

**Analisis Sentimen Pada Media Sosial X Terhadap  
Implementasi Kurikulum Merdeka Menggunakan Metode  
Fasttext Dan Long Short-Term Memory (LSTM)**

**SKRIPSI**



Oleh :

**Arif Fajar Pangestu**

**NPM. 20081010236**

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

**2024**

**Analisis Sentimen Pada Media Sosial X Terhadap  
Implementasi Kurikulum Merdeka Menggunakan Metode  
Fasttext Dan Long Short-Term Memory (LSTM)**

**SKRIPSI**

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



Oleh :

**Arif Fajar Pangestu**

**NPM. 20081010236**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
2024**

# LEMBAR PENGESAHAN SKRIPSI

**Judul** : Analisis Sentimen Pada Media Sosial X Terhadap Implementasi Kurikulum Merdeka Menggunakan Metode Fasttext Dan Long Short-Term Memory (LSTM)

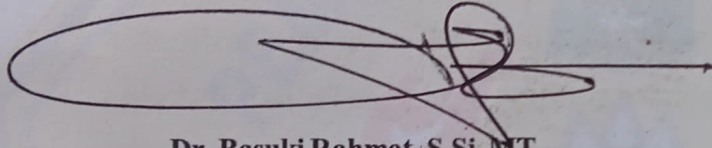
**Oleh** : Arif Fajar Pangestu

**NPM** : 20081010236

Telah Diseminarkan Dalam Ujian Skripsi Pada :  
Hari Senin, Tanggal 20 Mei 2024

## Mengetahui

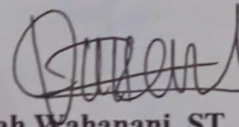
1. Dosen Pembimbing



Dr. Basuki Rahmat, S.Si, MT.  
NIP. 19690723 2021211 002

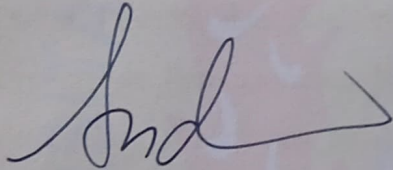
1.

Dosen Penguji



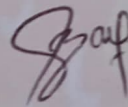
Henni Endah Wahanani, ST., M.Kom  
NIP. 19780922 2021212 005

2.



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

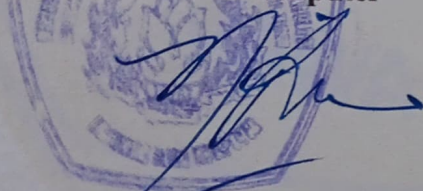
2.



Made Hanindia Prami Swari, S.Kom, M.Cs  
NIP. 19890205 2018032 001

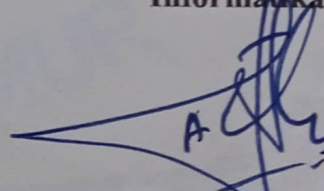
## Menyetujui

Dekan  
Fakultas Ilmu Komputer



Prof. Dr. Ir. Novifina Hendrasarie, MT  
NIP. 19681126 199403 2 001

Koordinator Program Studi  
Informatika



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

## Surat PERNYATAAN ANTI PLAGIAT

Saya, mahasiswa Program Studi Sarjana Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Arif Fajar Pangestu

NPM : 20081010236

Menyatakan dengan sesungguhnya bahwa Skripsi/Tugas Akhir yang saya kerjakan berjudul:

**“Analisis Sentimen Pada Media Sosial X Terhadap Implementasi Kurikulum Merdeka Menggunakan Metode Fasttext Dan Long Short-Term Memory (LSTM)”**

bukan merupakan plagiasi sebagian atau keseluruhan dari Skripsi/Tugas Akhir/Penelitian orang lain dari juga bukan merupakan produk dan software yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini secara keseluruhan adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional “Veteran” Jawa Timur maupun di Institut Pendidikan lain. Bukti hasil pengecekan plagiasi dokumen ini dapat ditelusuri melalui QR Code di bawah.

Apabila di kemudian hari terbukti bahwa dokumen ini merupakan plagiasi karya orang lain, saya sanggup menerima sanksi sesuai aturan yang berlaku.

Demikian atas perhatiannya disampaikan terima kasih.



Surabaya, 30 Mei 2024

Hormat saya,



Arif Fajar Pangestu

NPM. 20081010236

# **Analisis Sentimen Pada Media Sosial X Terhadap Implementasi Kurikulum Merdeka Menggunakan Metode Fasttext Dan Long Short-Term Memory (LSTM)**

**Nama Mahasiswa : Arif Fajar Pangestu**

**NPM : 20081010236**

**Program Studi : Informatika**

**Dosen Pembimbing : Dr. Basuki Rahmat, S.Si. MT.**

**Andreas Nugroho Sihananto, S.Kom., M.Kom.**

## **Abstrak**

Perubahan kurikulum adalah keharusan untuk mengikuti perkembangan zaman dan memastikan standar pendidikan terpenuhi. Namun, perubahan ini sering kali menimbulkan kebingungan di kalangan pendidik dan orang tua, yang mengganggu proses pendidikan. Kurikulum Merdeka, yang diperkenalkan sebagai inovasi penting dalam pendidikan Indonesia, menyediakan kerangka kerja yang lebih fleksibel dan sesuai dengan kebutuhan saat ini. Meskipun demikian, dengan meningkatnya jumlah peserta didik, tantangan yang dihadapi oleh sistem pendidikan Indonesia juga bertambah. Penelitian ini bertujuan untuk menganalisis opini yang muncul di media sosial X mengenai implementasi Kurikulum Merdeka, menggunakan metode word embedding FastText dan model klasifikasi Long Short-Term Memory (LSTM). Dua dataset uji coba digunakan dalam penelitian ini, yang pertama berisi 7.500 entri dan yang kedua 3.000 entri. Penelitian ini juga menguji delapan skenario yang berbeda, dengan kombinasi metode ekstraksi fitur Continuous Bag of Words (CBOW) dan Skip-Gram, serta variasi pemisahan data 80:20 dan 85:15. Hasilnya menunjukkan tingkat akurasi yang tinggi di semua skenario, di atas 85%. Temuan ini mengungkap dominasi sentimen negatif dalam setiap kategori yang diamati selama implementasi Kurikulum Merdeka, yang menunjukkan adanya beberapa tantangan atau hambatan dalam penerimaan dan aplikasi kurikulum tersebut di berbagai lingkungan pendidikan.

***Kata kunci:*** Analisis Sentimen, FastText, Kurikulum Merdeka, LSTM

## KATA PENGANTAR

Assalamu'alaikum Wr.Wb. Segala puji syukur bagi Allah SWT yang telah melimpahkan rahmat, taufik serta hidayah-Nya sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan judul:

### **"ANALISIS SENTIMEN KENDARAAN LISTRIK PADA TWITTER MENGGUNAKAN METODE LONG SHORT TERM MEMORY"**

Tujuan penyusunan skripsi ini adalah untuk memenuhi salah satu persyaratan dalam menyelesaikan Program Sarjana (S1) Jurusan Informatika di Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur.

Penulis mendapatkan berbagai dukungan, bantuan, bimbingan, dan nasehat dari berbagai pihak selama proses penulisan. Dengan rasa hormat dan ucapan terima kasih, penulis ingin menyampaikan apresiasi kepada semua yang telah membantu dalam setiap tahap penulisan skripsi, dari awal hingga akhir.

Penulis menyadari bahwa masih terdapat kekurangan dalam penelitian ini, mengingat adanya keterbatasan pengetahuan dan kemampuan. Oleh karena itu, penulis dengan tulus menerima segala bentuk kritik dan saran dari berbagai pihak untuk penyempurnaan laporan skripsi ini. Harapannya, masukan tersebut dapat memberikan manfaat bagi penulis dan para pembaca.

Surabaya, 30 Mei 2024  
Hormat saya,

Arif Fajar Pangestu  
NPM. 20081010236

## UCAPAN TERIMA KASIH

Laporan skripsi ini dapat terselesaikan dengan baik atas berkat dan rahmat Allah SWT yang telah memberikan limpahan kebaikan serta kesehatan bagi peneliti semua terutama bagi penulis agar dapat menyelesaikan skripsi sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Informatika.

Penulisan skripsi ini dapat terselesaikan tidak lepas dari bantuan, bimbingan, dukungan, dan nasehat berbagai pihak. Penulis juga berterima kasih pada semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini, baik secara langsung maupun tidak langsung.

Pada kesempatan ini penulis mengucapkan terima kasih yang setulusnya kepada:

1. Kedua Orang Tua dan Keluarga yang telah memberikan nasehat, doa, kasih sayang, dan semangat, serta memenuhi kebutuhan dan fasilitas penulis untuk memudahkan sesuatu yang berkaitan dengan perkuliahan.
2. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT, selaku Rektor Universitas Pembangunan Nasional "Veteran" Jawa Timur.
3. Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur
4. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom selaku Koordinator Program Studi Informatika Universitas Pembangunan Nasional "Veteran" Jawa Timur.
5. Bapak Dr. Basuki Rahmat, S.Si. MT. selaku Dosen Pembimbing Pertama yang telah membantu dalam penelitian skripsi serta memberikan masukan dan saran selama proses penulisan skripsi.
6. Bapak Andreas Nugroho Sihananto, S.Kom., M.Kom. selaku Dosen Pembimbing Kedua yang telah membantu dalam penelitian skripsi

serta memberikan masukan dan saran selama proses penulisan skripsi

7. Bapak dan Ibu Dosen Program Studi Informatika UPN "Veteran" Jawa Timur yang telah mendidik dan memberikan ilmunya sehingga penulis memiliki bekal untuk dapat melakukan penelitian ini.
8. Seluruh staff Fakultas Ilmu Komputer UPN Veteran Jawa Timur yang telah memberikan dukungan dalam penyelesaian rangkaian skripsi ini.
9. Teman-teman Informatika angkatan 2020 yang telah membantu, mendukung dan berjuang dalam dunia perkuliahan.
10. Pihak-pihak yang tidak dapat peneliti ucapkan menjadi penyemangat peneliti dalam menyelesaikan skripsi ini

Semoga Allah SWT memberikan balasan berupa kebaikan dunia dan akhirat kepada semua pihak yang telah peneliti sebutkan dalam menyelesaikan penelitian dan penyusunan laporan skripsi dengan baik.

Surabaya, 30 Mei 2024  
Hormat saya,

Arif Fajar Pangestu  
NPM. 20081010236



## Daftar Isi

Lembar Pengesahan SKRIPSI .....	i
Surat PERNYATAAN ANTI PLAGIAT .....	ii
KATA PENGANTAR .....	v
UCAPAN TERIMA KASIH .....	vi
Daftar Isi.....	viii
Daftar Tabel.....	xii
Daftar Gambar .....	xiv
BAB I PENDAHULUAN .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	5
1.3. Batasan Masalah .....	5
1.4. Tujuan.....	6
1.5. Manfaat.....	6
BAB II TINJAUAN PUSTAKA .....	7
2.1. Penelitian Pendukung .....	7
2.2. Analisis Sentimen.....	10
2.3. Kurikulum Merdeka .....	11
2.4. Media Sosial .....	11
2.5. <i>Web Scraping</i> .....	12
2.6. <i>Text Mining</i> .....	12
2.7. <i>Text Preprocessing</i> .....	13
2.7.1. <i>Cleaning</i> .....	13
2.7.2. <i>Case-Folding</i> .....	13

2.7.3.	<i>Tokenizing</i> .....	14
2.7.4.	<i>Stopword Removal</i> .....	14
2.7.5.	<i>Stemming</i> .....	14
2.8.	<i>Labeling</i> .....	14
2.9.	<i>Mean Absolute Error</i> .....	15
2.10.	<i>Word Embedding</i> .....	15
2.10.1.	<i>FastText</i> .....	16
2.11.	<i>Reccurent Neural Network</i> .....	18
2.12.	<i>Long Short-Term Memory</i> .....	19
2.13.	<i>Confusion Matrix</i> .....	23
<b>BAB III METODOLOGI PENELITIAN</b> .....		25
3.1.	<i>Alur Penelitian</i> .....	25
3.2.	<i>Scraping Data</i> .....	26
3.3.	<i>Preprocessing</i> .....	28
3.3.1.	<i>Cleaning</i> .....	28
3.3.2.	<i>Case Folding</i> .....	29
3.3.3.	<i>Normalization</i> .....	29
3.3.4.	<i>Stopword Removal</i> .....	30
3.3.5.	<i>Stemming</i> .....	31
3.3.6.	<i>Tokenizing</i> .....	31
3.4.	<i>Pembuatan Dataset</i> .....	32
3.5.	<i>Labelling Data</i> .....	33
3.6.	<i>Mean Absolute Error (MAE)</i> .....	34
3.7.	<i>Word Embedding FastText</i> .....	34
3.6.1.	<i>Subkata</i> .....	35

3.6.2.	N-gram.....	36
3.6.3.	Vektor representasi.....	37
3.7.	Data Spliting .....	40
3.8.	Klasifikasi Long Short-Term Memory.....	41
3.9.	Skema Pengujian.....	54
3.10.	Evaluasi Performa .....	56
3.11.	Hasil Nilai Sentimen .....	58
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>59</b>
4.1	Lingkungan Penelitian.....	59
4.1.1	Spesifikasi Perangkat Keras .....	60
4.1.2	Spesifikasi Perangkat Lunak .....	60
4.2	Scraping Data.....	60
4.3	Text Preprocessing .....	63
4.3.1	Cleaning .....	64
4.3.2	Case Folding.....	66
4.3.3	Normalization .....	67
4.3.4	Stopword Removal.....	69
4.3.5	Stemming .....	70
4.3.6	Tokenizing.....	71
4.4	Pembuatan Dataset .....	73
4.5	Pelabelan Dataset .....	75
4.6	Validasi Dataset .....	78
4.7	Ekstraksi Fitur FastText .....	80
4.7.1	Skip-Gram.....	82
4.7.2	Continuous Bag of Words .....	86

4.8	Data Splitting .....	89
4.9	Perancangan Model LSTM.....	94
4.9.1	Modeling LSTM .....	96
4.9.2	Hasil Model LSTM .....	112
4.10	Klasifikasi Data Kategori .....	113
<b>BAB V KESIMPULAN DAN SARAN.....</b>		<b>116</b>
5.1	Kesimpulan.....	116
5.2	Saran.....	116
<b>DAFTAR PUSTAKA .....</b>		<b>117</b>

## Daftar Tabel

Tabel 2. 1 Multi Class Confusion Matrix .....	23
Tabel 3. 1 Kata Kunci Kategori .....	26
Tabel 3. 2 Proses Labelling .....	34
Tabel 3. 3 Proses <i>Cleaning</i> .....	28
Tabel 3. 4 Proses <i>Case Folding</i> .....	29
Tabel 3. 5 Proses <i>Normalization</i> .....	30
Tabel 3. 6 Proses <i>Tokenizing</i> .....	32
Tabel 3. 7 Proses <i>Stopword Removal</i> .....	30
Tabel 3. 8 Proses <i>Stemming</i> .....	31
Tabel 3. 9 Proses subkata .....	35
Tabel 3. 10 Proses N-gram .....	36
Tabel 3. 11 Optimasi Hyperparamater LSTM .....	41
Tabel 3. 12 Contoh Evaluasi Sistem .....	57
Tabel 4. 1 Detail Spesifikasi Perangkat Keras .....	60
Tabel 4. 2 Detail Spesifikasi Perangkat Lunak .....	60
Tabel 4. 3 Contoh Hasil Scrapping .....	61
Tabel 4. 4 Jumlah Data Kurikulum Merdeka .....	62
Tabel 4. 5 Jumlah Hasil Filtering .....	63
Tabel 4. 6 Contoh Unprocessed Sentimen.....	64
Tabel 4. 7 Penerapan <i>Cleaning</i> .....	65
Tabel 4. 8 Penerapan <i>Case Folding</i> .....	66
Tabel 4. 9 Penerapan <i>Normalization</i> .....	68
Tabel 4. 10 Penerapan <i>Stopword Removal</i> .....	69
Tabel 4. 11 Penerapan <i>Stemming</i> .....	71
Tabel 4. 12 Penerapan <i>Tokenizing</i> .....	72
Tabel 4. 13 Jumlah Data Seimbang.....	74
Tabel 4. 14 Hasil Pelabelan Data .....	75
Tabel 4. 15 Jumlah Setiap Label Dataset 7.500.....	77
Tabel 4. 16 Jumlah label Dataset 3.000.....	78
Tabel 4. 17 Data MAE .....	79
Tabel 4. 18 Word Vector Data Pertama Skip-gram .....	84

Tabel 4. 19 Word Vector Data Pertama Continuous Bag of Words .....	88
Tabel 4. 20 Skenario.....	96
Tabel 4. 21 Perbandingan Performance Skenario .....	113
Tabel 4. 22 Hasil Analisis Sentiment Setiap kategori.....	115

## Daftar Gambar

Gambar 2. 1	Arsitektur Model FastText Continuous Bag-of-Word	17
Gambar 2. 2	Arsitektur Model FastText Skip Gram	18
Gambar 2. 3	Arsitektur Metode RNN	19
Gambar 2. 4	Arsitektur Metode LSTM	20
<u>Gambar 3. 1</u>	<u>Alur Penelitian</u>	25
<u>Gambar 3. 2</u>	<u>Alur Scraping</u>	26
<u>Gambar 3. 3</u>	<u>Media Sosial X</u>	27
<u>Gambar 3. 4</u>	<u>Proses Scraping Data</u>	27
<u>Gambar 3. 5</u>	<u>Data Ulasan</u>	33
<u>Gambar 3. 6</u>	<u>Alur <i>Preprocessing</i></u>	28
Gambar 3. 7	Alur arsitektur Fasttext	35
Gambar 3. 8	Vector N-gram	40
Gambar 3. 9	word embedding	40
Gambar 3. 10	Alur Pada Skenario Pertama	55
Gambar 3. 11	Alur Pada skenario Kedua	56
Gambar 3. 13	Evaluasi Performa	57
Gambar 4. 1	Visualisasi Jumlah Setiap Label Dataset 7.500	77
Gambar 4. 2	Visualisasi Jumlah Setiap Label Dataset 3.000	77
Gambar 4. 3	Visualisasi MAE pada label	80
Gambar 4. 4	Wordcloud Kata yang Sering keluar pada dataset 7.500	81
Gambar 4. 5	Wordcloud Kata yang Sering Keluar pada dataset 3.000	81
Gambar 4. 6	Visualisasi Label Dataset 7.500	82
Gambar 4. 7	Word Vectors Menggunakan Skip-gram	84
Gambar 4. 8	Pesebaran Kata Skip-gram	85
Gambar 4. 9	Convers Tokenizing Kata Skip-gram	86
Gambar 4. 10	Word Vectors Menggunakan Continuous bag of words	88
Gambar 4. 11	Pesebaran Kata Continuous Bag of Words	89
Gambar 4. 12	Visualisasi Jumlah setiap Label dataset 7.500 80:20	90
Gambar 4. 13	Visualisasi Jumlah setiap Label dataset 7.500 85:15	91
Gambar 4. 14	Visualisasi Jumlah setiap Label dataset 3.000 80:20	92
Gambar 4. 15	Visualisasi Jumlah setiap Label dataset 3.000 85:15	92

Gambar 4. 16 Visualisasi oversampling .....	93
Gambar 4. 17 Model Sumary Long Short Term Memory .....	95
Gambar 4. 18 Pelatian skenario pertama .....	97
Gambar 4. 19 Kurva Training Loss dan Accuracy Skenario Pertama .....	97
Gambar 4. 20 Nilai Skenario Pertama .....	98
Gambar 4. 21 Confusion Matrix Skenario Pertama .....	98
Gambar 4. 22 Pelatian skenario Kedua .....	99
Gambar 4. 23 Kurva Training Loss dan Accuracy Skenario Kedua .....	99
Gambar 4. 24 Nilai Skenario Kedua .....	100
Gambar 4. 25 Confusion Matrix Skenario Kedua .....	100
Gambar 4. 26 Pelatian skenario Ketiga .....	101
Gambar 4. 27 Kurva Training Loss dan Accuracy Skenario Ketiga .....	101
Gambar 4. 28 Nilai Skenario Ketiga .....	102
Gambar 4. 29 Confusion Matrix Skenario Ketiga .....	102
Gambar 4. 30 Pelatian skenario Keempat .....	103
Gambar 4. 31 Kurva Training Loss dan Accuracy Skenario Keempat .....	103
Gambar 4. 32 Nilai Skenario Keempat .....	104
Gambar 4. 33 Confusion Matrix Skenario Keempat .....	104
Gambar 4. 34 Pelatian skenario Kelima .....	105
Gambar 4. 35 Kurva Training Loss dan Accuracy Skenario Kelima .....	105
Gambar 4. 36 Nilai Skenario Kelima .....	106
Gambar 4. 37 Confusion Matrix Skenario Kelima .....	106
Gambar 4. 38 Pelatian skenario Keenam .....	107
Gambar 4. 39 Kurva Training Loss dan Accuracy Skenario Keenam .....	107
Gambar 4. 40 Nilai Skenario Keenam .....	108
Gambar 4. 41 Confusion Matrix Skenario Keenam .....	108
Gambar 4. 42 Pelatian skenario ketujuh .....	109
Gambar 4. 43 Kurva Training Loss dan Accuracy Skenario Ketujuh .....	109
Gambar 4. 44 Nilai Skenario Ketujuh .....	110
Gambar 4. 45 Confusion Matrix Skenario Ketujuh .....	110
Gambar 4. 46 Pelatian skenario kedelapan .....	111
Gambar 4. 47 Kurva Training Loss dan Accuracy Skenario Kedelapan .....	111
Gambar 4. 48 Nilai Skenario Kedelapan .....	112



Gambar 4. 49 Confusion Matrix Skenario Kedelapan .....	112
Gambar 4. 50 Visualisasi Jumlah data berdasarkan label di setiap kategori .....	115