

SKRIPSI



**ANALISIS SENTIMEN OPINI PENGGUNA MEDIA SOSIAL X TERHADAP
PEMILIHAN UMUM TAHUN 2024 MENGGUNAKAN METODE *NAIVE BAYES*
DAN *PARTICLE SWARM OPTIMIZATION***

KURNIAWAN DANIL

NIM. 2010511077

**PROGRAM STUDI S1 INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
TAHUN 2024**



**ANALISIS SENTIMEN OPINI PENGGUNA MEDIA SOSIAL X TERHADAP
PEMILIHAN UMUM TAHUN 2024 MENGGUNAKAN METODE *NAIVE BAYES*
DAN *PARTICLE SWARM OPTIMIZATION***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Mendapatkan Gelar
Sarjana Komputer**

KURNIAWAN DANIL

NIM. 2010511077

**PROGRAM STUDI S1 INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
TAHUN 2024**

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Tugas akhir ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Kurniawan Danil

NIM : 2010511077

Tanggal : 11 Juni 2024

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 11 Juni 2024

Yang menyatakan,



(Kurniawan Danil)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Kurniawan Danil

NIM : 2010511077

Fakultas : Ilmu Komputer

Program Studi : S1 Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

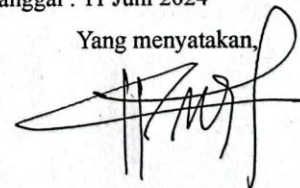
ANALISIS SENTIMEN OPINI PENGGUNA MEDIA SOSIAL X TERHADAP PEMILIHAN UMUM TAHUN 2024 MENGGUNAKAN METODE *NAÏVE BAYES* DAN *PARTICLE SWARM OPTIMIZATION*

Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian Pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 11 Juni 2024

Yang menyatakan,



(Kurniawan Danil)

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Skripsi diajukan oleh :

Nama : Kurniawan Danil

NIM : 2010511077

Program Studi : SI Informatika

Judul Skripsi : Analisis Sentimen Opini Pengguna Media Sosial X Terhadap Pemilihan Umum Tahun 2024 Menggunakan Metode *Naïve Bayes* Dan *Particle Swarm Optimization*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Theresia Wati, S.Kom., MTI.

Penguji 1



Zatin Niqotaini, S.Tr.Kom., M.Kom.

Penguji 2



Musthofa Galih Pradana, M.Kom.

Pembimbing 1



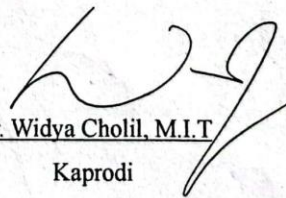
Novi Trisman Hadi, S.Pd., M.Kom.

Pembimbing 2



Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM

Dekan



Dr. Widya Cholil, M.I.T

Kaprodi

Ditetapkan di : Jakarta

Tanggal Ujian : 11 Juni 2024

ABSTRAK

Masyarakat Indonesia telah melaksanakan kegiatan rutin lima tahunan yaitu Pemilihan Umum (Pemilu) yang dilaksanakan pada tahun 2024. Pada media sosial X ramai opini masyarakat mengenai penyelenggaraan kegiatan ini, ada yang beropini positif maupun negatif. Kumpulan opini masyarakat tersebut selanjutnya akan dianalisa berdasarkan sentimennya untuk selanjutnya dapat digunakan dalam pengembangan model. Analisis ini bertujuan untuk mengetahui komposisi sentimen masyarakat terhadap Pemilihan Umum (Pemilu) Tahun 2024 dan mengukur kemampuan model dalam mengklasifikasi data tersebut. Hal ini berguna untuk mengetahui penerimaan masyarakat terhadap penyelenggaraan pemilu dan mengetahui performa model *Naive Bayes* bersama dengan *Particle Swarm Optimization* dalam menangani data opini tersebut. Sebelum pengembangan model, terlebih dahulu data diberi label sentimen, praproses, pembagian data, dan pembobotan kata menggunakan *TF-IDF*. Hasil penelitian ini adalah mayoritas masyarakat atau sebanyak 77% pengguna media sosial X memiliki sentimen positif terhadap penyelenggaraan Pemilihan Umum (Pemilu), serta klasifikasi data menggunakan model *Naive Bayes* bersama dengan *Particle Swarm Optimization* memiliki performa yang baik dengan nilai akurasi mencapai 83,9%.

Kata Kunci : Pemilu, X, *Naive Bayes*, *Particle Swarm Optimization*

ABSTRACT

The Indonesian people have carried out routine five-year activities, namely General Elections (Pemilu) which have been held in 2024. On social media X, there are many public opinions about the implementation of this activity, some are positive and negative. The collection of public opinion will then be analyzed based on its sentiment to be used in model development. This analysis aims to determine the composition of public sentiment towards the 2024 General Election and measure the model's ability to classify the data. This is useful for knowing the public's acceptance of the election and knowing the performance of the Naive Bayes model together with Particle Swarm Optimization in handling the opinion data. Before the development of the model, the data is first labeled with sentiment, preprocessing, data division, and word weighting using TF-IDF. The results of this study are the majority of people or as many as 77% of X social media users have positive sentiments towards the implementation of the General Election (Election), and data classification using the Naive Bayes model together with Particle Swarm Optimization has good performance with an accuracy value of 83.9%.

Keywords: *Election, X, Naive Bayes, Particle Swarm Optimization*

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Allah Swt. atas ridho-Nya saya dapat menyelesaikan penyusunan skripsi ini. Adapun judul skripsi yang saya ajukan adalah "Analisis Sentimen Opini Pengguna Media Sosial X Terhadap Pemilihan Umum Tahun 2024 Menggunakan Metode *Naive Bayes* dan *Particle Swarm Optimization*". Skripsi ini diajukan untuk memenuhi syarat kelulusan Mata Kuliah Tugas Akhir di Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta. Tidak dapat disangkal bahwa butuh usaha yang keras dalam penyelesaian pengerjaan skripsi ini. Namun, karya ini tidak akan selesai tanpa orang-orang tercinta di sekeliling saya yang mendukung dan membantu. Terima kasih saya sampaikan kepada:

1. Ayah dan Ibu dari Mahasiswa Kurniawan Danil (penulis)
2. Prof. Dr. Ir. Supriyanto, M.Sc., IPM., selaku Dekan FIK UPNVJ
3. Dr. Widya Cholil, S.Kom., M.I.T., selaku Kaprodi S1 Informatika
4. Yuni Widiastiwi, S.Kom., M.Si (almh), selaku Dosen Pembimbing Proposal
5. Musthofa Galih Pradana, M.Kom., selaku Dosen Pembimbing 1
6. Novi Trisman Hadi, S.Pd., M.Kom., selaku Dosen Pembimbing 2
7. Semua pihak yang telah membantu dan tidak dapat disebutkan satu per satu

Semoga segala kebaikan dan pertolongan semuanya mendapat berkah dari Allah Swt. dan akhirnya saya menyadari bahwa skripsi ini masih jauh dari kata sempurna, karena keterbatasan ilmu yang saya miliki. Untuk itu saya dengan kerendahan hati mengharapkan saran dan kritik yang sifatnya membangun dari semua pihak demi membangun laporan penelitian ini.

Jakarta, 06 Juni 2024

(Kurniawan Danil)

DAFTAR ISI

PERNYATAAN ORISINALITAS	i
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xii
DAFTAR PERSAMAAN	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Ruang Lingkup Penelitian	4
1.6 Luaran Yang Diharapkan	5
BAB II TINJAUAN PUSTAKA	6
2.1 Opini	6
2.2 Pemilihan Umum (Pemilu)	6
2.3 Sosial Media X (Twitter)	7
2.4 Pembelajaran Mesin (Machine Learning)	8
2.5 Penambangan Teks (Text Mining)	8
2.6 Analisis Sentimen	9
2.6.1 Case Folding	10
2.6.2 Data Cleaning	10
2.6.3 Normalization	10
2.6.4 Tokenization	10
2.6.5 Stop Words Removal	10

2.6.6	<i>Stemming</i>	10
2.7	Pembobotan Kata (<i>TF-IDF</i>)	11
2.8	<i>Particle Swarm Optimization</i>	12
2.9	Metode Klasifikasi <i>Naive Bayes</i>	13
2.10	Evaluasi Klasifikasi	14
2.10.1	Akurasi	15
2.10.2	<i>Confusion Matrix</i>	15
2.10.3	<i>Precision</i>	15
2.10.4	<i>Recall/ Sensitivity</i>	16
2.10.5	<i>Specificity / Selectivity</i>	16
2.10.6	<i>Fall Out</i>	16
2.10.7	<i>Miss Rate</i>	17
2.10.8	<i>F-1 Score</i>	17
2.10.9	<i>ROC Curve and AUC</i>	17
2.11	Penelitian Terdahulu	18
BAB III METODOLOGI PENELITIAN		21
3.1	Kerangka Penelitian	21
3.1.1.	Identifikasi Masalah	22
3.1.2.	Studi Literatur	22
3.1.3.	Akuisisi Data	22
3.1.4.	Pelabelan Data	22
3.1.5.	Praproses Data	23
3.1.6.	Pembobotan Kata	24
3.1.7.	Pembagian Data	25
3.1.8.	Model <i>Naive Bayes</i> dan <i>Particle Swarm Optimization</i>	25
3.1.9.	Evaluasi Model <i>Naive Bayes</i>	25
3.2	Perangkat Penelitian	26
3.2.1.	Perangkat Keras (<i>Hardware</i>)	26
3.2.2.	Perangkat Lunak (<i>Software</i>)	26
3.3	Waktu Penelitian	26
BAB IV HASIL DAN PEMBAHASAN		28
4.1	Akuisisi Data	28
4.2	Pelabelan Data	30
4.3	Praproses Data	37
4.3.1.	<i>Data Preparation</i>	37
4.3.2.	<i>Case Folding</i>	39
4.3.3.	<i>Data Cleaning</i>	40

4.3.4.	<i>Normalization</i>	46
4.3.5.	<i>Tokenization</i>	49
4.3.6.	<i>Stopwords Removal</i>	50
4.3.7.	<i>Stemming</i>	52
4.4	<i>Visualisasi Word Cloud</i>	54
4.5	<i>Pembagian Data</i>	57
4.6	<i>Pembobotan Kata TF-IDF</i>	57
4.7	<i>Model Naive Bayes</i>	60
4.8	<i>Model Naive Bayes dan Particle Swarm Optimization</i>	64
4.9	<i>Evaluasi Model Naive Bayes dan Particle Swarm Optimization</i>	69
4.9.1.	<i>Confusion Matrix</i>	69
4.9.2.	<i>Nilai Akurasi</i>	70
4.9.3.	<i>Precision</i>	70
4.9.4.	<i>Recall</i>	71
4.9.5.	<i>Specificity</i>	71
4.9.6.	<i>Fall Out</i>	72
4.9.7.	<i>Miss Rate</i>	73
4.9.8.	<i>Skor F1</i>	73
4.9.9.	<i>Kurva ROC</i>	74
BAB V KESIMPULAN DAN SARAN		76
5.1	Kesimpulan	76
5.2	Saran	76
DAFTAR PUSTAKA		78
LAMPIRAN		81

DAFTAR TABEL

Tabel 3. 1 Waktu Penelitian	27
Tabel 4. 1 Keterangan Kolom Data	29
Tabel 4. 2 Label Sentimen Danil	31
Tabel 4. 3 Label Sentimen Fernanda	31
Tabel 4. 4 Label Sentimen Alfaza	32
Tabel 4. 5 Rekap Sentimen Final.....	33
Tabel 4. 6 Nilai Null Data	37
Tabel 4. 7 Kolom Data Final	38
Tabel 4. 8 Rekap Data Duplikat	38
Tabel 4. 9 Tahap Case Folding.....	39
Tabel 4. 10 Penghapusan Emoji dan Tag	40
Tabel 4. 11 Penghapusan Hashtag, Angka, dan Simbol.....	42
Tabel 4. 12 Penghapusan Repetisi Kata dan Spasi	43
Tabel 4. 13 Penghapusan URL.....	44
Tabel 4. 14 Penghapusan Kata Tak Bermakna.....	45
Tabel 4. 15 Normalisasi Slang.....	47
Tabel 4. 16 Hapus Kata 3 Huruf.....	48
Tabel 4. 17 Tokenization	49
Tabel 4. 18 Proses Stopwords Removal	50
Tabel 4. 19 Hasil Stopwords Removal	51
Tabel 4. 20 Proses Stemming	52
Tabel 4. 21 Hasil Stemming	53
Tabel 4. 22 Pembagian Data.....	57
Tabel 4. 23 Dokumen Data TF-IDF	57
Tabel 4. 24 Perhitungan Pembobotan Kata TF-IDF.....	58
Tabel 4. 33 Perhitungan Naive Bayes 1	61
Tabel 4. 34 Perhitungan Naive Bayes 2	61
Tabel 4. 35 Perhitungan Naive Bayes 3	62
Tabel 4. 36 Perhitungan Naive Bayes 4	62
Tabel 4. 37 Perhitungan Naive Bayes 5	63
Tabel 4. 25 Hasil Model Naive Bayes.....	63
Tabel 4. 26 Perhitungan PSO 1	65
Tabel 4. 27 Perhitungan PSO 2	66

Tabel 4. 28 Perhitungan PSO 3	66
Tabel 4. 29 Perhitungan PSO 4	66
Tabel 4. 30 Perhitungan PSO 5	67
Tabel 4. 31 Perhitungan PSO 6	67
Tabel 4. 32 Perhitungan PSO 7	68
Tabel 4. 38 Hasil Keluaran NB + PSO	69
Tabel 4. 39 Nilai Confusion Matrix.....	69
Tabel 4. 40 Nilai Akurasi	70
Tabel 4. 41 Nilai Precision	70
Tabel 4. 42 Nilai Recall.....	71
Tabel 4. 43 Nilai Specificity.....	72
Tabel 4. 44 Nilai Fall Out.....	72
Tabel 4. 45 Nilai Miss Rate	73
Tabel 4. 46 Skor F1	73

DAFTAR GAMBAR

Gambar 1. 1 Tren Topik Pemilu (Getdaytrends, 2024).....	2
Gambar 1. 2 Tren Topik Penyelenggaraan Pemilu (Getdaytrends, 2024).....	2
Gambar 2. 1 Grafik Jumlah Pengguna Twitter di Indonesia (Rizaty, M.A, 2022).....	7
Gambar 2. 2 Aplikasi Penambangan Teks (Dasri, Y.B., et al., 2018).....	9
Gambar 2. 3 Perhitungan Confusion Matrix (Edlitera, 2023).....	15
Gambar 2. 4 Contoh Kurva ROC (Edlitera, 2023).....	18
Gambar 3. 1 Kerangka Penelitian.....	21
Gambar 4. 1 Dataset Pemilu.....	28
Gambar 4. 2 Nama Kolom Data.....	28
Gambar 4. 3 Distribusi Sentimen.....	36
<i>Gambar 4. 4 Distribusi Sentimen Berdasarkan Bulan.....</i>	<i>36</i>
Gambar 4. 5 Word Cloud Sentimen Positif.....	54
Gambar 4. 6 Word Cloud Sentimen Negatif.....	55
Gambar 4. 7 Word Cloud Keseluruhan.....	56
Gambar 4. 8 Kemunculan Kata Terbanyak.....	56
Gambar 4. 9 Hasil Pembobotan Kata TF-IDF.....	60
Gambar 4. 10 Kode Program PSO.....	64
Gambar 4. 11 Keluaran Program PSO.....	68
Gambar 4. 12 Kode Program Naive Bayes.....	68
Gambar 4. 13 Kurva ROC Konfigurasi 1.....	74
Gambar 4. 14 Kurva ROC Konfigurasi 2.....	74
Gambar 4. 15 Kurva ROC Konfigurasi 3.....	75

DAFTAR PERSAMAAN

Persamaan 2.1 Rumus Term Frequency	11
Persamaan 2.2 Ketentuan Term Frequency	11
Persamaan 2. 3 Rumus Inverse Document Frequency	11
Persamaan 2. 4 Rumus TF-IDF	12
Persamaan 2. 5 Rumus Kecepatan Partikel PSO	12
Persamaan 2. 6 Rumus Posisi Partikel PSO	13
Persamaan 2. 7 Teorema Bayes	13
Persamaan 2. 8 Rumus Naive Bayes 1	14
Persamaan 2. 9 Rumus Naive Bayes 2	14
Persamaan 2. 10 Rumus Akurasi.....	15
Persamaan 2. 11 Rumus Precision.....	16
Persamaan 2. 12 Rumus Recall	16
Persamaan 2. 13 Rumus Specificity	16
Persamaan 2. 14 Rumus Fall Out	17
Persamaan 2. 15 Rumus Miss Rate	17
Persamaan 2. 16 Rumus F1 Score 1	17
Persamaan 2. 17 Rumus F1 Score 2	17

DAFTAR LAMPIRAN

Lampiran 1 Kode Program	81
-------------------------------	----