

**PREDIKSI ELEKTABILITAS CALON BERDASARKAN
SENTIMEN DENGAN SUPPORT VECTOR MACHINE
(STUDI KASUS: PEMILIHAN PRESIDEN
DALAM PEMILU 2024)**



SKRIPSI

**Necoles
00000042589**

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

**PREDIKSI ELEKTABILITAS CALON BERDASARKAN
SENTIMEN DENGAN SUPPORT VECTOR MACHINE
(STUDI KASUS: PEMILIHAN PRESIDEN
DALAM PEMILU 2024)**



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

Necoless
00000042589

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

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Necoless
NIM : 00000042589
Program Studi : Informatika

Menyatakan dengan sesungguhnya bahwa Tesis/Skripsi/Tugas Akhir/Laporan Magang/MBKM saya yang berjudul:

Prediksi Elektabilitas Calon berdasarkan Sentimen dengan Support Vector Machine (Studi Kasus: Pemilihan Presiden dalam Pemilu 2024)

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain; Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan karya ilmiah, saya bersedia menerima konsekuensi untuk dinyatakan TIDAK LULUS. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 18 Juni 2024

UNIVERSIT
MULTIMEDIA
NUSANTARA



(Necoless)

HALAMAN PENGESAHAN

Skripsi dengan judul

**PREDIKSI ELEKTABILITAS CALON BERDASARKAN
SENTIMEN DENGAN SUPPORT VECTOR MACHINE
(STUDI KASUS: PEMILIHAN PRESIDEN
DALAM PEMILU 2024)**

oleh

Nama : Necoless
NIM : 00000042589
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika


Telah diujikan pada hari Selasa, 4 Juni 2024

Pukul 08.00 s/d 10.00 dan dinyatakan


LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang


(Suwito Pomalingo, S.Kom., M.Kom)
NIDN: 0911098201


Penguji


(Dennis Gunawan, S.Kom., M.Sc.)
NIDN: 0320059001

Pembimbing


(Moeljono Widjaja, B.Sc., M.Sc., Ph.D)
NIDN: 0311106903

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 : Necoless
NIM : 00000042589
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, 18 Juni 2024

Yang menyatakan

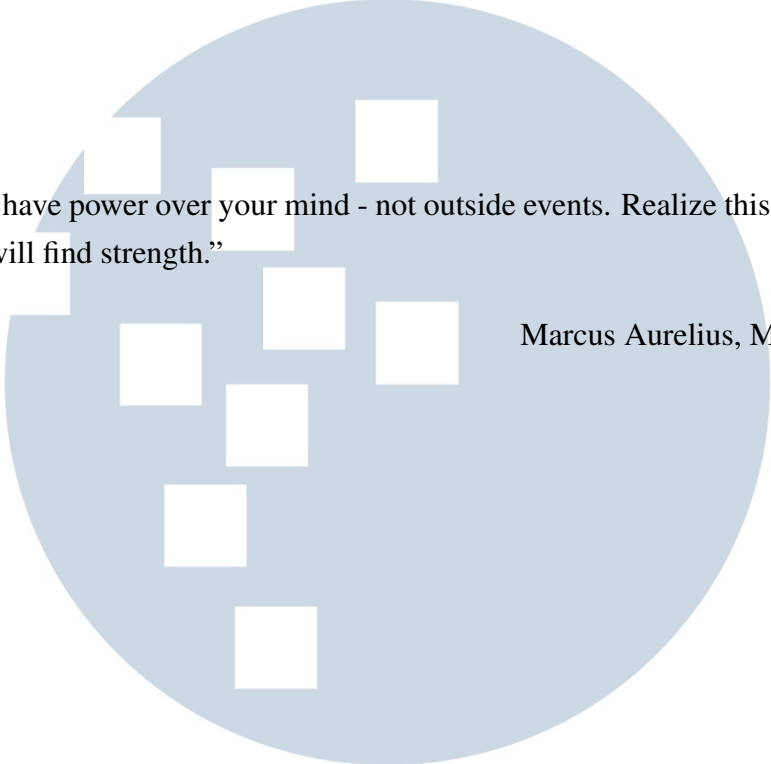


Necoless

U M 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

** 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



”You have power over your mind - not outside events. Realize this, and you will find strength.”

Marcus Aurelius, Meditations

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: Prediksi Elektabilitas Calon berdasarkan Sentimen dengan Support Vector Machine (Studi Kasus: Pemilihan Presiden dalam Pemilu 2024) 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. Bapak Dr. Ninok Leksono, 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 Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Pjs. Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Moeljono Widjaja, B.Sc., M.Sc., Ph.D, sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Orang Tua, teman dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

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

Tangerang, 18 Juni 2024



Necoless

**PREDIKSI ELEKTABILITAS CALON BERDASARKAN
SENTIMEN DENGAN SUPPORT VECTOR MACHINE
(STUDI KASUS: PEMILIHAN PRESIDEN
DALAM PEMILU 2024)**

Necoless

ABSTRAK

Indonesia merupakan negara demokratis yang melaksanakan pemilu. Salah satu pemilihan umum adalah pemilihan presiden pada pemilu 2024. Mengetahui elektabilitas adalah hal penting untuk calon maupun untuk masyarakat. Salah satu cara untuk mengetahui elektabilitas dalam masa pemilu adalah survei. Namun, survei dikatakan memakan banyak sumber daya. Oleh karena itu, dilakukan penelitian untuk memprediksi elektabilitas berdasarkan sentimen dari sosial media. Salah satu algoritma yang dapat digunakan adalah SVM. Model SVM tersebut dijadikan dasar untuk melakukan prediksi dengan tabulasi sederhana dan berdasarkan konsep pemrosesan informasi secara heuristik. Setelah dilakukan prediksi berdasarkan model tersebut, didapati hasil prediksi sangat berbeda jauh dalam hal persentase. Namun, urutan pemenang dari hasil prediksi sudah sesuai dengan hasil survei. Oleh karena itu, dapat dikatakan prediksi tidak dapat menggantikan survei tradisional sepenuhnya. Namun, prediksi bisa menjadi alternatif ataupun membantu survei tradisional.

Kata kunci: *Machine learning*, pemilu, prediksi, *sentiment analysis*, SVM

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

***Prediction of Candidate Electability based on Sentiment with Support Vector
Machine (Case Study: 2024 Presidential Elections)***

Necoless

ABSTRACT

Indonesia is a democratic country that holds elections. One of the general elections is the presidential election in 2024. Knowing electability is important for candidates and the public. One way to find out electability during the election period is a survey. However, surveys are said to be resource-intensive. Therefore, research was conducted to predict electability based on sentiment from social media. One algorithm that can be used is SVM. The SVM model is used as the basis for making predictions using simple tabulations and based on the concept of heuristic information processing. After making predictions based on this model, prediction results were obtained which were very different in terms of percentages. However, the order of winners from the predicted results is in accordance with the survey results. Therefore, it can be said that prediction cannot replace traditional monitoring completely. However, predictions can be an alternative or help traditional monitoring.

Keywords: *Election, machine learning, prediction, sentiment analysis, SVM*



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 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	5
BAB 2 LANDASAN TEORI	6
2.1 Pemilu	6
2.2 Natural Language Processing	6
2.3 Sentiment Analysis	6
2.4 Data Preprocessing	7
2.5 TextBlob	8
2.6 VADER	8
2.7 SMOTE	9
2.8 TF-IDF	9
2.9 Support Vector Machine	10
2.10 Hyperparameter	11
2.11 Grid Search	12
2.12 Confusion Matrix	12
2.13 Classification Report	13
2.13.1 Accuracy	14
2.13.2 Recall	14
2.13.3 Precision	14
2.13.4 F1 Score	14
2.14 Information Processing	15
BAB 3 METODOLOGI PENELITIAN	16
3.1 Spesifikasi Sistem	16
3.2 Metodologi Penelitian	16
3.3 Perancangan	17
3.3.1 Garis besar perancangan	17
3.3.2 Teknik pengumpulan data	18
3.3.3 Data preprocessing	18
3.3.4 Sentiment Analysis dengan SVM	19
3.3.5 Prediksi Pemilih	20

3.3.6	Perbandingan	21
BAB 4	HASIL DAN DISKUSI	22
4.1	Implementasi Sistem	22
4.1.1	Fetch Youtube Comments	22
4.1.2	Labelling	23
4.1.3	Preprocessing Data	26
4.1.4	TF-IDF	32
4.1.5	Data Splitting	33
4.1.6	SMOTE	33
4.1.7	Grid Search	34
4.1.8	Pelatihan Model SVM dan Evaluasi Model SVM	35
4.1.9	Confusion Matrix	39
4.1.10	Prediksi Pemilih	43
4.1.11	Perbandingan	48
BAB 5	SIMPULAN DAN SARAN	49
5.1	Simpulan	49
5.2	Saran	49
DAFTAR PUSTAKA		51



DAFTAR TABEL

Tabel 4.1	Hasil labelling	25
Tabel 4.2	Hasil cleaning	29
Tabel 4.3	Hasil normalization	29
Tabel 4.4	Hasil tokenization	30
Tabel 4.5	Hasil stopwords removal	31
Tabel 4.6	Hasil lemmatization	31
Tabel 4.7	Hasil translate ke Bahasa Inggris	32
Tabel 4.8	Perbandingan sebelum dan sesudah SMOTE	34
Tabel 4.9	Hasil grid search	35
Tabel 4.10	Perbandingan accuracy dengan labelling dengan TextBlob	37
Tabel 4.11	Perbandingan precision dengan labelling dengan TextBlob	37
Tabel 4.12	Perbandingan recall dengan labelling dengan TextBlob	38
Tabel 4.13	Perbandingan f1 score dengan labelling dengan TextBlob	38
Tabel 4.14	Hasil prediksi labelling dengan TextBlob	46
Tabel 4.15	Hasil prediksi labelling dengan VADER	47
Tabel 4.16	Hasil prediksi dengan labelling dengan TextBlob	47
Tabel 4.17	Hasil prediksi dengan labelling dengan VADER	48
Tabel 4.18	Perbandingan hasil prediksi dengan survei	48



DAFTAR GAMBAR

Gambar 2.1	Contoh grafik SVM	11
Gambar 2.2	Confusion matrix	13
Gambar 3.1	Flowchart garis besar perancangan	18
Gambar 3.2	Flowchart data preprocess	19
Gambar 4.1	Confusion matrix dengan labelling TextBlob	39
Gambar 4.2	Confusion matrix dengan labelling menggunakan VADER	40
Gambar 4.3	Confusion matrix data Anies labelling TextBlob	41



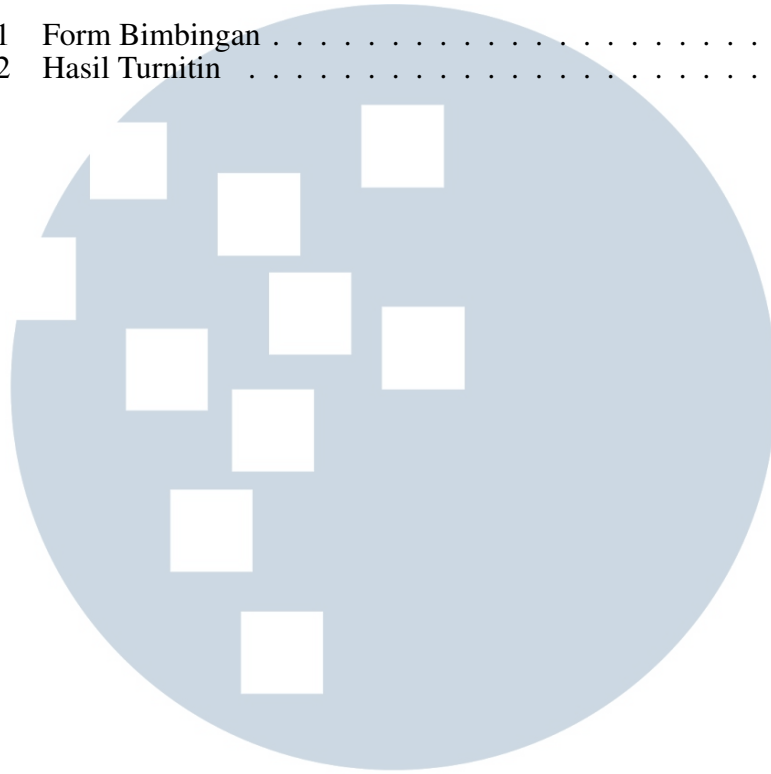
DAFTAR KODE

4.1	Potongan kode fungsi fetch comment	22
4.2	Potongan kode labelling dengan TextBlob	24
4.3	Potongan kode labelling dengan VADER	24
4.4	Potongan kode fungsi cleaning	26
4.5	Potongan kode fungsi normalisasi teks	27
4.6	Potongan kode fungsi tokenization	27
4.7	Potongan kode fungsi stopword removal	28
4.8	Potongan kode fungsi lemmatization	28
4.9	Potongan kode fungsi menerjemah ke bahasa Inggris	32
4.10	Potongan kode fungsi TF-IDF feature extraction	33
4.11	Potongan kode untuk data splitting	33
4.12	Potongan kode fungsi penerapan SMOTE	33
4.13	Potongan kode grid search	34
4.14	Potongan kode untuk pelatihan dan evaluasi model SVM	35
4.15	Potongan kode confusion matrix	39
4.16	Potongan kode label sentiment	43
4.17	Potongan kode penggabungan data	44
4.18	Potongan kode penghitungan sentimen	44
4.19	Potongan kode prediksi pemilih	45



DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan	56
Lampiran 2	Hasil Turnitin	59



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA