

**ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMINDAHAN IBU  
KOTA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR  
MACHINE DAN SELEKSI FITUR CHI SQUARE**



**SKRIPSI**

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

**Klaudius Carol Illona Banu**

**00000050432**

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

**2025**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMINDAHAN IBU  
KOTA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR  
MACHINE DAN SELEKSI FITUR CHI SQUARE**



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

**Klaudius Carol Illona Banu**

**00000050432**

**UNIVERSITAS  
MULTIMEDIA  
NUSANTARA**

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

**TANGERANG**

**2025**

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Klaudius Carol Illona Banu  
Nomor Induk Mahasiswa : 00000050432  
Program Studi : Informatika

Skripsi dengan judul:

**Analisis Sentimen Masyarakat Terhadap Pemandangan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine dan Seleksi Fitur Chi Square**

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis 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 maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 03 Januari 2025



(Klaudius Carol Illona Banu)

## HALAMAN PENGESAHAN

Skripsi dengan judul

### **ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMINDAHAN IBU KOTA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN SELEKSI FITUR CHI SQUARE**

oleh

Nama : Klaudius Carol Illona Banu  
NIM : 00000050432  
Program Studi : Informatika  
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Rabu, 08 Januari 2025  
Pukul 08.00 s/d 11.00 dan dinyatakan

**LULUS**

Dengan susunan penguji sebagai berikut

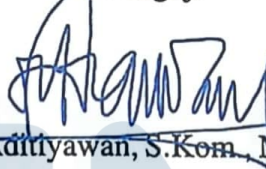
Ketua Sidang



(Dennis Gunawan, S.Kom., M.Sc.)

NIDN: 0320059001

Penguji



(Adityawan, S.Kom., M.Si.)

NIDN: 08994550022

Pembimbing



(Dr.Ir. P.M. Winarno, M.Kom.)

NIDN: 0330106002

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc.)

NIDN: 0315109103

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

Yang bertanda tangan di bawah ini:

Nama : Klaudius Carol Illona Banu  
NIM : 00000050432  
Program Studi : Informatika  
Jenjang : S1  
Judul Karya Ilmiah : Analisis Sentimen Masyarakat Terhadap Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine dan Seleksi Fitur Chi Square

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) \*\*.
- Lainnya, pilih salah satu:
  - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
  - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 03 Januari 2025

Yang menyatakan



Klaudius Carol Illona Banu

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

## Halaman Persembahan / Motto

"Success is not final, failure is not fatal: it is the courage to continue that counts."

Winston Churchill

UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## KATA PENGANTAR

Dengan penuh rasa syukur atas berkat dan karunia Tuhan Yang Maha Esa, penulisan laporan ini yang berjudul Analisis Sentimen Masyarakat Terhadap Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine dan Seleksi Fitur Chi Square berhasil diselesaikan. Laporan ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika di Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa penyelesaian skripsi ini tidak akan mungkin tercapai tanpa dukungan dan bimbingan dari berbagai pihak, sejak masa perkuliahan hingga tahap penyusunan skripsi ini. Oleh karena itu, saya ingin menyampaikan rasa terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Dr.Ir. P.M. Winarno, M.Kom., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Kepada orang tua, keluarga, dan teman-teman saya yang telah memberikan dukungan, baik secara material maupun moral, sehingga saya dapat menyelesaikan tesis ini.

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

Tangerang, 03 Januari 2025



Klaudius Carol Illona Banu

# ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMINDAHAN IBU KOTA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN SELEKSI FITUR CHI SQUARE

Klaudius Carol Illona Banu

## ABSTRAK

Penelitian ini dilakukan untuk menganalisis sentimen masyarakat mengenai pemindahan Ibu Kota Negara Indonesia melalui komentar pada platform YouTube, dengan menggunakan algoritma Support Vector Machine (SVM) dengan dan tanpa seleksi fitur Chi-Square. Algoritma SVM digunakan untuk melakukan prediksi dalam menganalisis sentimen dari komentar-komentar masyarakat terkait pemindahan Ibu Kota Negara. Data yang digunakan dalam penelitian ini berasal dari komentar pada empat video YouTube yang membahas tentang pemindahan Ibu Kota. Hasil penelitian menunjukkan bahwa penggunaan algoritma SVM tanpa seleksi fitur Chi-Square menghasilkan akurasi rata-rata sebesar 87% dengan tingkat presisi 88%, recall 85%, dan skor F1 sebesar 86%. Sementara itu, implementasi SVM dengan seleksi fitur Chi-Square menunjukkan peningkatan performa dengan akurasi mencapai 98% dengan presisi 98%, recall 98%, dan skor F1 sebesar 98%. Analisis sentimen mengungkapkan bahwa mayoritas komentar masyarakat menunjukkan sentimen negatif dengan persentase 59,05%, sementara sentimen positif sebesar 40,95%. Penelitian ini juga mengungkapkan bahwa penggunaan seleksi fitur Chi-Square dapat meningkatkan efisiensi komputasi dengan mengurangi dimensi fitur sebesar 40,05% tanpa mengorbankan akurasi model. Hasil ini menunjukkan bahwa masyarakat cenderung memiliki pandangan negatif terhadap pemindahan Ibu Kota Negara dalam platform media sosial *YouTube* per Oktober 2024, dan penggunaan seleksi fitur Chi-Square terbukti efektif dalam meningkatkan performa analisis sentimen.

**Kata kunci:** Analisis Sentimen, Komentar YouTube, Pemindahan Ibu Kota Negara, Seleksi fitur Chi-Square, Support Vector Machine.

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA



***PUBLIC SENTIMENT ANALYSIS ON INDONESIA'S CAPITAL CITY  
RELOCATION USING SUPPORT VECTOR MACHINE ALGORITHM AND  
CHI-SQUARE FEATURE SELECTION***

Klaudius Carol Illona Banu

***ABSTRACT***

*This research was conducted to analyze public sentiment regarding the relocation of Indonesia's Capital City through comments on the YouTube platform, using the Support Vector Machine (SVM) algorithm with and without Chi-Square feature selection. The SVM algorithm was employed to perform predictions in analyzing sentiment from public comments related to the capital city relocation. The data used in this research was derived from comments on four YouTube videos discussing the capital city relocation. The results showed that the implementation of the SVM algorithm without Chi-Square feature selection achieved an average accuracy of 87% with a precision rate of 88%, recall of 85%, and an F1 score of 86%. Meanwhile, the implementation of SVM with Chi-Square feature selection demonstrated improved performance with accuracy reaching 98%, precision of 98%, recall of 98%, and an F1 score of 98%. Sentiment analysis revealed that the majority of public comments exhibited negative sentiment with a percentage of 59,05%, while positive sentiment accounted for 40,95%. This research also revealed that the use of Chi-Square feature selection can enhance computational efficiency by reducing feature dimensions by 40,05% without compromising model accuracy. These results indicate that the public tends to have a negative view towards the capital city relocation on the YouTube social media platform as of October 2024, and the use of Chi-Square feature selection proved effective in improving sentiment analysis performance.*

***Keywords:*** *Capital City relocation, Chi-Square feature selection, Sentiment Analysis, Support Vector Machine, YouTube Comments.*

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

## 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 TABEL . . . . .	xi
DAFTAR GAMBAR . . . . .	xii
DAFTAR KODE . . . . .	xiii
DAFTAR RUMUS . . . . .	xiv
DAFTAR LAMPIRAN . . . . .	xv
BAB 1 PENDAHULUAN . . . . .	1
1.1 Latar Belakang Masalah . . . . .	1
1.2 Rumusan Masalah . . . . .	2
1.3 Batasan Permasalahan . . . . .	3
1.4 Tujuan Penelitian . . . . .	3
1.5 Manfaat Penelitian . . . . .	4
1.6 Sistematika Penulisan . . . . .	5
BAB 2 LANDASAN TEORI . . . . .	6
2.1 Analisis Sentimen . . . . .	6
2.2 Klasifikasi . . . . .	6
2.3 <i>Support Vector Machine</i> . . . . .	6
2.3.1 Pemilihan Data <i>Training</i> . . . . .	6
2.3.2 Penentuan Hyperplane . . . . .	7
2.3.3 Penerapan Kernel . . . . .	7
2.3.4 Evaluasi Model . . . . .	7
2.4 Seleksi Fitur <i>Chi Square</i> . . . . .	8
2.4.1 Implementasi . . . . .	8
2.5 Text-Preprocessing . . . . .	9
BAB 3 METODOLOGI PENELITIAN . . . . .	11
3.1 Alur Penelitian . . . . .	11
3.1.1 Studi Literatur . . . . .	12
3.1.2 Pengumpulan Data . . . . .	12
3.1.3 Perancangan Sistem dan Implementasi Algoritma . . . . .	12
3.1.4 Pengujian dan Evaluasi . . . . .	13
3.1.5 Laporan Hasil Analisis Dan Visualisasi Data . . . . .	13
3.2 Rancangan Sistem . . . . .	13
3.2.1 Pengumpulan Data . . . . .	14
3.2.2 <i>Text Pre-processing</i> . . . . .	14
3.2.3 <i>Term Frequency Inverse Document Frequency (TF-IDF)</i> . . . . .	15
3.2.4 Train-Test Split Data . . . . .	15
3.2.5 Penerapan <i>Chi-Square</i> Feature Selection . . . . .	16
3.2.6 Penerapan Model <i>Support Vector Machine</i> . . . . .	16
3.2.7 Pengujian Model dan Evaluasi . . . . .	16
3.2.8 Visualisasi Hasil . . . . .	16

BAB 4	HASIL DAN DISKUSI . . . . .	17
4.1	Spesifikasi Sistem . . . . .	17
4.2	Rancangan Sistem . . . . .	17
4.2.1	Pengumpulan Data . . . . .	18
4.2.2	Text PreProcessing . . . . .	21
4.2.3	Data Labeling . . . . .	28
4.2.4	TF-IDF Training . . . . .	31
4.2.5	<i>Training-Split</i> Data . . . . .	32
4.2.6	Penerapan Algoritma <i>Support Vector Machine</i> . . . . .	32
4.2.7	Peneraapn TF-IDF untuk <i>Chi-Square</i> . . . . .	33
4.2.8	Train-Test Split untuk SVM dan <i>Chi-Square</i> . . . . .	34
4.2.9	Evaluasi Perbandingan Model . . . . .	34
4.2.10	Analisis Tabel Kontingensi . . . . .	40
4.2.11	Visualisasi <i>Word Cloud</i> . . . . .	41
BAB 5	SIMPULAN DAN SARAN . . . . .	43
5.1	Simpulan . . . . .	43
5.2	Saran . . . . .	43
DAFTAR PUSTAKA	. . . . .	45



## DAFTAR TABEL

Tabel 2.1	Tabel contoh kontingensi <i>Chi-Square</i> untuk setiap kata . . .	9
Tabel 4.1	data jumlah <i>views</i> dan komentar pada video . . . . .	18
Tabel 4.2	Data komentar hasil <i>scraping</i> . . . . .	21
Tabel 4.3	Tabel Data Text Cleansing . . . . .	23
Tabel 4.4	Tabel Data Hasil Tokenisasi . . . . .	24
Tabel 4.5	Tabel Data hasil proses stemming . . . . .	25
Tabel 4.6	Tabel Classification Report SVM . . . . .	35
Tabel 4.7	Tabel Classification Report SVM + <i>Chi-Square</i> . . . . .	36
Tabel 4.8	Tabel Kontingensi Hasil Analisis Sentimen . . . . .	40



## DAFTAR GAMBAR

Gambar 3.1	Diagram alur mengenai langkah-langkah penelitian yang dilakukan. . . . .	11
Gambar 3.2	Diagram alur mengenai rancangan sistem yang dilakukan dalam penelitian . . . . .	14
Gambar 3.3	<i>flowchart</i> Tahapan <i>Text Pre-processing</i> . . . . .	15
Gambar 4.1	Grafik Perbandingan jumlah views dan Komentar pada video	19
Gambar 4.2	Contoh hasil dari custom stopwords-removal data . . . . .	24
Gambar 4.3	Contoh hasil <i>stemming data</i> . . . . .	26
Gambar 4.4	Contoh hasil dari custom stopwords-removal data . . . . .	28
Gambar 4.5	Contoh hasil <i>labeling data</i> . . . . .	31
Gambar 4.6	Perbandingan Confusion Matrix Original SVM dan <i>Chi-Square SVM</i> . . . . .	35
Gambar 4.7	Hasil Data Evaluasi Video Pertama . . . . .	37
Gambar 4.8	Hasil Data Evaluasi Video Kedua . . . . .	38
Gambar 4.9	Hasil Data Evaluasi Video Ketiga . . . . .	39
Gambar 4.10	Hasil Data Evaluasi Video Keempat . . . . .	40
Gambar 4.11	Hasil Data Word Cloud . . . . .	41



## DAFTAR KODE

Kode 4.1	Kode Data Komen Youtube . . . . .	19
Kode 4.2	Kode Data Scraping . . . . .	20
Kode 4.3	Kode Teks Pre-processing . . . . .	22
Kode 4.4	Kode Tokenisasi . . . . .	23
Kode 4.5	Kode Stemming . . . . .	25
Kode 4.6	Kode Normalisasi Data . . . . .	26
Kode 4.7	Kode Stopwords Removal Data . . . . .	27
Kode 4.8	Kode Labeling Dictionary . . . . .	28
Kode 4.9	Kode Labeling Data . . . . .	29
Kode 4.10	Kode TF-IDF Training Data . . . . .	31
Kode 4.11	Kode Traning-Split Data . . . . .	32
Kode 4.12	Kode penerapan SVM . . . . .	33
Kode 4.13	TF-IDF untuk <i>Chi-Square</i> . . . . .	33
Kode 4.14	Train-Test Split untuk <i>Chi-Square</i> . . . . .	34



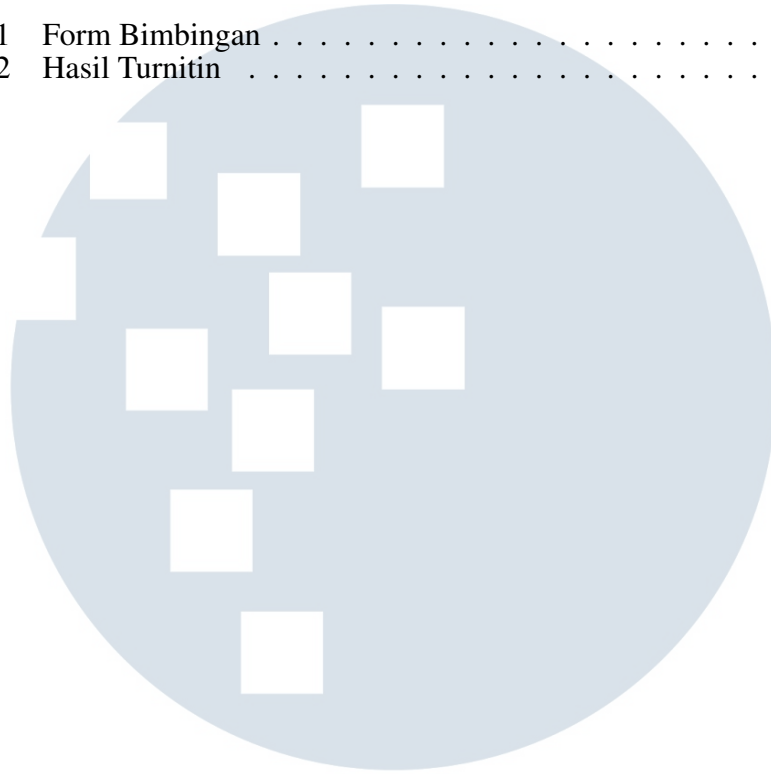
## DAFTAR RUMUS

Rumus 2.1 Rumus hyperplane Support Vector Machine .....	7
Rumus 2.2 Rumus accuracy dalam evaluasi model .....	7
Rumus 2.3 Rumus precision dalam evaluasi model .....	7
Rumus 2.4 Rumus recall dalam evaluasi model .....	7
Rumus 2.5 Rumus dasar Chi-Square .....	8
Rumus 2.6 Rumus matriks observasi Chi-Square .....	8
Rumus 2.7 Rumus nilai harapan Chi-Square .....	8
Rumus 2.8 Rumus Chi-Square untuk setiap kata .....	9



## DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan . . . . .	47
Lampiran 2	Hasil Turnitin . . . . .	48



UMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA