

SKRIPSI

**ANALISIS SENTIMEN TERHADAP FILM *DIRTY VOTE*
BERDASARKAN OPINI PENGGUNA X (*TWITTER*)
MENGUNAKAN METODE *CONVOLUTIONAL NEURAL
NETWORK***



*Mencerdaskan dan
Memartabatkan Bangsa*

AINUR RAFIQ ABDILLAH

1512620079

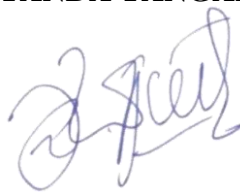

**PROGRAM STUDI
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA**

2024




**ANALISIS SENTIMEN TERHADAP FILM *DIRTY VOTE*
BERDASARKAN OPINI PENGGUNA X (TWITTER) MENGGUNAKAN
METODE *CONVOLUTIONAL NEURAL NETWORK***

Ainur Rafiq Abdillah, NIM 1512620079

HALAMAN PENGESAHAN

NAMA DOSEN	TANDA TANGAN	TANGGAL
Dr. Widodo, M.Kom NIP. 197203252005011002 (Dosen Pembimbing 1)		24 Juni 2024
Murien Nugraheni, S.T., M.C NIP. 198710112019032012 (Dosen Pembimbing 2)		24 Juni 2024

PENGESAHAN PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
Hamidillah Ajie, S.Si, M.T NIP. 197408242005011001 (Dosen Penguji I)		24 Juni 2024
Bambang Prasetya Adhi, S.Pd, M.Kom NIP. 198302252014041001 (Dosen Penguji II)		24 Juni 2024
Diat Nurhidayat, M.T.I NIP. 198308192018031001 (Dosen Penguji III)		24 Juni 2024

HALAMAN PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 24 Juni 2024
Yang membuat pernyataan



Ainur Rafiq Abdillah
No. Reg. 1512620079



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

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 : Ainur Rafiq Abdillah

NIM : 1512620079

Fakultas/Prodi : Teknik/Pendidikan Teknik Informatika dan Komputer

Alamat email : shirainnur@gmail.com

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

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

yang berjudul :

Analisis Sentimen Terhadap Film *Dirty Vote* Berdasarkan Opini Pengguna X (Twitter) Menggunakan Metode *Convolutional Neural Network*.

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan 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, 21 Juli 2024

Penulis

(Ainur Rafiq Abdillah)

KATA PENGANTAR

Puji Syukur saya panjatkan kehadiran Allah SWT atas segala rahmat, hidayah, dan karunia-Nya yang begitu besar. Dengan rendah hati dan penuh rasa syukur, saya menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada:

1. Untuk Umi dan Abi. Terima kasih atas segala doa, dukungan, nasihat, kasih sayang yang telah kalian berikan kepadaku. Semoga karya ini menjadi bukti dedikasi saya kepada Umi dan Abi;
2. Bapak Dr. Widodo S.Kom, M.Kom dan Ibu Murien Nugraheni, S.T, M.Cs. sebagai pembimbing saya. Terima kasih atas arahan, kesabaran, bimbingan dan kepedulian yang diberikan kepada saya dalam penulisan skripsi ini;
3. Bapak dan Ibu dosen Program Studi Pendidikan Teknik Informatika dan Komputer yang telah memberikan ilmu yang bermanfaat kepada saya;
4. Teman-teman terdekat yang telah memberikan dukungan moral, semangat, dan doa-doa yang tak terhingga kepada saya;
5. Terima kasih kepada diriku sendiri yang telah berjuang, bertahan, dan berusaha yang terbaik sampai detik ini.

Semoga Allah SWT membalas segala bentuk kebaikan yang telah diberikan oleh semua pihak.

Penulis menyadari bahwa skripsi ini memiliki kekurangan dan keterbatasan. Oleh karena itu, sangat diharapkan adanya masukan dan kritik yang konstruktif dari para pembaca untuk memperbaiki dan mengembangkan penelitian ini di masa yang akan datang.

Diharapkan juga bahwa skripsi ini dapat bermanfaat dan memberikan kontribusi yang signifikan bagi perkembangan ilmu pengetahuan, khususnya dalam aspek analisis sentimen dan *Convolutional Neural Network*. Semoga hasil dari penelitian ini dapat menjadi dasar yang berguna untuk penelitian selanjutnya.

**ANALISIS SENTIMEN TERHADAP FILM *DIRTY VOTE* BERDASARKAN
OPINI PENGGUNA X (TWITTER) MENGGUNAKAN METODE
*CONVOLUTIONAL NEURAL NETWORK***

Ainur Rafiq Abdillah

Dosen Pembimbing: Dr. Widodo M.Kom dan Murien Nugraheni S.T, M.Cs.

ABSTRAK

Penelitian ini berfokus pada Analisis Sentimen Terhadap Film *Dirty Vote* Berdasarkan Opini Pengguna X (Twitter) Menggunakan Metode *Convolutional Neural Network* (CNN), berdasarkan data *tweet* dari pengguna media sosial X (Twitter). Dalam penelitian ini, total data yang digunakan adalah 1010 *tweet* yang telah melalui proses pelabelan dan *pre-processing* dengan pembagian data latih dan data uji sebesar 80 : 20. Hasil dari penelitian ini menunjukkan bahwa terdapat nilai *True Positive* (TP) sebesar 44, *True Negative* (TN) sebesar 104, *False Positive* (FP) sebesar 26, *False Negative* (FN) sebesar 28. Nilai akurasi yang diperoleh adalah 73,27% dengan *precision* sebesar 80%, *recall* sebesar 78,8%, *specificity* sebesar 62,86% dan *f1-score* sebesar 79,4%. Analisis *word cloud* menunjukkan bahwa mayoritas opini terhadap film *Dirty Vote* cenderung positif, menandakan adanya dukungan atau apresiasi yang signifikan dari masyarakat. Berdasarkan hasil ini, disarankan untuk penelitian selanjutnya agar memperluas periode pengumpulan data, mengintegrasikan lebih banyak sumber data dari media sosial lain, serta membandingkan kinerja CNN dengan algoritma *machine learning* dan *deep learning* lainnya seperti KNN, SVM, Naïve Bayes, dan LSTM. Penyempurnaan pada tahap *pre-processing* juga direkomendasikan untuk meningkatkan akurasi model, terutama dalam mengatasi penggunaan bahasa gaul dan istilah *slang word* yang sering digunakan di media sosial.

Kata Kunci: Analisis Sentimen, *Convolutional Neural Network*, *Machine Learning*, *Deep Learning*, *Dirty Vote*, *pre-processing*

***SENTIMENT ANALYSIS OF DIRTY VOTE MOVIE BASED ON USER
OPINION X (TWITTER) USING CONVOLUTIONAL NEURAL NETWORK
METHOD***

Ainur Rafiq Abdillah

Advisors: Dr. Widodo S.Kom, M.Kom dan Murien Nugraheni S.T, M.Cs.

ABSTRACT

This research focuses on Sentiment Analysis of Dirty Vote Movies Based on X User Opinions (Twitter) Using the Convolutional Neural Network (CNN) Method, based on tweet data from X social media users (Twitter). In this study, the total data used is 1010 tweets that have gone through the labeling and pre-processing process with the division of training data and test data by 80: The results of this study show that there are True Positive (TP) values of 44, True Negative (TN) of 104, False Positive (FP) of 26, False Negative (FN) of 28. The accuracy value obtained is 73.27% with a precision of 80%, recall of 78.8%, specificity of 62.86% and f1-score of 79.4%. Word cloud analysis shows that the majority of opinions on the Dirty Vote movie tend to be positive, indicating significant support or appreciation from the public. Based on these results, it is recommended for future research to expand the data collection period, integrate more data sources from other social media, and compare CNN performance with other machine learning and deep learning algorithms such as KNN, SVM, Naïve Bayes, and LSTM. Improvements in the pre-processing stage are also recommended to increase model accuracy, especially in overcoming the use of slang and slang words that are often used on social media.

Keywords: Sentiment Analysis, Convolutional Neural Network, Machine Learning, Deep Learning, Pre-Processing.

DAFTAR ISI

HALAMAN PENGESAHAN.....	i
HALAMAN PERNYATAAN	ii
KATA PENGANTAR.....	iii
ABSTRAK.....	iv
DAFTAR ISI	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
BAB I.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah	2
1.3 Pembatasan Masalah	3
1.4 Perumusan Masalah	3
1.5 Tujuan Penelitian.....	3
1.6 Manfaat Penelitian	4
BAB II.....	5
2.1 <i>Text Mining</i>	5
2.2 Analisis Sentimen	5
2.3 X (Twitter).....	6
2.4 Film <i>Dirty Vote</i>	6
2.5 <i>Web Crawling</i>	7
2.6 <i>Confusion matrix</i>	8
2.7 <i>Pre-Processing Data</i>	10
2.7.1 <i>Case folding</i>	10
2.7.2 <i>Data Cleaning</i>	10
2.7.3 <i>Tokenization</i>	11
2.7.4 <i>Normalization</i>	14
2.7.5 <i>Stopwords Removal</i>	15
2.7.6 <i>Stemming</i>	15

2.8	<i>Convolutional Neural Network</i>	16
2.8.1	<i>Layer Pada Convolutional Neural Network</i>	17
2.8.2	<i>Arsitektur Convolutional Neural Network</i>	21
2.9	Data Latih dan Data Uji	23
2.10	Penelitian yang relevan	24
2.11	Kerangka Berpikir	32
BAB III		35
3.1	Tempat dan Waktu Penelitian	35
3.2	Alat dan Bahan Penelitian	35
3.3	Jenis Penelitian	35
3.4	Diagram Alir Penelitian	35
3.4.1	Identifikasi Masalah	36
3.4.2	Studi Literatur	37
3.4.3	Pengumpulan Data	37
3.4.4	<i>Cleaning Data</i>	37
3.4.5	Pelabelan Data	37
3.4.6	<i>Pre-Processing Data</i>	39
3.4.7	Klasifikasi Menggunakan <i>Convolutional Neural Network</i>	42
3.4.8	<i>Confusion Matrix</i>	43
3.4.9	Hasil dan Evaluasi	43
BAB IV		44
4.1	Hasil Penelitian	44
4.1.1	Pengumpulan Data	44
4.1.2	<i>Cleaning Data</i>	45
4.1.3	Pelabelan Data	46
4.1.4	<i>Pre-Processing Data</i>	48
4.1.5	Klasifikasi <i>Convolutional Neural Network</i>	54
4.2	Hasil dan Evaluasi	56
BAB V		61
5.1	Kesimpulan	61
5.2	Saran	62
DAFTAR PUSTAKA		63

LAMPIRAN.....65



DAFTAR TABEL

Nomor	Judul Tabel	Halaman
2.1	<i>Confusion matrix</i>	8
2.2	Contoh <i>Confusion Matrix</i>	9
2.3	Contoh <i>Case Folding</i>	10
2.4	Contoh <i>Data Cleaning</i>	11
2.5	Contoh <i>Tokenization</i>	11
2.6	Contoh <i>Normalization</i>	15
2.7	Contoh <i>Stopwords Removal</i>	15
2.8	Contoh <i>Stemming</i>	16
3.1	Daftar komponen pada laptop	35
3.2	Contoh data <i>tweet</i> yang sudah di <i>case folding</i>	40
3.3	Contoh data <i>tweet</i> yang sudah di <i>data cleaning</i>	40
3.4	Contoh data <i>tweet</i> yang sudah di <i>normalization</i>	41
3.5	Contoh data <i>tweet</i> yang sudah di <i>stopwords removal</i>	41
3.6	Contoh data <i>tweet</i> yang sudah di <i>stemming</i>	41
3.7	Contoh data <i>tweet</i> yang sudah di <i>tokenization</i>	42
4.1	Hasil Pelabelan Data	46
4.2	Hasil <i>Case Folding</i>	49
4.3	Hasil <i>Data Cleaning</i>	51
4.4	Hasil <i>Normalization</i>	52
4.5	Hasil <i>Stopwords Removal</i>	53
4.6	Hasil <i>Stemming</i>	54
4.7	Hasil <i>Tokenization</i>	54

DAFTAR GAMBAR

Nomor	Judul Gambar	Halaman
2.1	Bagian <i>Natural Language Processing</i> (NLP)	12
2.2	<i>Convolutional Layer</i>	18
2.3	<i>Max Pooling</i>	19
2.4	<i>LeNet-5 Architecture</i>	21
2.5	<i>AlexNet Architecture</i>	22
2.6	<i>GoogleNet Architecture</i>	22
2.7	<i>Architecture CNN</i> untuk Analisis Sentimen	23
2.8	Kerangka Berpikir	34
3.1	Diagram Alir Penelitian	36
4.1	<i>Code Crawling Data</i>	44
4.2	Hasil <i>Crawling Data</i>	45
4.3	<i>Code</i> Untuk Membersihkan Data Duplikat	45
4.4	Hasil <i>Cleaning Data</i>	45
4.5	Visualisasi Perbandingan Data	48
4.6	<i>Code Read Dataframe</i>	49
4.7	<i>Code Case Folding</i>	49
4.8	<i>Code Data Cleaning</i>	50
4.9	<i>Code Normalization</i>	51
4.10	<i>Code Stopwords Removal</i>	52
4.11	<i>Code Stemming</i>	53
4.12	<i>Code Tokenization</i>	54
4.13	<i>Code</i> Pembagian Data 80 : 20	55
4.14	<i>Code Model Convolutional Neural Network</i>	56
4.15	Hasil <i>Confusion Matrix Split Data</i> 80 : 20	57
4.16	Nilai <i>Confusion Matrix</i>	57
4.17	<i>Word Cloud</i> Sentimen Positif Terhadap <i>Dirty Vote</i>	60
4.18	<i>Word Cloud</i> Sentimen Negatif Terhadap <i>Dirty Vote</i>	60

DAFTAR LAMPIRAN

Nomor	Judul Lampiran	Halaman
1	Hasil <i>tweet</i> yang telah berhasil di <i>crawling</i>	65
2	Surat Pernyataan Dosen Pembimbing	66
3	Lembar Kelayakan	68
4	Surat Tugas	69
5	Lembar Konsultasi Dosen Pembimbing	70
6	Biodata, Permohonan, Persetujuan	72
7	Daftar Hasil Studi	79
8	Pra Transkrip	82

