

**ANALISIS SENTIMEN LEMBAGA SURVEI
PADA PEMILIHAN PRESIDEN 2024
MENGUNAKAN *SUPPORT VECTOR MACHINE***

Skripsi

**Disusun untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Statistika**

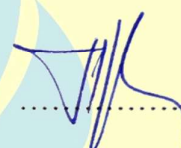


**Alisya Rahmawanti
1314620025**

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA
2026**

**LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI
ANALISIS SENTIMEN LEMBAGA SURVEI
PADA PEMILIHAN PRESIDEN 2024
MENGUNAKAN *SUPPORT VECTOR MACHINE***

Nama : Alisya Rahmawanti
NIM : 1314620025

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan :	<u>Dr. Hadi Nasbey, S.Pd., M.Si.</u> NIP. 197909162005011004		14/8/2026
Wakil Penanggung Jawab			
Wakil Dekan I :	<u>Dr. Meiliasari, S.Pd., M.Sc.</u> NIP. 197905042009122002		14/8/2026
Ketua Penguji :	<u>Prof. Dr. Dian Handayani, M.Si.</u> NIP. 197404151998032001		5/8/2026
Sekretaris :	<u>Dra. Widianti Rahayu, M.Si.</u> NIP. 196611032001122001		6/8/2026
Anggota			
Pembimbing I :	<u>Prof. Dr. Ir. Bagus Sumargo, M.Si.</u> NIP. 196309221986011001		28/7/2026
Pembimbing II :	<u>Qorry Meidianingsih, M.Si.</u> NIP. 199105192019032019		5/8/2026
Penguji Ahli :	<u>Faroh Ladayya, M.Si.</u> NIP. 199401282020122018		05/08/2026

Dinyatakan lulus ujian skripsi tanggal: 10 Februari 2026

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul **“Analisis Sentimen Lembaga Survei pada Pemilihan Presiden 2024 Menggunakan *Support Vector Machine*”** yang disusun sebagai syarat untuk memperoleh gelar Sarjana Statistika dari Program Studi Statistika Universitas Negeri Jakarta adalah karya ilmiah saya dengan arahan dari dosen pembimbing.

Sumber informasi yang diperoleh dari penulis lain yang telah dipublikasikan dan disebutkan dalam teks skripsi ini telah dicantumkan dalam Daftar Pustaka sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Jika di kemudian hari ditemukan Sebagian besar skripsi ini asli bukan hasil karya saya sendiri dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sanding dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Jakarta, 28 Januari 2026



Alisya Rahmawanti



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
PERPUSTAKAAN DAN KEARSIPAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Alisya Rahmawanti
NIM : 1314620025
Fakultas/Prodi : FMIPA/Statistika
Alamat email : alisyarahma1811@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan dan Kearsipan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Analisis Sentimen Lembaga Survei pada Pemilihan Presiden 2024 Menggunakan
Support Vector Machine

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 12 Agustus 2026

Penulis

(Alisya Rahmawanti)

ABSTRAK

ALISYA RAHMAWANTI. Analisis Sentimen Lembaga Survei pada Pemilihan Presiden 2024 Menggunakan *Support Vector Machine*. Skripsi, Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Januari 2026.

Media sosial menjadi salah satu sarana utama masyarakat dalam menyampaikan opini terhadap berbagai isu politik, termasuk terkait lembaga survei pada pemilihan presiden tahun 2024. Opini yang berkembang di media sosial dapat mencerminkan persepsi publik yang beragam, sehingga diperlukan analisis sentimen untuk mengelompokkan opini tersebut secara sistematis. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap lembaga survei Lembaga Survei Indonesia (LSI), *Center for Strategic and International Studies* (CSIS), dan Poltracking Indonesia berdasarkan data *tweet* dari media sosial X. Metode yang digunakan dalam penelitian ini adalah *text mining* dengan tahapan *text preprocessing*, pelabelan sentimen menggunakan pendekatan *lexicon-based*, serta ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF). Proses klasifikasi sentimen dilakukan menggunakan metode *Support Vector Machine* (SVM) *multi-class* dengan pendekatan *One Against All*. Untuk memperoleh parameter terbaik, digunakan *GridSearchCV*. Hasil penelitian menunjukkan bahwa sentimen positif mendominasi pembahasan terhadap ketiga lembaga survei, diikuti oleh sentimen negatif dan netral. SVM dengan kernel *Radial Basis Function* (RBF) menghasilkan performa yang lebih baik dibandingkan kernel linear berdasarkan nilai akurasi, presisi, *recall*, dan *F-score*. Dengan demikian, metode SVM dengan kernel RBF dinilai efektif dalam mengklasifikasikan sentimen masyarakat terhadap lembaga survei di media sosial X.

Kata kunci: Analisis sentimen, *Text mining*, TF-IDF, *Support Vector Machine*, Lembaga survei, Media sosial

ABSTRACT

ALISYA RAHMAWANTI. Sentiment Analysis of Survey Institutions in the 2024 Presidential Election Using Support Vector Machine. Thesis, Statistics Study Program, Faculty of Mathematics and Natural Sciences, State University of Jakarta. January 2026.

Social media has become one of the main platforms for the public to express opinions on various political issues, including survey institutions in 2024 presidential elections. Opinions expressed on social media reflect diverse public perceptions, making sentiment analysis necessary to systematically classify such opinions. This study aims to classify public sentiment toward survey institutions, namely Lembaga Survei Indonesia (LSI), Center for Strategic and International Studies (CSIS), and Poltracking Indonesia, based on tweet data from social media X. The method used in this study is text mining, which includes text preprocessing, sentiment labeling using a lexicon-based approach, and feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF). Sentiment classification is performed using a multi-class Support Vector Machine (SVM) with a One Against All approach. To obtain the optimal parameters, GridSearchCV is applied. The results show that positive sentiment dominates discussions related to the three survey institutions, followed by negative and neutral sentiments. Based on model performance evaluation, SVM with a Radial Basis Function (RBF) kernel achieves better performance than the linear kernel in terms of accuracy, precision, recall, and F-score. Therefore, the SVM method with an RBF kernel is considered effective for classifying public sentiment toward survey institutions on social media X.

Keywords: Sentiment analysis, Text mining, TF-IDF, Support Vector Machine, Survey institutions, Social media

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa, karena atas rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Lembaga Survei pada Pemilihan Presiden 2024 Menggunakan *Support Vector Machine*” dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa penyelesaian penelitian tidak terlepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT atas segala rahmat dan kemudahan yang diberikan sehingga penulis dapat menyelesaikan skripsi ini.
2. Kedua orang tua penulis, Bapak Ali Fatkhi dan Ibu Syafa'atun, Adik (Putri) dan Kakak (Mas Olan) serta keluarga besar dari penulis yang telah senantiasa mendukung dan memberikan semangat kepada penulis.
3. Ibu Prof. Dr. Dian Handayani, M.Si. selaku koordinator Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.
4. Bapak Prof. Dr. Ir. Bagus Sumargo, M.Si. dan Ibu Qorry Meidianingsih, M.Si., selaku dosen pembimbing I dan pembimbing II yang telah meluangkan waktu, tenaga dan pikiran untuk memberikan arahan, bimbingan, kritik dan saran yang sangat berarti kepada penulis selama proses penyusunan skripsi sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
5. Bapak/Ibu dosen dan seluruh civitas akademika yang telah memberikan ilmu dan pengalaman selama masa perkuliahan serta staf administrasi di lingkup Program Studi Statistika dan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Jakarta atas bantuannya selama pelaksanaan Pendidikan hingga penyusunan skripsi.

6. Teman-teman penulis selama merantau di Jakarta baik yang dari UNJ seperti Najwa, Ade, Fareka, Fatimah, Nathaya, Esty, Alif, Diva, Ishmah, Indah, Debo dan Kak Gadis serta teman-teman penulis dari instansi magang dan tempat kerja penulis seperti, Kak Ajis, Kak Rehan, Kak Zalfa, Dwi, Mba Olla, Mba Nisa, Mba Ido, Sarah, Putra, Ratu, Nina, Kak Kiara, Kak Muti, Kak Nindy, Kak Seha, Janna dan teman berbagi cerita yaitu Joli serta tak lupa pula teman-teman grup di SMA penulis yaitu Diah, Hayigha, Wanda, Ira, Ina, Selvi dan Adi yang senantiasa memberikan semangat dan motivasi kepada penulis.
7. Rekan-rekan mahasiswa/i Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta atas kekeluargaan dan pertemanan yang diberikan selama ini.
8. Kerabat dan orang-orang terdekat yang senantiasa memberikan doa, dukungan moral, dan motivasi kepada penulis.
9. Seluruh pihak yang telah membantu penulis, baik secara langsung maupun tidak langsung, yang tidak dapat disebutkan satu per satu.
10. Selain itu, penulis juga ingin berterima kasih kepada diri sendiri yang telah berusaha, selalu bangkit dari kesulitan, tetap melangkah dan tidak menyerah hingga tahap akhir penyusunan skripsi ini sehingga penulis dapat menyelesaikan studi di Program Studi Statistika Universitas Negeri Jakarta.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan demi penyempurnaan penelitian ini di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat dan menjadi referensi bagi penelitian selanjutnya.

Jakarta, 28 Januari 2026



Alisya Rahmawanti

DAFTAR ISI

LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI.....	i
LEMBAR PERNYATAAN KEASLIAN SKRIPSI.....	ii
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iii
ABSTRAK.....	iii
ABSTRACT.....	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	7
1.3 Batasan Masalah.....	7
1.4 Tujuan Penelitian	8
1.5 Manfaat Penelitian	8
BAB II TINJAUAN PUSTAKA.....	9
2.1 Analisis Sentimen	9
2.1.1 Pendekatan <i>Lexicon-Based</i> (InSet)	10
2.2 <i>Text Mining</i>	10
2.2.1 <i>Text Preprocessing</i>	11
2.3 Pembobotan Kata Menggunakan <i>Term Frequency-Inverse Document Frequency</i> (TF-IDF).....	13
2.3.1 Contoh Perhitungan <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	15

2.4	<i>Support Vector Machine (SVM)</i>	19
2.4.1	SVM pada <i>Linearly Separable Data</i>	20
2.4.2	SVM pada <i>Nonlinearly Separable Data</i>	25
2.4.3	<i>Multi Class SVM</i>	30
2.5	<i>K-fold Cross Validation</i>	35
2.6	<i>Grid Search</i>	36
2.7	Performa Metode Klasifikasi	37
2.8	<i>Quick Count</i> atau Hitung Cepat	39
2.9	Pemilihan Presiden.....	40
2.10	Lembaga Survei	40
BAB III METODOLOGI PENELITIAN.....		42
3.1	Data	42
3.2	Prosedur dan Analisis Data	42
BAB IV HASIL DAN PEMBAHASAN		46
4.1	<i>Crawling Data Lembaga Survei</i>	46
4.2	<i>Filtering dan Labelling Data</i>	52
4.3	<i>Text Preprocessing</i>	54
4.4	<i>Term Frequency–Inverse Document Frequency (TF-IDF)</i>	56
4.5	<i>Splitting Data dan K-Fold Cross Validation</i>	57
4.6	Pemodelan dengan <i>Multi Class SVM</i>	58
4.7	Pembentukan <i>Hyperplane</i>	59
4.8	Margin Maksimum.....	60
4.9	Fungsi Keputusan (<i>Decision Function</i>)	60
4.10	Prediksi dan Evaluasi Model.....	62
4.11	Hasil dan Interpretasi	65
BAB V KESIMPULAN DAN SARAN.....		70

5.1	Kesimpulan	70
5.2	Saran.....	70
DAFTAR PUSTAKA		72
LAMPIRAN.....		76
RIWAYAT HIDUP.....		79



DAFTAR TABEL

Tabel 2.1 Contoh Hasil Proses <i>Text Preprocessing</i>	13
Tabel 2.2 Contoh Data Dokumen.....	15
Tabel 2.3 Contoh Hasil Perhitungan TF.....	16
Tabel 2.4 Contoh Hasil Perhitungan IDF	17
Tabel 2.5 Contoh Hasil Perhitungan TF-IDF	18
Tabel 2.6 Contoh Empat SVM Biner dengan Metode <i>One-Against-All</i>	31
Tabel 2.7 Contoh Enam SVM Biner dengan Metode <i>One-Against-One</i>	32
Tabel 2.8 Contoh Enam SVM Biner dengan Metode DAGSVM	34
Tabel 2.9 <i>Confusion Matrix</i> 2×2	37
Tabel 2.10 <i>Confusion Matrix</i> 3×3	38
Tabel 2.11 Kriteria Nilai <i>F-score</i>	39
Tabel 4.1 Konfigurasi yang Digunakan untuk Crawling	46
Tabel 4.2 Akun dengan Jumlah Engagement Terbanyak.....	48
Tabel 4.3 Contoh Hasil Filtering Data	53
Tabel 4.4 Contoh Hasil Labelling Data.....	54
Tabel 4.5 Hasil Proses Text Preprocessing	55
Tabel 4.6 TF-IDF Term Penting.....	56
Tabel 4.7 Kombinasi Parameter dengan Rata-rata Skor tertinggi.....	58
Tabel 4.8 Hasil Pembentukan Hyperplane.....	59
Tabel 4.9 Confusion Matrix Model dengan Kernel RBF	62
Tabel 4.10 Confusion Matrix Model dengan Kernel Linear	63
Tabel 4.11 Evaluasi Performa Model SVM.....	64

DAFTAR GAMBAR

Gambar 1.1 10 Negara dengan Pengguna X Terbanyak di Dunia (Juli 2023)	2
Gambar 1.2 Keakuratan Lembaga Survei dalam Melakukan <i>Quick Count</i> pada Pilpres 2019 dan 2024.....	4
Gambar 2.1 Bentuk <i>Hyperplane</i> : (a) <i>Alternative Hyperplane</i> (b) <i>Optimal Hyperplane</i> pada SVM.....	19
Gambar 2.2 Perbedaan SVM dari <i>Linearly Separable Data</i> dan <i>Nonlinearly Separable Data</i>	25
Gambar 2.3 <i>Soft Margin Hyperplane</i>	26
Gambar 2.4 Transformasi dari <i>Input Space</i> ke <i>Feature Space</i>	27
Gambar 2.5 Ilustrasi <i>Hyperplane</i> pada <i>Linear Kernel</i>	29
Gambar 2.6 Ilustrasi <i>Hyperplane</i> pada <i>Polynomial Kernel</i>	29
Gambar 2.7 Ilustrasi <i>Hyperplane</i> pada RBF Kernel.....	30
Gambar 2.8 Contoh Klasifikasi dengan Metode <i>One-Against-All</i>	32
Gambar 2.9 Contoh Klasifikasi dengan Metode <i>One-Against-One</i>	33
Gambar 2.10 Contoh Klasifikasi dengan Metode DAGSVM	34
Gambar 2.11 Gambaran <i>10-fold cross validation</i>	35
Gambar 2.12 <i>Grid Search</i>	36
Gambar 3.1 Diagram Alir Penelitian.....	45
Gambar 4.1 Sebaran Tweet Berdasarkan Lembaga Survei	47
Gambar 4.2 Jumlah Unggahan Harian Tiga Lembaga Survei.....	47
Gambar 4.3 Total Engagement Harian Tiga Lembaga Survei	48
Gambar 4.4 Word Cloud Tweet Lembaga Survei Indonesia (LSI).....	49
Gambar 4.5 (a) Unggahan dan (b) Engagement Harian LSI	49
Gambar 4.6 Word Cloud Tweet CSIS	50
Gambar 4.7 (a) Unggahan dan (b) Engagement Harian CSIS.....	51
Gambar 4.8 Word Cloud Tweet Poltracking	52
Gambar 4.9 (a) Unggahan dan (b) Engagement Harian Poltracking	52
Gambar 4.10 Sebaran Tweet Berdasarkan Kategori Sentimen	54
Gambar 4.11 Sebaran Tweet Berdasarkan Kategori Sentimen	65
Gambar 4.12 Sebaran Tweet LSI Berdasarkan Kategori Sentimen	66

Gambar 4. 13 (a) Word Cloud Sentimen Positif LSI (b) Word Cloud Sentimen Negatif LSI.....	66
Gambar 4.14 Sebaran Tweet CSIS Berdasarkan Kategori Sentimen.....	67
Gambar 4.15 (a) Word Cloud Sentimen Positif CSIS (b) Word Cloud Sentimen Negatif CSIS	67
Gambar 4.16 Sebaran Tweet Poltracking Berdasarkan Kategori Sentimen.....	68
Gambar 4.17 (a) Word Cloud Sentimen Positif Poltracking (b) Word Cloud Sentimen Negatif Poltracking	68

