

**ANALISIS SENTIMEN TOKOH POLITIK SEBAGAI KAJIAN  
KANDIDAT CAPRES DAN CAWAPRES MENGGUNAKAN  
FASTTEXT EMBEDDINGS DAN CONVOLUTIONAL  
NEURAL NETWORK PADA 10 WILAYAH DI INDONESIA**

**SKRIPSI**



**Oleh:**

**FAHRI IZZUDDIN ZULKARNAEN**

**NPM.19081010046**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR**

**2023**

**ANALISIS SENTIMEN TOKOH POLITIK SEBAGAI KAJIAN  
KANDIDAT CAPRES DAN CAWAPRES MENGGUNAKAN  
FASTTEXT EMBEDDINGS DAN CONVOLUTIONAL  
NEURAL NETWORK PADA 10 WILAYAH DI INDONESIA**

**SKRIPSI**

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar Sarjana  
Komputer Program Studi Informatika



Oleh:

**FAHRI IZZUDDIN ZULKARNAEN**

**NPM.19081010046**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR**

**2023**

**LEMBAR PENGESAHAN**  
**SKRIPSI**

**Judul : ANALISIS SENTIMEN TOKOH POLITIK SEBAGAI KAJIAN KANDIDAT CAPRES DAN CAWAPRES MENGGUNAKAN FASTTEXT EMBEDDINGS DAN CONVOLUTIONAL NEURAL NETWORK PADA 10 WILAYAH DI INDONESIA**

**Oleh : Fahri Izzuddin Zulkarnaen**

**NPM : 19081010046**

**Telah Diseminarkan Dalam Ujian Skripsi Pada:**

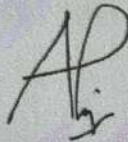
**Hari Selasa, Tanggal 23 Mei 2023**

**Mengetahui:**

**Dosen Pembimbing**

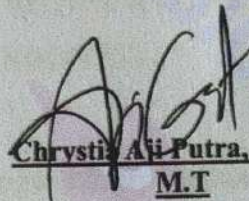
**Dosen Penguji**

1.



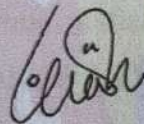
**Dr. Eng. Ir. Angraini Puspita Sari, S.T., M.T.**  
NPT. 222198 60 816400

1.



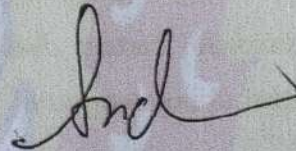
**Chrystia Aji Putra, S.Kom., M.T.**  
NIP. 19861008 2021211 001

2.



**Agung Mustika Rizki, S.Kom., M.Kom.**  
NIP. 19930725 202203 1008

2.

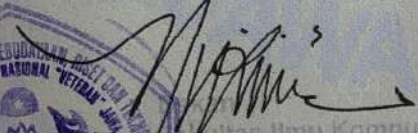


**Andreas Nugroho Sihananto, S.Kom., M.Kom.**  
NPT. 211199 00 412271

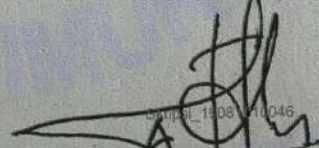
**Menyetujui:**

**Dekan  
Fakultas Ilmu Komputer**

**Koordinator Program Studi  
Informatika,**



**Dr. Ar. Novirina Hendrasarie, M.T.**  
NIP. 19681126 199403 2 001



**Fetty Tri Anggraeny, S.Kom., M.Kom.**  
NIP. 19820211 2021212 005



## **SURAT PERNYATAAN ANTI PLAGIAT**

Saya, mahasiswa program studi Informatika UPN “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Fahri Izzuddin Zulkarnaen

NPM : 19081010046

Menyatakan bahwa judul skripsi yang saya ajukan dan kerjakan yang berjudul

**“ANALISIS SENTIMEN TOKOH POLITIK SEBAGAI KAJIAN  
KANDIDAT CAPRES DAN CAWAPRES MENGGUNAKAN FASTTEXT  
EMBEDDINGS DAN CONVOLUTIONAL NEURAL NETWORK PADA 10  
WILAYAH DI INDONESIA”**

Bukan merupakan plagiat dari skripsi/tugas akhir/penelitian orang lain dan juga bukan merupakan produk dan atau software yang saya beli dari pihak lain. Saya juga menyatakan bahwa skripsi ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi Pendidikan lain.

Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka saya siap menerima segala konsekuensinya.

Surabaya, 23 Mei 2023

Penulis



**FAHRI IZZUDDIN ZULKARNAEN**

**NPM. 19081010046**

# ANALISIS SENTIMEN TOKOH POLITIK SEBAGAI KAJIAN KANDIDAT CAPRES DAN CAWAPRES MENGGUNAKAN FASTTEXT EMBEDDINGS DAN CONVOLUTIONAL NEURAL NETWORK PADA 10 WILAYAH DI INDONESIA

Nama Mahasiswa : Fahri Izzuddin Zulkarnaen  
NPM : 19081010046  
Program Studi : Informatika  
Dosen Pembimbing : Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.  
Agung Mustika Rizki, S.Kom, M.Kom.

---

## ABSTRAK

Internet telah menjadi sarana untuk menyimpan dan mengakses berbagai informasi secara publik. Masyarakat sendiri cenderung menggunakan media sosial untuk menyatakan pendapat mereka terhadap suatu produk, kebijakan politik, dan politisi baik secara partai maupun personal. Analisis sentimen kandidat capres dan cawapres dilakukan dengan tujuan untuk memperoleh pemahaman tentang pandangan masyarakat terhadap masing-masing kandidat. Berdasarkan analisis kualitatif popularitas dan sentimen masyarakat Indonesia pada media sosial, dapat diketahui nama-nama kandidat yang berpotensi unggul sebagai capres atau cawapres pada pemilihan presiden yang masih lebih dari satu tahun lagi. Analisis sentimen dilakukan dengan cara melakukan pengambilan data dari media sosial Twitter.

Penelitian ini membagi sentimen menjadi 4 kelas, yaitu bahagia, sedih, cinta, marah. Pelabelan dilakukan dengan menggunakan bantuan *pre-trained Roberta model*. Pembuatan model dilakukan dengan membandingkan performansi FastText *embeddings*, Keras *embeddings* dan FastText *average word vector* dengan menggunakan algoritma *convolutional neural network*. Perbandingan dilakukan dengan melihat *evaluation matrix* dari nilai terbaik yang dihasilkan setiap model. Sebagai hasil, FastText *embeddings* unggul dengan nilai *f1-score* 0,9510 selanjutnya Keras *embeddings* memiliki nilai *f1-score* terbaik 0,9443 dan FastText *average word vector* dengan nilai *f1-score* 0,7861.

Perhitungan sentimen dilaksanakan dengan melakukan perhitungan berdasarkan poin. Perhitungan poin tersebut meliputi bahagia 2 poin, cinta 1 poin, sedih -1 poin, marah -2 poin. Hasil yang didapatkan Ganjar Pranowo unggul dalam 9 bagian wilayah dengan total poin 42.216 poin sehingga menjadikan Ganjar Pranowo sebagai kandidat kuat menjadi calon presiden. Pada kandidat calon wakil presiden sendiri, Erick Thohir unggul menjadi kandidat kuat sebagai cawapres dengan unggul pada 5 bagian wilayah dengan total poin yang didapatkan ialah 55.464 poin.

**Kata kunci:** Analisis Sentimen, *Convolution Neural Network*, Pemilu, *Deep Learning*.

## **KATA PENGANTAR**

Assalamu'alaikum Wr.Wb. Segala puji bagi Allah SWT yang telah melimpahkan rahmat, taufik serta hidayah-Nya sehingga peneliti dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Tokoh Politik Sebagai Kajian Kandidat Capres dan Cawapres Menggunakan FastText Embeddings dan Convolutional Neural Network Pada 10 Wilayah di Indonesia”. Tidak lupa Shalawat dan salam semoga senantiasa ditunjukkan bagi Rasulullah SAW.

Penulis mengucapkan terima kasih sebanyak-banyaknya kepada pihak yang telah memberikan bantuan baik secara moril, nasihat, semangat, saran maupun materil kepada pihak-pihak yang telah terlibat dalam penyusunan skripsi.

Teruntuk pihak-pihak yang telah membantu penulis dalam pengerjaan penelitian dan skripsi tersebut, semoga Allah membalas dengan balasan sebaik-baiknya. Penulis menyadari masih terdapat kekurangan dalam penelitian ini, baik pada kode program, analisis, maupun penulisan laporan skripsi. Namun, penulis berharap skripsi yang telah tersusun ini dapat bermanfaat bagi perkembangan ilmu pengetahuan di dunia.

## UCAPAN TERIMA KASIH

Laporan skripsi ini dapat terselesaikan dengan baik atas berkat dan rahmat Allah SWT yang telah memberikan limpahan kebaikan serta kesehatan bagi kita semua terutama bagi penulis agar dapat menyelesaikan skripsi sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Informatika. Penyelesaian penulisan skripsi ini tidak lepas dari bimbingan, bantuan dan pengarahan dari berbagai pihak, untuk itu segala hormat penulis mengucapkan terimakasih kepada :

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Dr. Ir. Novirina Hendrasarie, M.T., selaku Dekan Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom., selaku Koordinator Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT. selaku dosen pembimbing 1 (satu) skripsi ini yang selalu membimbing dengan baik, memberikan masukan dan pemecahan masalah yang tepat dan membimbing dengan sepenuh hati sehingga keseluruhan rangkaian penyusunan skripsi ini selesai tepat pada waktunya.
5. Bapak Agung Mustika Rizki, S.Kom, M.Kom. selaku dosen pembimbing 2 (dua) skripsi ini yang selalu memberikan motivasi, mengingatkan saya untuk selalu teliti dan tepat waktu dan memberikan arahan yang baik sehingga penyusunan kode program ini dapat berjalan dengan lancar.
6. Seluruh anggota dosen pengajar yang tidak dapat saya sebutkan satu per satu, yang telah menginspirasi dan memotivasi penulis untuk mendapatkan ilmu yang sangat berguna sekali untuk kedepannya.
7. Terimakasih yang tak terhingga untuk kedua orang tua Ayah Pra Budi Zulkarnaen dan Bunda Mia Finggary Anigta yang telah banyak sekali memberikan dukungan dalam keseluruhan rangkaian penyusunan skripsi ini.
8. Seluruh staff Fakultas Ilmu Komputer UPN Veteran Jawa Timur yang telah memberikan dukungan dalam penyelesaian rangkaian skripsi ini.

9. Kepada semua pihak yang telah membantu dan yang tidak dapat disebutkan oleh penulis satu per satu.

Penulis menyadari bahwa masih banyak kekurangan dalam keseluruhan rangkaian skripsi ini, meskipun telah diusahakan sebaik mungkin. Hal ini semata-mata disebabkan oleh keterbatasan kemampuan dan kekhilafan dari penulis, namun penulis sangat berharap semoga tesis ini dapat memiliki banyak pemanfaat bagi peneliti dan kelanjutan perkembangan ilmu pengetahuan di Dunia.

Surabaya, 23 Mei 2023

Fahri Izzuddin Zulkarnaen



## DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI .....	i
SURAT PERNYATAAN ANTI PLAGIAT .....	ii
ABSTRAK .....	iii
KATA PENGANTAR.....	iv
UCAPAN TERIMA KASIH .....	v
DAFTAR ISI .....	vii
DAFTAR GAMBAR .....	x
DAFTAR TABEL.....	xii
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 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1 Penelitian Terdahulu.....	6
2.2 Demokrasi .....	7
2.3 Pemilu dan Kampanye .....	8
2.4 Analisis Sentimen.....	9
2.5 Poltracking Indonesia.....	10
2.6 Twitter .....	10
2.7 <i>Web Scrapping</i> .....	11
2.8 <i>Text Preprocessing</i> .....	11
2.9 <i>Neural Network</i> .....	12
2.10 <i>Deep Learning</i> .....	13
2.11 FastText .....	14
2.12 <i>Convolutional Neural Network</i> .....	15
2.12.1 <i>Word Embedding</i> .....	16
2.12.2 <i>Convolutional Layer</i> .....	17
2.12.3 <i>ReLU Activation</i> .....	19
2.12.4 <i>Pooling Layer</i> .....	20
2.12.5 <i>Dropout Layer</i> .....	21
2.12.6 <i>Softmax Activation</i> .....	23

2.13 <i>Confusion Matrix</i> .....	25
BAB III METODE PENELITIAN.....	28
3.1 <i>Scrapping</i> Data Kandidat .....	28
3.2 <i>Text Preprocessing</i> .....	31
3.2.1 <i>Case Folding</i> .....	32
3.2.2 <i>Remove All Unusual Character</i> .....	32
3.2.3 <i>Remove non-alphabet</i> .....	32
3.2.4 Normalisasi Kata Alay .....	33
3.2.5 <i>Stemming</i> .....	33
3.2.6 <i>Stopwords Removal</i> .....	33
3.2.7 <i>Remove Sentence &lt; 5 Kata</i> .....	33
3.3 Pembuatan <i>Dataset</i> .....	34
3.4 Pelabelan <i>Dataset</i> .....	34
3.5 Pembuatan Model Deep Learning.....	35
3.6 Evaluasi Model Terbaik dan Pembuatan Fungsi Klasifikasi.....	45
3.7 Analisis Hasil Popularitas dan Sentimen.....	45
3.7.1 Analisis Popularitas .....	46
3.7.2 Analisis Sentimen.....	46
BAB IV HASIL DAN PEMBAHASAN.....	48
4.1 Lingkungan Penelitian .....	48
4.1.1 Spesifikasi Perangkat Keras .....	48
4.1.2 Spesifikasi Perangkat Lunak.....	48
4.2 <i>Scrapping</i> Data Kandidat .....	49
4.3 <i>Text Preprocessing</i> .....	53
4.3.1 <i>Case Folding</i> .....	54
4.3.2 <i>Remove All Unusual Character</i> .....	55
4.3.3 <i>Remove Non-Alphabet</i> .....	55
4.3.4 Normalisasi Kata Alay .....	56
4.3.5 <i>Stemming</i> .....	57
4.3.6 <i>Stopwords Removal</i> .....	57
4.3.7 <i>Remove Sentence &lt; 5 Kata</i> .....	58
4.4 Pembuatan <i>Dataset</i> .....	61
4.5 Pelabelan <i>Dataset</i> .....	61
4.6 Pembuatan Model <i>Deep Learning</i> .....	64
4.6.1 Persiapan <i>Model Library</i> .....	64
4.6.2 Pembuatan <i>Model Library</i> .....	68

4.6.2 Evaluasi <i>Model Library</i> .....	71
4.7 Pembuatan model tanpa menggunakan <i>library</i> .....	76
4.7.1 Persiapan Data dan Pembuatan <i>Model Tanpa Library</i> .....	76
4.7.2 Pelatihan dan Evaluasi Model.....	89
4.8 Evaluasi Model Terbaik dan Pembuatan Fungsi Klasifikasi.....	95
4.8.1 Evaluasi Model Terbaik .....	95
4.8.2 Pembuatan Fungsi Klasifikasi.....	97
4.9 Analisis Hasil Popularitas dan Sentimen.....	100
4.9.1 Analisis Popularitas .....	100
4.9.2 Analisis Sentimen.....	101
4.9.3 Kesimpulan Analisis .....	105
BAB V KESIMPULAN .....	108
5.1 Kesimpulan .....	108
5.2 Saran.....	108
DAFTAR PUSTAKA .....	110
BIODATA PENULIS .....	117

## DAFTAR GAMBAR

Gambar 2.1 Visualisasi Neural Network .....	12
Gambar 2.2 Penggambaran N-Grams Pada FastText.....	15
Gambar 2.3 Visualisasi Text CNN .....	15
Gambar 2.4 Visualisasi Convolution Layer .....	18
Gambar 2.5 Grafik Fungsi Aktivasi ReLU .....	20
Gambar 2.6 Visualisasi Pooling Layer.....	21
Gambar 2.7 Visualisasi Dropout Layer.....	22
Gambar 2.8 Rumus Dropout Layer.....	22
Gambar 2.9 Visualisasi Dropout .....	23
Gambar 3.1 Tahapan Penelitian .....	28
Gambar 3.2 Tahapan Text Preprocessing.....	31
Gambar 3.3 Arsitektur Convolutional Neural Network.....	35
Gambar 3.4 N-Grams Pada FastText .....	36
Gambar 3.5 Sentence Embeddings .....	37
Gambar 4.1 Tabel Jumlah Data .....	66
Gambar 4.2 Diagram Batang Jumlah Data .....	66
Gambar 4.3 Visualisasi Jumlah Data Seimbang .....	67
Gambar 4.4 Model Summary.....	70
Gambar 4.5 Training Output .....	71
Gambar 4.6 Evaluation metrics Data latih .....	72
Gambar 4.7 Grafik Evaluation metrics Setiap Iterasi .....	73
Gambar 4.8 Confusion Matrix Test Data .....	74
Gambar 4.9 Evaluation metrics Prediksi Sentimen.....	75
Gambar 4.10 Evaluation metrics Terbaik.....	75
Gambar 4.11 Evaluation metrics Keras Embedding .....	75
Gambar 4.12 Output Data latih .....	91
Gambar 4.13 Nilai Evaluation metrics Rata-Rata dan Last Training.....	92
Gambar 4.14 Grafik Evaluation metrics Setiap Epoch .....	93
Gambar 4.15 Confusion Matrix Test Data .....	94
Gambar 4.16 <i>Evaluation</i> Metrics Test Data.....	95

Gambar 4.17 Hasil Labeling Dataset Ganjar Pranowo .....	99
Gambar 4.18 Jumlah Sentimen Ganjar Pranowo .....	100

## DAFTAR TABEL

Tabel 3.1 Kata Kunci Kandidat .....	29
Tabel 3.2 Kata Kunci Wilayah .....	30
Tabel 3.3 Contoh Hasil Scrapping .....	31
Tabel 3.4 Inisialisasi Hyperparameter .....	45
Tabel 4.1 Detail Spesifikasi Perangkat Keras .....	48
Tabel 4.2 Detail Spesifikasi Perangkat Lunak .....	49
Tabel 4.3 Contoh Hasil Scrapping Data Kandidat .....	50
Tabel 4.4 Detail Jumlah Data Ganjar Pranowo .....	51
Tabel 4.5 Detail Jumlah Tweets Anies Baswedan .....	52
Tabel 4.6 Detail Jumlah Tweets Prabowo .....	52
Tabel 4.7 Detail Jumlah Tweets Erick Thohir .....	52
Tabel 4.8 Detail Jumlah Tweets Ridwan Kamil .....	53
Tabel 4.9 Contoh Unprocessed Sentimen .....	53
Tabel 4.10 Penerapan Case Folding .....	54
Tabel 4.11 Penerapan Remove All Unusual Character .....	55
Tabel 4.12 Penerapan Remove non-alphabet .....	56
Tabel 4.13 Penerapan Normalisasi Kata Alay .....	56
Tabel 4.14 Penerapan Stemming .....	57
Tabel 4.15 Penerapan Stopwords Removal .....	58
Tabel 4.16 Penerapan Remove Sentence > 4 Kata .....	58
Tabel 4.17 Penerapan Remove Sentence < 5 Kata .....	58
Tabel 4.18 Contoh Hasil Text Preprocessing .....	60
Tabel 4.19 Jumlah Preprocessed Tweets masing-masing Kandidat .....	60
Tabel 4.20 Contoh Hasil Pelabelan Data .....	62
Tabel 4.21 Hasil Pelabelan Dataset 2.500 .....	63
Tabel 4.22 Hasil Pelabelan Dataset 10.000 .....	64
Tabel 4.23 Hasil Hyperparameter Tuning .....	95
Tabel 4.24 Perbandingan <i>F1-Score</i> Model Terbaik .....	96
Tabel 4.25 Perbandingan <i>f1-score</i> FastText dan Keras <i>embedding</i> .....	96
Tabel 4.26 Jumlah Tweet Kandidat .....	101
Tabel 4.27 Hasil Analisis Sentimen Ganjar Pranowo .....	102

Tabel 4.28 Hasil Analisis Sentimen Anies Baswedan .....	102
Tabel 4.29 Hasil Analisis Sentimen Prabowo Subianto .....	103
Tabel 4.30 Hasil Analisis Sentimen Erick Thohir .....	104
Tabel 4.31 Hasil Analisis Sentimen Ridwan Kamil .....	104
Tabel 4.32 Kandidat Capres Unggul Dalam Bagian Wilayah .....	105
Tabel 4.33 Kandidat Cawapres Unggul Dalam Bagian Wilayah .....	106