

# Analisis Sentimen Pada Ulasan Aplikasi Ferizy Menggunakan Metode Support Vector Machine

Oleh:

Rizky Septia Putra

Rohman Dijaya

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Oktober, 2023

# Pendahuluan

- Dalam upaya untuk meningkatkan kualitas pelayanan kepada pelanggan di sektor transportasi laut, PT. ASDP Indonesia Ferry telah mengambil langkah penting dengan menerapkan transformasi digital pada proses pemesanan dan pembelian tiket feri secara daring melalui peluncuran sistem aplikasi dan situs web resmi bernama Ferizy pada bulan Mei 2020. Platform aplikasi dan situs web Ferizy dibuat dengan maksud untuk menyederhanakan proses pemesanan dan pembelian tiket feri. Saat ini aplikasi Ferizy telah tersedia di Google Play Store dan dapat diunduh serta pengguna dapat memberikan ulasan, kritik, dan saran terhadap fungsi maupun kinerja dari keseluruhan aplikasi.
- Menurut informasi yang diperoleh dari Google Play Store pada bulan Agustus 2023, aplikasi Ferizy telah diunduh oleh lebih dari 1 juta pengguna dan menerima rating sebesar 3,4. Ulasan yang tersedia di Google Play Store adalah sumber informasi berharga yang berjumlah besar dan bersifat tidak terstruktur sehingga diperlukan suatu teknik untuk mengetahui bagaimana ulasan pengguna terhadap aplikasi tersebut. Analisis sentimen dapat digunakan untuk mengidentifikasi ulasan pengguna menjadi ulasan sentimen positif dan sentimen negatif. Salah satu metode untuk mengklasifikasikan data ulasan di bidang text mining yaitu Support Vector Machine (SVM). Support Vector Machine (SVM) adalah algoritma yang terkenal karena kemampuannya menghasilkan solusi optimal dalam konteks klasifikasi. Algoritma ini dikembangkan oleh Vapnik sebagai model machine learning yang menggunakan kernel untuk tugas klasifikasi dan regresi.

# Rumusan Masalah

- Bagaimana mengembangkan sistem identifikasi analisis sentimen ke dalam ulasan kelas positif dan negatif pada aplikasi Ferizy ?
- Bagaimana mengukur tingkat akurasi analisis sentimen yang dihasilkan menggunakan metode Support Vector Machine ?”

# Metode

01

Pengumpulan Data

Dataset diperoleh dari Google Play Store. Dataset berupa kumpulan ulasan pengguna aplikasi Ferizy yang berjumlah 1500 data ulasan. Pelabelan data berdasarkan rating.

02

Text Preprocessing

- Case Folding
- Tokenizing
- Normalisasi Kata
- Filtering StopWord
- Stemming

03

Pembobotan Kata

Pembobotan TF-IDF (Term Frequency-Inverse Document Frequency)

04

Klasifikasi Data

Klasifikasi menggunakan Metode Support Vector Machine

05

Pengujian

Pengujian Hasil Akurasi menggunakan Confusion Matrix

# Hasil

Berdasarkan hasil pengujian, ditemukan bahwa penggunaan kernel RBF pada penelitian ini menghasilkan akurasi tertinggi dibandingkan dengan kernel lainnya. Oleh karena itu, penggunaan metode Support Vector Machine dengan total data sebanyak 1.500 data berhasil mencapai tingkat akurasi yang signifikan, yaitu mencapai 90,71% dan evaluasi hasil pengujian dengan menggunakan Confusion Matrix juga memvalidasi nilai akurasi sebesar 91% pada pembagian dataset dengan nilai rasio perbandingan 90% : 10%.

# Pembahasan

## Hasil Preprocessing

### Case Folding

Input	Output
Susah di akses	susah di akses
Kadang suka susah mau bayar	kadang suka susah mau bayar
Aplikasinya bagus dan bermanfaat	aplikasinya bagus dan bermanfaat

### Tokenizing

Input	Output
susah di akses	['susah', 'di', 'akses']
kadang suka susah mau bayar	['kadang', 'suka', 'susah', 'mau', 'bayar']
aplikasinya bagus dan bermanfaat	['aplikasinya', 'bagus', 'dan', 'bermanfaat']

# Pembahasan

## Hasil Preprocessing

### Normalisasi Kata

Input	Output
['susah', 'di', 'akses']	['susah', 'di', 'akses']
['kadang', 'suka', 'susah', 'mau', 'bayar']	['kadang', 'suka', 'susah', 'mau', 'bayar']
['aplikasinya', 'bagus', 'dan', 'bermanfaat']	['aplikasinya', 'bagus', 'dan', 'bermanfaat']

### Filtering Stopword

Input	Output
['susah', 'di', 'akses']	['susah', 'akses']
['kadang', 'suka', 'susah', 'mau', 'bayar']	['kadang', 'suka', 'susah', 'bayar']
['aplikasinya', 'bagus', 'dan', 'bermanfaat']	['aplikasi', 'bagus', 'bermanfaat']

# Pembahasan

Hasil Preprocessing

Stemming

Input	Output
['susah', 'akses']	['susah', 'akses']
['kadang', 'suka', 'susah', 'bayar']	['kadang', 'suka', 'susah', 'bayar']
['aplikasi', 'bagus', 'bermanfaat']	['aplikasi', 'bagus', 'manfaat']

Setelah tahap preprocessing data selesai, langkah selanjutnya adalah menyimpan hasil preprocessing ini ke dalam sebuah file baru. File tersebut akan digunakan sebagai dataset dalam proses klasifikasi.



# Pembahasan

## Hasil Pengujian Menggunakan Kernel Linear

Rasio Perbandingan	Akurasi
Data Latih 70% : Data Uji 30%	87,07%
Data Latih 80% : Data Uji 20%	87,16%
Data Latih 90% : Data Uji 10%	87,43%

## Hasil Pengujian Menggunakan Kernel Polynomial

Rasio Perbandingan	Akurasi
Data Latih 70% : Data Uji 30%	80,15%
Data Latih 80% : Data Uji 20%	84,15%
Data Latih 90% : Data Uji 10%	85,25%

# Pembahasan

## Hasil Pengujian Menggunakan Kernel RBF

Rasio Perbandingan	Akurasi
Data Latih 70% : Data Uji 30%	87,98%
Data Latih 80% : Data Uji 20%	90,16%
Data Latih 90% : Data Uji 10%	90,71%

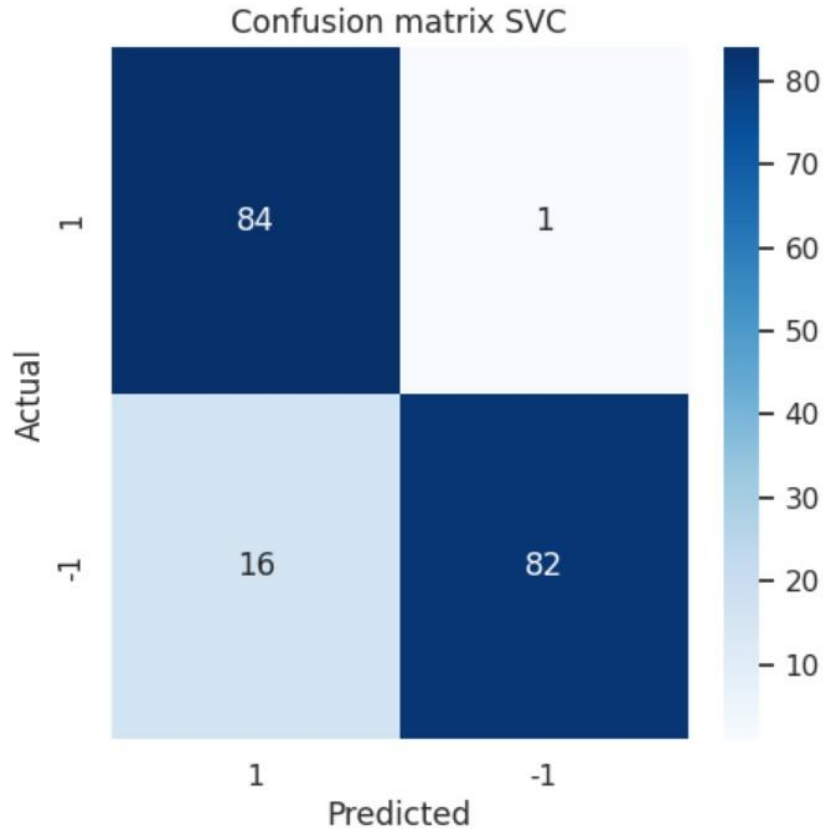
## Hasil Pengujian Menggunakan Kernel Sigmoid

Rasio Perbandingan	Akurasi
Data Latih 70% : Data Uji 30%	86,52%
Data Latih 80% : Data Uji 20%	87,71%
Data Latih 90% : Data Uji 10%	87,43%

Dari hasil pengujian menggunakan kernel yang ada di metode Support Vector Machine, akurasi tertinggi didapatkan menggunakan kernel RBF dengan besar akurasi 90,71%

# Pembahasan

- Pengujian Menggunakan Confusion Matrix



	precision	recall	f1-score	support
-1	0.84	0.99	0.91	85
1	0.99	0.84	0.91	98
accuracy			0.91	183
macro avg	0.91	0.91	0.91	183
weighted avg	0.92	0.91	0.91	183

# Pembahasan

- Evaluasi Hasil Pengujian

1. Perhitungan Nilai Precision :

$$\text{Precision Sentimen Positif} = \frac{TP}{TP + FP} = \frac{84}{84 + 1} = 0,9882$$

$$\text{Precision Sentimen Negatif} = \frac{TN}{TN + FN} = \frac{82}{82 + 16} = 0,8367$$

2. Perhitungan Nilai Recall :

$$\text{Recall Sentimen Positif} = \frac{TP}{TP + FN} = \frac{84}{84 + 16} = 0,84$$

$$\text{Recall Sentimen Negatif} = \frac{TN}{TN + FP} = \frac{82}{82 + 1} = 0,988$$

3. Perhitungan Nilai F<sub>1</sub>-Score :

$$F_1\text{-Score Sentimen Positif} = 2 \times \frac{(\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} = 2 \times \frac{0,84 \times 0,9882}{0,84 + 0,9882} = 0,908$$

$$F_1\text{-Score Sentimen Negatif} = 2 \times \frac{(\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} = 2 \times \frac{0,988 \times 0,8367}{0,988 + 0,8367} = 0,9061$$

4. Perhitungan Nilai Akurasi :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{84 + 82}{84 + 82 + 1 + 16} = 0,9071 \times 100\% = 90,71\%$$

# Temuan Penting Penelitian

- Metode Support Vector Machine terbukti dapat digunakan untuk klasifikasi data.
- Akurasi tertinggi dalam penelitian ini didapatkan melalui evaluasi dengan menggunakan metode confusion matrix mampu memperoleh akurasi sebesar 90,71%.

# Manfaat Penelitian

1. Manfaat bagi Penulis  
Penulis dapat menambah pengetahuan dan pemahaman dalam topik text mining khususnya metode Support Vector Machine. Selain itu, penelitian ini dapat berperan sebagai wadah untuk mengaplikasikan pengetahuan dan keterampilan yang telah diperoleh selama proses perkuliahan.
2. Manfaat bagi Universitas  
Penelitian ini dapat melengkapi sumber literatur di perpustakaan dan dapat digunakan sebagai referensi bagi mahasiswa serta calon peneliti lain sebagai panduan dalam pengembangan dan peninjauan metode Support Vector Machine.
3. Manfaat bagi Pembaca  
Penelitian ini dapat menjadi rujukan bagi pembaca yang berpotensi untuk dikembangkan lebih lanjut dalam penelitian-penelitian selanjutnya.

# Referensi

- [1] C. Malisi, "ANALISIS KINERJA SISTEM RESERVASI DAN PEMBAYARAN TIKET FERI (FERIZY) DENGAN METODE IMPORTANCE PERFORMANCE ANALYSIS (IPA) PADA PT. ASDP INDONESIA FERRY," 2021.
- [2] S. Fransiska dan A. Irham Gufroni, "Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method," *Scientific Journal of Informatics*, vol. 7, no. 2, hlm. 2407–7658, 2020, [Daring]. Tersedia pada: <http://journal.unnes.ac.id/nju/index.php/sji>
- [3] T. M. Permata Aulia, N. Arifin, dan R. Mayasari, "Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19," *SINTECH (Science and Information Technology) Journal*, vol. 4, no. 2, hlm. 139–145, 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [4] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, dan S. Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine," *CESS (Journal of Computer Engineering, System and Science)*, vol. 5, no. 2, hlm. 293, 2020, doi: 10.24114/cess.v5i2.18186.
- [5] Z. Alhaq, A. Mustopa, S. Mulyatun, dan J. D. Santoso, "PENERAPAN METODE SUPPORT VECTOR MACHINE UNTUK ANALISIS Informatika Universitas Amikom Yogyakarta Abstraksi Keywords : Pendahuluan Tinjauan Pustaka," *Jurnal of Information System Management*, vol. Vol. 3, no. 2, hlm. 44–49, 2021, [Daring]. Tersedia pada: <https://jurnal.amikom.ac.id/index.php/joism/article/view/558>
- [6] M. I. H. A. D. Akbari, A. Novianty, dan C. Setianingsih, "Analisis Sentimen Menggunakan Metode Learning Vector Quantization Sentiment Analysis Using Learning Vector Quantization Method," 2017.
- [7] L. Ardiani, H. Sujaini, dan T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 8, no. 2, hlm. 183, Apr 2020, doi: 10.26418/justin.v8i2.36776.
- [8] A. C. Pradikdo dan A. Ristyawan, "Model Klasifikasi Abstrak Skripsi Menggunakan Text Mining Untuk Pengkategorian Skripsi Sesuai Bidang Kajian," *Jurnal SIMETRIS*, vol. 9, no. 2, hlm. 1091–1098, 2018.
- [9] B. B. Baskoro, I. Susanto, dan S. Khomsah, "Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR)," *INISTA (Journal of Informatics Information System Software Engineering and Applications)*, vol. 3, no. 2, hlm. 21–29, 2021, doi: 10.20895/INISTA.V3I2.

# Referensi

- [10] S. Khairunnisa, A. Adiwijaya, dan S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 5, no. 2, hlm. 406, Apr 2021, doi: 10.30865/mib.v5i2.2835.
- [11] D. Alita, Y. Fernando, dan H. Sulistiani, "IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER," Jurnal TEKNOKOMPAK, vol. 14, no. 2, hlm. 86, 2020.
- [12] U. Khaira, R. Johanda, P. E. P. Utomo, dan T. Suratno, "Sentiment Analysis Of Cyberbullying On Twitter Using SentiStrength," Indonesian Journal of Artificial Intelligence and Data Mining, vol. 3, no. 1, hlm. 21, Mei 2020, doi: 10.24014/ijaidm.v3i1.9145.
- [13] M. Nurjannah dan I. Fitri Astuti, "PENERAPAN ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK TEXT MINING," Jurnal Informatika Mulawarman, vol. 8, no. 3, hlm. 110–113, 2013.
- [14] N. L. Ratniasih, M. Sudarma, dan N. Gunantara, "Penerapan Text Mining Dalam Spam Filtering Untuk Aplikasi Chat," Majalah Ilmiah Teknologi Elektro, vol. 16, no. 3, hlm. 13, 2017, doi: 10.24843/mite.2017.v16i03p03.
- [15] V. I. Santoso, G. Virginia, dan Y. Lukito, "PENERAPAN SENTIMENT ANALYSIS PADA HASIL EVALUASI DOSEN DENGAN METODE SUPPORT VECTOR MACHINE," 2017.
- [16] F. F. Irfani, M. Triyanto, A. D. Hartanto, dan Kusnawi, "Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine," JBMI (Jurnal Bisnis, Manajemen, dan Informatika), vol. 16, no. 3, hlm. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.
- [17] D. P. Daryfayi Edyt dan I. Asror, "Sentimen Analisis pada Ulasan Google Play Store Menggunakan Metode Naïve Bayes," Sentimen Analisis pada Ulasan Google Play Store Menggunakan Metode Naïve Bayes, vol. 7, no. Ulasan Pada Google Play Store, hlm. 11, 2020.
- [18] F. D. Ananda dan Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine," MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, vol. 20, no. 2, hlm. 407–416, Mei 2021, doi: 10.30812/matrik.v20i2.1130.



