

**ANALISIS SENTIMEN PADA KOMENTAR YOUTUBE
MENGENAI MOBIL LISTRIK MENGGUNAKAN
ALGORITMA LSTM DAN CNN**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Skripsi

Zefanya Febrina Valentine Mehe

0000061664

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

2025

**ANALISIS SENTIMEN PADA KOMENTAR YOUTUBE
MENGENAI MOBIL LISTRIK MENGGUNAKAN
ALGORITMA LSTM DAN CNN**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer

Zefanya Febrina Valentine Mehe

00000061664

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

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Zefanya Febrina Valentine Mehe

Nomor Induk Mahasiswa : 00000061664

Program Studi : Sistem Informasi

Skripsi dengan judul:

Analisis Sentimen pada Komentar Youtube Mengenai
Mobil Listrik Menggunakan Algoritma
LSTM dan CNN

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis 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 maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 05 Juni 2025



(Zefanya Febrina Valentine Mehe)

HALAMAN PERSETUJUAN

Skripsi dengan judul

Analisis Sentimen pada Komentar Youtube Mengenai
Mobil Listrik Menggunakan Algoritma
LSTM dan CNN

Oleh

Nama : Zefanya Febrina Valentine Mehe
NIM : 00000061664
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah disetujui untuk diajukan pada
Sidang Ujian Skripsi Universitas Multimedia Nusantara
Tangerang, 05 Juni 2025

Pembimbing



Dr. Erick Fernando, S.Kom, M.S.I
1029118501

Ketua Program Studi Sistem Informasi



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

HALAMAN PENGESAHAN

Skripsi dengan judul
Analisis Sentimen pada Komentar Youtube Mengenai
Mobil Listrik Menggunakan Algoritma
LSTM dan CNN

Oleh

Nama : Zefanya Febrina Valentine Mehe
NIM : 00000061664
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Senin, 23 Juni 2025
Pukul 08.00 s.d 10.00 dan dinyatakan
LULUS
Dengan susunan penguji sebagai berikut.

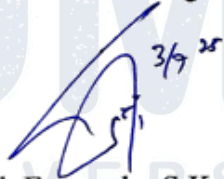
Ketua Sidang

Penguji



Melissa Indah Fianty, S.Kom., MMSI.
0313019201

Pembimbing



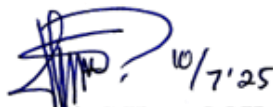
Dr. Erick Fernando, S.Kom, M.S.I
1029118501

Ketua Program Studi Sistem Informasi



3/7/2023

Ahmad Faza, S.Kom., M.T.I.
0312019501



Ririn Ikana Desanti, S.Kom., M.Kom.
313058001

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Zefanya Febrina Valentine Mehe
NIM : 00000061664
Program Studi : Sistem Informasi
Jenjang : D3/S1/S2* (pilih salah satu)
Judul Karya Ilmiah : Analisis Sentimen pada Komentar Youtube
Mengenai Mobil Listrik Menggunakan
Algoritma LSTM dan CNN

Menyatakan dengan sesungguhnya bahwa saya bersedia* (pilih salah satu):

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 05 Juni 2025


(Zefanya Febrina Valentine Mehe)

* Pilih salah satu

** Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji Syukur atas selesainya penulisan Laporan Skripsi ini dengan judul: “Analisis Sentimen pada Komentar Youtube Mengenai Mobil Listrik Menggunakan Algoritma LSTM dan CNN” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Jurusan Sistem Informasi Pada Fakultas Teknik & Informatika Universitas Multimedia Nusantara. Penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi penulis untuk menyelesaikan tugas akhir ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., 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. Bapak Dr. Erick Fernando, S.Kom, M.S.I, sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya skripsi ini.
5. Keluarga yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan skripsi ini.
6. Rekan di kampus serta orang-orang terdekat yang telah memberikan masukan, dukungan, serta informasi yang mendukung sehingga penulis dapat menyelesaikan laporan skripsi ini.

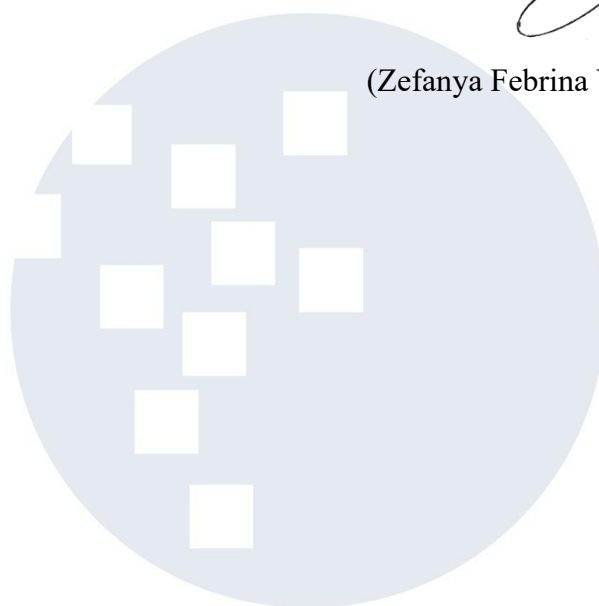
Penulis menyadari bahwa masih terdapat ketidaksempurnaan dalam laporan ini. Oleh karena itu, penulis menerima segala masukan, kritik, dan saran agar penulis dapat membuat karya tulis yang lebih baik kedepannya.

Semoga laporan ini dapat menjadi acuan, pembelajaran, serta memberikan informasi atau pengetahuan yang bermanfaat bagi para pembaca.

Tangerang, 05 Juni 2025



(Zefanya Febrina Valentine Mehe)



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

ANALISIS SENTIMEN PADA KOMENTAR YOUTUBE MENGENAI MOBIL LISTRIK MENGGUNAKAN ALGORITMA LSTM DAN CNN

(Zefanya Febrina Valentine Mehe)

ABSTRAK

Perubahan iklim global meningkatkan perhatian pada emisi karbon dari kendaraan berbahan bakar fosil. Mobil listrik hadir sebagai solusi ramah lingkungan karena mampu mengurangi emisi gas rumah kaca dan polusi udara. Namun, adopsi di Indonesia masih menghadapi tantangan seperti infrastruktur pengisian daya yang terbatas dan harga yang relatif tinggi, meskipun penjualannya menunjukkan peningkatan. Persepsi masyarakat di media sosial, khususnya YouTube, penting untuk dianalisis lebih lanjut.

Penelitian ini menggunakan algoritma *Long Short-Term Memory* (LSTM) dan *Convolutional Neural Network* (CNN) untuk analisis sentimen komentar YouTube tentang mobil listrik. LSTM unggul dalam memproses data sekuensial, sedangkan CNN efektif menangkap fitur lokal. Untuk mengatasi ketidakseimbangan kelas data, digunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE).

Hasil penelitian menunjukkan model CNN dengan SMOTE memberikan kinerja terbaik, dengan akurasi 78,08% dan *macro f1-score* sebesar 0,77, mengungguli LSTM yang hanya mencapai 74,04%. Sentimen masyarakat mencerminkan pandangan positif, seperti efisiensi energi dan pengurangan emisi. Namun, hambatan utama tetap pada harga tinggi dan kurangnya infrastruktur. Studi ini memberikan wawasan bagi produsen, pembuat kebijakan, serta konsumen dalam mendorong adopsi mobil listrik di Indonesia.

Kata kunci: Analisis Sentimen, CNN, LSTM, Mobil Listrik

SENTIMENT ANALYSIS ON YOUTUBE COMMENTS

ABOUT ELECTRIC VEHICLES USING

LSTM AND CNN

(Zefanya Febrina Valentine Mehe)

ABSTRACT (English)

Global climate change has increased awareness of carbon emissions from fossil fuel-powered vehicles. Electric vehicles (EVs) have emerged as an environmentally friendly solution, capable of reducing greenhouse gas emissions and air pollution. However, EV adoption in Indonesia still faces challenges such as limited charging infrastructure and relatively high prices, although sales have shown an upward trend. Public perception on social media, particularly YouTube, is important to analyze further.

This study employs Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) algorithms to analyze sentiment in YouTube comments regarding electric vehicles. LSTM excels at processing sequential data, while CNN is effective in capturing important local features. To address class imbalance in the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) is applied.

The results show that the CNN model with SMOTE achieved the best performance, with a test accuracy of 78.08% and a macro F1-score of 0.77, outperforming LSTM, which only reached 74.04%. Public sentiment reflects positive views such as energy efficiency and emission reduction. However, key obstacles remain in high prices and lack of infrastructure. This study provides valuable insights for EV manufacturers, policymakers, and consumers to support the adoption of electric vehicles in Indonesia.

Keywords: CNN, Electric Vehicle, LSTM, Sentiment Analysis

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
<i>ABSTRACT (English)</i>	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan dan Manfaat Penelitian	4
1.4.1 Tujuan Penelitian.....	4
1.4.2 Manfaat Penelitian.....	5
1.5 Sistematika Penulisan	5
BAB II LANDASAN TEORI	7
2.1 Penelitian Terdahulu.....	7
2.2 Teori tentang Topik Skripsi	14
2.2.1 Analisis Sentimen	14
2.2.2 Mobil Listrik	15
2.2.3 Youtube.....	16
2.2.4 CRISP-DM	17
2.2.5 API Data <i>Extraction</i>	18
2.2.6 <i>Text Preprocessing</i>	19
2.2.7 <i>Semi Supervised Learning</i>	20
2.2.8 <i>Term Frequency – Inverse Document Frequency (TF-IDF)</i>	21

2.2.9 <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	22
2.2.10 <i>Confusion Matrix</i>	22
2.2.11 <i>User Acceptance Test</i>	23
2.3 Teori tentang Framework/Algoritma yang digunakan	24
2.3.1 Streamlit	24
2.3.2 <i>Long Short-Term Memory (LSTM)</i>	25
2.3.3 <i>Convolutional Neural Network (CNN)</i>	27
2.4 Teori tentang tools/software yang digunakan.....	28
2.4.1 Google Colaboratory	28
2.4.2 Visual Studio Code	28
BAB III METODOLOGI PENELITIAN	30
3.1 Gambaran Umum Objek Penelitian.....	30
3.2 Metode Penelitian	32
3.2.1 <i>Business Understanding</i>	33
3.2.2 <i>Data Understanding</i>	33
3.2.3 <i>Data Preparation</i>	35
3.2.4 <i>Modelling</i>	39
3.2.5 <i>Evaluation</i>	39
3.2.6 <i>Deployment</i>	40
3.2.7 <i>User Acceptance Test</i>	40
3.3 Teknik Pengumpulan Data.....	42
3.4 Teknik Analisis Data	43
BAB IV ANALISIS DAN HASIL PENELITIAN	44
4.1 <i>Business Understanding</i>	44
4.2 <i>Data Understanding</i>	44
4.3 <i>Data Preparation</i>	46
4.3.1 <i>Text Cleaning & Case Folding</i>	46
4.3.2 <i>Tokenization</i>	47
4.3.3 <i>Stemming</i>	48
4.3.4 <i>Handle Slang</i>	49
4.3.5 <i>Stopword Removal</i>	50
4.3.6 <i>Labeling</i>	51
4.3.7 <i>Data Splitting</i>	55

4.3.8 Penerapan TF-IDF Untuk Ekstraksi Fitur	55
4.4 <i>Modelling</i>	56
4.4.1 <i>Modelling</i> dengan Algoritma LSTM	56
4.4.2 <i>Modelling</i> dengan Algoritma LSTM & SMOTE	58
4.4.3 <i>Modelling</i> dengan Algoritma CNN.....	59
4.4.4 <i>Modelling</i> dengan Algoritma CNN & SMOTE.....	60
4.5 <i>Evaluation</i>	61
4.5.1 <i>Evaluation</i> Model LSTM.....	61
4.5.2 <i>Evaluation</i> Model LSTM & SMOTE	63
4.5.3 <i>Evaluation</i> Model CNN.....	66
4.5.4 <i>Evaluation</i> Model CNN & SMOTE.....	68
4.5.5 Penyimpanan Model dan Ekstraksi Fitur	70
4.6 <i>Deployment</i>	70
4.7 <i>User Acceptance Test</i>	76
4.8 Hasil Penelitian	78
4.9 Pembahasan Hasil Penelitian	81
BAB V SIMPULAN DAN SARAN	85
5.1 Simpulan.....	85
5.2 Saran.....	86
DAFTAR PUSTAKA	88
LAMPIRAN.....	94



DAFTAR TABEL

Tabel 2.1 Tabel Jurnal <i>Text Mining & Sentiment Analysis</i>	7
Tabel 2.2 Tabel Jurnal <i>Naive Bayes & Community Detection</i>	7
Tabel 2.3 Tabel Jurnal Sentimen Insentif Pajak EV	8
Tabel 2.4 Tabel Jurnal Sentimen Analisis CNN	9
Tabel 2.5 Tabel Jurnal VADER & LDA <i>Topic Modeling</i>	10
Tabel 2.6 Tabel Jurnal Perbandingan SVM dan <i>Naive Bayes</i>	10
Tabel 2.7 Tabel Jurnal Analisis Sentimen dengan <i>Naive Bayes</i>	11
Tabel 2.8 Tabel Jurnal Sentimen EV <i>FastText & IndoBERT</i>	12
Tabel 2.9 Tabel Jurnal Akurasi <i>Logistic Regression</i> dengan PCA	12
Tabel 2.10 Tabel Jurnal Evaluasi LSTM dalam Analisis Sentimen EV	13
Tabel 4.1 Tabel Kamus <i>Slang</i>	49
Tabel 4.2 Tabel Daftar Pertanyaan dan Hasil Jawaban UAT	76
Tabel 4.3 Tabel Hasil Evaluasi Model	78
Tabel 4.4 Tabel Perbandingan Model	81



DAFTAR GAMBAR

Gambar 3.1 Info Sumber Data Yang Digunakan	31
Gambar 3.2 Alur Metode Penelitian	32
Gambar 3.3 Contoh Komen Netral	34
Gambar 3.4 Tampilan Beberapa Komentar Asli.....	42
Gambar 4.1 Fungsi <i>Scraping Dataset</i>	44
Gambar 4.2 Hasil <i>Scraping Dataset</i>	45
Gambar 4.3 <i>Code Text Cleaning</i> dan <i>Case Folding</i>	46
Gambar 4.4 Hasil <i>Text Cleaning</i> dan <i>Case Folding</i>	47
Gambar 4.5 <i>Code Tokenization</i>	47
Gambar 4.6 Hasil <i>Tokenization</i>	48
Gambar 4.7 <i>Code Stemming</i>	48
Gambar 4.8 Hasil <i>Stemming</i>	48
Gambar 4.9 <i>Code Handle Slang</i>	49
Gambar 4.10 Hasil <i>Handle Slang</i>	50
Gambar 4.11 <i>Code Stopword Removal</i>	50
Gambar 4.12 Hasil <i>Stopword Removal</i>	51
Gambar 4.13 <i>Labeling Self-Training XGBoost</i>	51
Gambar 4.14 Histogram <i>Confidence Score</i>	52
Gambar 4.15 <i>Bar Chart</i> Distribusi Sentimen.....	53
Gambar 4.16 <i>Wordcloud</i> Positif.....	54
Gambar 4.17 <i>Wordcloud</i> Negatif.....	54
Gambar 4.18 <i>Train & Test Split Data</i>	55
Gambar 4.19 Penerapan TF-IDF.....	56
Gambar 4.20 <i>Modelling LSTM</i>	57
Gambar 4.21 <i>Modelling LSTM</i> dengan SMOTE.....	58
Gambar 4.22 <i>Modelling CNN</i>	59
Gambar 4.23 <i>Modelling CNN</i> dengan SMOTE.....	60
Gambar 4.24 Hasil Evaluasi Model LSTM.....	62
Gambar 4.25 Grafik Kinerja Model LSTM	63
Gambar 4.26 Hasil Evaluasi Model LSTM dengan SMOTE	64
Gambar 4.27 Grafik Kinerja Model LSTM dengan SMOTE	65
Gambar 4.28 Hasil Evaluasi Model CNN.....	66
Gambar 4.29 Grafik Kinerja Model CNN.....	67
Gambar 4.30 Hasil Evaluasi Model CNN dengan SMOTE.....	68
Gambar 4.31 Grafik Kinerja Model CNN dengan SMOTE	69
Gambar 4.32 <i>Save Model</i> dan TF-IDF.....	70
Gambar 4.33 Tampilan <i>Input Link</i> dan Selama Proses.....	71
Gambar 4.34 Tampilan Tabel Data Yang Sudah di Proses.....	72
Gambar 4.35 Tampilan Visualisasi pada <i>Deploy</i>	73
Gambar 4.36 Tampilan Hasil Analisis Sentimen Positif Dengan SahabatAI.....	74
Gambar 4.37 Tampilan Hasil Analisis Sentimen Negatif Dengan SahabatAI.....	75

DAFTAR RUMUS

Rumus 2.1 Rumus <i>Term Frequency</i>	21
Rumus 2.2 Rumus <i>Inverse Document Frequency</i>	21
Rumus 2.3 Rumus TF-IDF.....	21
Rumus 2.4 Rumus <i>Accuracy</i>	22
Rumus 2.5 Rumus <i>Precision</i>	23
Rumus 2.6 Rumus <i>Recall</i>	23
Rumus 2.7 Rumus <i>F1-Score</i>	23
Rumus 2.8 Rumus <i>Forget Gate</i>	25
Rumus 2.9 Rumus <i>Input Gate</i>	26
Rumus 2.10 Rumus <i>Update Cell State</i>	26
Rumus 2.11 Rumus <i>Output Gate</i>	26
Rumus 2.12 Rumus <i>Convolutional Layer</i>	27
Rumus 2.13 Rumus <i>Activation Layer</i>	27
Rumus 2.14 Rumus <i>Pooling Layer</i>	28
Rumus 2.15 Rumus <i>Fully Connected Layer</i>	28
Rumus 2.16 Rumus <i>Output Layer</i>	28



DAFTAR LAMPIRAN

Lampiran A Hasil Turnitin..... 94
Lampiran B Formulir Bimbingan 105



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA