



**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
PASANGAN CAPRES CAWAPRES NO 1 PADA KOLOM  
KOMENTAR TWITTER MENGGUNAKAN METODE K-  
NEAREST NEIGHBORS (KNN)**

**SKRIPSI**

**Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Program Studi Teknik Informatika**

**Oleh :**

**Farras Ananta**

**2020.11.0027**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER DAN SAINS  
UNIVERSITAS INDO GLOBAL MANDIRI 2025**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
PASANGAN CAPRES CAWAPRES NO 1 PADA KOLOM  
KOMENTAR TWITTER MENGGUNAKAN METODE K-  
NEAREST NEIGHBORS (KNN)**



**SKRIPSI**

**Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Program Studi Teknik Informatika**

**Oleh :**

**Farras Ananta**

**2020.11.0027**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER DAN SAINS  
UNIVERSITAS INDO GLOBAL MANDIRI 2025**

# SURAT LEMBAR PENGESAHAN SKRIPSI

## LEMBAR PENGESAHAN SKRIPSI

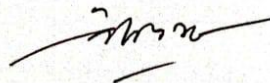
ANALISIS SENTIMEN MASYARAKAT TERHADAP  
PASANGAN CAPRES CAWAPRES NO. 1 PADA KOLOM  
KOMENTAR TWITTER MENGGUNAKAN METODE K-  
NEAREST NEIGHBOURS (K-NN)

Oleh

Farras Ananta  
NPM : 2020.11.0027

Palembang, 14 Februari 2025

Pembimbing I



Dr. Rendra Gustriansyah, S.T., M.Kom  
NIK : 1999.01.0006

Pembimbing II




Lastri Widya Astuti, M.Kom  
NIK: 2003.01.0063

Mengetahui,

Dekan Fakultas Ilmu Komputer dan Sains

FAKULTAS ILMU KOMPUTER DAN SAINS

**UIGM**



Rudi Heriansyah, S.T., M.Eng. Ph.D.  
NIK:2022.01.0315

## SURAT LEMBAR PERSETUJUAN DEWAN PENGUJI

### LEMBAR PERSETUJUAN DEWAN PENGUJI

Pada hari Selasa tanggal 11 Februari 2025 telah dilaksanakan ujian sidang skripsi :

Nama : Farras ananta

NPM : 2020.11.0027

Judul : Analisis Sentimen Masyarakat Terhadap Pasangan Capres Dan Cawapres No 1 Pada Kolom Komentar Twitter Menggunakan Metode K-Nearest Neighbours

Oleh Prodi Teknik Informatika Fakultas Ilmu Komputer dan Sains Universitas Indo Global Mandiri Palembang

Palembang, 11 Februari 2025

Penguji 1,



Dr. Gasim, S.Kom., M.Si

NIK: 2023.01.0340

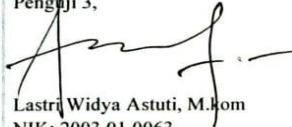
Penguji 2,



Ir. Mustafa Ramadhan, M.T.

NIK: 2002.03.0172

Penguji 3,



Lastris Widya Astuti, M.Kom

NIK: 2003.01.0063

Menyetujui,  
Ka. Prodi Teknik Informatika



Zaid Romegar Mair, S.T., M.Cs

NIK: 2021.01.0307

## SURAT KETERANGAN REVISI SKRIPSI



SURAT KETERANGAN REVISI SKRIPSI  
PROGRAM STUDI TEKNIK INFORMATIKA (SI)  
FASILKOM DAN SAINS UNIVERSITAS INDO GLOBAL MANDIRI

Kami yang bertanda tangan dibawah ini, menerangkan bahwa :

Nama : Farras Ananta  
NPM : 2020.11.0027  
Judul : Analisis Sentimen Masyarakat Terhadap Pasangan Capres  
Dan Cawapres No 1 Pada Kolom Komentar Twitter  
Menggunakan Metode K-Nearest Neighbours

Mahasiswa yang namanya tercantum diatas, telah selesai merevisi penulisan SKRIPSI

Palembang, 14 Februari 2025

Penguji 1,

Dr. Gasim, S.Kom., M.Si  
NIK: 2023.01.0340

Penguji 2,

Ir. Mustafa Ramadhan, M.T.  
NIK: 2002.03.0172

Penguji 3,

Lastri Widya Astuti, M.Kom  
NIK: 2003.01.0063

Menyetujui,  
Ka. Prodi Teknik Informatika

Zaid Romegar Mar, S.T., M.Cs  
NIK: 2021.01.0307

**ANALISIS SENTIMEN MASYARAKAT PADA TWITTER  
TERHADAP PASANGAN CAPRES CAWAPRES NO 1  
MENGUNAKAN METODE K-NEAREST NEIGHBORS (KNN)**

**ABSTRAK**

Twitter, sebagai lautan pendekatan kreatif dan pemikiran singkat, menjadi Media perdebatan digital di mana sentimen masyarakat terhadap calon presiden (capres) dan calon wakil presiden (cawapres) saling berbenturan. Dengan karakteristik uniknya, pesan singkat atau "*tweet*" menjadi kanvas bagi warga dunia daring untuk mengekspresikan dukungan, kritik, dan opini mereka. Analisis sentimen, sebagai cabang utama dalam ilmu *data mining*, menjadi penelitian yang substansial dalam ekstraksi dan interpretasi data tekstual. Fokus utamanya adalah pada pemahaman mendalam mengenai opini, evaluasi, sikap, emosi, penilaian, dan sentimen individu terhadap berbagai entitas, seperti produk, tokoh, organisasi, serta isu-isu yang tengah beredar. Pada penelitian ini dilakukan analisis terhadap sentiment masyarakat terhadap Calon presiden (Capres), dan Calon wakil presiden (Cawapres) secara spesifik terhadap Pasangan Nomor urut 1 (satu) yang dikumpulkan dari kolom komentar thread baik dari berita mengenai paslon no 1, ataupun thread kampanye paslon tersebut.

Kata kunci: Analisis, sentimen, *Twitter*, *KNN*, *masyarakat*

# **ANALYSIS OF PUBLIC SENTIMENT ON TWITTER TOWARDS PRESIDENTIAL CANDIDATE PAIR NO. 1 USING THE K-NEAREST NEIGHBORS (KNN) METHOD**

## **ABSTRACT**

Twitter, as a sea of creative approaches and short thoughts, has become a medium for digital debates where public sentiments towards presidential candidates (presidential candidates) and vice presidential candidates (vice presidential candidates) clash with each other. With its unique characteristics, short messages or “tweets” become a canvas for online citizens to express their support, criticism, and opinions. Sentiment analysis, as a major branch of data mining, has become a substantial research in the extraction and interpretation of textual data. Its main focus is on deeply understanding the opinions, evaluations, attitudes, emotions, judgments, and sentiments of individuals towards various entities, such as products, personalities, organizations, and issues. In this study, an analysis of public sentiment towards presidential candidates (Capres), and vice presidential candidates (Cawapres) specifically towards Pair Number 1 (one) was carried out, which was collected from the thread comment column either from news about candidate number 1, or the candidate's campaign thread.

Keyword: Analysis, Sentiment, Twitter, KNN, Public

## KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan proposal penelitian tepat pada waktunya, tidak lupa juga shalawat beserta salam kepada Nabi Muhammad SAW beserta pengikutnya hingga dan insyaallah kita semua dapat bertemu di surga nanti pada akhir zaman. Proposal penelitian yang penulis buat dengan judul “ Analisis sentimen masyarakat terhadap pasangan capres cawapres no. 1 pada kolom komentar twitter menggunakan metode K-nearest neighbors (KNN)” disusun guna memenuhi syarat untuk lanjut pada sidang skripsi program studi Teknik Informatika, Fakultas Ilmu Komputer dan Sains, Universitas Indo Global Mandiri Palembang.

Terimakasih juga tidak lupa diucapkan atas bantuan yang diberikan selama penyusunan skripsi ini kepada:

1. Dr. Marzuki Alie, SE., MM, selaku Rektor Universitas Indo Global Mandiri Palembang
2. Rudi Heriansyah, ST., M.Eng. Ph.D selaku Dekan Fakultas Ilmu Komputer dan Sains Universitas Indo Global Mandiri.
3. Zaid Romegar Mair, S.T., M.Cs sebagai Ketua Program Studi Teknik Informatika Universitas Indo Global Mandiri
4. Septa Cahyani, S.Kom., M.Cs. sebagai Dosen Pembimbing Akademik.
5. Dr. Rendra Gustriyansyah, S.T., M.Kom. sebagai Dosen Pembimbing I.
6. Lastri Widya Astuti, S.Kom., M.Kom. sebagai Dosen Pembimbing II.
7. Bapak/Ibu Dosen Fakultas Ilmu Komputer dan Sains dan Karyawan /Karyawati Universitas Indo Global Mandiri.
8. Keluarga dan teman-teman yang telah memberikan dukungan materi maupun moral sehingga penulis bisa menyelesaikan proposal ini

Penulis menyadari bahwa proposal ini masih jauh dari kata sempurna, oleh karena itu kritik beserta saran untuk perbaikan dan pengembangan sangat dibutuhkan.



Akhir kata, semoga proposal penelitian ini bermanfaat bagi semua pihak. Penulis mengucapkan terima kasih.

Palembang, 21 November 2024

Penulis

Farras ananta

NPM. 2020110027

## DAFTAR ISI

SURAT LEMBAR PENGESAHAN SKRIPSI .....	i
SURAT LEMBAR PERSETUJUAN DEWAN PENGUJI .....	ii
SURAT KETERANGAN REVISI SKRIPSI .....	iii
ABSTRAK .....	iv
ABSTRACT .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI .....	vi
DAFTAR GAMBAR .....	viii
DAFTAR TABEL .....	ix
DAFTAR LAMPIRAN .....	xi
DAFTAR RUMUS .....	xii
BAB I PENDAHULUAN .....	1
1.1    Latar belakang .....	1
1.2    Batasan masalah .....	4
1.3    Rumusan masalah .....	4
1.4    Tujuan dan Manfaat penelitian .....	5
1.4.1    Tujuan Penelitian .....	5
1.4.2    Manfaat penelitian .....	5
1.5    Sistematika penulisan .....	5
BAB II TINJAUAN PUSTAKA .....	7
2.1    Sentimen analisis .....	7
2.1.1    Masyarakat .....	7
2.2 <i>Twitter</i> .....	8
2.3    Opini Publik .....	9
2.4 <i>Machine learning</i> .....	10
2.5 <i>Text Mining</i> .....	13
2.5.1 <i>Preprocessing text</i> .....	13
2.6 <i>Preprocessing data</i> .....	15

2.7	<i>Natural Language Processing (NLP)</i> .....	16
2.8	Metode <i>elbow</i> .....	16
2.9	<i>K-Nearest Neighbors</i> .....	17
2.10	<i>Confusion matrix</i> .....	21
2.11	Python .....	23
2.12	Penelitian terdahulu .....	23
BAB III METODOLOGI PENELITIAN .....		26
3.1	Tahapan penelitian .....	26
3.2	Studi literatur .....	27
3.3	Pengumpulan data .....	27
3.4	Kebutuhan perangkat lunak .....	27
3.5	<i>Data Preprocessing</i> .....	28
3.6	Analisis sentimen .....	33
3.7	Hasil pengujian .....	40
BAB IV HASIL DAN PEMBAHASAN .....		41
4.1	Pengumpulan data .....	41
4.2	<i>Data Preprocessing</i> .....	42
4.2.1	<i>Cleaning</i> .....	42
4.2.2	<i>Case folding</i> .....	44
4.2.3	<i>Steaming</i> .....	45
4.2.4	<i>Stopward removal</i> .....	47
4.2.5	<i>Tokenizing</i> .....	48
4.2.6	TF-IDF .....	50
4.3	Uji coba model <i>K-nearest neighbors</i> .....	52
4.4	Hasil dan pembahasan .....	57
BAB V PENUTUP .....		59
5.1	Kesimpulan .....	59
5.2	Saran .....	59
Daftar Pustaka .....		61
Lampiran .....		65

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Gambaran gambar sederhana dari algoritma <i>K-Nearest Neighbor</i> (KNN).....	20
<b>Gambar 3.1</b> Tahapan Penelitian.....	26
<b>Gambar 3.2</b> Proses pengumpulan data .....	27
<b>Gambar 3.3</b> Proses <i>preprocessing data</i> .....	28
<b>Gambar 4.1</b> Pemanggilan data.....	42
<b>Gambar 4.2</b> <i>Cleaning</i> .....	42
<b>Gambar 4.3</b> <i>Case folding</i> .....	44
<b>Gambar 4.4</b> Proses <i>Steaming</i> .....	45
<b>Gambar 4.5</b> Proses <i>stopward removal</i> .....	47
<b>Gambar 4.6</b> Proses <i>tokenizing</i> .....	48
<b>Gambar 4.7</b> Proses TF-IDF.....	50
<b>Gambar 4.8</b> Proses pelabelan TF-IDF.....	51
<b>Gambar 4.9</b> Proses <i>Splitting data</i> .....	53
<b>Gambar 4.10</b> Proses metode <i>elbow</i> .....	53
<b>Gambar 4.11</b> Hasil Metode <i>elbow</i> .....	54
<b>Gambar 4.12</b> <i>Define pipeline</i> .....	54
<b>Gambar 4.13</b> Proses latih data.....	55
<b>Gambar 4.14</b> Matrik evaluasi pelatihan data.....	55
<b>Gambar 4.15</b> Proses pengujian data dengan <i>data testing</i> .....	56
<b>Gambar 4. 16</b> Hasil evaluasi matrik dan <i>confusion matrik data testing</i> .....	57

## DAFTAR TABEL

<b>Tabel 2.1</b> Tabel <i>confusion matrix</i> .....	21
<b>Tabel 2.2</b> Penelitian terdahulu.....	23
<b>Tabel 2.3</b> Penelitian terdahulu.....	24
<b>Tabel 3. 1</b> <i>case folding</i> .....	29
<b>Tabel 3. 2</b> <i>Cleaning</i> .....	30
<b>Tabel 3.3</b> <i>stemming</i> .....	30
<b>Tabel 3.4</b> <i>stopword removal</i> .....	31
<b>Tabel 3.5</b> <i>Tokenizing</i> .....	32
<b>Tabel 3. 6</b> TF-IDF.....	33
<b>Tabel 3.7</b> TF-IDF.....	34
<b>Tabel 3.8</b> Metode <i>elbow</i> .....	36
<b>Tabel 3.9</b> Metode <i>elbow</i> .....	37
<b>Tabel 3.10</b> Metode <i>elbow</i> .....	38
<b>Tabel 3.11</b> Proses <i>Splitting data</i> .....	39
<b>Tabel 3.12</b> Proses Model K-NN.....	39
<b>Tabel 3. 13</b> Proses K-NN.....	40
<b>Tabel 4.1</b> <i>Dataset</i> .....	41
<b>Tabel 4.2</b> sebelum <i>cleaning</i> .....	43
<b>Tabel 4.3</b> Hasil setelah <i>cleaning</i> .....	43
<b>Tabel 4.4</b> Hasil setelah <i>casefolding</i> .....	44
<b>Tabel 4.5</b> Hasil setelah <i>casefolding</i> .....	45
<b>Tabel 4.6</b> Sebelum proses <i>stemming</i> .....	46
<b>Tabel 4.7</b> setelah <i>steamming</i> .....	46
<b>Tabel 4.8</b> sebelum <i>stopward removal</i> .....	47
<b>Tabel 4.9</b> Hasil <i>stopward removal</i> .....	48
<b>Tabel 4.10</b> Sebelum <i>tokenizing</i> .....	49
<b>Tabel 4.11</b> Hasil <i>tokenizing</i> .....	49

<b>Tabel 4.12</b> Hasil pelabelan TF-IDF .....	51
<b>Tabel 4.13</b> Hasil TF-IDF .....	52

## **DAFTAR LAMPIRAN**

Lampiran 1 Daftar riwayat hidup.....	65
Lampiran 2 Surat pernyataan tidak plagiat.....	66
Lampiran 3 Kartu bimbingan.....	69

## DAFTAR RUMUS

Rumus 2.1.....	15
Rumus 2.2.....	16
Rumus 2.3.....	17
Rumus 2.4.....	18
Rumus 2.5.....	21
Rumus 2.6.....	21
Rumus 2.7.....	21
Rumus 2.8.....	21