

**ANALISIS SENTIMEN PUBLIK TERKAIT DEFORESTASI DI
INDONESIA MENGGUNAKAN KLASIFIKASI *SUPPORT
VECTOR MACHINE* (SVM)**



SKRIPSI

Amandaselvi Aulia

2103040128

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN SAINS
UNIVERSITAS MUHAMMADIYAH PURWOKERTO
MEI 2025**

**ANALISIS SENTIMEN PUBLIK TERKAIT DEFORESTASI DI
INDONESIA MENGGUNAKAN KLASIFIKASI *SUPPORT
VECTOR MACHINE* (SVM)**



SKRIPSI

Diajukan untuk memenuhi salah satu syarat memperoleh gelar sarjana Komputer

Amandaselvi Aulia

2103040128

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN SAINS
UNIVERSITAS MUHAMMADIYAH PURWOKERTO
MEI 2025**

HALAMAN PERSETUJUAN

Skripsi yang diajukan oleh:

Nama : Amandaselvi Aulia

NIM : 2103040128

Program Studi : Teknik Informatika

Fakultas : Teknik Dan Sains

Perguruan Tinggi : Universitas Muhammadiyah Purwokerto

Judul : Analisis Sentimen Publik Terkait Deforestasi di Indonesia Menggunakan Klasifikasi *Support Vector Machine* (SVM)

telah diterima dan disetujui

Purwokerto, 19 Mei 2025

PEMBIMBING

Maulida Ayu F., S.Kom., M.Cs.

NIK. 2160823

HALAMAN PENGESAHAN

Skripsi yang diajukan oleh:

Nama : Amandaselvi Aulia

NIM : 2103040128

Program Studi : Teknik Informatika

Fakultas : Teknik dan Sains

Perguruan Tinggi : Universitas Muhammadiyah Purwokerto

Judul : Analisis Sentimen Publik Terkait Deforestasi Di
Indonesia Menggunakan Klasifikasi *Support Vector
Machine* (SVM)

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian
persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer (S.Kom.)
pada Program Studi Teknik Informatika, Fakultas Teknik dan Sains, Universitas
Muhammadiyah Purwokerto.

DEWAN PENGUJI

Penguji 1 (Pembimbing) : Maulida Ayu Fitriani, S.Kom.,M.Cs.

Penguji 2 : Muhammad Hamka, S.T., M.Kom.

Penguji 3 : Elindra Ambar Pambudi, S.Kom.,M.Kom.

Ditetapkan di : Purwokerto

Tanggal : 14 Mei 2025

Mengetahui

Dekan Fakultas Teknik dan Sains



Dr. Ir. Iskahar, S.T., M.T.

NIK. 2160207

HALAMAN PERNYATAAN ORISINALITAS

Saya yang bertandatangan di bawah ini,

Nama : Amandaselvi Aulia

NIM. : 2103040128

Program Studi : Teknik Informatika

Fakultas : Teknik dan Sains

Perguruan Tinggi : Universitas Muhammadiyah Purwokerto

Menyatakan dengan sebenar-benarnya bahwa skripsi ini adalah hasil karya saya dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta bukan hasil penjiplakan dari karya orang lain.

Demikian pernyataan ini saya buat dan apabila kelak di kemudian hari terbukti ada unsur penjiplakan, saya bersedia mempertanggungjawabkan sesuai dengan ketentuan yang berlaku.

Purwokerto, 14 Mei 2025

uat pernyataan



Amandaselvi Aulia

NIM. 2103040128

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Muhammadiyah Purwokerto dan demi pengembangan ilmu pengetahuan, saya yang bertanda tangan di bawah ini:

Nama : Amandaselvi Aulia
NIM. : 2103040128
Program Studi : Teknik Informatika
Fakultas : Teknik dan Sains
Perguruan Tinggi : Universitas Muhammadiyah Purwokerto
Jenis karya : Skripsi

menyetujui untuk memberikan Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right) kepada Universitas Muhammadiyah Purwokerto atas karya ilmiah saya yang berjudul:

“Analisis Sentimen Publik Terkait Deforestasi di Indonesia Menggunakan
Klasifikasi *Support Vector Machine* (SVM)”

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Muhammadiyah Purwokerto berhak menyimpan, mengalihmedia/ mengalihformatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan skripsi saya dengan tetap mencantumkan nama saya sebagai penulis/ pencipta dan sebagai pemilik Hak Cipta.

Pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Purwokerto

Pada tanggal: 14 Mei 2025



menyatakan

Amandaselvi Aulia

KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Publik Terkait Deforestasi di Indonesia Menggunakan Klasifikasi *Support Vector Machine* (SVM)”.

Laporan skripsi ini diajukan untuk memenuhi syarat yang diperlukan untuk memperoleh gelar Sarjana Komputer (S.Kom) pada program studi Teknik Informatika Fakultas Teknik dan Sains Universitas Muhammadiyah Purwokerto.

Selama proses penyusunan laporan ini, penulis telah menerima banyak dukungan, bimbingan, serta motivasi dari berbagai pihak. Penulis menyadari bahwa laporan ini masih jauh dari sempurna, baik dari segi isi maupun penyajiannya. Oleh karena itu, penulis sangat mengharapkan saran dan kritik dari pembaca agar laporan skripsi nantinya dapat menjadi laporan yang lebih baik lagi.

Purwokerto, 14 Mei 2025

Amandaselvi Aulia

HALAMAN PERSEMBAHAN

Dalam menyelesaikan skripsi ini, penulis tidak lepas dari bantuan berupa dukungan dan bimbingan dari berbagai pihak sehingga skripsi ini dapat terselesaikan. Dengan segala kerendahan hati, dan rasa hormat melalui tulisan ini, penulis ingin mengucapkan terimakasih kepada:

1. Prof. Jebul Suroso, selaku Rektor Universitas Muhammadiyah Purwokerto.
2. Bapak Dr. Ir. Iskahar, S.T., M.T. selaku Dekan Fakultas Teknik Dan Sains, Universitas Muhammadiyah purwokerto.
3. Bapak Agung Purwo Wicaksono, S.T., M.Kom. selaku Ketua Program Studi Teknik Informatika, Universitas Muhammadiyah Purwokerto.
4. Ibu Maulida Ayu Fitriani, S.Kom.,M.Cs. selaku dosen pembimbing yang telah membantu penulis untuk mengarahkan dan membimbing penulis dalam penyusunan skripsi.
5. Seluruh dosen dan karyawan Program Studi Teknik Informatika yang telah banyak memberikan bimbingan dan ilmu pengetahuan selama penulis mengikuti perkuliahan sehingga dapat menyelesaikan skripsi ini.
6. Kedua orang tua tercinta, Bapak Danto dan Ibu Sugiarti. Terima kasih atas segala pengorbanan, kasih sayang, doa yang tak pernah putus, dan dukungan tanpa batas yang telah diberikan sehingga penulis dapat menyelesaikan penyusunan laporan skripsi ini.
7. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu penulis dalam menyelesaikan laporan skripsi ini.

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN ORISINALITAS	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	vi
KATA PENGANTAR.....	vii
HALAMAN PERSEMBAHAN	viii
DAFTAR ISI	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
ABSTRAK	xiii
<i>ABSTRACT</i>	xiv
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Rumusan Masalah	3
C. Batasan Masalah.....	3
D. Tujuan Penelitian.....	3
E. Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
A. Peneliti Terdahulu	5
B. Landasan Teori	8
BAB III METODE PENELITIAN.....	15
A. Jenis Penelitian.....	15
B. Sumber Data.....	15
C. Langkah-Langkah Penelitian	15
BAB IV HASIL DAN PEMBAHASAN.....	26
A. Pengumpulan Data	26
B. <i>Preprocessing</i>	27

C. Pelabelan Data.....	39
D. Visualisasi Data.....	42
E. Pembobotan Kata (TF-IDF).....	47
F. Klasifikasi <i>Support Vector Machine</i>	50
G. Evaluasi Model.....	53
BAB V PENUTUP.....	61
A. Kesimpulan	61
B. Saran.....	61
DAFTAR PUSTAKA.....	63
LAMPIRAN.....	70

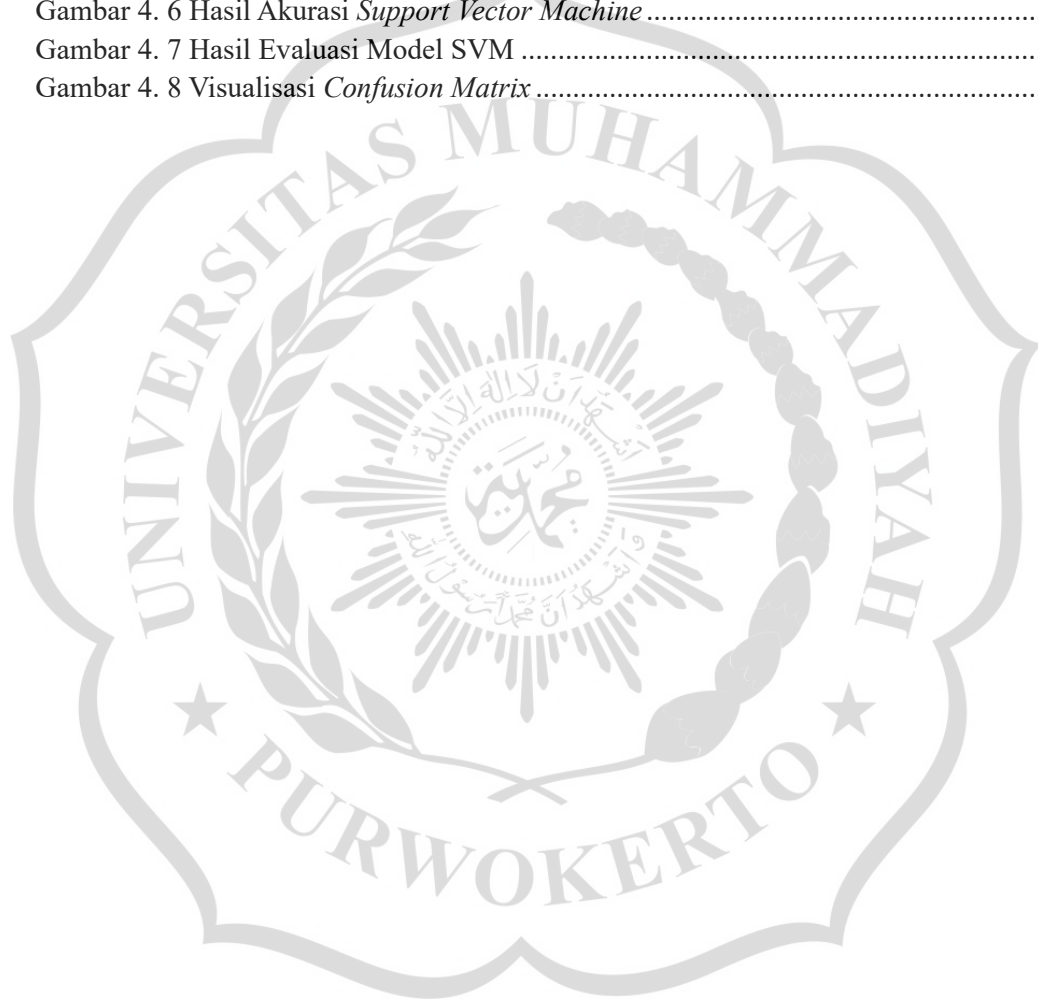


DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	13
Tabel 3. 1 Contoh <i>Tweet</i> Hasil <i>Cleaning</i>	18
Tabel 3. 2 Contoh <i>Tweet</i> Hasil <i>Case Folding</i>	18
Tabel 3. 3 Contoh <i>Tweet</i> Hasil Normalisasi	19
Tabel 3. 4 Contoh <i>Tweet</i> Hasil <i>Tokenizing</i>	20
Tabel 3. 5 Contoh <i>Tweet</i> Hasil <i>Stopword Removal</i>	21
Tabel 3. 6 Contoh <i>Tweet</i> Hasil <i>Stemming</i>	22
Tabel 3. 7 Bobot Sentimen Pada VADER.....	23
Tabel 3. 8 Contoh Hasil Perhitungan <i>Confusion Matrix</i>	25
Tabel 4. 1 Hasil <i>Crawling</i> Data.....	26
Tabel 4. 2 Hasil Penghapusan Data Duplikat.....	28
Tabel 4. 3 Hasil <i>Cleaning</i>	29
Tabel 4. 4 Hasil <i>Case Folding</i>	31
Tabel 4. 5 Hasil Normalisasi	32
Tabel 4. 6 Hasil <i>Tokenizing</i>	34
Tabel 4. 7 Hasil <i>Stopword Removal</i>	35
Tabel 4. 8 Hasil <i>Stemming</i>	37
Tabel 4. 9 Hasil <i>Data Translation</i>	38
Tabel 4. 10 Hasil Pelabelan Data	39
Tabel 4. 11 Hasil Perhitungan Manual <i>Compound Score</i>	41
Tabel 4. 12 Jumlah Pelabelan Data	41
Tabel 4. 13 Data <i>Tweet Wordcloud</i> Positif.....	44
Tabel 4. 14 Data <i>Tweet Wordcloud</i> Negatif	45
Tabel 4. 15 Data <i>Tweet Wordcloud</i> Netral	47
Tabel 4. 16 Contoh Data Untuk TF-IDF	48
Tabel 4. 17 Perhitungan TF dan TF Normalisasi	48
Tabel 4. 18 Perhitungan IDF dan TF-IDF	49
Tabel 4. 19 Perhitungan manual SVM.....	51
Tabel 4. 20 Hasil Splitting Data	52

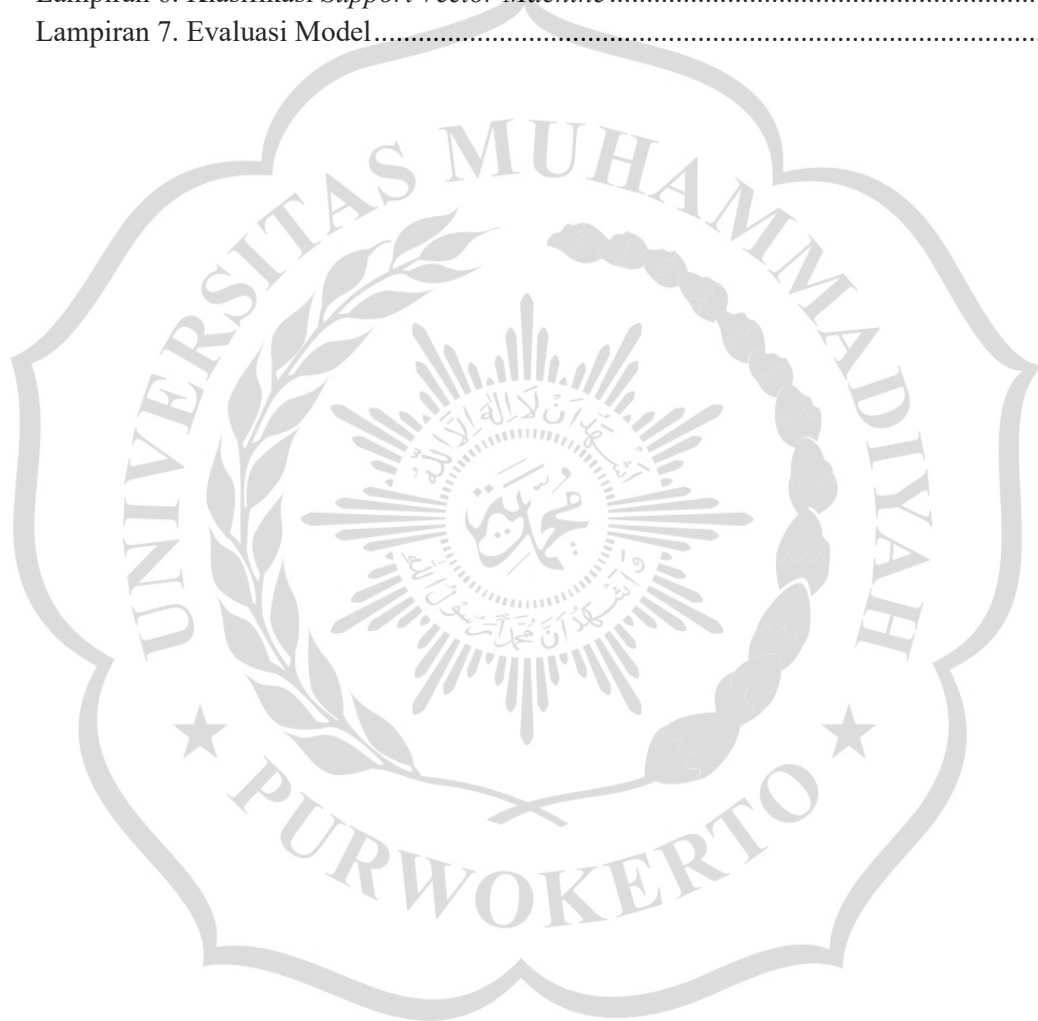
DAFTAR GAMBAR

Gambar 3. 1 <i>Flowchart</i> Penelitian	16
Gambar 4. 1 Visualisasi <i>Pie Chart</i>	42
Gambar 4. 2 Visualisasi <i>Wordcloud</i> Semua Data.....	43
Gambar 4. 3 Visualisasi <i>Wordcloud</i> Sentimen Positif.....	43
Gambar 4. 4 Visualisasi <i>Wordcloud</i> Sentimen Negatif	45
Gambar 4. 5 Visualisasi <i>Wordcloud</i> Sentimen Netral	46
Gambar 4. 6 Hasil Akurasi <i>Support Vector Machine</i>	53
Gambar 4. 7 Hasil Evaluasi Model SVM	54
Gambar 4. 8 Visualisasi <i>Confusion Matrix</i>	54



DAFTAR LAMPIRAN

Lampiran 1. <i>Crawling</i> Data	70
Lampiran 2. <i>Preprocessing</i> Data.....	70
Lampiran 3. Pelabelan Data	77
Lampiran 4. Visualisasi Data	77
Lampiran 5. Pembobotan Kata TF-IDF	79
Lampiran 6. Klasifikasi <i>Support Vector Machine</i>	81
Lampiran 7. Evaluasi Model.....	82



ANALISIS SENTIMEN PUBLIK TERKAIT DEFORESTASI DI INDONESIA MENGGUNAKAN KLASIFIKASI *SUPPORT VECTOR MACHINE* (SVM)

Amandaselvi Aulia, Maulida Ayu Fitriani, S.Kom.,M.Cs.

ABSTRAK

Deforestasi merupakan proses penggundulan atau penghilangan tutupan hutan di suatu area yang disebabkan oleh konversi lahan untuk keperluan pertanian, perkebunan, pertambangan, atau pembangunan infrastruktur. Deforestasi di Indonesia ini memunculkan beragam pandangan di masyarakat baik yang mendukung maupun yang menentang. Oleh karena itu penelitian ini bertujuan untuk menganalisis sentimen publik terkait deforestasi di Indonesia menggunakan klasifikasi *Support Vector Machine*. Data diperoleh dari media sosial *X* pada 1 Januari 2021 hingga 1 Oktober 2024. Data dikumpulkan melalui teknik *crawling* dengan kata kunci “deforestasi”, “penggundulan hutan”, “pembukaan lahan”, dan “kebakaran hutan”. Data yang telah dikumpulkan kemudian melalui serangkaian tahap *preprocessing* yang digunakan untuk membersihkan elemen, karakter, atau kata yang kurang relevan, data yang digunakan dalam penelitian ini berjumlah 8.359. Data tersebut dilabeli menggunakan *lexicon* VADER. Visualisasi data yang dilakukan menggunakan visualisasi pie chart dan wordcloud. Pembobotan kata dihitung menggunakan teknik TF-IDF untuk menentukan tingkat pentingnya setiap kata dalam dokumen. Data dibagi menjadi data latih dan data dengan proporsi 80:20, dimana algoritma *Support Vector Machine* digunakan untuk mengklasifikasikan sentimen. Hasil penelitian menunjukkan sentimen positif sebesar 36,3% atau 3.035 data, sentimen negatif sebesar 35,9% atau 2.999 data, dan sentimen netral sebesar 27,8% atau 2.325 data. Hasil evaluasi menunjukkan penggunaan algoritma *Support Vector Machine* mendapatkan nilai akurasi sebesar 70%, nilai *precision*, *recall*, dan *f1-score* sebesar 70%.

Kata kunci: deforestasi, analisis sentimen, *support vector machine*

**SENTIMENT ANALYSIS PUBLIC CONCERNING DEFORESTATION IN
INDONESIA USING SUPPORT VECTOR MACHINE (SVM)
CLASSIFICATION**

Amandaselvi Aulia, Maulida Ayu Fitriani, S.Kom.,M.Cs.

ABSTRACT

Deforestation is the process of deforestation or removal of forest cover in an area caused by land conversion for agricultural, plantation, mining or infrastructure development purposes. Deforestation in Indonesia has given rise to various views in society, both for and against. Therefore, this research aims to analyze public sentiment regarding deforestation in Indonesia using the Support Vector Machine classification. Data was obtained from X social media on January 1, 2021 to October 1, 2024. Data was collected through crawling technique with the keywords "deforestation", "deforestation", "land clearing", and "forest fires". The data that has been collected then goes through a series of preprocessing stages which are used to clean elements, characters or words that are less relevant. The data used in this research amounted to 8,359. The data is labeled using the VADER lexicon. Data visualization was carried out using pie chart and wordcloud visualization. Word weighting is calculated using the TF-IDF technique to determine the level of importance of each word in the document. The data is divided into training data and data with a proportion of 80:20, where the Support Vector Machine algorithm is used to classify sentiment. The research results showed positive sentiment of 36.3% or 3,035 data, negative sentiment of 35.9% or 2,999 data, and neutral sentiment of 27.8% or 2,325 data. The evaluation results show that using the Support Vector Machine algorithm gets an accuracy value of 70%, a precision, recall and f1-score value of 70%.

Keywords: *deforestation, sentiment analysis, support vector machine*