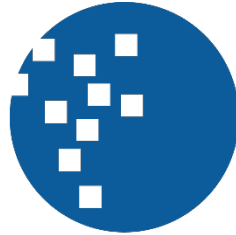


**ANALISIS SENTIMEN *TWEET* MASYARAKAT TERHADAP
PEMERINTAHAN JOKOWI MENGGUNAKAN METODE
LSTM DAN CNN**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Davina Febryanthi Kurniawan

0000047956

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

**ANALISIS SENTIMEN *TWEET* MASYARAKAT TERHADAP
PEMERINTAHAN JOKOWI MENGGUNAKAN METODE
LSTM DAN CNN**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer (S.Kom)

Davina Febryanthi Kurniawan
00000047956

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Davina Febryanthi Kurniawan

Nomor Induk Mahasiswa : 00000047956

Program studi : Sistem Informasi

Skripsi dengan judul:

ANALISIS SENTIMEN *TWEET* MASYARAKAT TERHADAP PEMERINTAHAN JOKOWI MENGGUNAKAN METODE LSTM DAN CNN

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 15 Mei 2024



Davina Febryanthi Kurniawan

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

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN *TWEET* MASYARAKAT TERHADAP PEMERINTAHAN JOKOWI MENGGUNAKAN METODE LSTM DAN CNN

Oleh

Nama : Davina Febryanthi Kurniawan
NIM : 00000047956
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

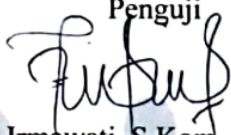
Telah diujikan pada hari Kamis, 30 Mei 2024
Pukul 09.00 s.d 11.00 dan dinyatakan
LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang


Ririn Ikana Desanti, S.Kom., M.Kom.
0313058001

Penguji


Dr. Irmawati, S.Kom., M.M.S.I.
0805097703

Pembimbing


Ir. Raymond Sunardi Detama, M.C.I.S.
0328046803

Ketua Program Studi Sistem Informasi


Ririn Ikana Desanti, S.Kom., M.Kom.
0313058001

**LEMBAR PERSETUJUAN PUBLIKASI
KARYA ILMIAH MAHASISWA**

Yang bertanda tangan dibawah ini:

Nama : Davina Febryanthi Kurniawan

Nomor Induk Mahasiswa : 00000047956

Program Studi : Sistem Informasi

Jenjang : ~~S2~~ / S1 / ~~D3~~

Judul Karya Ilmiah :

**ANALISIS SENTIMEN *TWEET* MASYARAKAT TERHADAP
PEMERINTAHAN JOKOWI MENGGUNAKAN METODE LSTM DAN
CNN**

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, dikarenakan:

Dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) *.

Tangerang, 30 Mei 2024

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Davina Febryanthi Kurniawan

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat-Nya sehingga penulis berhasil menyelesaikan laporan skripsi yang berjudul "Analisis Sentimen *Tweet* Masyarakat Terhadap Pemerintahan Jokowi Menggunakan Metode LSTM dan CNN" tepat pada waktunya. Pembuatan tugas akhir ini merupakan salah satu syarat untuk menyelesaikan pendidikan program strata satu di Universitas Multimedia Nusantara.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya dukungan, bantuan, dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Pada kesempatan ini, penulis menyampaikan terima kasih setulus-tulusnya kepada:

1. Dr. Ninok Leksono, M.A., selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, ST, M.SC., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ir. Raymond Sunardi Oetama, M.C.I.S., sebagai Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Kedua orang tua penulis, yang selalu memberikan kasih sayang, doa, nasehat, serta dukungan sehingga penulis dapat menyelesaikan tugas akhir ini.

Penulis juga mengucapkan terima kasih kepada seluruh sahabat, teman, rekan kerja, dan kepada seluruh pihak yang tidak dapat disebutkan satu-persatu yang telah mendukung kelangsungan pengerjaan hingga terealisasinya laporan skripsi ini. Akhir kata, penulis menyadari bahwa penulisan laporan ini masih jauh dari kata sempurna. Oleh karena itu, diharapkan adanya kritik dan saran yang konstruktif demi perbaikan laporan skripsi ini di masa mendatang. Dengan adanya laporan

skripsi ini, diharapkan dapat memberikan manfaat bagi pembaca, serta memberikan kontribusi pada dunia riset dan pendidikan Indonesia.

Tangerang, 15 Mei 2024



Davina Febryanthi Kurniawan



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ANALISIS SENTIMEN *TWEET* MASYARAKAT TERHADAP PEMERINTAHAN JOKOWI MENGGUNAKAN METODE LSTM DAN CNN

Davina Febryanthi Kurniawan

ABSTRAK

Pemerintahan saat ini, yang dipimpin oleh Presiden Jokowi, mendapatkan perhatian publik yang signifikan melalui platform media sosial X, menimbulkan beragam respons mulai dari pujian hingga kritik dan sindiran. Dengan mengumpulkan data dari *tweet* publik, penelitian ini bertujuan untuk memahami dan memberikan wawasan tentang sentimen positif dan negatif publik terhadap pemerintahan dan diteliti juga terkait lima prioritas kerja pemerintahan Jokowi. Penelitian ini melakukan analisis dengan metode LSTM dan CNN menggunakan dua fitur ekstraksi, yaitu TF-IDF dan *Word Embedding* serta membandingkan tujuh *optimizer*, yaitu *Adam*, *RMSprop*, *SGD*, *Adagrad*, *Adadelta*, *Adamax*, dan *Nadam*. Pada tahap terakhir penelitian ini menggunakan *confusion matrix* untuk evaluasi kinerja model LSTM dan CNN yang telah dilakukan. Model LSTM menggunakan *Word Embedding* dengan *optimizer AdaMax* menunjukkan akurasi dan tingkat *loss* yang konsisten, mencapai akurasi pelatihan 91.43% dengan *loss* 3.22% dan akurasi validasi 91.21% dengan *loss* 3,38% serta diperoleh hasil *classification report* dengan nilai akurasi sebesar 91.2%, nilai presisi 90.6%, nilai *recall* 93.4%, dan nilai *f1-score* 92%. Demikian pula, model CNN mencapai hasil terbaik dengan menggunakan fitur *Word Embedding* dan *optimizer Adamax*, menghasilkan akurasi pelatihan 98.23% dengan *loss* 0.079% dan akurasi validasi 91.47% dengan *loss* 3.57% serta diperoleh hasil *classification report* dengan nilai akurasi sebesar 91.4%, nilai presisi 91.5%, nilai *recall* 92.8%, dan nilai *f1-score* 92.1%.

Kata kunci: Analisis Sentimen, *Convolutional Neural Network*, Jokowi, *Long Short-Term Memory*, X

U 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

SENTIMENT ANALYSIS OF PUBLIC TWEETS TOWARDS JOKOWI'S GOVERNMENT USING LSTM AND CNN METHODS

Davina Febryanthi Kurniawan

ABSTRACT

The current government, led by President Jokowi, has gained significant public attention through social media platform X, eliciting a wide range of responses ranging from praise to criticism and sarcasm. By collecting data from public tweets, this study aims to understand and provide insight into the positive and negative public sentiment towards the government and is also researched related to the five work priorities of the Jokowi administration. This study conducted an analysis using the LSTM and CNN methods using two extraction features, namely TF-IDF and Word Embedding and compared seven optimizers, namely Adam, RMSprop, SGD, Adagrad, Adadelta, Adamax, and Nadam. In the last stage of this study, a confusion matrix was used to evaluate the performance of the LSTM and CNN models that had been carried out. The LSTM model using Word Embedding with AdaMax optimizer showed consistent accuracy and loss rate, achieving training accuracy of 91.43% with a loss of 3.22% and validation accuracy of 91.21% with a loss of 3.38% and obtained classification report results with an accuracy value of 91.2%, a precision value of 90.6%, a recall value 93.4%, and the F1-score score is 92%. Similarly, the CNN model achieved the best results by using the Word Embedding feature and the Adamax optimizer, resulting in a training accuracy of 98.23% with a loss of 0.079% and a validation accuracy of 91.47% with a loss of 3.57% and obtained the results of a classification report with an accuracy value of 91.4%, a precision value of 91.5%, a recall value of 92.8%, and F1-Score score 92.1%.

Keywords: Convolutional Neural Network, Jokowi, Long Short-Term Memory, Sentiment Analysis, X

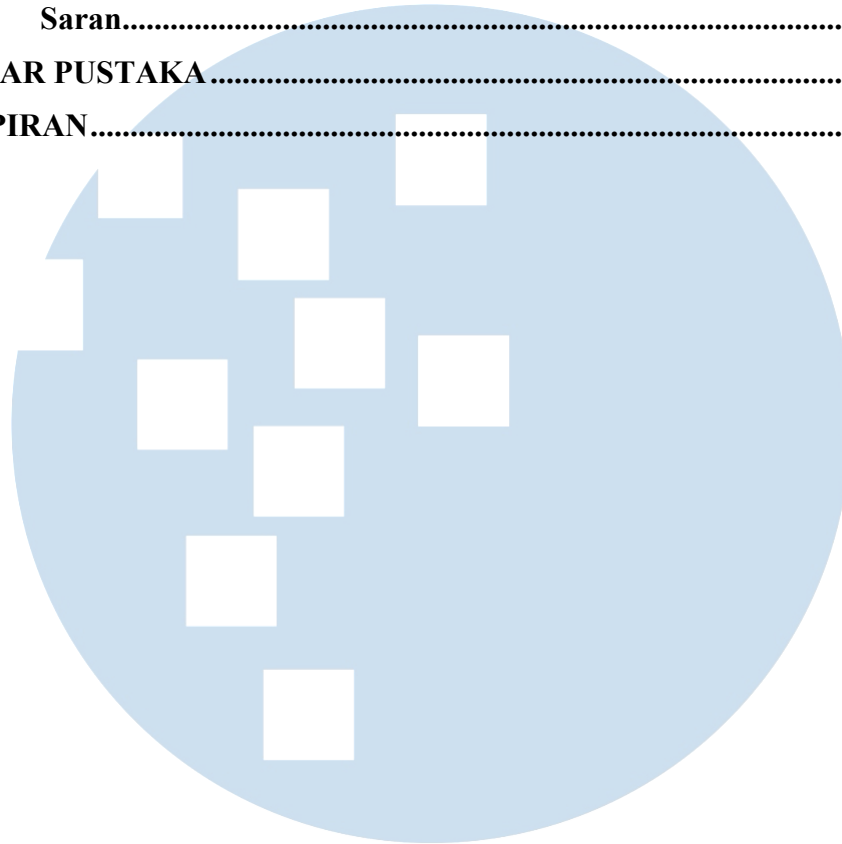
U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN	iii
LEMBAR PERSETUJUAN PUBLIKASI	iii
KATA PENGANTAR.....	v
ABSTRAK	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah	5
1.4 Tujuan dan Manfaat Penelitian.....	6
1.4.1 Tujuan Penelitian	6
1.4.2 Manfaat Penelitian	6
1.5 Sistematika Penulisan	7
BAB II LANDASAN TEORI	9
2.1 Penelitian Terdahulu	9
2.2 Analisis Sentimen	13
2.2.1 Natural Language Processing (NLP).....	14
2.2.2 Deep Learning	16
2.2.3 Web Scraping	17
2.3 <i>Cross Industry Standard Process for Data Mining (CRISP-DM)</i>.....	18
2.3.1 Business Understanding.....	19
2.3.2 Data Understanding	20
2.3.3 Data Preparation.....	20
2.3.4 Modeling.....	26
2.3.5 Evaluation	33
2.4 <i>Tools</i>	35

2.4.1	X.....	35
2.4.2	Python	35
2.4.3	Visual Studio Code.....	36
2.4.4	Tweet-harvest	37
BAB III METODOLOGI PENELITIAN		38
3.1	Gambaran Umum Objek Penelitian.....	38
3.2	Metode Penelitian.....	41
3.3	Teknik Pengumpulan Data.....	41
3.4	Teknik Analisis Data.....	42
3.4.1	Business Understanding.....	42
3.4.2	Data Understanding	42
3.4.3	Data Preparation.....	43
3.4.3.1	Data Preprocessing.....	43
3.4.3.2	Feature Extraction	45
3.4.3.3	Data Splitting	46
3.4.4	Modeling.....	46
3.4.5	Evaluation	49
BAB IV ANALISIS DAN HASIL PENELITIAN		50
4.1	<i>Business Understanding</i>	50
4.2	<i>Data Understanding</i>	50
4.2.1	<i>Data Collection</i>	51
4.2.2	<i>Data Labeling</i>	53
4.3	<i>Data Preparation</i>	55
4.3.1	<i>Data Preprocessing.....</i>	55
4.3.2	<i>Feature Extraction</i>	65
4.3.3	<i>Data Splitting.....</i>	67
4.4	<i>Modeling</i>	68
4.4.1	<i>Long Short-Term Memory.....</i>	68
4.4.2	<i>Convolutional Neural Network.....</i>	73
4.5	<i>Evaluation.....</i>	80
4.6	<i>Hasil dan Pembahasan.....</i>	83
BAB V SIMPULAN DAN SARAN.....		95

5.1	Simpulan	95
5.2	Saran.....	96
	DAFTAR PUSTAKA.....	97
	LAMPIRAN.....	103



UMMN

**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**

DAFTAR TABEL

Tabel 4.1 Atribut Data.....	52
Tabel 4.2 Perbandingan Sebelum dan Sesudah Case Folding dan Cleaning	56
Tabel 4.3 Perbandingan Sebelum dan Sesudah Tokenization.....	58
Tabel 4.4 Perbandingan Sebelum dan Sesudah Normalization.....	59
Tabel 4.5 Perbandingan Sebelum dan Sesudah Stopwords Removal.....	61
Tabel 4.6 Perbandingan Sebelum dan Sesudah Stemming.....	63
Tabel 4.7 Teks Final diberi Sentimen dan Skor Sentimen.....	65
Tabel 4.8 Arsitektur LSTM (TF-IDF) yang digunakan.....	69
Tabel 4.9 Perbandingan Optimizer pada LSTM (TF-IDF).....	69
Tabel 4.10 Arsitektur LSTM (Word Embedding) yang digunakan.....	72
Tabel 4.11 Perbandingan Optimizer pada LSTM (Word Embedding).....	72
Tabel 4.12 Arsitektur CNN (TF-IDF) yang digunakan.....	74
Tabel 4.13 Perbandingan Optimizer pada CNN (TF-IDF).....	75
Tabel 4.14 Arsitektur CNN (Word Embedding) yang digunakan.....	78
Tabel 4.15 Perbandingan Optimizer pada CNN (Word Embedding).....	78
Tabel 4.16 Perbandingan dengan Penelitian Terdahulu.....	84
Tabel 4. 17 Hasil Terbaik LSTM.....	86
Tabel 4. 18 Hasil Terbaik CNN.....	88
Tabel 4.19 Contoh Tweet Setiap Keyword.....	94

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 1.1 Survei Pengguna Sosial Media di Indonesia 2024 oleh We Are Social.....	1
Gambar 1.2 Jumlah Pengguna X di dunia 2023 oleh laporan We Are Social. 2	
Gambar 2.1 Metodologi CRISP-DM [33]	19
Gambar 2.2 Arsitektur LSTM [40]	27
Gambar 2.3 Arsitektur CNN [42]	31
Gambar 3.1 Alur Penelitian	40
Gambar 4.1 Tweet Harvest	51
Gambar 4.2 Scraping Data.....	51
Gambar 4.3 Hasil Scraping	52
Gambar 4.4 Data Labeling	54
Gambar 4.5 Hasil Labeling	54
Gambar 4.6 Proses Case Folding dan Data Cleaning.....	56
Gambar 4.7 Proses Tokenization.....	58
Gambar 4.8 Proses Normalization.....	59
Gambar 4.9 Proses Stopwords Removal.....	61
Gambar 4.10 Proses Stemming.....	63
Gambar 4.11 Proses Sentiment Labeling.....	64
Gambar 4.12 Proses TF-IDF	66
Gambar 4.13 Proses Word Embedding	66
Gambar 4.14 Proses Data Splitting.....	67
Gambar 4.15 Grafik History Training LSTM (TF-IDF) dengan Optimizer 70	
Gambar 4.16 Grafik History Training LSTM (Word Embedding) dengan Optimizer	73
Gambar 4.17 Grafik History Training CNN (TF-IDF) dengan Optimizer ...	76
Gambar 4.18 Grafik History Training CNN (Word Embedding) dengan Optimizer	79
Gambar 4.19 Classification Report LSTM.....	80
Gambar 4.20 Confusion Matrix LSTM	81
Gambar 4.21 Uji Statistik Kinerja Optimizer pada Model LSTM	81
Gambar 4.22 Classification Report CNN.....	82
Gambar 4.23 Confusion Matrix CNN	83
Gambar 4.24 Uji Statistik Kinerja Optimizer pada Model CNN.....	83
Gambar 4.25 Grafik History Training LSTM menggunakan Word Embedding dan AdaMax.....	87
Gambar 4.26 Grafik History Training CNN menggunakan Word Embedding dan AdaMax	88
Gambar 4.27 Persebaran Tweet Keyword "Pembangunan SDM"	89
Gambar 4.28 Top Words keyword "Pembangunan SDM"	89

Gambar 4.29 Persebaran Tweet Keyword "Pembangunan Infrastruktur" .	90
Gambar 4.30 Top Words Keyword "Pembangunan Infrstruktur"	90
Gambar 4.31 Persebaran Tweet Keyword "Kendala Regulasi"	91
Gambar 4.32 Top Words Keyword "Kendala Regulasi"	91
Gambar 4.33 Persebaran Tweet dengan Keyword "Transformasi Ekonomi"	92
Gambar 4.34 Top Words Keyword "Transformasi Ekonomi"	92
Gambar 4.35 Persebaran Tweet dengan Keyword "Penyederhanaan Birokrasi"	93
Gambar 4.36 Top Words Keyword "Penyerdehanaan Birokrasi"	93

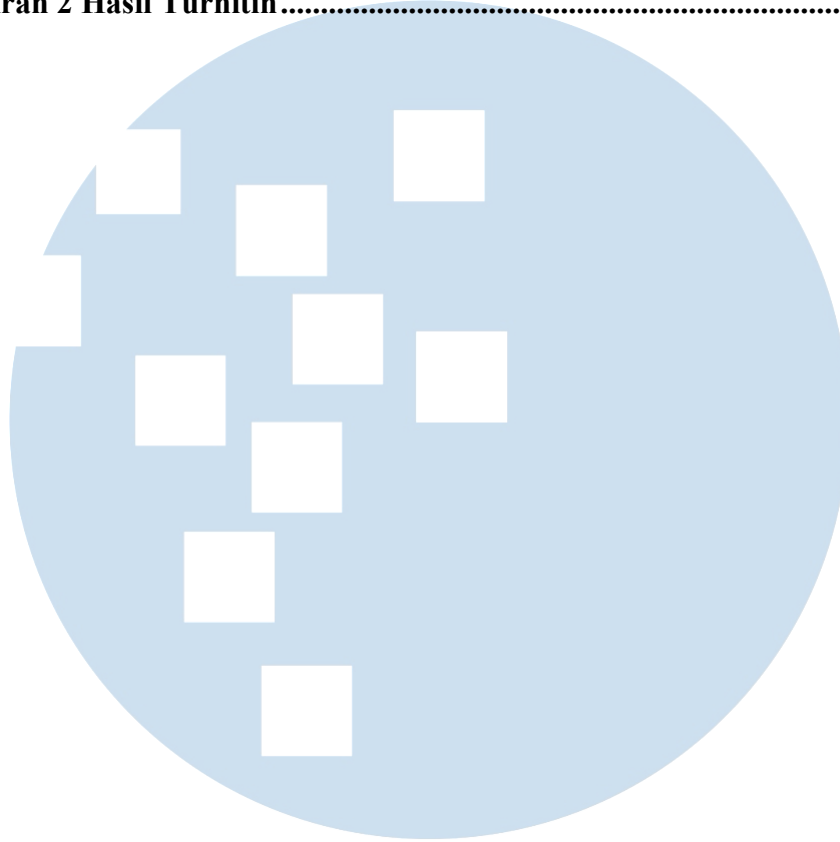


UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1 Form Bimbingan Skripsi..... 103
Lampiran 2 Hasil Turnitin..... 104



UMMN

**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**