

**ANALISIS SENTIMEN PENGGUNA APLIKASI X TERHADAP  
PENGUNSI ROHINGYA DI ACEH MENGGUNAKAN  
METODE SUPPORT VECTOR MACHINE**



**UMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

**SKRIPSI**

**Niko Adriano**  
**00000047735**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG  
2024**

**ANALISIS SENTIMEN PENGGUNA APLIKASI X TERHADAP  
PENGUNSI ROHINGYA DI ACEH MENGGUNAKAN  
METODE SUPPORT VECTOR MACHINE**



**SKRIPSI**

Diajukan sebagai salah satu syarat untuk memperoleh  
Gelar Sarjana Komputer (S.Kom.)

**Niko Adriano**

**00000047735**

**UMN**

**UNIVERSITAS**

**MULTIMEDIA**

**NUSANTARA**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA**

**TANGERANG**

**2024**

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Niko Adriano  
Nomor Induk Mahasiswa : 00000047735  
Program Studi : Informatika

Skripsi dengan judul:

**Analisis Sentimen Pengguna Aplikasi X terhadap Pengungsi Rohingya di Aceh Menggunakan Metode Support Vector Machine**

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 15 Mei 2024



UNIVERSITAS (Niko Adriano)  
MULTIMEDIA  
NUSANTARA

**HALAMAN PENGESAHAN**

Skripsi dengan judul

**ANALISIS SENTIMEN PENGGUNA APLIKASI X TERHADAP  
PENGUNGS ROHINGYA DI ACEH MENGGUNAKAN  
METODE SUPPORT VECTOR MACHINE**

oleh

Nama : Niko Adriano  
NIM : 00000047735  
Program Studi : Informatika  
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 4 Juni 2024

Pukul 10.00 s/s 12.00 dan dinyatakan

**LULUS**

Dengan susunan penguji sebagai berikut

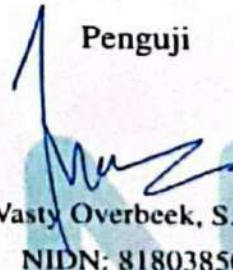
Ketua Sidang



(Angga Aditya Permana, S.Kom., M.Kom.)

NIDN: 0407118901

Penguji



(Marlinda Vasty Overbeek, S.Kom, M.Kom)

NIDN: 818038501

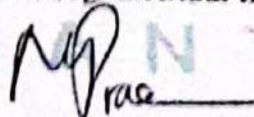
Pembimbing



(Suwito Pomalingo, S.Kom., M.Kom.)

NIDN: 0911098201

Pjs. Ketua Program Studi Informatika.



(Dr. Eng. Niki Prastomo, S.T., M.Sc.)

NIDN: 0419128203

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK  
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Niko Adriano

NIM : 00000047735

Program Studi : Informatika

Jenjang : S1

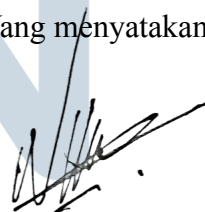
Jenis Karya :

Skripsi Menyatakan dengan sesungguhnya  
bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)\*\*.

Tangerang, 15 Mei 2024

Yang menyatakan

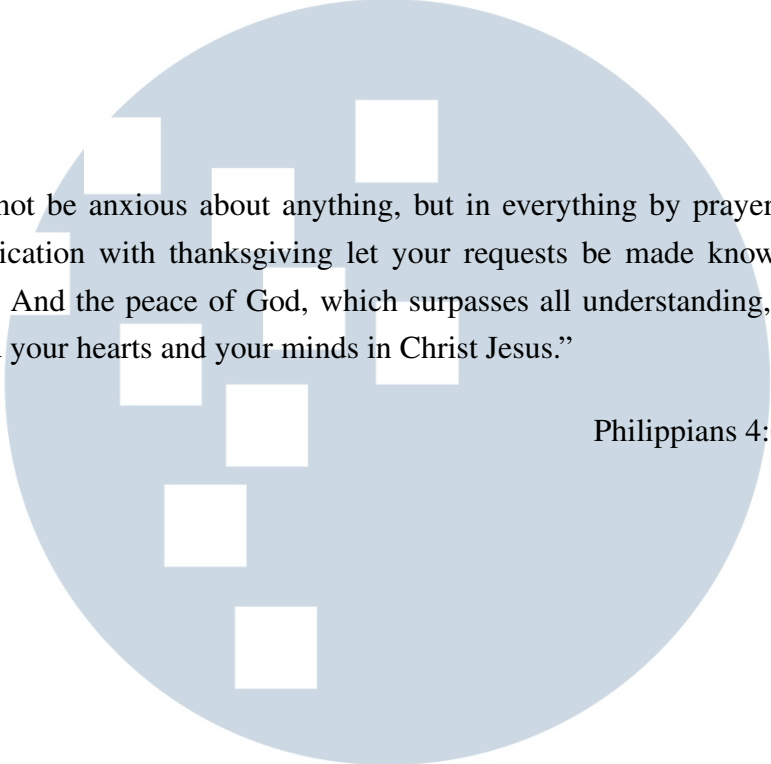


Niko Adriano

UMMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

\*\* Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

## Halaman Persembahan / Motto



”Do not be anxious about anything, but in everything by prayer and supplication with thanksgiving let your requests be made known to God. And the peace of God, which surpasses all understanding, will guard your hearts and your minds in Christ Jesus.”

Philippians 4:6-7 (ESV)

UMMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Analisis Sentimen Pengguna Aplikasi X terhadap Pengungsi Rohingya di Aceh Menggunakan Metode Support Vector Machine dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika dan Ketua Program Studi Informatika Universitas Multimedia Nusantara.
3. Bapak Suwito Pomalingo, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
4. Keluarga saya khususnya orang tua yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.
5. Diri saya sendiri yang telah berjuang, pantang menyerah, dan mampu untuk menyelesaikan tesis ini

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 15 Mei 2024



Niko Adriano

**ANALISIS SENTIMEN PENGGUNA APLIKASI X TERHADAP  
PENGUNGS ROHINGYA DI ACEH MENGGUNAKAN  
METODE SUPPORT VECTOR MACHINE**

Niko Adriano

**ABSTRAK**

Permasalahan sosial yang terus berkembang antara pengungsi Rohingya dan masyarakat lokal Aceh memicu berbagai opini dari masyarakat Indonesia. Opini-opini tersebut banyak diungkapkan melalui media sosial salah satunya X, yang pada akhirnya menimbulkan sentimen terhadap pengungsi Rohingya di Aceh. Perbedaan sentimen masyarakat khususnya pada 'post' X menjadi dasar penelitian untuk melakukan klasifikasi dan analisis sentimen. Pengambilan data dari X melalui metode scraping menghasilkan 12778 data yang kemudian dibersihkan menjadi 10196 data. Data tersebut kemudian diklasifikasikan menjadi tiga kategori sentimen, yaitu positif, netral, dan negatif dengan pemodelan Support Vector Machine (SVM) menggunakan seleksi fitur Chi Square. Hasil klasifikasi menunjukkan polaritas sebesar 61.3% sentimen negatif, 20.5% sentimen positif, dan 18.2% sentimen netral. Sentimen didominasi oleh sentimen negatif dengan pembahasan utama terkait kritik ketidakpuasan terhadap penanganan pemerintah dan keprihatinan terhadap pengungsi Rohingya. Hasil evaluasi pemodelan SVM dengan seleksi fitur Chi Square serta kombinasi parameter kernel rbf,  $C = 10$ , dan  $\gamma = 1$  menunjukkan tingkat akurasi sebesar 81%, presisi sebesar 80%, dan recall sebesar 81%. dan F1-score sebesar 80%.

**Kata kunci:** Analisis Sentimen, *Chi Square*, *Confussion Matrix*, Pengungsi Rohingya, *Support Vector Machine*

U I M N  
U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A



***Sentiment Analysis of X Application User towards Rohingya Refugees in Aceh  
Using the Support Vector Machine Method***

Niko Adriano

***ABSTRACT***

*The ongoing social issues between Rohingya refugees and the local Acehnese community have sparked various opinions from the Indonesian public. These opinions are widely expressed through social media, including on X, which has ultimately led to sentiments toward Rohingya refugees in Aceh. The differences in public sentiment, particularly on X posts, form the basis for research to classify and analyze sentiment. Data collection from X through scraping methods resulted in 12,778 data points, which were then cleaned down to 10,196 data points. The data were then classified into three sentiment categories: positive, neutral, and negative, using Support Vector Machine (SVM) modeling with Chi-Square feature selection. The classification results showed a polarity of 61.3% negative sentiment, 20.5% positive sentiment, and 18.2% neutral sentiment. The sentiment was dominated by negative sentiment, with the main discussions related to criticism of government handling and concerns for the Rohingya refugees. The evaluation results of the SVM model with Chi-Square feature selection and a combination of rbf kernel parameters,  $C = 10$ , and  $\gamma = 1$ , showed an accuracy rate of 81%, precision of 80%, recall of 81%, and an F1-score of 80%.*

***Keywords:*** *Sentiment Analysis, Chi Square, Confussion Matrix, Rohingya Refugees, Support Vector Machine*



## DAFTAR ISI

HALAMAN JUDUL . . . . .	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT . . . . .	ii
HALAMAN PENGESAHAN . . . . .	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH . . . . .	iv
HALAMAN PERSEMBAHAN/MOTO . . . . .	v
KATA PENGANTAR . . . . .	vi
ABSTRAK . . . . .	vii
ABSTRACT . . . . .	viii
DAFTAR ISI . . . . .	ix
DAFTAR GAMBAR . . . . .	xi
DAFTAR TABEL . . . . .	xii
DAFTAR KODE . . . . .	xiii
DAFTAR LAMPIRAN . . . . .	xiv
BAB 1 PENDAHULUAN . . . . .	1
1.1 Latar Belakang Masalah . . . . .	1
1.2 Rumusan Masalah . . . . .	3
1.3 Batasan Permasalahan . . . . .	3
1.4 Tujuan Penelitian . . . . .	4
1.5 Manfaat Penelitian . . . . .	4
1.6 Sistematika Penulisan . . . . .	4
BAB 2 LANDASAN TEORI . . . . .	6
2.1 Teori tentang Topik . . . . .	6
2.1.1 Analisis Sentimen . . . . .	6
2.1.2 Pengungsi Rohingya . . . . .	6
2.2 Teori tentang Algoritma . . . . .	7
2.2.1 Support Vector Machine . . . . .	7
2.2.2 TF-IDF . . . . .	10
2.2.3 Chi Square . . . . .	11
2.2.4 Confussion Matrix . . . . .	12
BAB 3 METODOLOGI PENELITIAN . . . . .	15
3.1 Alur Diagram Penelitian . . . . .	15
3.2 Studi Literatur . . . . .	15
3.3 Data Scraping . . . . .	16
3.4 Text Preprocessing . . . . .	17
3.5 Data Labeling . . . . .	19
3.6 Visualisasi dan Analisa Hasil Labeling . . . . .	20
3.7 Implementasi Feature Extraction dengan TF-IDF . . . . .	21
3.8 Implementasi Feature Selection dengan Chi Square . . . . .	21
3.9 Implementasi Modeling dengan Algoritma SVM . . . . .	22
BAB 4 HASIL DAN DISKUSI . . . . .	24
4.1 Spesifikasi Sistem . . . . .	24
4.2 Scraping Data . . . . .	24
4.3 Text Preprocessing . . . . .	26
4.3.1 Cleansing . . . . .	28
4.3.2 Case Folding . . . . .	30
4.3.3 Tokenization . . . . .	30
4.3.4 Normalization . . . . .	31
4.3.5 Stopword Removal . . . . .	34

4.3.6	Stemming . . . . .	35
4.4	Data Labeling . . . . .	36
4.5	Visualisasi dan Analisa Hasil Labeling . . . . .	38
4.6	Feature Extraction dengan TF-IDF . . . . .	41
4.7	Feature Selection dengan Chi Square . . . . .	42
4.8	Implementasi Modeling SVM . . . . .	43
4.8.1	Train-Test Split Data . . . . .	43
4.8.2	Pencarian Parameter SVM Terbaik . . . . .	44
4.8.3	Modeling SVM . . . . .	45
4.8.4	Evaluasi . . . . .	46
4.8.5	Prediksi Teks . . . . .	47
BAB 5	SIMPULAN DAN SARAN . . . . .	49
5.1	Simpulan . . . . .	49
5.2	Saran . . . . .	49
	DAFTAR PUSTAKA . . . . .	50



## DAFTAR GAMBAR

Gambar 1.1	Pengungsi Rohingnya di Bluka Teubai, Aceh . . . . .	1
Gambar 2.1	Hyperplane Terbaik . . . . .	7
Gambar 3.1	Alur Metodologi Penelitian . . . . .	15
Gambar 3.2	Alur Data <i>Scraping</i> . . . . .	16
Gambar 3.3	Alur <i>Text Preprocessing</i> . . . . .	17
Gambar 3.4	Alur <i>Normalization</i> . . . . .	18
Gambar 3.5	Alur <i>Labeling</i> . . . . .	20
Gambar 3.6	Alur <i>Feature Extraction</i> dengan TF-IDF . . . . .	21
Gambar 3.7	Alur Seleksi Fitur dengan Chi-Square . . . . .	22
Gambar 3.8	Alur <i>Modeling</i> dengan Algoritma SVM . . . . .	23
Gambar 4.1	Distribusi Sentimen . . . . .	39
Gambar 4.2	Frekuensi Data Sentimen Positif . . . . .	39
Gambar 4.3	Frekuensi Data Sentimen Netral . . . . .	40
Gambar 4.4	Frekuensi Data Sentimen Negatif . . . . .	40
Gambar 4.5	Hasil Pencarian Parameter SVM terbaik . . . . .	45
Gambar 4.6	Classification Report Hasil <i>Modeling</i> . . . . .	47
Gambar 4.7	Confussion Matrix Hasil <i>Modeling</i> . . . . .	47
Gambar 4.8	Hasil Prediksi Kelas dari Teks Manual . . . . .	48



## DAFTAR TABEL

Tabel 2.1	Perbandingan berbagai jenis kernel dalam SVM . . . . .	10
Tabel 2.2	Confusion Matrix pada Label Negatif . . . . .	12
Tabel 2.3	Confusion Matrix pada Label Netral . . . . .	13
Tabel 2.4	Confusion Matrix pada Label Positif . . . . .	13
Tabel 4.1	Rincian Pengumpulan <i>Dataset</i> . . . . .	25
Tabel 4.2	Perbandingan Data Sebelum dan Sesudah <i>Penyesuaian</i> . . . . .	27
Tabel 4.3	Perbandingan Teks Sebelum dan Sesudah <i>Cleansing</i> . . . . .	29
Tabel 4.4	Perbandingan Teks Sebelum dan Sesudah <i>Case Folding</i> . . . . .	30
Tabel 4.5	Perbandingan Teks Sebelum dan Sesudah <i>Tokenization</i> . . . . .	31
Tabel 4.6	Perbandingan Teks Sebelum dan Sesudah <i>Normalization</i> . . . . .	34
Tabel 4.7	Perbandingan Teks Sebelum dan Sesudah <i>Stopword Removal</i> . . . . .	35
Tabel 4.8	Perbandingan Teks Sebelum dan Sesudah <i>Stemming</i> . . . . .	36
Tabel 4.9	Hasil Data <i>Labeling</i> . . . . .	38



## DAFTAR KODE

4.1	Pseudocode Scraping Data dari X	24
4.2	Pseudocode Penggabungan Data	26
4.3	Pseudocode Penyesuaian Data dari Bias	26
4.4	Pseudocode Cleansing	28
4.5	Pseudocode Case Folding	30
4.6	Pseudocode Tokenization	31
4.7	Pseudocode Normalization	32
4.8	Pseudocode Pengecekan Kata Indikasi Singkatan Terbanyak	32
4.9	Pseudocode Stop Word Removal	34
4.10	Pseudocode Stemming	35
4.11	Pseudocode Data Labeling	36
4.12	Pseudocode Feature Extraction TF-IDF	41
4.13	Pseudocode Feature Selection Chi Square	42
4.14	Pseudocode Train Test Split Data	43
4.15	Pencarian Parameter SVM Terbaik dengan GridSearchCV	44
4.16	Pseudocode Modeling SVM	45
4.17	Pseudocode Evaluasi Model	46
4.18	Pseudocode Prediksi Kelas Teks Baru	47

UMMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan . . . . .	53
Lampiran 2	Source Code Crawling Data . . . . .	54
Lampiran 3	Source Code Combined Data . . . . .	54
Lampiran 4	Source Code Penyesuaian Data . . . . .	54
Lampiran 5	Source Code Cleansing Data . . . . .	55
Lampiran 6	Source Code Case Folding Data . . . . .	55
Lampiran 7	Source Code Tokenization Data . . . . .	56
Lampiran 8	Source Code Normalization Data . . . . .	56
Lampiran 9	Source Code Cek 100 Kata Indikasi Singkatan Terbanyak . . . . .	57
Lampiran 10	Source Code Normalization Manual Data . . . . .	58
Lampiran 11	Source Code Stopword Removal . . . . .	58
Lampiran 12	Source Code Stemming . . . . .	59
Lampiran 13	Source Code Data Labeling Bagian 1 . . . . .	59
Lampiran 14	Source Code Data Labeling Bagian 2 . . . . .	60
Lampiran 15	Source Code Data Labeling Bagian 3 . . . . .	61
Lampiran 16	Source Code Data Labeling Bagian 4 . . . . .	62
Lampiran 17	Source Code Data Labeling Bagian 5 . . . . .	62
Lampiran 18	Source Code Visualisasi Pie Chart . . . . .	63
Lampiran 19	Source Code Visualisasi WordCloud . . . . .	63
Lampiran 20	Source Code TF-IDF, Chi Square, Split Data, dan GridSearchCV . . . . .	64
Lampiran 21	Source Code Modeling SVM . . . . .	65
Lampiran 22	Source Code Prediksi Teks Baru . . . . .	66

