

PREDIKSI HARGA SAHAM MENGGUNAKAN MODEL LONG SHORT-TERM MEMORY

Tugas Akhir

**Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat
Sarjana Komputer**



Dibuat Oleh:

Rocky Sito Andilolo

210711286

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2025**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

PREDIKSI HARGA SAHAM MENGGUNAKAN MODEL LONG SHORT-TERM MEMORY

yang disusun oleh

Rocky Sito Andilolo

210711286

dinyatakan telah memenuhi syarat pada tanggal 27 Oktober 2025

		Keterangan
Dosen Pembimbing 1	: Prof. Dr. Ir. Alb. Joko Santoso, M.T.	Telah Menyetujui
Dosen Pembimbing 2	: Dra. Emawati, M.T., Ph.D.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Prof. Dr. Ir. Alb. Joko Santoso, M.T.	Telah Menyetujui
Penguji 2	: Irya Wisnubhadra, S.T., M.T., Ph.D., IPM	Telah Menyetujui
Penguji 3	: Eddy Julianto, S.T., M.T.	Telah Menyetujui

Yogyakarta, 27 Oktober 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Prof. Dr. Ir. Parama Kartika Dewa SP., S.T., M.T., IPU, ASEAN Eng.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH

Saya yang bertanda tangan di bawah ini :

Nama Lengkap : Rocky Sito Andilolo
NPM : 210711286
Program Studi : Informatika
Fakultas : Teknologi Industri
Judul Penelitian : Prediksi Harga Saham Menggunakan Model Long
Short-Term Memory

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini di buat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 08 Juli 2025
Yang menyatakan,

Rocky Sito Andilolo
210711286

HALAMAN PERSEMBAHAN

Dengan penuh syukur dan kerendahan hati, kupersembahkan karya sederhana ini sebagai pucuk perjalanan akademisku. Pertama-tama, kepada kedua orang tuaku tercinta yang menjadi tiang penyangga hidupku. Terima kasih atas doa tanpa henti, pengorbanan tak terhitung, dan cinta tanpa syarat yang mengalir dalam setiap napas kalian. Setiap langkahku adalah buah kasih sayang yang tak pernah kering, kekuatan yang mengangkatku ketika lelah, dan keyakinan yang menguatkanmu saat ragu. Tanpa dukungan kalian, tak mungkin aku berdiri di titik ini.

Kepada kekasihku, Liony Evania, sahabat jiwa yang menjadi cahaya dalam perjalananku. Terima kasih untuk setiap senyuman yang menghangatkan hari-hari panjang, kesabaranmu mendengarkan keluh kesahku, dan cinta yang tak pernah berhenti menyemangati. Keberanianku menyelesaikan tantangan ini lahir dari keyakinan bahwa kau selalu ada di sisiku, percaya padaku bahkan saat aku tak percaya pada diriku sendiri. Perjalanan ini terasa lebih bermakna karena kita melewatinya bersama.

Tak lupa kepada saudara-saudaraku yang menjadi tempatku berbagi tawa dan cerita. Terima kasih untuk candaan yang menghilangkan penat, dukungan tanpa syarat, dan kebersamaan yang mengingatkanmu pada arti keluarga. Juga kepada sahabat dan teman seperjuangan di bangku kuliah, terima kasih atas diskusi seru hingga larut malam, kerja sama mengerjakan tugas, dan bahu-membahu melewati masa-masa genting. Kebersamaan kita mengubah beban menjadi kenangan indah.

Secara khusus, kupanjatkan terima kasih kepada para dosen pembimbing, penguji, dan seluruh staf akademik yang telah membimbingku dengan ilmu dan kebijaksanaan. Terima kasih untuk pintu yang selalu terbuka, kritik yang membangun, dan bimbingan yang tak kenal lelah. Akhirnya, kepada Sang Pencipta yang mengizinkanmu menyelesaikan perjalanan ini, segala puji dan

syukur kupanatkan atas rahmat dan karunia-Mu.

Semoga setiap kebaikan yang telah diberikan kepadaku berbalas berlipat ganda, dan karya kecil ini dapat menjadi manfaat bagi sesama.

Rocky Sito Andilolo
Yogyakarta, 18 Juli 2025

KATA PENGANTAR

Puji dan syukur penulis haturkan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan pembuatan tugas akhir “Prediksi Harga Saham Menggunakan Model Long Short-Term Memory” ini dengan baik.

Penulisan tugas akhir ini bertujuan untuk memenuhi salah satu syarat untuk mencapai derajat sarjana komputer dari Program Studi Informatika, Fakultas Teknologi Industri di Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam pembuatan tugas akhir ini penulis telah mendapat bantuan, bimbingan, dan dorongan dari banyak pihak. Untuk itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yesus Kristus yang selalu membimbing dalam iman-Nya, memberikan berkat-Nya, dan menyertai penulis selalu.
2. Bapak Dr. Ir. Parama Kartika Dewa SP., S.T., M.T. IPU, selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
3. Bapak Prof. Dr. Ir. Albertus Joko Santoso, M.T., selaku dosen pembimbing I yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
4. Ibu Dra. Ernawati., M.T., Ph.D., selaku dosen pembimbing II yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.

Demikian laporan tugas akhir ini dibuat, dan penulis mengucapkan terima kasih kepada semua pihak. Semoga laporan ini dapat bermanfaat bagi pembaca.

Yogyakarta, 08 Juli 2025

Rocky Sito Andilolo

210711286

DAFTAR ISI

PREDIKSI HARGA SAHAM MENGGUNAKAN MODEL LONG SHORT-TERM MEMORY	i
LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH.....	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
DAFTAR ISTILAH	xi
INTISARI	xvi
BAB I.....	1
PENDAHULUAN	1
A. Latar Belakang.....	1
B. Rumusan Masalah.....	2
C. Batasan Masalah	3
D. Tujuan Penelitian	3
E. Metode Penelitian	4
F. Sistematika Penulisan	8
BAB II.....	9
TINJAUAN PUSTAKA	9
BAB III	15
LANDASAN TEORI.....	15
A. Metode Prediksi	15
B. Integrasi Sentimen	18
C. Fitur Teknikal dan Normalisasi	19
D. Metrik Evaluasi.....	20
BAB IV	21
DATASET DAN PEMROSESAN DATA	21
A. Deskripsi Masalah	21
B. Dataset	22
C. Pra-Pemrosesan Data	24
1. Pembersihan Data	24

2. Fitur Teknikal	27
3. Normalisasi	32
4. Pembagian Dataset	35
D. Ringkasan Akhir	38
BAB V	43
IMPLEMENTASI MODEL DAN PENGUJIAN SISTEM	43
A. Implementasi Model	43
B. Pengujian Sistem	60
1. Konsistensi Output	60
2. Pengaruh Sentimen	Error! Bookmark not defined.
BAB VI	65
PENUTUP	65
A. Kesimpulan	65
B. Saran	67
DAFTAR PUSTAKA	68

DAFTAR GAMBAR

Gambar 1.1. Alur Metode Penelitian	4
Gambar 4.1. Diagram Alur Pemrosesan Data.....	22
Gambar 4.2. Visualisasi Tanpa Resample dan Interpolasi.....	26
Gambar 4.3. Visualisasi dengan Resample dan Interpolasi	26
Gambar 4.4. Contoh Visualisasi Overlay Harga Saham dan Sentimen	27
Gambar 4.4. Visualisasi MA 5 dan MA 20.....	30
Gambar 4.5. Visualisasi RSI 14 dan Volatility 5	31
Gambar 4.6. Heatmap Korelasi Fitur	42
Gambar 5.1. Code Pengambilan Data Saham	43
Gambar 5.2. Code Pengambilan Data Sentimen Bagian 1.....	45
Gambar 5.3. Code Pengambilan Data Sentimen Bagian 2.....	46
Gambar 5.4. Code Resample dan Interpolasi.....	47
Gambar 5.5. Code Menggabungkan Sentimen Dengan Data Saham dan Shifting Data Sentimen (Lag_1).....	47
Gambar 5.6. Code Pembuatan Fitur Lag dan Teknikal.....	48
Gambar 5.7. Code Normalisasi Data	49
Gambar 5.8. Code Split Data Train/Test.....	49
Gambar 5.9. Code Pembagian Data Per Periode	50
Gambar 5.10. Code Implementasi Arsitektur	51
Gambar 5.11. Code Stateful Transfer Learning	52
Gambar 5.12. Code Mekanisme Hyperparameter Tuning	52
Gambar 5.13. Code Evaluasi Kinerja Model	53
Gambar 5.14. Code Evaluasi Kinerja Model	56
Gambar 5.15. Visualisasi Rata – Rata Prediksi LSTM.....	57
Gambar 5.16. Visualisasi Rata – Rata Prediksi ARIMA	59

DAFTAR TABEL

Tabel 2.1. Tabel Perbandingan Penelitian	12
Tabel 4.1. Tabel Preview Sampel Data Saham	23
Tabel 4.2. Tabel Preview Data Sentimen.....	23
Tabel 4.3. Tabel Preview Sampel Gabungan Data Saham dan Sentimen.....	25
Tabel 4.4. Tabel Preview Sampel Fitur Teknikal	29
Tabel 4.5. Tabel Preview Normalisasi	34
Tabel 4.6. Tabel Distribusi Data Per Periode.....	36
Tabel 4.7. Tabel Preview Sampel Data Per Periode	37
Tabel 4.8. Tabel Preview Dataset Akhir	40
Tabel 5.1. Hasil Rata – Rata Evaluasi Model	55
Tabel 5.2. Statistik Mean & Deviasi Standar	61
Tabel 5.3. Statistik Analisis Kontribusi Fitur Sentimen pada LSTM	62
Tabel 5.4. Statistik Analisis Kontribusi Fitur Sentimen pada ARIMA.....	63

DAFTAR ISTILAH

Istilah	Penjelasan singkat
AAPL	Ticker saham untuk Apple Inc., objek studi pada penelitian.
AIC (<i>Akaike Information Criterion</i>)	Kriteria pemilihan model statistik yang menyeimbangkan kecocokan dan kompleksitas (nilai lebih kecil lebih baik).
Adam (optimizer)	Optimizer berbasis adaptive moment estimation yang umum dipakai untuk melatih jaringan saraf.
Analisis kontribusi fitur sentimen (studi ablasi / ablation study)	Pengujian komparatif antara model yang menyertakan fitur sentimen dan yang tidak, untuk mengukur kontribusi fitur sentimen terhadap kinerja model. (Dokumen menyebutnya sebagai “analisis kontribusi fitur sentimen” / “studi ablation”).
ARIMA (<i>Autoregressive Integrated Moving Average</i>)	Model deret waktu klasik yang menggabungkan komponen AR, I (differencing), dan MA; digunakan sebagai baseline.
auto_arima()	Fungsi/algoritme otomatis untuk memilih parameter (p,d,q) ARIMA, biasanya berdasar AIC.
Autosave / json.dump	Mekanisme menyimpan artefak eksperimen (model/hyperparams) ke disk (mis. via <code>json.dump</code>) agar eksperimen dapat direproduksi.
Batch / Epoch	<i>Batch</i> : subset data yang diproses tiap iterasi; <i>Epoch</i> : satu kali lewat seluruh data pelatihan.
Bayesian Optimization / Bayesian Tuning	Metode optimasi hyperparameter yang menggunakan model probabilistik (surrogate) untuk mencari kombinasi parameter terbaik secara efisien.
Bullish / Bearish	<i>Bullish</i> adalah ekspektasi harga naik (optimis) sedangkan <i>Bearish</i> adalah ekspektasi harga turun (pesimis)
Close (harga penutupan)	Harga terakhir suatu saham pada sesi perdagangan, target prediksi penelitian.
Clone_model / Weight cloning	Teknik menyalin arsitektur dan bobot (weights) model utama untuk fine-

	tuning pada subset data (digunakan di strategi stateful transfer).
<i>Compound score (VADER)</i>	Skor gabungan VADER yang mewakili polaritas sentimen teks (nilai compound sering digunakan sebagai skor sentimen harian).
Correlation (Korelasi)	Ukuran hubungan linier antar variabel (rentang -1 sampai $+1$).
Determinisme ketat	Praktik menetapkan seed, versi library, dan pengaturan lingkungan agar eksperimen dapat direproduksi.
<i>Differencing (I pada ARIMA)</i>	Operasi untuk membuat deret waktu stasioner dengan mengurangi nilai sebelumnya (operator ∇ atau $(1-L)$).
<i>Dense layer</i>	Lapisan fully-connected pada jaringan saraf, biasa dipakai sebagai layer akhir untuk regresi.
<i>Death Cross</i>	<i>moving average</i> jangka pendek memotong ke bawah <i>moving average</i> jangka panjang yang menandakan potensi tren turun jangka panjang.
Forward fill (ffill)	Teknik mengisi nilai yang hilang dengan mengambil nilai terakhir yang tersedia (dipakai saat resample tanggal non-trading).
GARCH	Model untuk memodelkan heteroskedastisitas/volatilitas bersyarat (muncul di tinjauan pustaka).
Heatmap	Visualisasi matriks (mis. matriks korelasi antar fitur) yang menggunakan skema warna untuk merepresentasikan kekuatan hubungan antarvariabel; muncul sebagai Gambar 4.6 Heatmap Korelasi Fitur.
<i>Huber loss</i>	Fungsi loss yang robust terhadap outlier (kombinasi MSE dan MAE) disebut sebagai pilihan loss di tuning.
<i>Hyperparameter tuning</i> / Keras Tuner / HyperModel / hp.Int / hp.Float	Proses pencarian konfigurasi hyperparameter (contoh: unit LSTM, learning rate). Dokumen menggunakan Keras Tuner dan HyperModel untuk otomatisasi tuning.
<i>In-sample</i> / <i>Out-of-sample</i>	<i>In-sample</i> : evaluasi pada data yang dipakai pelatihan; <i>Out-of-sample</i> :

	evaluasi pada data yang belum dilihat model (uji).
<i>Interpolate</i> / Interpolasi linier	Mengisi nilai yang hilang di antara dua titik dengan metode linear interpolation (dipakai setelah forward-fill).
Lag / Lagging / <i>Sentiment_Lag1</i>	Pergeseran nilai ke belakang (mis. <i>lag_1</i> = nilai $t-1$). <i>Sentiment_Lag1</i> adalah skor sentimen digeser satu hari agar terhindar data leakage.
<i>Learning rate</i> (lr)	Parameter optimisasi yang mengatur besaran langkah pembaruan bobot selama training.
LSTM (<i>Long Short-Term Memory</i>)	Varian RNN dengan gate (forget/input/output) dan cell state untuk menangkap ketergantungan jangka panjang; arsitektur utama penelitian.
(LSTM) Gates: forget, input, output, cell state	Komponen LSTM yang mengontrol penyimpanan/luahan informasi (dijelaskan matematis di Bab III).
MA, MA_5, MA_20, MA_5_20_ratio	Moving Average (rata-rata bergerak) dan rasio MA5/MA20 sebagai indikator tren.
MAE / MAPE / MSE / R^2	Metrik evaluasi: Mean Absolute Error, Mean Absolute Percentage Error, Mean Squared Error, Koefisien Determinasi.
<i>MinMaxScaler</i> / Normalisasi Min-Max	Metode scaling membawa fitur ke rentang [0,1] untuk menyamakan skala input LSTM.
Momentum_5D (indikator)	Indikator teknikal yang mengukur percepatan harga selama 5 hari.
<i>Missing Value Handling</i> (fillna, zero filling)	Penanganan data hilang: <code>fillna(0)</code> untuk sentimen hari tanpa tweet, forward fill, atau interpolasi.
<i>Overbought</i>	Kondisi di mana harga dianggap telah naik berlebihan dalam jangka pendek sehingga kemungkinan koreksi turun meningkat; sering diindikasikan $RSI > 70$.
<i>Oversold</i>	Kondisi di mana harga dianggap turun berlebihan dalam jangka pendek

	sehingga kemungkinan rebound naik meningkat; sering diindikasikan $RSI < 30$.
Overlay plot (<i>overlay</i>)	Teknik visualisasi yang menumpuk dua atau lebih deret waktu pada satu grafik (mis. harga saham dan skor sentimen) untuk melihat korelasi/komovemen antar seri. Dokumen menyebutnya sebagai <i>overlay harga saham dan sentimen</i> .
Pandas: reindex / resample / shift / interpolate / ffill	Fungsi Pandas yang dipakai untuk sinkronisasi kalender, ubah frekuensi ke harian, pergeseran (shift), dan pengisian nilai.
Periode COVID-19 (Segmentasi)	Definisi periode eksperimen: <i>Sebelum COVID</i> (9 Apr 2017–29 Feb 2020), <i>Selama COVID</i> (1 Mar 2020–31 Des 2022), <i>Setelah COVID</i> (1 Jan 2023–5 Aug 2025).
Return_1D	Return harian: persentase perubahan harga penutupan dari satu hari ke hari berikutnya.
Resample('D')	Operasi mengubah frekuensi data ke harian penuh (dipakai agar data saham sejajar dengan sentimen harian).
RSI_14 (Relative Strength Index)	Indikator momentum yang mengukur rasio kenaikan vs penurunan rata-rata selama 14 hari.
RSI_MA5	Fitur komposit yang menggabungkan nilai RSI_14 dengan MA_5 (RSI_14 dibagi MA_5) dipakai untuk menangkap interaksi momentum dengan tren. (Muncul di Tabel 4.4 / 4.8).
Runtun Waktu / <i>Time Series</i>	Data yang berurut menurut waktu (konteks analisis deret waktu).
Scraping / Selenium / Twitter_Scraper / Autentikasi	Teknik pengambilan data dari web: Selenium digunakan pada proses scraping tweet dengan autentikasi.
Sentiment / Sentiment_Lag1 / SentimentIntensityAnalyzer	Sentimen = skor numerik dari analisis teks (VADER). <i>Sentiment_Lag1</i> adalah versi digeser satu hari.

	SentimentIntensityAnalyzer adalah kelas VADER yang menghasilkan skor compound.
Shuffling (shuffle=False)	Pengaturan shuffle=False pada pembagian train/test menjaga urutan kronologis agar tidak terjadi data leakage.
SimpleLSTMModel (HyperModel)	Template/kelas LSTM yang digunakan Keras Tuner untuk eksplorasi hyperparameter (disebut di implementasi).
<i>Stateful Transfer Learning</i> / Stateful LSTM	Teknik melanjutkan pelatihan model pada subset data sambil mempertahankan state/optimizer agar transfer pengetahuan terjadi.
Study / Experiment Reproducibility	Praktik teknis (seed, pencatatan hyperparams, penyimpanan model) agar eksperimen dapat diulang.
<i>Technical Indicators</i> (Indikator Teknikal)	Kumpulan indikator turunan harga (MA, RSI, Volatility, Momentum, rasio) yang menjadi fitur input model.
Volatility_5D	Volatilitas 5 hari: standar deviasi return selama periode 5 hari.
VADER / SentimentIntensityAnalyzer	Lexicon-based sentiment analyzer (Valence Aware Dictionary and sEntiment Reasoner) yang memberi skor compound untuk teks pendek seperti tweet.
Windowing / Sliding Window	Teknik membuat sequence input untuk LSTM berdasarkan jendela waktu (sequence length).
yfinance	Library Python untuk mengunduh data pasar (Yahoo Finance) digunakan untuk pengambilan data AAPL.
zero filling (fillna(0))	Mengisi tanggal tanpa tweet dengan 0 agar kalender sentimen lengkap.

INTISARI

PREDIKSI HARGA SAHAM MENGGUNAKAN MODEL LONG SHORT-TERM MEMORY

Intisari

Rocky Sito Andilolo
210711286

Prediksi harga saham merupakan tantangan yang kompleks karena dipengaruhi oleh berbagai faktor seperti kinerja fundamental perusahaan, kondisi ekonomi global, kebijakan moneter, hingga sentimen pasar. Model tradisional seperti ARIMA sering digunakan dalam analisis runtun waktu, namun memiliki keterbatasan dalam menangkap pola non-linear dan volatilitas ekstrem, khususnya pada periode krisis seperti pandemi COVID-19. Selain itu, pemanfaatan data sentimen sebagai salah satu faktor utama yang memengaruhi pergerakan harga saham masih jarang dieksplorasi secara mendalam.

Untuk menjawab tantangan tersebut, penelitian ini menggunakan model *Long Short-Term Memory* (LSTM) dengan integrasi data sentimen dari Twitter menggunakan teknik *scraping* dan analisis VADER. Data sentimen diselaraskan dengan data harga saham *Apple Inc.* (AAPL) melalui proses *resampling*, *interpolasi*, serta pemberian *lag* satu hari. Selain sentimen, penelitian ini juga menambahkan indikator teknikal. Data kemudian dinormalisasi, sementara optimasi model dilakukan melalui *Bayesian Tuning* dan *Stateful Transfer Learning*. Sebagai pembanding, model ARIMA digunakan sebagai baseline.

Hasil penelitian menunjukkan bahwa LSTM secara konsisten unggul dibandingkan ARIMA. Pada periode pandemi COVID-19, LSTM tetap stabil dengan MAPE 0,0923%, sementara ARIMA mengalami penurunan performa dengan MAPE 0,5303%. Berdasarkan analisis kontribusi fitur sentimen juga, penelitian ini membuktikan efektivitas integrasi data sentimen dan model non-linear dalam meningkatkan prediksi harga saham modern.

Kata Kunci : Prediksi Harga Saham, LSTM, ARIMA, Sentimen Twitter, VADER.

Dosen Pembimbing I : Prof. Dr. Ir. Albertus Joko Santoso, M.T.

Dosen Pembimbing II : Dra. Ernawati., M.T., Ph.D.

Jadwal Sidang Tugas Akhir :