

**ANALISIS SENTIMEN PENDAPAT MASYARAKAT
TERHADAP BMKG SEBAGAI LAYANAN PENYEDIA
INFORMASI CUACA DAN IKLIM MENGGUNAKAN
METODE ALGORITMA *LSTM***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

RIAN ANTONO

21.11.4007

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN PENDAPAT MASYARAKAT TERHADAP
BMKG SEBAGAI LAYANAN PENYEDIA INFORMASI CUACA
DAN IKLIM MENGGUNAKAN METODE ALGORITMA *LSTM***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

RIAN ANTONO

21.11.4007

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PENDAPAT MASYARAKAT TERHADAP
BMKG SEBAGAI LAYANAN PENYEDIA INFORMASI CUACA
DAN IKLIM MENGGUNAKAN METODE ALGORITMA *LSTM***

yang disusun dan diajukan oleh

Rian Antono

21.11.4007

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Juli 2025

Dosen Pembimbing,

Rizqi Sukma Kharisma, S.Kom., M.Kom.
NIK. 190302215

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PENDAPAT MASYARAKAT TERHADAP
BMKG SEBAGAI LAYANAN PENYEDIA INFORMASI CUACA
DAN IKLIM MENGGUNAKAN METODE ALGORITMA *LSTM***

yang disusun dan diajukan oleh

Rian Antono

21.11.4007

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Bayu Setiaji, M.Kom.
NIK. 190302216

Irina Rofni Wulandari, S.Pd., M.Eng.
NIK. 190302329

Rizqi Sukma Kharisma, S.Kom., M.Kom.
NIK. 190302215

Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rian Antono
NIM : 21.11.4007

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN PENDAPAT MASYARAKAT TERHADAP BMKG SEBAGAI LAYANAN PENYEDIA INFORMASI CUACA DAN IKLIM MENGUNAKAN METODE ALGORITMA LSTM

Dosen Pembimbing : Rizqi Sukma Kharisma, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 Juli 2025

Yang Menyatakan,



Rian Antono

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam kepada Tuhan Yang Maha Esa (Allah SWT), karya sederhana ini kupersembahkan untuk:

1. Terima kasih Bapak dan Ibu tercinta, atas, doa, pengorbanan, dan dukungan yang tiada henti. Terima kasih telah menjadi sumber kekuatan dan semangat dalam setiap langkah hidupku.
2. Keluarga besarku, yang selalu memberikan doa dan motivasi selama proses penyusunan skripsi ini.
3. Dosen pembimbing, yang telah sabar membimbing dan memberikan ilmu serta arah selama penyusunan skripsi ini.
4. Teman-teman seperjuangan WDM dimas, alif, dika, rio, hizqil dan andri yang selalu hadir memberi semangat, tawa, dan cerita selama masa kuliah
5. Diriku sendiri, yang telah berjuang, menghabiskan banyak kopi, kurang tidur, dan hingga pada akhirnya sampai di titik ini. Semoga karya ini dapat bermanfaat dan menjadi awal dari langkah-langkah besar berikutnya.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Rizqi Sukma Kharisma, S.Kom., M.Kom., selaku dosen pembimbing yang telah dengan sabar membimbing, memberikan arahan, serta membantu penulis dalam penyusunan skripsi ini.
2. Tim Dosen Penguji Skripsi, yang telah memberikan masukan dan saran yang sangat berharga bagi perbaikan dan penyempurnaan karya ini.
3. Orang tua dan keluarga penulis, atas doa, dukungan, dan kasih sayang yang tiada henti dalam mendukung proses studi penulis hingga penyusunan skripsi ini.
4. Pihak-pihak lain yang terkait secara langsung, yang telah memberikan bantuan dan dukungan dalam proses penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis membuka diri terhadap segala bentuk masukan yang membangun. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak yang membacanya.

Yogyakarta, 11 Juli 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xiii
DAFTAR LAMBANG DAN SINGKATAN	xiv
DAFTAR ISTILAH	xv
INTISARI	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori	15
2.2.1 Analisis Sentimen	15

2.2.2 <i>Natural Language Processing (NLP)</i>	15
2.2.3 <i>Deep Learning</i>	16
2.2.4 <i>Exploratory Data Analysis (EDA)</i>	16
2.2.5 <i>Word Embedding</i>	16
2.2.6 <i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	17
2.2.6 <i>Recurrent Neural Network (RNN)</i>	17
2.2.7 <i>Long Short-Term Memory (LSTM)</i>	18
2.2.8 <i>Confusion Matrix</i>	21
BAB III METODE PENELITIAN	23
3.1 <i>Objek Penelitian</i>	23
3.2 <i>Alur Penelitian</i>	23
3.2.1 <i>Scraping Data Twitter</i>	25
3.2.2 <i>Preprocessing Data</i>	25
3.2.3 <i>Labeling</i>	26
3.2.4 <i>EDA</i>	26
3.2.5 <i>Pre-Modeling</i>	27
3.2.6 <i>Modeling LSTM</i>	28
3.2.7 <i>Model Evaluation</i>	31
3.2.8 <i>Deployment</i>	31
3.3 <i>Alat dan Bahan</i>	31
3.3.1 <i>Bahan Penelitian</i>	31
3.3.2 <i>Alat Penelitian</i>	32
BAB IV HASIL DAN PEMBAHASAN	33
4.1 <i>Scraping Data X</i>	33
4.2 <i>Preprocessing Data</i>	34
4.2.1 <i>Text Cleaning</i>	34
4.2.2 <i>Case Folding</i>	36

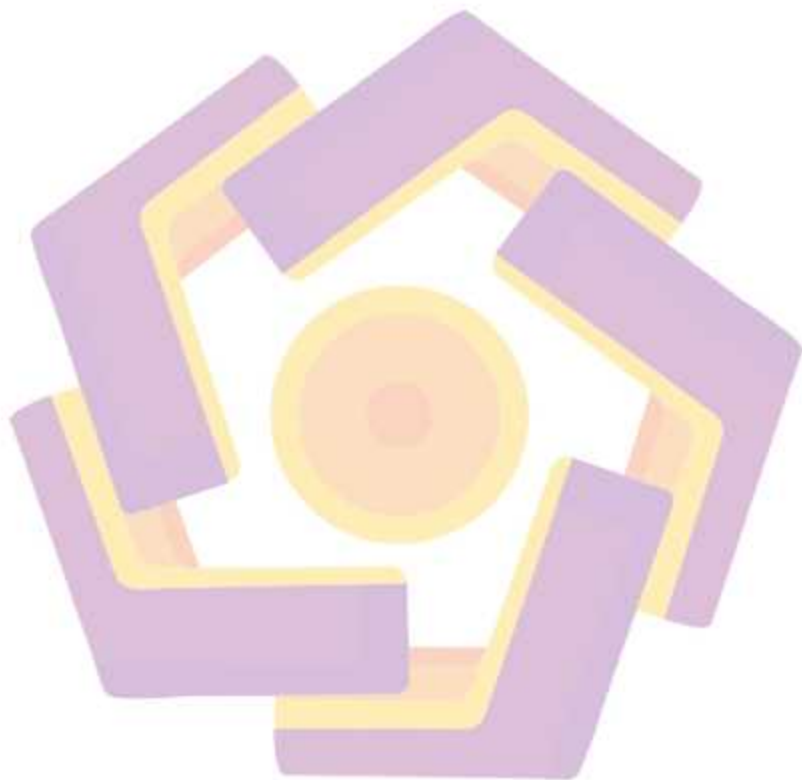
4.2.3 <i>Normalized</i>	37
4.2.4 <i>Stopword Removal</i>	39
4.2.5 <i>Tokenizing</i>	40
4.2.6 <i>Stemming</i>	41
4.3 <i>Labeling</i>	43
4.4 <i>EDA</i>	44
4.4.1 <i>Distribusi Sentimen</i>	44
4.4.2 <i>Word Cloud Keseluruhan Data</i>	46
4.4.3 <i>Word Cloud Sentimen Positif dan Negatif</i>	47
4.4.4 <i>Distribusi Panjang Tweet</i>	49
4.5 <i>Pre modeling</i>	49
4.5.1 <i>Splitting Dataset</i>	50
4.5.2 <i>Tokenized & Padding</i>	51
4.5.3 <i>Word Embedding</i>	52
4.5.4 <i>SMOTE</i>	54
4.6 <i>Modeling LSTM</i>	56
4.7 <i>Model Evaluation</i>	59
4.8 <i>Deployment</i>	61
4.9 <i>Perbandingan Penelitian</i>	62
BAB V PENUTUP.....	65
5.1 <i>Kesimpulan</i>	65
5.2 <i>Saran</i>	65
REFERENSI.....	67
LAMPIRAN.....	70

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	9
Tabel 2. 2 <i>Confusion Matrix</i>	21
Tabel 3.1 Alat Penelitian	32
Tabel 4. 1 Hasil Scraping <i>Dataset X</i>	33
Tabel 4. 2 Data Sebelum <i>Text_Cleaning</i>	35
Tabel 4. 3 Data Sesudah <i>Text Cleaning</i>	36
Tabel 4. 4 Hasil <i>Case Folding</i>	37
Tabel 4. 5 Hasil <i>Normalized</i>	38
Tabel 4. 6 Hasil <i>Stopword Removal</i>	39
Tabel 4. 7 Hasil <i>Tokenizing</i>	41
Tabel 4. 8 Hasil <i>Stemming</i>	42
Tabel 4. 9 Hasil <i>Labeling</i>	43
Tabel 4. 10 Hasil <i>Tokenized & Padding</i>	52
Tabel 4.11 Hasil <i>Word Embedding</i>	53
Tabel 4. 12 Hasil <i>Training & Validation LSTM</i>	57
Tabel 4. 13 <i>Evaluasi model LSTM</i>	60
Tabel 4. 14 Perbandingan dengan Penelitian Sebelumnya	63

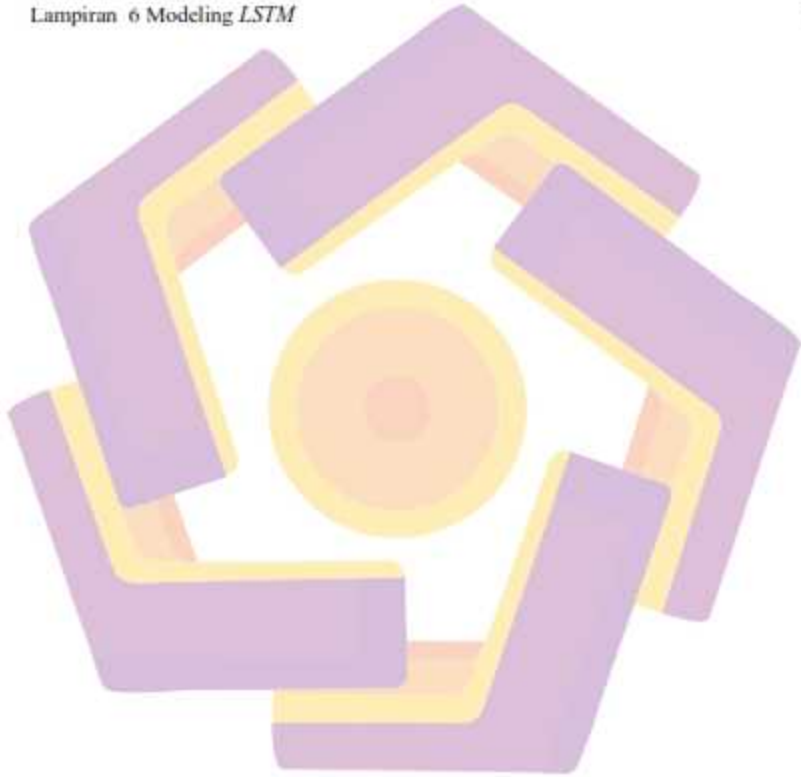
DAFTAR GAMBAR

Gambar 2. 1 <i>RNN Architecture</i>	18
Gambar 2. 2 <i>LSTM Architecture</i>	18
Gambar 3.1 Alur Penelitian	24
Gambar 4. 1 Kode <i>Text Cleaning</i>	35
Gambar 4. 2 Kode <i>Case Folding</i>	37
Gambar 4. 3 Kode <i>Normalized</i>	38
Gambar 4. 4 Kode <i>Stopword Removal</i>	39
Gambar 4. 5 Kode <i>Tokenizing</i>	40
Gambar 4. 6 Kode <i>Stemming</i>	42
Gambar 4. 7 Kode <i>Labeling</i>	43
Gambar 4. 8 Kode Distribusi Sentimen	45
Gambar 4. 9 Gambar Distribusi Sentimen	45
Gambar 4. 10 Kode <i>Word Cloud</i> Keseluruhan Data	46
Gambar 4. 11 <i>Word Cloud</i> Keseluruhan Data	47
Gambar 4. 12 Kode <i>Word Cloud</i> Sentimen Positif dan Negatif	47
Gambar 4. 13 <i>Word Cloud</i> Sentimen Positif	48
Gambar 4. 14 <i>Word Cloud</i> Sentimen Negatif	49
Gambar 4. 15 Distribusi Panjang Tweet	49
Gambar 4. 16 Kode Splitting Dataset	50
Gambar 4. 17 Split Data <i>Train & Test</i>	51
Gambar 4. 18 Kode <i>Tokenized & Padding</i>	51
Gambar 4. 19 Kode <i>Word Embedding</i>	53
Gambar 4. 20 Kode Distribusi Sebelum <i>SMOTE</i>	54
Gambar 4. 21 Distribusi data sebelum <i>SMOTE</i>	55
Gambar 4. 22 Kode Setelah <i>SMOTE</i>	55
Gambar 4. 23 Distribusi data setelah <i>SMOTE</i>	56
Gambar 4. 24 Kode <i>Modeling LSTM</i>	57
Gambar 4. 25 Tingkat Akurasi dari <i>Epoch LSTM</i>	58
Gambar 4. 26 Tingkat <i>Loss</i> dari <i>Epoch LSTM</i>	59
Gambar 4. 27 Kode <i>Model Evaluation</i>	59
Gambar 4. 28 <i>Confusion Matrix LSTM</i>	61
Gambar 4. 29 Halaman Awal	62



DAFTAR LAMPIRAN

Lampiran 1 Scraping Data X	70
Lampiran 2 <i>Preprocessing</i>	71
Lampiran 3 <i>Labeling</i>	72
Lampiran 4 <i>EDA</i>	74
Lampiran 5 <i>Premodeling</i>	77
Lampiran 6 Modeling <i>LSTM</i>	78



DAFTAR LAMBANG DAN SINGKATAN

σ	Fungsi aktivasi sigmoid
Tanh	Fungsi aktivasi hyperbolic tangent
f_t	Nilai forget gate pada waktu ke-t
i_t	Nilai input gate pada waktu ke-t
o_t	Nilai output gate pada waktu ke-t
$\bar{c}_t$	Informasi baru untuk cell state pada waktu ke-t
c_t	Cell state pada waktu ke-t
h_t	Hidden state pada waktu ke-t
$h_{(t-1)}$	Hidden state dari waktu sebelumnya
x_t	Input pada waktu ke-t
W_f, W_i, W_o	Bobot untuk forget gate, input gate, output gate
b_f, b_i, b_o	Bias untuk forget gate, input gate, output gate
API	Application Programming Interface
CSV	Comma-Separated Values
EDA	Exploratory Data Analysis
FN	False Negative
FP	False Positive
LSTM	Long Short-Term Memory
NLP	Natural Language Processing
OOV	Out of Vocabulary
RNN	Recurrent Neural Network
SMOTE	Synthetic Minority Oversampling Technique
TP	True Positive
TN	True Negative
TF-IDF	Term Frequency-Inverse Document Frequency

DAFTAR ISTILAH



Accuracy	Persentase prediksi yang benar
Automatic Labeling	Pelabelan otomatis tanpa manusia.
Bias	Nilai tambahan dalam jaringan saraf
Confusion matrix	Tabel evaluasi hasil prediksi model
Epoch	Satu siklus pelatihan penuh pada data
Embedding Layer	Lapisan untuk mengubah token menjadi vektor representasi
<i>F1-score</i>	Rata-rata harmonik <i>precision</i> dan <i>recall</i>
Labeling	Pemberian label pada data
Lexicon	Kamus kata dengan sentimen
Numerik	Representasi data dalam bentuk angka
Overfitting	Model terlalu cocok dengan data latih
Precision	Akurasi data positif yang benar
<i>Recall</i>	Kemampuan mendeteksi data positif
Sequence	Urutan token yang digunakan sebagai input
Token	Unit terkecil dari teks, biasanya berupa kata atau sub-kata
Vektor	besaran yang mempunyai arah

INTISARI

Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) merupakan lembaga yang aktif menggunakan media sosial X (sebelumnya Twitter) sebagai sarana komunikasi publik terkait informasi cuaca dan iklim. Tingginya aktivitas masyarakat di media sosial memunculkan berbagai opini yang tidak terstruktur terhadap layanan BMKG, sehingga menyulitkan proses evaluasi secara manual. Ketidak terstruktur opini ini dapat berdampak pada keterbatasan instansi dalam memahami persepsi publik secara objektif, yang pada akhirnya berpengaruh terhadap efektivitas perbaikan layanan publik.

Penelitian ini menggunakan metode *Long Short-Term Memory (LSTM)* untuk mengklasifikasikan sentimen masyarakat terhadap BMKG yang diperoleh dari 3.400 data tweet hasil scraping. Langkah-langkah penelitian meliputi *preprocessing* data (*text cleaning, tokenizing, stemming*), pelabelan sentimen berbasis *lexicon*, penyesimbangan data menggunakan teknik *Synthetic Minority Over-sampling Technique (SMOTE)*, serta pelatihan dan evaluasi model *LSTM* menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Model kemudian diimplementasikan ke dalam aplikasi web berbasis Streamlit.

Hasil penelitian menunjukkan bahwa model *LSTM* mampu mencapai akurasi sebesar 88% dengan *F1-score* tertinggi 0,92 untuk sentimen negatif dan 0,75 untuk sentimen positif. Sistem ini berkontribusi dalam menyediakan klasifikasi opini publik secara otomatis yang dapat dimanfaatkan oleh BMKG maupun instansi pemerintah lainnya dalam mengevaluasi kualitas layanan secara lebih efektif dan berbasis data. Penelitian lanjutan disarankan untuk mengeksplorasi arsitektur yang lebih kompleks seperti *Bi-LSTM* atau *BERT* agar hasil klasifikasi lebih optimal.

Kata kunci: Analisis Sentimen, *LSTM*, BMKG, Media Sosial, *SMOTE*

ABSTRACT

The Meteorology, Climatology, and Geophysics Agency (BMKG) is a public institution that actively utilizes social media platform X (formerly Twitter) to communicate weather and climate information to the public. However, the high volume of unstructured public opinions shared through social media makes it difficult to conduct manual sentiment analysis, potentially hindering the institution's ability to evaluate service performance objectively. This challenge may impact the effectiveness of public service improvements if public perception is not well understood.

This study employs the Long Short-Term Memory (LSTM) method to classify public sentiment towards BMKG based on 3,400 tweets collected through web scraping. The research involves several stages, including data preprocessing (text cleaning, tokenization, stemming), sentiment labeling using a lexicon-based approach, data balancing with the Synthetic Minority Over-sampling Technique (SMOTE), and model training and evaluation using accuracy, precision, recall, and F1-score as performance metrics. The final model is deployed as a web-based application using the Streamlit framework.

The results show that the LSTM model achieved an accuracy of 88%, with the highest F1-score of 0.92 for negative sentiment and 0.75 for positive sentiment. This system contributes to the development of an automated sentiment classification tool that can be utilized by BMKG and other government institutions to improve public services through data-driven insights. Future research is recommended to explore more advanced architectures such as Bi-LSTM or BERT to enhance classification performance.

Keywords: Sentiment Analysis, LSTM, BMKG, Social Media, SMOTE