF mampu merepresentasikan teks dalam bentuk numerik secara efektif. Berdasarkan latar belakang tersebut, penelitian ini mengimplementasikan analisis sentimen ulasan pengguna Shopee menggunakan SVM serta mengembangkan sistem dalam bentuk aplikasi web agar mudah digunakan.

## METODA

Penelitian ini menggunakan pendekatan machine learning dengan tahapan sebagai berikut:

### 1. DATASET

Dataset yang digunakan berasal dari Kaggle dan berisi ulasan pengguna Shopee dalam bahasa Indonesia. Dataset memiliki atribut utama berupa teks ulasan (content) dan skor penilaian (score). Skor tersebut digunakan untuk membentuk label sentimen, di mana skor  $\leq 2$  dikategorikan sebagai sentimen negatif dan skor  $\geq 4$  sebagai sentimen positif.

### 2. TEXT PREPROCESSING

Tahapan preprocessing dilakukan untuk membersihkan data teks, meliputi:

- Case folding (mengubah teks menjadi huruf kecil)
- Penghapusan tanda baca, angka, dan karakter khusus
- Stopword removal untuk menghilangkan kata tidak bermakna
- Normalisasi teks

### 3. FEATURE EXTRACTION

Ekstraksi fitur dilakukan menggunakan metode TF-IDF untuk mengubah teks menjadi vektor numerik. Metode ini dipilih karena mampu menekan pengaruh kata yang sering muncul namun kurang informatif,

### 4. KLASIFIKASI DENGAN SVM

Model SVM dilatih menggunakan dua kernel, yaitu Linear dan Radial Basis Function (RBF). Dataset dibagi menjadi data latih (80%) dan data uji (20%). Selain itu, dilakukan penyesuaian parameter untuk menghindari overfitting dan underfitting.

## 5. EVALUASI

Evaluasi model dilakukan menggunakan confusion matrix dan metrik akurasi. Visualisasi WordCloud digunakan untuk mengidentifikasi kata-kata dominan pada sentimen positif dan negatif.

## 6. IMPLEMENTASI APLIKASI

Model yang telah dilatih diintegrasikan ke dalam aplikasi web berbasis Streamlit yang mendukung unggah dataset dan prediksi sentimen secara langsung.

## HASIL DAN PEMBAHASAN

Hasil pelatihan model menunjukkan bahwa SVM dengan kernel Linear menghasilkan akurasi yang lebih tinggi dibandingkan kernel RBF. Hal ini disebabkan oleh karakteristik data teks yang memiliki dimensi fitur sangat tinggi sehingga lebih cocok ditangani oleh kernel Linear.

Confusion matrix menunjukkan bahwa model mampu mengklasifikasikan sebagian besar ulasan dengan benar. Kesalahan prediksi umumnya terjadi pada ulasan yang bersifat ambigu atau sarkasme. WordCloud yang dihasilkan memperlihatkan bahwa kata-kata seperti *bagus*, *mantap*, dan *puas* dominan pada sentimen positif, sedangkan kata *kecewa*, *buruk*, dan *masalah* sering muncul pada sentimen negatif.

Aplikasi Streamlit yang dikembangkan memungkinkan pengguna untuk mengunggah dataset, melatih model, dan menguji prediksi sentimen secara interaktif baik melalui perangkat desktop maupun mobile.

## SIMPULAN DAN SARAN

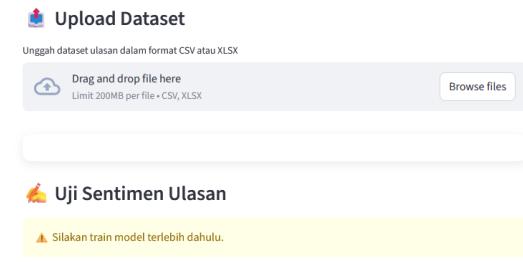
Berdasarkan hasil penelitian, dapat disimpulkan bahwa metode Support Vector Machine dengan ekstraksi fitur TF-IDF mampu mengklasifikasikan sentimen ulasan pengguna Shopee dengan baik. Kernel Linear terbukti memberikan performa yang lebih optimal dibandingkan kernel RBF pada data teks. Sistem yang dikembangkan dalam bentuk aplikasi web memberikan kemudahan bagi pengguna dalam melakukan analisis sentimen secara mandiri.

Sebagai saran, penelitian selanjutnya dapat mengembangkan model dengan menambahkan kelas sentimen netral, menggunakan algoritma deep learning, atau mengintegrasikan sistem dengan API untuk analisis data secara real-time.

## OUTPUT HASIL PROGRAM



Gambar 1. Tampilan Program SVM



Gambar 2. Tampilan Program SVM

**Upload Dataset**

Unggah dataset ulasan dalam format CSV atau XLSX

Drag and drop file here  
Limit 200MB per file • CSV, XLSX

Browse files

Data ulasan Shopee tentang COD.csv 0.5MB

Dataset berhasil diupload

Lihat contoh dataset

**Ketentuan Dataset**

- Memiliki kolom teks ulasan
- Memiliki kolom skor (1-5)
- Skor 1-2 → Negatif
- Skor 4-5 → Positif
- Skor 3 → Tidak digunakan

**Uji Sentimen Ulasan**

Masukkan ulasan pengguna

Aplikasi sangat membantu saya

Prediksi Sentimen

Sentimen POSITIF

Gambar 3. Upload Dataset Dan Train\_Model

**Konfigurasi Dataset**

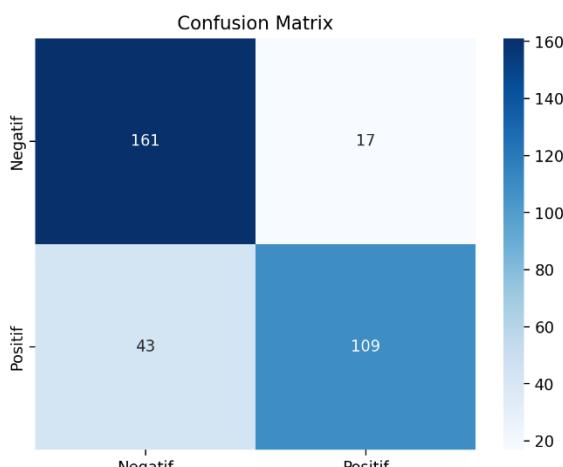
Kolom teks ulasan  
content

Kolom skor  
score

Train Model

Training selesai | Akurasi: 0.82

Gambar 4. Train\_Model Dari Dataset



Gambar 5. Hasil Confusion Matrix

## PUSTAKA ACUAN

- Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval.*
- Joachims, T. (1998). *Text Categorization with Support Vector Machines. European Conference on Machine Learning.*
- Pedregosa, F., et al. (2011). *Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.*
- Kaggle. *Dataset Analisis Sentimen Ulasan Shopee.*

## LINK PROJECT/PROGRAM

### GitHub:

<https://github.com/Ardikamal/sentiment-svm-shopee.git>

### Deploy Streamlit:

<https://sentiment-svm-shopee-ardikamalkarima.streamlit.app/>

### Sumber Dataset:

<https://www.kaggle.com/code/alvianardiansyah/analisis-sentimen-pengguna-shopee-dengan-svm/input>

### Link Video Penjelasan Program:

[https://www.tiktok.com/@ardikamal/video/7589364144491613460?is\\_from\\_webapp=1&sender\\_device=pc&web\\_id=7579701110202910215](https://www.tiktok.com/@ardikamal/video/7589364144491613460?is_from_webapp=1&sender_device=pc&web_id=7579701110202910215)