

SKRIPSI

**KLASIFIKASI SENTIMEN KOMENTAR YOUTUBE
MALAKA PROJECT MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE (SVM)
(STUDI KASUS: VIDEO "*BISNIS GELAP DOKTER &
PERUSAHAAN FARMASI*")**



LAILA ROHMATUL I'ZZAH

NIM. 22106010007

**STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA**

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2026

**KLASIFIKASI SENTIMEN KOMENTAR YOUTUBE
MALAKA PROJECT MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE (SVM)
(STUDI KASUS: VIDEO "*BISNIS GELAP DOKTER &
PERUSAHAAN FARMASI*")**

Skripsi

Sebagai salah satu syarat memperoleh derajat

Sarjana S-1 Matematika



LAILA ROHMATUL I'ZZAH
NIM. 22106010007
PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2026



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Laila Rohmatul I'zzah

NIM : 22106010007

Judul Skripsi : Klasifikasi Sentimen Komentar YouTube Malaka Project Menggunakan Algoritma *Support Vector Machine* (SVM)

Studi Kasus: Video "*Bisnis Gelap Dokter & Perusahaan Farmasi*"

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Pembimbing I

Arif Munandar, M.Sc.

NIP. 19920721 201903 1 013

Yogyakarta, 22 Mei 2026

Pembimbing II

Muhamad Rashif Hilmi, S.Si., M.Sc.

NIP. 19920309 202012 1 001



PENGESAHAN TUGAS AKHIR

Nomor : B-1209/Un.02/DST/PP.00.9/06/2026

Tugas Akhir dengan judul : Klasifikasi Sentimen Komentar YouTube Malaka Project Menggunakan Algoritma Support Vector Machine (SVM), (Studi Kasus: Video "Bisnis Gelap Dokter & Perusahaan Farmasi")

yang dipersiapkan dan disusun oleh:

Nama : LAILA ROHMATUL IZZAH
Nomor Induk Mahasiswa : 22106010007
Telah diujikan pada : Rabu, 03 Juni 2026
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang
Arif Munandar, M.Sc.
SIGNED

Valid ID: 6a223e9a9924d



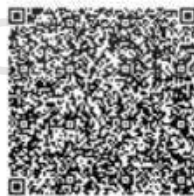
Penguji I
Muhamad Rashif Hilmi, S.Si., M.Sc.
SIGNED

Valid ID: 6a224e67b52c3



Penguji II
Sri Utami Zuliana, S.Si., M.Sc., Ph.D.
SIGNED

Valid ID: 6a21f75741b9c



Yogyakarta, 03 Juni 2026
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi
Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 6a226190c10da

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Laila Rohmatul I'zzah
NIM : 22106010007
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 22 Mei 2026



Laila Rohmatul I'zzah

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

*Dengan penuh rasa syukur dan kerendahan hati, skripsi ini
saya persembahkan kepada
kedua orang tua tercinta,
kakak dan adik tersayang,
keluarga besar,
sahabat,
serta teman-teman seperjuangan.
Terima kasih atas doa dan dukungannya sehingga karya ini terwujud.*



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN MOTTO

*“Kemudian, apabila engkau telah membulatkan tekad,
bertawakallah kepada Allah.”*

(Q.S. Ali-‘Imran: 159)

*“Apabila hamba-hamba-Ku bertanya kepadamu (Muhammad) tentang Aku,
maka (jawablah), sesungguhnya aku dekat.”*

(Q.S. Al-Baqarah: 186)

*Semua jatuh bangunmu hal yang biasa.
Angan dan pertanyaan, waktu yang menjawabnya.
Berikan tenggat waktu, bersedihlah secukupnya.
Rayakan perasaanmu sebagai manusia.*

(Mata Air - Hindia)

*Melamban bukanlah hal yang tabu. Kadang itu yang kau butuh.
Bersandar, hibahkan bebanmu. Tak perlu kau berhenti kurasi.
Ini hanya sementara. Bukan ujung dari rencana.*

(33x – Perunggu)

KATA PENGANTAR

Segala puji bagi Allah Swt. yang senantiasa melimpahkan karunia, petunjuk, dan pertolongan-Nya kepada penulis sehingga dapat menyelesaikan penyusunan skripsi berjudul “*Klasifikasi Sentimen Komentar YouTube Malaka Project Menggunakan Algoritma Support Vector Machine (SVM) dengan Studi Kasus Video Berjudul Bisnis Gelap Dokter & Perusahaan Farmasi*”. Sholawat dan salam semoga tetap tercurah kepada Nabi Muhammad Saw. yang kita nantikan syafaatnya di hari kiamat nanti.

Skripsi ini disusun sebagai salah satu prasyarat dalam menyelesaikan studi dan mencapai gelar Sarjana Matematika. Selama proses penulisan skripsi ini, penulis melewati berbagai tahapan yang memerlukan ketekunan, kesungguhan, dan kesabaran. Penulis menyadari bahwa pencapaian ini tidak terlepas dari dukungan berbagai pihak yang memberikan doa, bimbingan, maupun bantuan moral maupun teknis. Untuk itu, pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih yang tulus kepada:

1. Almameter UIN Sunan Kalijaga Yogyakarta yang telah memberikan ruang dan kesempatan kepada penulis untuk terus belajar dan berkembang.
2. Bapak Prof. Noorhaidi Hasan, S.Ag., M.A., M.Phil., Ph.D. selaku Rektor UIN Sunan Kalijaga Yogyakarta.
3. Ibu Prof. Dr. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sekaligus Dosen Pembimbing Akademik (DPA) yang memberikan arahan, masukan, dan nasihat kepada penulis selama di bangku perkuliahan.
4. Ibu Dr. Epha Diana Supandi, S.Si., M.Sc., selaku Ketua Program Studi Matematika UIN Sunan Kalijaga Yogyakarta.
5. Bapak Arif Munandar, M.Sc., selaku Dosen Pembimbing I yang selalu memberikan bimbingan dan nasihat dalam penyusunan penelitian ini.

6. Bapak Muhamad Rashif Hilmi, S.Si., M.Sc., selaku Dosen Pembimbing II atas kesediaannya untuk membimbing dengan penuh dedikasi, kesabaran, dan ketelitian dalam penyusunan penelitian ini.
7. Seluruh tenaga pendidik dan staf Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta yang selalu memberikan ilmu dan pelayanan yang baik selama perkuliahan.
8. Kedua orang tua tercinta, Bapak Muhammad Mundir dan Ibu Iin Fadhilah yang selalu memberikan doa, dukungan, nasihat, dan cinta yang tak ternilai harganya. Bangga rasanya mampu mewujudkan salah satu mimpi Ibu yang ingin anaknya menempuh pendidikan tinggi. Semoga karya kecil ini bisa memberikan kehangatan di hati mereka, menjadi kabar bahagia bahwa keringatnya selama ini tak pernah sia-sia.
9. Seluruh keluarga besar Pondok Pesantren Chujjatul Islam Mlangi, khususnya Kyai Murobik Aziz dan Bu Nyai Nur Afifah yang dengan tulus memberikan bimbingan, pengajaran, dan nasihat kepada penulis.
10. Kakak dan adik kandung penulis, Kharis dan Rivan yang selalu mendoakan penulis agar dapat menyelesaikan skripsi ini.
11. Diri penulis sendiri, Laila Rohmatul I'zzah. Apresiasi sebesar-besarnya karena telah berjuang menyelesaikan apa yang telah dimulai. Terima kasih sudah bertahan hingga detik ini, *“segala doa yang baik adanya, untukmu dan mimpimu yang mulia”*.
12. Sahabat penulis *“Jos Gandos”*, Sinta, Liani, Nafisa, Nila, Alya, Alina, Tarisha, Derrida, yang membersamai selama perkuliahan dan menjadi rumah kedua di perantauan.
13. Kelima *“Otak Encer”*, Alifya, Mila, Hanin, Dhila, dan Riska, yang menjadi sahabat dan teman penulis dalam menjalani kuliah di konsentrasi Statistika
14. Teman-teman satu bimbingan Pak Rashif, Bunga, Viga, Icha, Dhila, dan Dzakiya yang telah membersamai selama bimbingan skripsi.
15. Sobat KANUARA (KKN Nusantara Kanoman Dua) yang selalu di hati. Lisna, Niken, Adib, dan Irul (keempat teman penulis dari UIN Sunan Kalijaga Yogyakarta), Anti dan Sekar dari UIN Purwokerto, Ayin dan Viki

dari UIN Kediri, Ratna dari UIN Malang, Misbah dari UIN Tulungagung, Irwan dari UIN Pare-Pare, dan Ghozali dari UIN Jember.

16. Teman-teman Matematika 2022, terutama pada konsentrasi Statistika yang sudah kebersamai selama perkuliahan dan menyemangati satu sama lain.
17. Seluruh pihak yang tidak dapat disebutkan satu per satu yang telah membantu penulis dalam menyelesaikan skripsi ini, semoga Allah Swt. membalas kebaikannya.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, sehingga penulis mengharapkan kritik dan saran dari semua pihak untuk kebaikan dari skripsi ini.

Yogyakarta, 5 Juni 2026

Penulis



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN TUGAS AKHIR.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN	iv
HALAMAN PERSEMBAHAN	v
HALAMAN MOTTO	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMBANG	xvi
INTISARI.....	xvii
ABSTRACT.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.6 Tinjauan Pustaka	6
1.7 Sistematika Penelitian	9
BAB II DASAR TEORI	11
2.1 Data Mining.....	11
2.2 Machine Learning	11
2.3 Klasifikasi Teks.....	13
2.4 Analisis Sentimen.....	15
2.5 Preprocessing	16
2.6 Kamus Lexicon	19
2.7 Ekstraksi Fitur	21
2.7.1 <i>Bag of Words (BoW)</i>	21
2.7.2 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	22

2.8	Vektor pada Ruang Euclides (\mathbb{R}^n).....	25
2.8.1	Operasi Vektor.....	26
2.9	Jarak Euclidean	27
2.10	Matriks	28
2.10.1	Operasi Matriks	29
2.10.2	Jenis-jenis Matriks.....	31
2.11	Quadratic Programming	32
2.12	Metode Pengali Lagrange.....	33
2.13	Syarat Karush-Kuhn-Tucker	34
BAB III METODE PENELITIAN.....		36
3.1	Jenis Penelitian	36
3.1	Metode Pengumpulan Data	36
3.3	Variabel Penelitian	37
3.4	Alat Pengolahan Data.....	38
3.5	Metode Analisis Data	38
3.6	Flowchart Penelitian.....	40
BAB IV PEMBAHASAN.....		41
4.1	Support Vector Machine (SVM)	41
4.1.1	Hyperplane	42
4.1.2	SVM pada <i>Linearly Separable Data</i>	44
4.1.3	SVM pada <i>Non-Linearly Separable Data</i>	46
4.1.4	<i>Kernel Trick</i>	48
4.1.5	<i>Simplified Sequential Minimal Optimization (Simplified SMO)</i>	51
4.1.6	SVM Multikelas dengan Metode <i>One-Against-All</i>	54
4.2	Evaluasi Model.....	56
4.3	Contoh Perhitungan Support Vector Machine (SVM) Multiclass	58
4.3.1	Pelatihan Model 1 dengan Kernel Linear.....	59
4.3.2	Pengujian Model 1 dengan Kernel Linear.....	71
BAB V STUDI KASUS.....		73
5.1	Pengumpulan Data	73
5.2	Data Preprocessing	73
5.3	Pelabelan Data.....	79
5.4	Data Splitting.....	80
5.5	Ekstraksi Fitur	81

5.6 Pembentukan Model.....	88
5.7 Evaluasi Model.....	90
5.8 Perbandingan Model	97
5.9 Implementasi dengan Model Terbaik.....	100
BAB VI PENUTUP	103
6.1 Kesimpulan.....	103
6.2 Saran.....	104
DAFTAR PUSTAKA	105
LAMPIRAN.....	108



DAFTAR TABEL

Tabel 1.1 Tinjauan Pustaka.....	7
Tabel 4.2 Contoh 3 Model SVM dengan Metode One-Against-All.....	55
Tabel 4.3 Tabel Confusion Matrix.....	57
Tabel 4.4 Dataset Perhitungan Manual.....	58
Tabel 4.5 Target Model Positif vs Bukan Positif.....	59
Tabel 4.6 Matriks Kernel Linear.....	60
Tabel 4.7 Hasil Error Awal pada Data ke- i	61
Tabel 4.8 Kondisi KKT.....	62
Tabel 4.9 Error Data Pembanding (E_j).....	63
Tabel 4.10 Nilai alpha data i dan j	63
Tabel 4.11 Nilai L dan H (target sama).....	64
Tabel 4.12 Nilai L dan H (Kondisi).....	64
Tabel 4.13 Perhitungan eta (η).....	65
Tabel 4.14 Nilai Alpha Baru data ke- j	66
Tabel 4.15 Nilai Alpha Data j pada Segmen Garis.....	67
Tabel 4.16 Kondisi Nilai Absolut.....	68
Tabel 4.17 Nilai Alpha Data Baru ke- i	69
Tabel 4.18 Nilai Alpha dan Bias pada Data Pelatihan.....	70
Tabel 5.1 Hasil Proses Data Cleaning.....	74
Tabel 5.2 Hasil Case Folding.....	75
Tabel 5.3 Hasil Proses Tokenization.....	76
Tabel 5.4 Hasil Proses Normalization.....	77
Tabel 5.5 Hasil Stopword Removal.....	77
Tabel 5.6 Hasil Proses Data Stemming.....	78
Tabel 5.7 Hasil Pelabelan dengan Kamus Lexicon.....	80
Tabel 5.8 Hasil Ekstraksi Fitur TF-IDF.....	82
Tabel 5.9 Representasi Vektor TF-IDF.....	83
Tabel 5.10 Hasil Confusion Matrix Kernel Linear.....	91

Tabel 5.11 Hasil Evaluasi Model Kernel Linear.....	92
Tabel 5.12 Hasil Confusion Matrix Kernel Polinomial	93
Tabel 5.13 Hasil Evaluasi Model Kernel Polinomial.....	94
Tabel 5.14 Hasil Confusion Matrix Model Kernel RBF.....	95
Tabel 5.15 Hasil Evaluasi Model Kernel RBF.....	96
Tabel 5.16 Perbandingan Nilai Presisi Antar Kernel	97
Tabel 5.17 Perbandingan Nilai Recall Antar Kernel	98
Tabel 5.18 Perbandingan Nilai F1-Score Antar Kernel.....	99
Tabel 5.19 Perbandingan Nilai Akurasi Antar Kernel.....	99

DAFTAR GAMBAR

Gambar 3.1 Flowchart Penelitian.....	40
Gambar 4.2 Hyperplane berupa garis lurus dan bidang datar.....	42
Gambar 4.3 Visualisasi Margin SVM.....	43
Gambar 4.4 Pemetaan Data ke <i>Feature Space</i>	49
Gambar 4.5 Contoh Klasifikasi dengan Metode <i>One-Against-All</i>	56
Gambar 5.1 Grafik Kata Khas Antar Kelas	86
Gambar 5.2 Hasil Implementasi Model Kernel Linear	101

DAFTAR LAMBANG

w	=	Vektor bobot
b	=	Bias
x_i	=	Data ke- i
y_i	=	Kelas data ke- i
α_i	=	Parameter Lagrange
ξ_i	=	Variabel <i>slack</i>
C	=	Parameter regularisasi
$K(x_i, x_j)$	=	Fungsi kernel
$f(x)$	=	Fungsi keputusan
t	=	<i>Term</i>
d	=	Dokumen
λ, μ	=	Pengali Lagrange
$\nabla f(x)$	=	Gradien fungsi
$H(x)$	=	Matriks Hessian
W	=	<i>Wolfe Dual</i>
L	=	Fungsi Lagrange
p	=	Derajat polinomial
γ	=	Parameter gamma
n	=	Jumlah data
N	=	Jumlah seluruh dokumen
ϕ	=	Fungsi pemetaan kernel
η	=	Turunan kedua fungsi objektif

INTISARI

KLASIFIKASI SENTIMEN KOMENTAR YOUTUBE MALAKA PROJECT MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE* (SVM) (STUDI KASUS: VIDEO "*BISNIS GELAP DOKTER & PERUSAHAAN FARMASI*")

Oleh
LAILA ROHMATUL I'ZZAH
NIM. 22106010007

Komentar YouTube merupakan data teks tidak terstruktur yang mengandung beragam opini publik dan memerlukan algoritma klasifikasi yang optimal bekerja pada data teks berdimensi tinggi. Penelitian ini bertujuan untuk membandingkan performa kernel linear, polinomial, dan *radial basis function* (RBF) pada algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen komentar YouTube Malaka Project. Dataset penelitian diperoleh melalui hasil *scraping* menggunakan YouTube Data API v3 pada video berjudul "*Bisnis Gelap Dokter & Perusahaan Farmasi*" dengan jumlah data sebanyak 3.428 komentar yang diperoleh pada Desember 2025. Data diklasifikasikan menjadi tiga kelas sentimen, yaitu positif, negatif, dan netral. Tahapan penelitian meliputi *preprocessing* data, pelabelan sentimen menggunakan kamus InSet Lexicon, pembagian data dengan rasio 80:20, ekstraksi fitur menggunakan TF-IDF, klasifikasi menggunakan algoritma SVM *multiclass* dengan pendekatan *One-Against-All* (OvR), serta evaluasi model menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan *confusion matrix*. Hasil penelitian menunjukkan bahwa kernel linear dan RBF memiliki nilai akurasi yang sama sebesar 0,73 sedangkan kernel polinomial lebih rendah yaitu 0,60. Namun, kernel linear memperoleh nilai presisi dan *F1-score* tertinggi sebesar 0,77 pada kelas positif. Dengan demikian, dapat disimpulkan bahwa kernel linear dapat mengklasifikasikan sentimen komentar YouTube Malaka Project dengan lebih baik dibandingkan kedua kernel lainnya.

Kata kunci: klasifikasi sentimen, komentar YouTube Malaka Project, *Support Vector Machine* (SVM), kernel SVM.

ABSTRACT

SENTIMENT CLASSIFICATION OF MALAKA PROJECT YOUTUBE COMMENTS USING THE SUPPORT VECTOR MACHINE (SVM) ALGORITHM (CASE STUDY: THE VIDEO “THE DARK BUSINESS OF DOCTORS AND PHARMACEUTICAL COMPANIES”)

By
LAILA ROHMATUL I'ZZAH
NIM. 22106010007

YouTube comments are unstructured textual data that contain diverse public opinions and require a classification algorithm capable of effectively handling high-dimensional text data. This study aims to compare the performance of linear, polynomial, and Radial Basis Function (RBF) kernels in the Support Vector Machine (SVM) algorithm for sentiment classification of comments on the Malaka Project YouTube channel. The dataset was collected through scraping using the YouTube Data API v3 from a video entitled “The Dark Business of Doctors and Pharmaceutical Companies,” resulting in 3,428 comments collected in December 2025. The data were classified into three sentiment categories: positive, negative, and neutral. The research procedure consisted of data preprocessing, sentiment labeling using the InSet Lexicon dictionary, data splitting with an 80:20 ratio, feature extraction using TF-IDF, classification using a multiclass SVM algorithm with the One-Against-All (OvR) approach, and model evaluation using accuracy, precision, recall, F1-score, and confusion matrix metrics. The results show that the linear and RBF kernels achieved the same accuracy score of 0.73, while the polynomial kernel obtained a lower accuracy score of 0.60. However, the linear kernel achieved the highest precision and F1-score of 0.77 for the positive sentiment class. Therefore, it can be concluded that the linear kernel performs better in classifying sentiment in Malaka Project YouTube comments compared to the other two kernels.

Keywords: sentiment classification, Malaka Project YouTube comments, Support Vector Machine (SVM), SVM kernels.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital di Indonesia mengalami peningkatan yang signifikan dalam beberapa tahun terakhir. Berdasarkan laporan resmi Global Digital Report, pada akhir tahun 2025 jumlah pengguna internet di Indonesia mencapai 230 juta individu atau setara dengan 80,5% dari total populasi (DataReportal, 2025). Di antara berbagai platform digital yang digunakan, YouTube menjadi salah satu platform dengan jangkauan pengguna mencapai 151 juta akun (DataReportal, 2025). Jumlah tersebut meningkat sebesar 5,6% dibandingkan tahun sebelumnya.

Tingginya jumlah pengguna tersebut berimplikasi pada meningkatnya interaksi pengguna melalui kolom komentar yang memuat beragam respons terhadap suatu konten. Jumlah komentar yang dihasilkan pada suatu video dapat mencapai skala yang besar sehingga proses identifikasi opini atau sentimen pengguna menjadi sulit dilakukan secara manual. Oleh karena itu, diperlukan metode klasifikasi otomatis untuk mengelompokkan komentar ke dalam kategori sentimen tertentu sehingga informasi yang terkandung dalam komentar dapat dianalisis secara lebih efisien.

Salah satu pendekatan yang umum digunakan untuk melakukan klasifikasi sentimen adalah algoritma *machine learning*. Namun, komentar yang dihasilkan pengguna masih berbentuk data teks tidak terstruktur (Aggarwal & Zhai, 2012). Oleh karena itu, data teks perlu direpresentasikan terlebih dahulu ke dalam bentuk numerik sebelum dapat diproses pada tahap klasifikasi (Manning et al., 2008). Bentuk representasi yang banyak digunakan adalah model ruang vektor (*vector space model*), di mana setiap dokumen direpresentasikan sebagai vektor berdasarkan kemunculan kata-kata penyusunnya (Aggarwal & Zhai, 2012). Dokumen yang dimaksud pada penelitian ini adalah komentar. Pada model ini, bobot setiap kata dihitung menggunakan skema pembobotan seperti *Term*

Frequency-Inverse Document Frequency (TF-IDF) yang mengukur relevansi suatu *term* berdasarkan frekuensi kemunculannya dalam dokumen serta tingkat distribusi terhadap keseluruhan korpus, di mana korpus merupakan kumpulan dari beberapa dokumen (Manning et al., 2008).

Representasi fitur menggunakan TF-IDF menghasilkan ruang dengan jumlah dimensi yang sebanding dengan banyaknya kosakata unik dalam korpus. Pada data komentar YouTube, variasi penggunaan kata yang tinggi menyebabkan terbentuknya ruang fitur berdimensi tinggi (*high-dimensional space*) (Manning et al., 2008). Selain itu, setiap dokumen umumnya hanya memuat sebagian kecil dari keseluruhan kosakata yang tersedia sehingga mayoritas komponen vektor bernilai nol dan menghasilkan representasi yang bersifat *sparse* (Aggarwal & Zhai, 2012). Karakteristik data tersebut dapat menyulitkan proses klasifikasi karena banyaknya fitur yang harus diproses dan dominasi nilai nol pada representasi data. Oleh karena itu, diperlukan metode klasifikasi yang mampu memisahkan kelas secara optimal pada ruang fitur berdimensi tinggi.

Terdapat beberapa algoritma *machine learning* yang umum digunakan untuk klasifikasi teks, seperti *Naïve Bayes*, *K-Nearest-Neighbor* (KNN), *Decision Tree*, Regresi Logistik, hingga *Support Vector Machine* (SVM) (Heryadi & Wahyono, 2020). Masing-masing algoritma memiliki karakteristik dan pendekatan berbeda dalam membangun model klasifikasi. Oleh karena itu, pemilihan algoritma perlu disesuaikan dengan karakteristik data dan representasi fitur yang digunakan.

Salah satu algoritma yang banyak digunakan untuk menangani data berdimensi tinggi adalah *Support Vector Machine* (SVM). Algoritma ini dikenal memiliki performa yang baik karena kemampuannya dalam memaksimalkan margin antar kelas pada ruang fitur yang kompleks. Hal ini diperkuat dengan studi komparatif yang dilakukan oleh (Hsu, 2020) terhadap beberapa algoritma *supervised learning* pada tugas klasifikasi teks, yang menunjukkan bahwa algoritma SVM, Regresi Logistik, dan *Multilayer Perceptron* memiliki performa lebih baik dibandingkan metode lainnya, dengan SVM menunjukkan hasil paling

unggul secara umum. Temuan tersebut memperkuat pertimbangan pemilihan *Support Vector Machine* (SVM) sebagai metode klasifikasi dalam penelitian ini.

Pada dasarnya, algoritma SVM dirancang sebagai metode klasifikasi biner untuk memisahkan dua kelas. Sementara itu, untuk menangani data *multiclass*, dikembangkan pendekatan *One-Against-All* (OvR) yang bekerja dengan membangun sejumlah model biner, di mana setiap model dilatih untuk membedakan satu kelas terhadap seluruh kelas lainnya. Berdasarkan pendekatan tersebut, penelitian ini menerapkan algoritma SVM *multiclass* untuk melakukan klasifikasi komentar YouTube ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral.

Dalam implementasinya, pembentukan model SVM melibatkan penyelesaian permasalahan optimisasi kuadratik (*Quadratic Programming*). Salah satu metode yang umum digunakan untuk menyelesaikan optimisasi tersebut adalah *Sequential Minimal Optimization* (SMO), yaitu metode yang memecah permasalahan kuadratik menjadi sub-masalah yang melibatkan dua parameter Lagrange pada setiap iterasi (Platt, 1998). Pada penelitian ini, konsep *Simplified SMO* digunakan untuk mendukung penjelasan perhitungan manual dalam pembentukan model SVM *multiclass*.

Pemilihan platform YouTube Malaka Project sebagai objek studi difokuskan pada video yang terdapat dalam kategori “Populer” pada platform tersebut, yaitu video berjudul "*Bisnis Gelap Dokter & Perusahaan Farmasi*" yang diambil pada tanggal 5 Desember 2025. Video ini dipilih karena mengangkat isu yang bersifat aktual dan kontroversial mengenai praktik industri medis di Indonesia sehingga memunculkan beragam tanggapan dari masyarakat. Keragaman tersebut ditunjukkan dengan komentar yang memuat opini, kritik, maupun pandangan yang berbeda terhadap topik yang dibahas. Selain itu, tingginya atensi publik terhadap video tersebut ditunjukkan oleh jumlah tayangan yang mencapai 3.645.871 kali dan 3.900 komentar pada saat pengambilan data.

Jumlah komentar yang relatif besar menyediakan dataset yang memadai untuk klasifikasi sentimen berbasis algoritma *machine learning*. Di sisi lain, keberagaman opini yang muncul pada kolom komentar menghasilkan data teks dengan variasi sentimen yang cukup beragam. Hal tersebut menjadikan dataset ini relevan untuk menguji performa kernel linear, polinomial, dan RBF dalam algoritma SVM pada kasus klasifikasi sentimen *multiclass*.

Sejumlah penelitian sebelumnya telah membahas penerapan algoritma *Support Vector Machine* (SVM) dalam klasifikasi dan analisis sentimen data teks. Penelitian oleh (Riki et al., 2025) menunjukkan bahwa SVM dengan optimasi SMO memiliki performa yang lebih unggul dibandingkan *Naïve Bayes* dalam mengklasifikasikan ulasan produk *e-commerce*, sehingga dinilai lebih efektif dalam menangani data sentimen yang kompleks. Selanjutnya, (Praghakusma & Charibaldi, 2021) melakukan komparasi antara beberapa fungsi kernel pada SVM dan memperoleh hasil bahwa kernel linear menghasilkan akurasi tertinggi sebesar 83,06% dibanding kernel polinomial dan sigmoid pada analisis sentimen media sosial.

Kemudian, (Anam et al., 2025) mengembangkan model *ensemble soft-voting* SVM berbasis kombinasi beberapa kernel dan memperoleh performa sangat tinggi mencapai 98% pada semua metrik evaluasi, menunjukkan efektivitas integrasi kernel dalam menangkap pola linear dan nonlinear pada data sentimen. Sementara itu, penelitian (Akbar, 2015) menerapkan SVM dengan optimasi *Simplified SMO* dan pendekatan *One-Against-All* (OvR) pada klasifikasi *multiclass* data numerik, dengan rata-rata akurasi terbaik sebesar 0,8939 melalui pengujian *cross validation*.

Penelitian ini membandingkan performa beberapa fungsi kernel SVM dalam klasifikasi sentimen komentar YouTube menggunakan pendekatan *multiclass*. Oleh karena itu, penelitian ini difokuskan untuk menguji dan membandingkan performa kernel linear, polinomial, dan *Radial Basis Function* (RBF) menggunakan algoritma *Support Vector Machine* (SVM) dalam

mengklasifikasikan sentimen komentar video YouTube Malaka Project berjudul “*Bisnis Gelap Dokter & Perusahaan Farmasi*”.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah dalam penelitian ini disusun sebagai berikut:

1. Bagaimana proses klasifikasi sentimen komentar YouTube Malaka Project menggunakan algoritma *Support Vector Machine* (SVM)?
2. Bagaimana performa algoritma *Support Vector Machine* (SVM) dengan kernel linear, polinomial, dan *Radial Basis Function* (RBF) dalam melakukan klasifikasi sentimen komentar YouTube Malaka Project?

1.3 Batasan Masalah

Untuk menjaga fokus penelitian, ditetapkan batasan masalah sebagai berikut:

1. Data yang digunakan pada penelitian ini adalah komentar dari satu video YouTube Malaka Project berjudul “*Bisnis Gelap Dokter & Perusahaan Farmasi*”.
2. Analisis hanya dilakukan pada komentar berbahasa Indonesia.
3. Klasifikasi sentimen dibatasi pada tiga kelas, yaitu positif, negatif, dan netral.
4. Algoritma yang digunakan adalah *Support Vector Machine* (SVM) dengan optimasi *Sequential Minimal Optimization* (SMO) dan pendekatan *One-Against-All* untuk klasifikasi multikelas.
5. Evaluasi performa difokuskan pada perbandingan tiga fungsi kernel SVM, yaitu Linear, Polinomial, dan *Radial Basis Function* (RBF); menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F-1 Score*, dan *Confusion Matrix*.
6. *Software* yang digunakan pada penelitian ini adalah Python versi 3.12.13.

1.4 Tujuan Penelitian

Tujuan penulisan penelitian ini adalah sebagai berikut:

1. Menjelaskan proses klasifikasi sentimen komentar YouTube Malaka Project menggunakan algoritma *Support Vector Machine* (SVM).
2. Menganalisis kinerja algoritma *Support Vector Machine* (SVM) dengan kernel linear, polinomial, dan *Radial Basis Function* (RBF) dalam melakukan klasifikasi sentimen komentar YouTube Malaka Project

1.5 Manfaat Penelitian

Manfaat yang diharapkan dalam penyusunan penelitian ini adalah sebagai berikut:

1. Memberikan pemahaman terkait proses klasifikasi sentimen komentar YouTube menggunakan algoritma *Support Vector Machine* (SVM).
2. Memberikan pemahaman tentang analisis kinerja algoritma *Support Vector Machine* (SVM) dengan kernel linear, polinomial, dan *Radial Basis Function* (RBF) dalam melakukan klasifikasi sentimen pada data teks berdimensi tinggi.

1.6 Tinjauan Pustaka

Beberapa penelitian terdahulu yang relevan dengan penerapan algoritma *Support Vector Machine* (SVM) pada klasifikasi dan analisis sentimen telah ditinjau sebagai landasan dalam penelitian ini. Tinjauan tersebut meliputi aspek metode yang digunakan, strategi penanganan klasifikasi *multiclass*, optimasi algoritma, pemilihan fungsi kernel, serta karakteristik data yang dianalisis. Penelaahan terhadap beberapa penelitian tersebut bertujuan untuk mengidentifikasi persamaan dan perbedaan pendekatan yang digunakan, sehingga posisi penelitian ini dapat dirumuskan secara lebih jelas.

Tabel 1.1 Tinjauan Pustaka

No.	Judul Penelitian	Metode	Studi Kasus	Persamaan dan Perbedaan
1.	(Riki et al., 2025)	<ol style="list-style-type: none"> 1. Algoritma SVM. 2. Membandingkan performa model Naïve Bayes dan SVM. 3. Klasifikasi sentimen <i>multiclass</i> (positif, negatif, netral) 	Data teks berupa ulasan produk <i>E-Commerce</i> Shopee.	<p>Persamaan: Menggunakan SVM pada klasifikasi sentimen <i>multiclass</i> berbasis data teks.</p> <p>Perbedaan: Tidak melakukan perbandingan fungsi kernel SVM dan tidak menerapkan pendekatan <i>One-Against-All</i> (OvR).</p>
2.	(Praghakusma & Charibaldi, 2021)	<ol style="list-style-type: none"> 1. Menggunakan algoritma SVM. 2. Membandingkan kernel linear, polinomial, dan sigmoid. 3. Analisis sentimen dengan <i>multiclass</i> 	Data teks berupa komentar Instagram dan cuitan Twitter tentang KPK	<p>Persamaan: Membandingkan fungsi kernel SVM pada analisis sentimen <i>multiclass</i>.</p> <p>Perbedaan: Menggunakan kernel sigmoid</p>

No.	Judul Penelitian	Metode	Studi Kasus	Persamaan dan Perbedaan
		(positif, negatif, netral) 4. Evaluasi model menggunakan metrik <i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , dan <i>F-1 Score</i> .		sebagai kernel pembanding dan tidak menerapkan pendekatan <i>One-Against-All</i> (OvR).
3.	(Anam et al., 2025)	1. Menggunakan algoritma SVM. 2. Membandingkan kernel linear, polinomial, sigmoid, dan RBF. 3. Analisis sentimen biner (positif dan negatif). 4. Evaluasi model menggunakan metrik <i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , dan <i>F-1 Score</i> .	Data teks berupa cuitan Twitter tentang Debat Naturalisasi Pemain Timnas Sepak Bola Indonesia.	Persamaan: Menggunakan beberapa fungsi kernel pada algoritma SVM. Perbedaan: Menggunakan pendekatan ensemble beberapa kernel untuk klasifikasi biner.
4.	(Akbar, 2015).	1. Menggunakan algoritma SVM <i>multiclass</i> (normal, rentan, mengkhawatirkan). 2. Menggunakan pendekatan <i>One-</i>	Data numerik pasien stroke dari RSUD Dr. Saiful	Persamaan: Menggunakan pendekatan <i>One-Against-All</i> (OvR) pada klasifikasi <i>multiclass</i> .

No.	Judul Penelitian	Metode	Studi Kasus	Persamaan dan Perbedaan
		<i>Against-All</i> untuk kasus <i>multiclass</i> . 3. Evaluasi menggunakan metrik <i>Accuracy</i> dan 10-Fold Cross Validation.	Anwar Malang.	Perbedaan: Menggunakan data numerik dan tidak membandingkan fungsi kernel SVM.

1.7 Sistematika Penelitian

Sistematika penulisan skripsi ini terdiri dari enam bab, di antaranya yaitu bab pendahuluan, landasan teori, metode penelitian, pembahasan, studi kasus, dan penutup. Adapun uraian setiap bab adalah sebagai berikut:

1. BAB I PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, tinjauan pustaka, dan sistematika penulisan.

2. BAB II LANDASAN TEORI

Bab ini berisi teori-teori yang menjadi dasar penelitian, mencakup konsep dasar matematika, konsep data, hingga konsep dasar yang menjadi penunjang algoritma pada penelitian ini.

3. BAB III METODE PENELITIAN

Bab ini menjelaskan jenis penelitian, metode pengumpulan data, variabel penelitian, alat pengolahan data, metode analisis data, dan *flowchart* penelitian.

4. BAB IV PEMBAHASAN

Bab ini membahas mengenai algoritma yang digunakan pada penelitian, yaitu algoritma *Support Vector Machine* (SVM), pembahasan mengenai evaluasi model, serta contoh sederhana penerapan algoritma.

5. BAB V STUDI KASUS

Bab ini membahas penerapan algoritma *Support Vector Machine* (SVM) pada klasifikasi multikelas komentar YouTube Malaka Project.

6. BAB VI PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian yang dilakukan dan saran untuk penelitian atau pengembangan selanjutnya.



BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil penelitian tentang klasifikasi sentimen komentar YouTube Malaka Project menggunakan algoritma *Support Vector Machine* (SVM) pada video berjudul “*Bisnis Gelap Dokter & Perusahaan Farmasi*”, maka dapat diambil kesimpulan sebagai berikut:

1. Proses klasifikasi sentimen komentar YouTube dilakukan melalui beberapa tahapan, yaitu pengumpulan data menggunakan YouTube Data API v3, *data preprocessing* yang meliputi *data cleaning*, *case folding*, *tokenization*, *normalization*, *stopword removal*, dan *stemming*. Kemudian, pelabelan sentimen menggunakan metode *lexicon-based* dengan kamus InSet Lexicon. Selanjutnya, data dibagi menjadi data latih dan data uji dengan rasio 80:20, kemudian dilakukan ekstraksi fitur menggunakan TF-IDF. Proses berikutnya adalah pembentukan model kernel linear, Polinomial, dan *Radial Basis Function* (RBF) menggunakan algoritma *Support Vector Machine* (SVM) multiclass dengan pendekatan *One-Against-All* (OvR). Pada tahap akhir, dilakukan evaluasi model menggunakan *confusion matrix*, akurasi, presisi, *recall*, dan *F1-score*.
2. Hasil evaluasi menunjukkan bahwa algoritma *Support Vector Machine* (SVM) dengan kernel linear, Polinomial, dan *Radial Basis Function* (RBF) mampu melakukan klasifikasi sentimen komentar YouTube Malaka Project ke dalam tiga kelas, yaitu positif, negatif, dan netral. Berdasarkan pengujian menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*, diperoleh bahwa kernel linear memperoleh nilai akurasi sebesar 0,73 yang setara dengan RBF, tetapi memiliki nilai presisi dan *F1-score* yang lebih tinggi pada sebagian besar kelas sentimen dibandingkan kedua kernel lainnya. Hal ini menunjukkan bahwa

kernel linear memiliki kemampuan yang lebih baik dalam melakukan klasifikasi sentimen pada data teks komentar YouTube yang berdimensi tinggi.

6.2 Saran

Berdasarkan hasil penelitian, berikut beberapa saran untuk penelitian selanjutnya:

1. Melakukan penyempurnaan pada tahap *preprocessing*, seperti penambahan *stopword* khusus yang sesuai konteks data dan peningkatan normalisasi agar lebih mampu menangani variasi bahasa tidak baku.
2. Metode pelabelan leksikon pada penelitian ini masih memiliki keterbatasan dalam memahami konteks kalimat, sehingga penulis sarankan untuk melakukan pelabelan secara manual untuk menghasilkan analisis sentimen yang lebih akurat.
3. Penelitian selanjutnya dapat menggunakan dataset yang lebih beragam dengan topik luas agar dapat mengurangi dominasi kata dengan topik tertentu dan meningkatkan kualitas model klasifikasi.

DAFTAR PUSTAKA

- Aggarwal, C. C., & Zhai, C. X. (2012). Mining text data. In *Mining Text Data* (Vol. 9781461432). Springer. <https://doi.org/10.1007/978-1-4614-3223-4>
- Akbar, A. L. (2015). *Implementasi Algoritma SVM untuk Mengetahui Tingkat Resiko Penyakit Stroke*. Universitas Brawijaya.
- Amna, S. W., Sudipa, I. G. I., Putra, T. A. E., Wahidin, A. J., Syukrilla, W. A., Wardhani, A. K., Heryana, N., Indriyani, T., & Santoso, L. W. (2023). Data Mining. In *Mining of Massive Datasets* (Vol. 2, Issue January 2013). PT Global Eksekutif Teknologi.
- Anam, M. K., Lestari, T. P., Efrizoni, L., Handayani, N. S., & Andhika, I. (2025). Sentiment Analysis Optimization Using Ensemble of Multiple SVM Kernel Functions. *Resti*, 9(4), 905–914. <https://doi.org/https://doi.org/10.29207/resti.v9i4.6708> Received:
- Andrew, N. (2009). *Autumn 2009 The Simplified SMO Algorithm*. CS 229. <https://cs229.stanford.edu/materials/smo.pdf>
- Anton, H., & Rorres, C. (2013). *Elementary linear algebra : Applications Version 11th Edition*. Wiley.
- Bahri, S. (2017). *Teori Pengoptimuman* (Vol. 1). Program Studi Matematika UNRAM.
- Basriati, S., Safitri, E., Ulfa, N., & Matematika, J. (2020). *Optimasi Rata-Rata Produksi Kelapa di Kabupaten Indragiri Hilir Menggunakan Metode Wolfe*. 6(1), 90–97.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Boser, B. E., Guyon, I. M., & Vladimir N Vapnik. (1992). A training algorithm for optimal margin classifiers. *Gautampendse.Com*, 144–152.
- Cholissodin, I., Sutrisno, Soebroto, A. A., Hasanah, U., & Febiola, Y. I. (2020). *AI, Machine Learning & Deep Learning*. Universitas Brawijaya.
- Christina, S., & Ronaldo, D. (2016). Peran Opinion Mining Dan Sentiment Analysis Untuk Mengidentifikasi Sentimen Publik Dalam Sistem E-Governance. *J. Teknol*, 69.
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1111/j.1747-0285.2009.00840.x>
- DataReportal, S. K. (2025). *Digital 2026: Indonesia*. Data Reportal. <https://datareportal.com/reports/digital-2026-indonesia>
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced*

- Approaches in Analyzing Unstructured Data* (1st editio). Cambridge University Press.
- Heryadi, Y., & Wahyono, T. (2020). *Machine Learning Konsep dan Implementasi* (Cetakan I). Penerbit Gava Media.
- Hsu, B. M. (2020). Comparison of supervised classification models on textual data. *Mathematics*, 8(5). <https://doi.org/10.3390/MATH8050851>
- Jurafsky, D., & Martin, J. H. (2025). *Speech and Language Processing*.
- Koto, F., & Rahmaningtyas, G. Y. (2017). *InSet Lexicon : Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs*. 391–394.
- Kowalczyk, A. (2014). *Linear Kernel: Why is it recommended for text classification ? SVM Tutorial*. <https://www.svm-tutorial.com/2014/10/svm-linear-kernel-good-text-classification/>
- Kowalczyk, A. (2017). Support Vector Machines. In C. Wright (Ed.), *Succinctly* (1st ed.). Syncfusion Inc.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers. <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>
- Machali, I. (2021). *Metode Penelitian Kuantitatif*. Fakultas Ilmu Tarbiyah dan Keguruan Universitas Islam Negeri (UIN) Sunan Kalijaga Yogyakarta.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Nocedal, J., & Wright, S. J. (2006). *Numerical Optimization*. Springer.
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). *Support Vector Machine – Teori dan Aplikasinya dalam Bioinformatika*. <https://doi.org/10.18411/trnio-12-2023-769>
- Nur Hayati, D. (2024). *Machine Learning* (M. K. Neny Rosmawarni, S.Kom (ed.); Cetakan Pe). PT Penamuda Media.
- Panji Bintoro, Ratnasari, Edy Wihardjo, I. P. P., & Asari, A. (2024). *Pengantar Machine Learning* (A. Asari (ed.); 1st ed.). PT Mafy Media Literasi Indonesia.
- Platt, J. C. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. *Microsoft Research, July 1998*, 1–21. <https://www.microsoft.com/en-us/research/wp-content/uploads/1998/04/sequential-minimal-optimization.pdf>
- Praghakusma, A. Z., & Charibaldi, N. (2021). Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi). *JSTIE (Jurnal Sarjana Teknik Informatika) (E-Journal)*, 9(2), 88. <https://doi.org/10.12928/jstie.v9i2.20181>

- Prasetyo, Y. A., Utami, E., & Yaqin, A. (2024). *Pengaruh Komposisi Split Data Terhadap Performa Akurasi Analisis Sentimen Algoritma Naïve Bayes dan SVM*. 6(2), 382–390. <https://doi.org/10.33650/jeeecom.v4i2>
- Putra, I. G. A. J., Asih, N. M., & Widana, I. N. (2015). Optimalisasi Penjualan Kain Endek Dengan Metode Karush-Kuhn-Tucker (Kkt). *E-Jurnal Matematika*, 4(4), 158. <https://doi.org/10.24843/mtk.2015.v04.i04.p105>
- Ressan, M. B., & Hassan, R. F. (2022). Naïve-Bayes family for sentiment analysis during COVID-19 pandemic and classification tweets. *Indonesian Journal of Electrical Engineering and Computer Science*, 28(1), 375–383. <https://doi.org/10.11591/ijeecs.v28.i1.pp375-383>
- Riki, Eliesse Dameria, S., Hermawan, A., Junaedi, & Kurnia, Y. (2025). Optimizing Sentiment Classification of E-Commerce Product Reviews: A Comparative Study of Naïve Bayes and SVM with SMO. *JUITA: Jurnal Informatika*, 13(3), 287–295. <https://doi.org/10.30595/juita.v13i3.26642>
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>.
- Sembiring, K. (2007). *Penerapan Teknik Support Vector Machine Untuk Pendeteksian Intrusi Pada Jaringan* [Institut Teknologi Bandung]. https://digilib.itb.ac.id/gdl/view_data/penerapan-teknik-support-vector-machine-untuk-pendeteksian-intrusi-pada-jaringan
- Sudipa, I. G. I. (2023). *Representation Knowledge Data Mining* (D. Ediana (ed.); 1st ed.). PT Global Eksekutif Teknologi.
- Sunyoto, A., & Fatta, H. Al. (2023). *Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier*. VII(September), 31–40.
- Taha, H. A. (2017). *Operations Research An Introduction Tenth Edition*. Pearson.