

**ANALISIS SENTIMEN PUBLIK TERHADAP ANJLOKNYA
IHSG PERIODE MARET-APRIL 2025 DI X (TWITTER)
MENGUNAKAN LSTM DAN INDOBERT**



SKRIPSI

**MUHAMAD FARHAN FADILLAH
00000061559**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISIS SENTIMEN PUBLIK TERHADAP ANJLOKNYA
IHSG PERIODE MARET-APRIL 2025 DI X (TWITTER)
MENGUNAKAN LSTM DAN INDOBERT**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**MUHAMAD FARHAN FADILLAH
00000061559**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Muhamad Farhan Fadillah
Nomor Induk Mahasiswa : 00000061559
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Publik terhadap Anjloknya IHSG Periode Maret–April 2025 di X (Twitter) Menggunakan LSTM dan IndoBERT

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, 4 Juli 2025



A handwritten signature in black ink, appearing to read 'Farhan'.

(Muhamad Farhan Fadillah)

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 PUBLIK TERHADAP ANJLOKNYA IHSG
PERIODE MARET-APRIL 2025 DI X (TWITTER) MENGGUNAKAN
LSTM DAN INDOBERT**

oleh

Nama : Muhamad Farhan Fadillah
NIM : 00000061559
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Rabu, 9 Juli 2025

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Dr. Sy. Yuliani Yakub, S.Kom., M.T)

NIDN: 0411037904

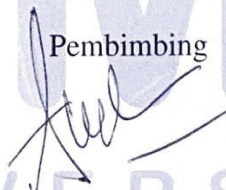
Penguji



(Dr. Ir. Winarno, M.Kom.)

NIDN: 0330106002

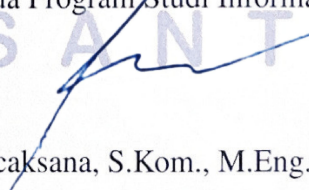
Pembimbing



(Anak Agung Ngurah Ananda Kusuma, B.Eng., M.Eng., Ph.D.)

NIDK: 08984101024

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Muhamad Farhan Fadillah
NIM : 00000061559
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen Publik terhadap
Anjloknya IHSG Periode Maret–April
2025 di X (Twitter) Menggunakan
LSTM dan IndoBERT

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 tiga tahun.

Tangerang, 4 Juli 2025

Yang menyatakan



Muhamad Farhan Fadillah

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



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PERSEMBAHAN / MOTTO

"It's just an empty cup."

- Hudson Hornet

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa, atas rahmat-Nya penulis dapat menyelesaikan tugas akhir berjudul **“Analisis Sentimen Publik terhadap Anjloknya IHSG Periode Maret-April 2025 di X (Twitter) Menggunakan LSTM dan IndoBERT”**. Penulisan ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Komputer di Universitas Multimedia Nusantara.

Penyelesaian tugas akhir ini tidak terlepas dari bimbingan dan dukungan berbagai pihak. Untuk itu, penulis mengucapkan terima kasih yang tulus kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika.
4. Bapak Anak Agung Ngurah Ananda Kusuma, B.Eng., M.Eng., Ph.D., selaku dosen pembimbing, atas segala bimbingan dan arahan yang diberikan.
5. Keluarga tercinta atas segala dukungan moril dan materiel.
6. Seluruh sahabat dan rekan: Karen, Adri, Cahyo, Agil, Kafi, Kevin, dan Grup "Manusia Kuat", atas dukungan dan kerja samanya.

Penulis menyadari sepenuhnya bahwa tugas akhir ini masih jauh dari sempurna dan terbuka untuk setiap kritik serta saran yang membangun. Semoga karya ini dapat bermanfaat bagi para pembaca.

Tangerang, 4 Juli 2025



Muhamad Farhan Fadillah

**ANALISIS SENTIMEN PUBLIK TERHADAP ANJLOKNYA IHSG
PERIODE MARET-APRIL 2025 DI X (TWITTER) MENGGUNAKAN
LSTM DAN INDOBERT**

Muhamad Farhan Fadillah

ABSTRAK

Pergerakan Indeks Harga Saham Gabungan (IHSG) sering memicu diskusi publik di media sosial. Penelitian ini menganalisis sentimen publik terhadap anjloknya IHSG pada Maret-April 2025 dan membandingkan kinerja model *deep learning* Bi-LSTM dengan IndoBERT. Menggunakan kerangka CRISP-DM, model Bi-LSTM dilatih pada ~5.000 data *pseudo-labeled*, sementara IndoBERT di-*fine-tuning* pada 529 data berlabel manual. Hasil menunjukkan pendekatan *transfer learning* dengan IndoBERT (akurasi 79,25%) secara signifikan lebih unggul daripada Bi-LSTM (akurasi 68,87%). Analisis kualitatif juga menemukan dualisme persepsi publik antara kekhawatiran (kata "turun", "anjlok") dan optimisme (kata "kuat", "beli"). Disimpulkan bahwa penggunaan model *pre-trained* seperti IndoBERT pada data berkualitas tinggi merupakan pendekatan yang lebih efektif untuk analisis sentimen keuangan di Indonesia.

Kata kunci: Analisis Sentimen, IHSG, IndoBERT, LSTM, Twitter



***PUBLIC SENTIMENT ANALYSIS ON THE DECLINE OF IHSG DURING
MARCH—APRIL 2025 ON X (TWITTER) USING LSTM AND INDOBERT***

Muhamad Farhan Fadillah

ABSTRACT

The movement of the Indonesia Stock Exchange Index (IHSG) often sparks public discussion on social media. This study analyzes public sentiment toward the IHSG downturn during March–April 2025 and compares the performance of the Bi-LSTM deep learning model with IndoBERT. Using the CRISP-DM framework, the Bi-LSTM model was trained on approximately 5,000 pseudo-labeled data, while IndoBERT was fine-tuned on 529 manually labeled data. The results show that the transfer learning approach with IndoBERT (accuracy of 79.25%) significantly outperformed Bi-LSTM (accuracy of 68.87%). A qualitative analysis also revealed a duality in public perception, balancing between concern (words like “drop,” “plunge”) and optimism (words like “strong,” “buy”). It is concluded that leveraging pre-trained models such as IndoBERT on high-quality data is a more effective approach for financial sentiment analysis in Indonesia.

Keywords: IHSG, IndoBERT, LSTM, Sentiment Analysis, Twitter



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	vi
KATA PENGANTAR	vii
KATA PENGANTAR	vii
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Penelitian Terdahulu	6
2.2 Analisis Sentimen	12
2.3 <i>Natural Language Processing</i> (NLP)	12
2.3.1 <i>Word Cloud</i>	14
2.4 <i>Deep Learning</i> untuk NLP	14
2.5 Metodologi CRISP-DM	16
2.5.1 <i>Business Understanding</i> (Pemahaman Masalah)	17
2.5.2 <i>Data Understanding</i> (Pemahaman Data)	17
2.5.3 <i>Data Preparation</i> (Persiapan Data)	18
2.5.4 <i>Modeling</i> (Pemodelan)	18
2.5.5 <i>Evaluation</i> (Evaluasi)	18
2.6 Metode Pemodelan	19
2.6.1 <i>Bidirectional Long Short-Term Memory</i> (Bi-LSTM)	19
2.6.2 IndoBERT (Arsitektur Berbasis Transformer)	21
2.7 Metode Evaluasi	23
BAB 3 METODOLOGI PENELITIAN	25
3.1 Kerangka dan Alur Penelitian	25
3.2 Perangkat Penelitian dan Sumber Data	27
3.2.1 Spesifikasi Perangkat Keras	27
3.2.2 Sumber Data	27
3.2.3 Perangkat Lunak Pendukung	27
3.3 Tahapan Penelitian	28
3.3.1 Pengumpulan Data	28
3.3.2 Persiapan Data	30
3.3.3 Pemodelan Komparatif	36
3.3.4 Metode Evaluasi	44

BAB 4	HASIL DAN PEMBAHASAN	45
4.1	Hasil Fase Data Understanding	45
4.1.1	Hasil Pengumpulan Data	45
4.1.2	Hasil Analisis Kualitatif dengan Word Cloud	46
4.2	Hasil Data Preparation	48
4.2.1	Hasil Data Preprocessing	48
4.2.2	Hasil Pelabelan dan Distribusi Sentimen	50
4.2.3	Hasil Pembagian Dataset	51
4.3	Hasil Eksperimen Model Bidirectional LSTM	52
4.3.1	Kinerja Model Bi-LSTM Awal (Baseline)	52
4.3.2	Optimasi dan Kinerja Model Bi-LSTM Final	54
4.3.3	Kinerja Bi-LSTM pada Skenario Uji Sensitivitas (70:30).	57
4.4	Hasil Eksperimen Model IndoBERT	58
4.4.1	Kinerja Model IndoBERT Awal (Baseline)	58
4.4.2	Optimasi dan Kinerja Model IndoBERT Final	59
4.4.3	Kinerja IndoBERT pada Skenario Uji Sensitivitas (70:30).	61
4.5	Pembahasan	62
4.5.1	Analisis Perbandingan Kinerja Model	62
BAB 5	SIMPULAN DAN SARAN	64
5.1	Simpulan	64
5.2	Saran	65
DAFTAR PUSTAKA		66



DAFTAR TABEL

Tabel 2.1	Perbandingan Penelitian Terdahulu dan Kontribusi Penelitian (1)	7
Tabel 2.2	Perbandingan Penelitian Terdahulu dan Kontribusi Penelitian (2)	8
Tabel 2.3	Perbandingan Penelitian Terdahulu dan Kontribusi Penelitian (3)	9
Tabel 2.4	Perbandingan Penelitian Terdahulu dan Kontribusi Penelitian (4)	10
Tabel 3.1	Contoh Hasil Setiap Tahapan Preprocessing	32
Tabel 3.2	Arsitektur Final Model Bi-LSTM Hasil Optimasi	40
Tabel 4.1	Atribut Data Mentah Hasil Pengumpulan	46
Tabel 4.2	Contoh Hasil Case Folding dan Cleaning	49
Tabel 4.3	Contoh Hasil Tokenization	49
Tabel 4.4	Contoh Hasil Stopword Removal	49
Tabel 4.5	Contoh Hasil Stemming	50
Tabel 4.6	Contoh Hasil Pelabelan Sentimen Otomatis pada Sampel Data	50
Tabel 4.7	Hasil Pembagian Dataset (Skenario Uji Utama 80:20)	52
Tabel 4.8	Hasil Pembagian Dataset (Skenario Uji Sensitivitas 70:30)	52
Tabel 4.9	Rangkuman Arsitektur Model Bi-LSTM Awal	53
Tabel 4.10	Parameter Pelatihan untuk Model Bi-LSTM Awal	53
Tabel 4.11	Laporan Klasifikasi Model Bi-LSTM Awal	54
Tabel 4.12	Ruang Pencarian Hiperparameter untuk Optimasi Bi-LSTM	55
Tabel 4.13	Hiperparameter Terbaik Hasil Optimasi Bi-LSTM	55
Tabel 4.14	Parameter Pelatihan untuk Model Bi-LSTM Final	56
Tabel 4.15	Laporan Klasifikasi Model Bi-LSTM Final (Setelah Tuning)	56
Tabel 4.16	Laporan Klasifikasi Model Bi-LSTM Final (Rasio 70:30)	57
Tabel 4.17	Parameter untuk Proses Fine-Tuning IndoBERT Awal	58
Tabel 4.18	Laporan Klasifikasi Model IndoBERT Awal	58
Tabel 4.19	Ruang Pencarian Hiperparameter untuk Fine-Tuning IndoBERT	59
Tabel 4.20	Hiperparameter Terbaik Hasil Optimasi Fine-Tuning IndoBERT	60
Tabel 4.21	Parameter Pelatihan untuk Model IndoBERT Final	60
Tabel 4.22	Laporan Klasifikasi Model IndoBERT Final (Setelah Tuning)	60
Tabel 4.23	Laporan Klasifikasi Model IndoBERT Final (Rasio 70:30)	61
Tabel 4.24	Perbandingan Kinerja Akhir Model pada Kedua Skenario Uji	62

DAFTAR GAMBAR

Gambar 2.1	Ilustrasi Alur Kerja Umum pada Proyek NLP [1]	14
Gambar 2.2	Ilustrasi Arsitektur Deep Neural Network (sumber: [2]) . .	15
Gambar 2.3	Siklus Metodologi CRISP-DM [3]	17
Gambar 2.4	Diagram Arsitektur Sel LSTM [4]	19
Gambar 2.5	Diagram Arsitektur Lengkap Model Transformer [5]	21
Gambar 2.6	Ilustrasi Konsep Dasar Confusion Matrix (2x2) [6]	23
Gambar 3.1	Diagram Alur Penelitian	26
Gambar 3.2	Diagram Alur Proses Pengumpulan Data	29
Gambar 3.3	Diagram Alur Proses Preprocessing Teks	31
Gambar 3.4	Diagram Alur Proses Pelabelan dan Ekstraksi Fitur	34
Gambar 3.5	Diagram Alur Proses Eksperimen Model Bi-LSTM	36
Gambar 3.6	Diagram Alur Umum Proses Optimasi Hiperparameter . .	38
Gambar 3.7	Diagram Alur Proses Eksperimen Model IndoBERT	41
Gambar 4.1	Tampilan Sebagian Data Mentah Hasil Scraping	45
Gambar 4.2	Word Cloud Keseluruhan Sentimen Terkait IHSG	47
Gambar 4.3	Word Cloud Berdasarkan Kategori Sentimen (Positif, Netral, Negatif)	47
Gambar 4.4	Distribusi Sentimen Publik Hasil Pseudo-Labeling	51
Gambar 4.5	Confusion Matrix untuk Model Bi-LSTM Awal	54
Gambar 4.6	Confusion Matrix untuk Model Bi-LSTM Final	56
Gambar 4.7	Confusion Matrix Model Bi-LSTM Final (Rasio 70:30) . .	57
Gambar 4.8	Confusion Matrix untuk Model IndoBERT Awal	59
Gambar 4.9	Confusion Matrix untuk Model IndoBERT Final	61
Gambar 4.10	Confusion Matrix Model IndoBERT Final (Rasio 70:30) .	62



DAFTAR KODE

Kode 3.1	Eksekusi Pengambilan Data dengan Tweet-Harvest	29
Kode 3.2	Implementasi Lengkap Alur Preprocessing Teks	32
Kode 3.3	Pembangunan dan Pelatihan Model Bi-LSTM Awal	37
Kode 3.4	Definisi Ruang Pencarian untuk Keras Tuner	39
Kode 3.5	Pelatihan Model IndoBERT Awal	41
Kode 3.6	Definisi Ruang Pencarian dan Eksekusi Hyperparameter Search .	43



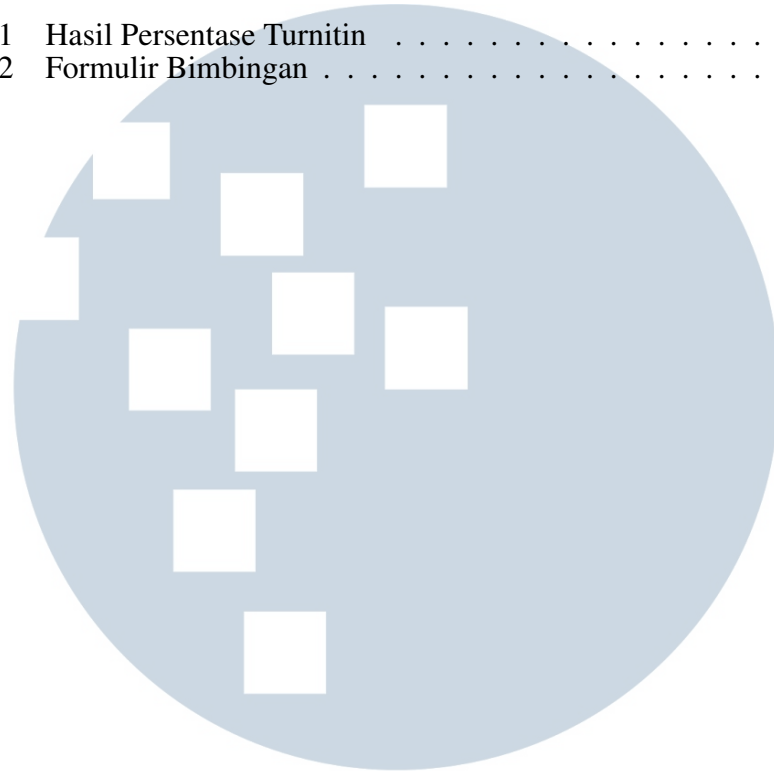
DAFTAR RUMUS

Rumus 2.1	<i>Neuron</i>	15
Rumus 2.2	<i>Forget Gate</i>	19
Rumus 2.3	<i>Input Gate</i>	20
Rumus 2.4	<i>Output Gate</i>	20
Rumus 2.5	<i>Bi-LSTM Output</i>	21
Rumus 2.6	<i>Scaled Dot-Product Attention</i>	22
Rumus 2.7	<i>Accuracy</i>	24
Rumus 2.8	<i>Precision</i>	24
Rumus 2.9	<i>Recall</i>	24
Rumus 2.10	<i>F1-Score</i>	24



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	71
Lampiran 2	Formulir Bimbingan	72



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA