

BAB II

LANDASAN TEORI

2.1 Penelitian Terdahulu

Berdasarkan penelitian terdahulu yang telah disajikan pada Tabel 2.1, dapat diketahui bahwa prediksi harga saham merupakan permasalahan yang kompleks karena data pasar keuangan memiliki karakteristik fluktuatif, non-linear, noisy, serta dipengaruhi oleh perubahan kondisi pasar dari waktu ke waktu. Beberapa penelitian menunjukkan bahwa pendekatan machine learning dan deep learning dapat digunakan untuk mempelajari pola historis pada data finansial dan membantu meningkatkan kemampuan prediksi dibandingkan pendekatan konvensional [3], [4].

Penelitian terdahulu juga menunjukkan bahwa analisis saham tidak hanya berkaitan dengan data historis harga, tetapi juga dapat dipengaruhi oleh faktor fundamental dan faktor eksternal. Penelitian [2] menunjukkan bahwa rasio keuangan dapat digunakan sebagai variabel penting dalam proses pemilihan saham berbasis machine learning. Selain itu, penelitian [3] menjelaskan bahwa karakteristik perusahaan, faktor makroekonomi, variabel pasar, dan sentimen investor dapat mendukung prediksi return saham. Hal tersebut menunjukkan bahwa pergerakan harga saham dipengaruhi oleh banyak faktor, sehingga prediksi harga saham menjadi permasalahan yang kompleks. Namun, penelitian ini membatasi variabel input pada data historis harga saham dan indikator teknikal agar fokus penelitian tetap terarah pada pengembangan model prediksi berbasis time series dan feature engineering teknikal.

Dari sisi metode, model Long Short-Term Memory (LSTM) banyak digunakan dalam penelitian prediksi saham karena memiliki kemampuan untuk mempelajari hubungan temporal pada data deret waktu. Penelitian [8] menunjukkan bahwa model LSTM yang dioptimasi dapat meningkatkan performa prediksi harga saham. Selain itu, penelitian [9] menunjukkan bahwa pendekatan multi-input LSTM dapat membantu model memanfaatkan lebih banyak informasi dalam proses prediksi. Hal

ini menunjukkan bahwa LSTM relevan digunakan untuk data saham karena harga saham memiliki pola historis yang tersusun berdasarkan urutan waktu.

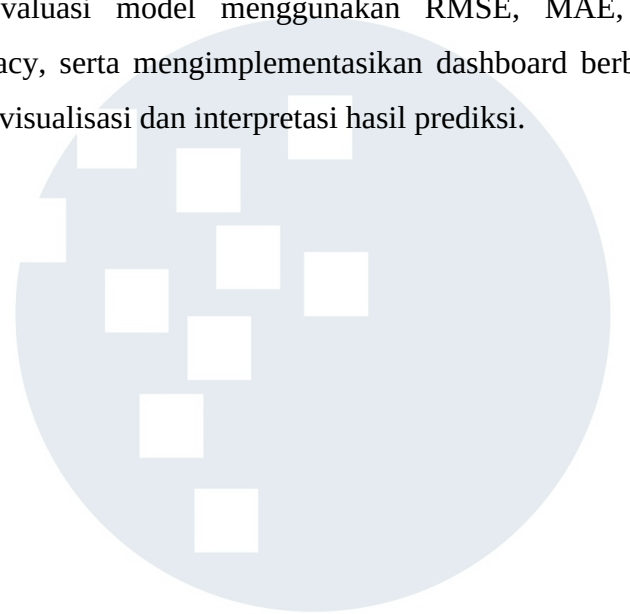
Selain LSTM, XGBoost juga banyak digunakan dalam penelitian prediksi saham karena mampu menangkap hubungan non-linear antarfitur pada data tabular. Penelitian [7] menunjukkan bahwa XGBoost dapat digunakan dalam sistem prediksi dan perdagangan saham berbasis data historis. Kemampuan XGBoost dalam mempelajari hubungan antarfitur menjadi relevan ketika data telah melalui proses feature engineering, seperti penambahan indikator teknikal. Indikator seperti SMA, RSI, dan Bollinger Bands dapat membantu model memahami tren, momentum, dan volatilitas harga saham [10].

Meskipun LSTM dan XGBoost memiliki keunggulan masing-masing, penggunaan model secara terpisah masih memiliki keterbatasan. LSTM mampu mempelajari pola temporal, tetapi hasil prediksinya dapat kurang responsif ketika terjadi perubahan harga yang tajam. Sebaliknya, XGBoost mampu menangkap hubungan non-linear antarfitur, tetapi tidak secara khusus mempelajari urutan waktu pada data time series. Oleh karena itu, beberapa penelitian mulai mengembangkan pendekatan hybrid untuk menggabungkan keunggulan lebih dari satu model dalam satu kerangka pembelajaran [6].

Penelitian dalam negeri juga menunjukkan bahwa penerapan machine learning dan deep learning pada data saham Indonesia semakin berkembang. Penelitian [11] membahas prediksi harga saham Indonesia dengan mengintegrasikan indikator teknikal dan fundamental. Penelitian [14] dan [15] menggunakan pendekatan LSTM untuk memprediksi harga saham, sedangkan penelitian [18] menerapkan model Hybrid LSTM-XGBoost pada prediksi harga penutupan saham BBRI. Selain itu, penelitian dari Universitas Multimedia Nusantara juga menunjukkan penggunaan artificial intelligence dan LSTM pada aset finansial, baik pada prediksi harga saham maupun analisis data cryptocurrency yang memiliki karakteristik volatil [12], [13].

Berdasarkan uraian tersebut, penelitian ini diarahkan untuk mengembangkan pendekatan Hybrid Stacking LSTM-XGBoost pada saham dengan volatilitas tinggi

di pasar modal Indonesia. LSTM digunakan untuk mempelajari pola temporal dari data historis saham, sedangkan output prediksi LSTM digunakan sebagai meta-feature yang diproses oleh XGBoost sebagai meta-learner untuk menghasilkan prediksi akhir. Penelitian ini diterapkan pada saham emiten Grup Bakrie, yaitu BUMI sebagai objek pengujian utama dan DEWA sebagai objek pengujian tambahan. Selain itu, penelitian ini menggunakan indikator teknikal sebagai fitur tambahan, mengevaluasi model menggunakan RMSE, MAE, MAPE, dan Directional Accuracy, serta mengimplementasikan dashboard berbasis Streamlit untuk mendukung visualisasi dan interpretasi hasil prediksi.



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Referensi ke-1

Keterangan	Uraian
Judul	Stock Selection Using Machine Learning Based on Financial Ratios
Nama Penulis	P.-F. Tsai, C.-H. Gao, S.-M. Yuan
Sumber Jurnal	Mathematics
Tahun	2023
Permasalahan	Pemilihan saham membutuhkan analisis yang tidak hanya mengandalkan harga historis, tetapi juga mempertimbangkan rasio keuangan perusahaan sebagai indikator fundamental.
Framework / Model	Machine learning berbasis financial ratios
Temuan	Rasio keuangan dapat digunakan sebagai variabel penting dalam proses pemilihan saham berbasis machine learning karena mampu merepresentasikan kondisi fundamental perusahaan.
Relevansi	Menjadi dasar penguatan latar belakang bahwa analisis saham tidak hanya dipengaruhi data teknikal, tetapi juga faktor fundamental seperti rasio keuangan.

Referensi ke-2

Keterangan	Uraian
Judul	Stock Return Prediction with Multiple Measures Using Neural Network Models
Nama Penulis	C. Wang
Sumber Jurnal	Financial Innovation
Tahun	2024
Permasalahan	Prediksi return saham dipengaruhi oleh berbagai faktor, seperti karakteristik perusahaan, kondisi makroekonomi, variabel pasar, dan sentimen investor.
Framework / Model	Neural network models
Temuan	Penggunaan beberapa jenis variabel seperti firm-specific characteristics, macroeconomic factors, market variables, dan investor sentiment dapat membantu meningkatkan prediksi return saham.
Relevansi	Menguatkan bahwa pergerakan saham dipengaruhi oleh faktor fundamental dan eksternal, meskipun penelitian ini dibatasi pada data historis harga dan indikator teknikal.

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

Referensi ke-3

Keterangan	Uraian
Judul	Forecasting Relative Returns for S&P 500 Stocks Using Machine Learning
Nama Penulis	H. H. Htun, M. Biehl, N. Petkov
Sumber Jurnal	Financial Innovation
Tahun	2024
Permasalahan	Prediksi return saham memiliki tantangan karena data pasar saham bersifat kompleks, dinamis, dan dipengaruhi banyak faktor.
Framework / Model	Machine learning
Temuan	Machine learning dapat digunakan untuk mempelajari pola dari data historis dan membantu proses prediksi return saham.
Relevansi	Mendukung urgensi penggunaan machine learning dalam prediksi saham modern, khususnya pada data yang kompleks dan fluktuatif.



Referensi ke-4

Keterangan	Uraian
Judul	A Hybrid Model for Stock Price Prediction Based on Multi-View Heterogeneous Data
Nama Penulis	W. Long, J. Gao, K. Bai, Z. Lu
Sumber Jurnal	Financial Innovation
Tahun	2024
Permasalahan	Prediksi harga saham membutuhkan pemanfaatan berbagai jenis data karena satu sumber informasi saja belum tentu cukup untuk menangkap kompleksitas pasar.
Framework / Model	Hybrid model berbasis multi-view heterogeneous data
Temuan	Pendekatan hybrid dapat memanfaatkan berbagai sumber informasi dan meningkatkan kemampuan prediksi harga saham dibanding pendekatan tunggal.
Relevansi	Relevan karena penelitian ini juga menggunakan pendekatan hybrid untuk menggabungkan keunggulan LSTM dan XGBoost.

Referensi ke-5

Keterangan	Uraian
Judul	A Machine Learning Trading System for the Stock Market Based on N-Period Min-Max Labeling Using XGBoost
Nama Penulis	Y. Han, J. Kim, D. Enke
Sumber Jurnal	Expert Systems with Applications
Tahun	2023
Permasalahan	Sistem prediksi dan trading saham membutuhkan metode yang mampu menangkap pola pergerakan harga dan hubungan antarfitur secara efektif.
Framework / Model	XGBoost dengan N-period Min-Max labeling
Temuan	XGBoost dapat digunakan dalam sistem prediksi/trading saham karena mampu mempelajari pola dari data harga dan menghasilkan performa prediksi yang baik.
Relevansi	Menguatkan pemilihan XGBoost sebagai model yang mampu menangkap hubungan non-linear pada data saham berbentuk tabular.

Referensi ke-6

Keterangan	Uraian
Judul	Stock Price Prediction with Optimized Deep LSTM Network with Artificial Rabbits Optimization Algorithm
Nama Penulis	B. Gülmez
Sumber Jurnal	Expert Systems with Applications
Tahun	2023
Permasalahan	Prediksi harga saham membutuhkan model deep learning yang mampu menangkap pola temporal pada data historis dan meningkatkan akurasi prediksi.
Framework / Model	Optimized deep LSTM
Temuan	Model LSTM yang dioptimasi mampu meningkatkan performa prediksi harga saham dibandingkan pendekatan LSTM standar.
Relevansi	Mendukung penggunaan LSTM dalam penelitian ini karena LSTM mampu mempelajari pola time series pada data saham.

Referensi ke-7

Keterangan	Uraian
Judul	Stock Movement Prediction: A Multi-Input LSTM Approach
Nama Penulis	P. Tang, C. Tang, K. Wang
Sumber Jurnal	Journal of Forecasting
Tahun	2024
Permasalahan	Prediksi pergerakan saham membutuhkan model yang mampu memanfaatkan lebih dari satu jenis input agar informasi historis dapat dipelajari secara lebih baik.
Framework / Model	Multi-input LSTM
Temuan	Pendekatan multi-input LSTM dapat membantu model memanfaatkan berbagai informasi input untuk memprediksi pergerakan saham.
Relevansi	Relevan dengan penggunaan LSTM sebagai model pembaca pola temporal pada data historis saham.



Referensi ke-8

Keterangan	Uraian
Judul	Key Technical Indicators for Stock Market Prediction
Nama Penulis	S. M. Mostafavi, A. R. Hooman
Sumber Jurnal	Machine Learning with Applications
Tahun	2025
Permasalahan	Prediksi pasar saham membutuhkan pemilihan indikator teknikal yang tepat agar model dapat memahami tren, momentum, dan volatilitas harga.
Framework / Model	Machine learning dan technical indicators
Temuan	Indikator teknikal dapat digunakan sebagai fitur penting dalam proses prediksi pasar saham.
Relevansi	Menguatkan penggunaan indikator teknikal seperti SMA, RSI, dan Bollinger Bands sebagai fitur tambahan dalam penelitian ini.



Referensi ke-9

Keterangan	Uraian
Judul	Comparative Analysis of LSTM and GRU Models for Indonesian Stock Price Prediction with Integrated Technical and Fundamental Indicators
Nama Penulis	G. Adiatmaja, R. Indraswari
Sumber Jurnal	Proceedings of 2024 International Conference on Information Technology Systems and Innovation (ICITSI)
Tahun	2024
Permasalahan	Prediksi harga saham Indonesia membutuhkan model time series yang mampu memanfaatkan indikator teknikal dan fundamental secara terpadu.
Framework / Model	LSTM dan GRU dengan indikator teknikal dan fundamental
Temuan	Integrasi indikator teknikal dan fundamental dapat digunakan untuk mendukung prediksi harga saham Indonesia.
Relevansi	Menjadi pembanding penelitian dalam negeri dan memperkuat bahwa penelitian saham di Indonesia dapat memanfaatkan indikator teknikal maupun fundamental.

Referensi ke-10

Keterangan	Uraian
Judul	How Can We Predict Transportation Stock Prices Using Artificial Intelligence? Findings from Experiments with Long Short-Term Memory Based Algorithms
Nama Penulis	D. A. Kristiyanti, W. B. N. Pramudya, S. A. Sanjaya
Sumber Jurnal	International Journal of Information Management Data Insights
Tahun	2024
Permasalahan	Prediksi harga saham sektor transportasi membutuhkan pendekatan artificial intelligence yang mampu menangkap pola historis pada data harga saham.
Framework / Model	LSTM-based algorithms
Temuan	Algoritma berbasis LSTM dapat digunakan untuk memprediksi harga saham sektor transportasi di Indonesia.
Relevansi	Memenuhi sitasi dosen UMN dan relevan karena sama-sama membahas prediksi harga saham menggunakan pendekatan LSTM.

Referensi ke-11

Keterangan	Uraian
Judul	Optimizing Long Short-Term Memory Hyperparameter for Cryptocurrency Sentiment Analysis with Swarm Intelligence Algorithms
Nama Penulis	K. Ekachandra, D. A. Kristiyanti
Sumber Jurnal	IAES International Journal of Artificial Intelligence (IJ-AI)
Tahun	2025
Permasalahan	Analisis sentimen cryptocurrency membutuhkan optimasi model deep learning karena aset kripto memiliki karakteristik volatil dan dipengaruhi oleh opini publik.
Framework / Model	LSTM dengan optimasi hyperparameter berbasis swarm intelligence
Temuan	Optimasi hyperparameter LSTM dapat membantu meningkatkan performa analisis sentimen pada data cryptocurrency.
Relevansi	Memenuhi sitasi dosen UMN dan relevan karena membahas aset finansial volatil serta penggunaan LSTM dalam konteks data finansial digital.

Referensi ke-12

Keterangan	Uraian
Judul	Prediksi Harga Penutupan Saham BBRI dengan Model Hybrid LSTM-XGBoost
Nama Penulis	N. Selayanti, D. A. Putri, T. Trimono, M. Idhom
Sumber Jurnal	Informatika: Jurnal Teknik Informatika dan Multimedia
Tahun	2025
Permasalahan	Prediksi harga penutupan saham membutuhkan model yang dapat memanfaatkan pola temporal sekaligus hubungan antarfitur agar hasil prediksi lebih akurat.
Framework / Model	Hybrid LSTM-XGBoost
Temuan	Model hybrid LSTM-XGBoost digunakan untuk memprediksi harga penutupan saham BBRI dan menunjukkan potensi dalam meningkatkan hasil prediksi.
Relevansi	Sangat relevan karena sama-sama menggunakan pendekatan Hybrid LSTM-XGBoost pada prediksi saham Indonesia.



Berdasarkan penelitian terdahulu yang telah disajikan, dapat diketahui bahwa prediksi harga saham merupakan permasalahan yang kompleks karena data pasar keuangan memiliki karakteristik fluktuatif, non-linear, noisy, serta dipengaruhi oleh perubahan kondisi pasar dari waktu ke waktu. Beberapa penelitian menunjukkan bahwa pendekatan machine learning dan deep learning dapat digunakan untuk mempelajari pola historis pada data finansial serta membantu meningkatkan kemampuan prediksi dibandingkan pendekatan konvensional. Selain itu, penelitian terbaru juga menunjukkan bahwa analisis saham tidak hanya bergantung pada data historis harga, tetapi juga dapat dipengaruhi oleh faktor fundamental seperti rasio keuangan, karakteristik perusahaan, kondisi makroekonomi, dan sentimen investor.

Dari sisi analisis fundamental, beberapa penelitian menggunakan financial ratios dan firm-specific characteristics untuk mendukung proses prediksi return maupun pemilihan saham. Hal ini menunjukkan bahwa faktor fundamental memiliki peran penting dalam menjelaskan kondisi dan prospek perusahaan. Meskipun demikian, penelitian ini tidak menggunakan data fundamental sebagai variabel input model karena ruang lingkup penelitian dibatasi pada data historis harga saham dan indikator teknikal. Pembahasan fundamental digunakan sebagai penguat latar belakang bahwa pergerakan harga saham dipengaruhi oleh banyak faktor, sehingga prediksi harga saham menjadi permasalahan yang kompleks.

Model deep learning, khususnya Long Short-Term Memory (LSTM), banyak digunakan dalam penelitian prediksi saham karena memiliki kemampuan untuk mempelajari hubungan temporal pada data deret waktu. LSTM mampu mempertahankan informasi historis dalam jangka panjang sehingga sesuai untuk memodelkan pola pergerakan harga saham yang bergantung pada urutan waktu. Beberapa penelitian terbaru juga menunjukkan bahwa penggunaan multi-input, optimasi parameter, dan pengembangan arsitektur LSTM dapat meningkatkan performa prediksi karena model memperoleh informasi yang lebih kaya dari data historis dan fitur tambahan.

Di sisi lain, algoritma Extreme Gradient Boosting (XGBoost) banyak digunakan karena mampu menangani hubungan non-linear antarfitur dan memiliki kemampuan generalisasi yang baik pada data tabular. XGBoost sesuai digunakan ketika data telah melalui proses feature engineering, seperti penambahan indikator teknikal. Indikator teknikal seperti Simple Moving Average (SMA), Relative Strength Index (RSI), dan Bollinger Bands dapat membantu model menangkap informasi mengenai tren, momentum, dan volatilitas harga saham, sehingga memperkaya input dalam proses pembelajaran model.

Meskipun LSTM dan XGBoost memiliki keunggulan masing-masing, kedua model tersebut tetap memiliki keterbatasan apabila digunakan secara terpisah. LSTM unggul dalam menangkap pola temporal, tetapi prediksi yang dihasilkan dapat kurang responsif ketika terjadi perubahan harga yang tajam. Sebaliknya, XGBoost efektif dalam memodelkan hubungan non-linear antarfitur, tetapi tidak secara eksplisit mempelajari urutan waktu pada data time series. Kondisi tersebut menunjukkan bahwa integrasi kedua model diperlukan agar proses prediksi dapat memanfaatkan kemampuan temporal learning dan non-linear feature learning secara bersamaan.

Beberapa penelitian terdahulu telah menunjukkan bahwa pendekatan hybrid dan stacking dapat digunakan untuk menggabungkan lebih dari satu model dalam satu kerangka pembelajaran. Pendekatan stacking memungkinkan output dari model dasar digunakan sebagai informasi tambahan bagi model berikutnya untuk menghasilkan prediksi akhir. Dalam konteks penelitian ini, LSTM digunakan untuk mempelajari pola temporal dari data historis saham, sedangkan hasil prediksi LSTM digunakan sebagai meta-feature yang diproses oleh XGBoost sebagai meta-learner. Dengan demikian, model Hybrid Stacking LSTM-XGBoost diharapkan dapat menghasilkan prediksi yang lebih adaptif terhadap karakteristik saham dengan volatilitas tinggi.

Berdasarkan penelitian luar negeri, model machine learning, deep learning, dan hybrid telah banyak dikembangkan untuk prediksi harga saham dengan memanfaatkan data historis, indikator teknikal, indikator fundamental, maupun

data eksternal lainnya. Sementara itu, penelitian dalam negeri menunjukkan bahwa model LSTM, GRU, XGBoost, dan hybrid mulai banyak diterapkan pada data saham Indonesia, terutama pada sektor perbankan, transportasi, dan saham-saham yang memiliki pola pergerakan harga dinamis. Selain itu, riset dosen Universitas Multimedia Nusantara juga menunjukkan bahwa pendekatan artificial intelligence, LSTM, dan analisis berbasis data finansial relevan digunakan pada aset yang memiliki karakteristik volatil, seperti saham dan cryptocurrency.

Berdasarkan kajian tersebut, penelitian ini memiliki perbedaan dibandingkan penelitian sebelumnya. Penelitian ini berfokus pada penerapan Hybrid Stacking LSTM-XGBoost pada saham emiten Grup Bakrie, yaitu BUMI sebagai objek pengujian utama dan DEWA sebagai objek pengujian tambahan. Pemilihan objek tersebut dilakukan karena kedua saham memiliki karakteristik pergerakan harga yang fluktuatif dan sesuai untuk menguji performa model pada kondisi volatilitas tinggi. Penelitian ini juga mengintegrasikan indikator teknikal sebagai fitur tambahan, menggunakan evaluasi multi-metrik seperti RMSE, MAE, MAPE, dan Directional Accuracy, serta mengimplementasikan dashboard visualisasi berbasis Streamlit untuk mendukung proses analisis dan interpretasi hasil prediksi secara lebih komprehensif.



2.2 Teori yang berkaitan

Bagian ini memaparkan konsep dan teori yang menjadi landasan dalam penelitian, meliputi karakteristik pasar modal, data *time series*, serta metode *Machine Learning* dan *deep learning* yang digunakan dalam proses prediksi harga saham.

2.2.1 Teori Pasar Modal dan Harga Saham

Pasar modal merupakan instrumen penting dalam sistem keuangan yang berfungsi sebagai sarana penghimpunan dana jangka panjang melalui perdagangan efek, termasuk saham. Harga saham mencerminkan nilai suatu perusahaan yang terbentuk melalui mekanisme permintaan dan penawaran di pasar. Pergerakan harga saham bersifat dinamis dan dipengaruhi oleh berbagai faktor, baik internal maupun eksternal, seperti kinerja perusahaan, kondisi ekonomi makro, kebijakan pemerintah, serta sentimen investor [1].

Dalam praktiknya, harga saham memiliki karakteristik volatilitas yang tinggi dan cenderung tidak stabil. Fluktuasi ini menunjukkan bahwa pola pergerakan harga saham tidak hanya bersifat linear, tetapi juga mengandung pola kompleks yang sulit dimodelkan menggunakan pendekatan konvensional. Oleh karena itu, diperlukan metode analisis yang mampu menangkap hubungan non-linear dan dinamika temporal dalam data historis harga saham [2].

2.2.2 Teori Data Time Series dalam Prediksi Keuangan

Data *time series* merupakan data yang dikumpulkan secara berurutan berdasarkan waktu, di mana setiap observasi memiliki ketergantungan terhadap nilai sebelumnya. Dalam konteks keuangan, data harga saham termasuk dalam kategori *time series* yang memiliki karakteristik seperti tren, musiman, volatilitas, serta *noise* yang tinggi [3].

Salah satu tantangan utama dalam pemodelan *time series* adalah adanya pola non-linear dan perubahan distribusi data secara dinamis (*non-stationary*). Hal ini menyebabkan metode statistik konvensional seringkali kurang mampu memberikan

hasil prediksi yang optimal, sehingga diperlukan pendekatan berbasis *Machine Learning* yang lebih adaptif [4].

Secara konseptual, teori yang berkaitan dengan *Machine Learning* dan deep learning dalam penelitian ini dapat dijelaskan melalui beberapa poin berikut.

2.2.3 *Machine Learning* dalam Prediksi Saham

Machine Learning merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan sistem komputer untuk mempelajari pola dari data secara otomatis dan menghasilkan prediksi tanpa pemrograman eksplisit. Dalam konteks prediksi harga saham, *Machine Learning* digunakan untuk membangun model regresi yang mampu memperkirakan nilai harga saham pada periode berikutnya berdasarkan data historis serta pola pergerakan pasar yang terbentuk dari data time series.

Prediksi harga saham termasuk dalam kategori *supervised learning*, yaitu metode pembelajaran yang menggunakan data berlabel. Variabel input (*features*) dalam penelitian ini berupa data historis saham seperti *Open*, *High*, *Low*, *Close*, dan *Volume*, serta indikator teknikal seperti *Simple Moving Average* (SMA), *Relative Strength Index* (RSI), dan *Bollinger Bands*. Sementara itu, variabel output berupa harga penutupan (*Close*) pada periode berikutnya. Permasalahan ini termasuk dalam kategori regresi karena menghasilkan nilai numerik sebagai output [4].

Pendekatan *Machine Learning* memiliki keunggulan dalam menangani hubungan non-linear serta interaksi kompleks antar variabel, sehingga lebih adaptif dibandingkan metode statistik konvensional dalam memodelkan data finansial. Selain itu, penggunaan indikator teknikal sebagai proses *feature engineering* juga membantu model dalam menangkap pola tren, momentum, dan volatilitas pasar secara lebih efektif, khususnya pada saham dengan karakteristik volatilitas tinggi.

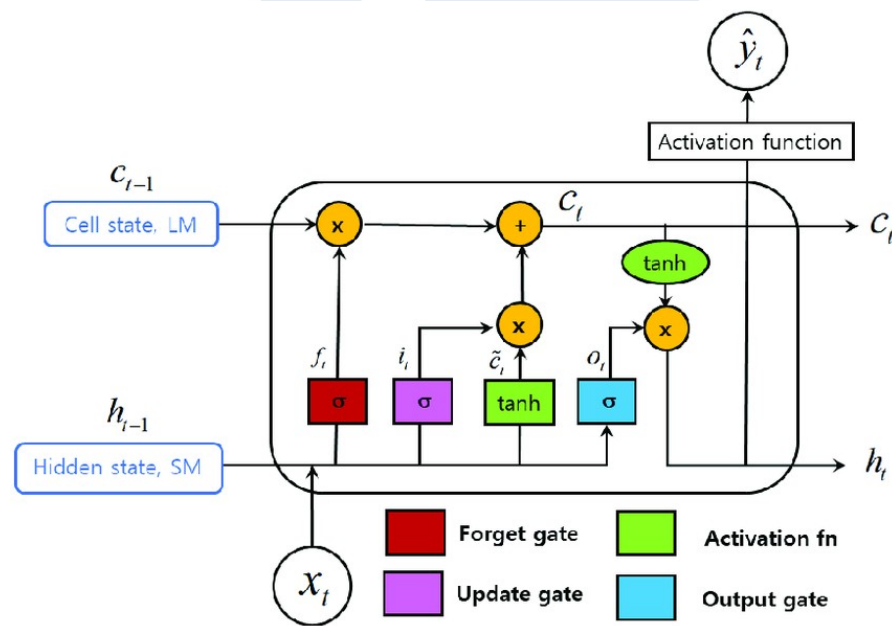
2.2.4 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) merupakan model komputasi yang terinspirasi dari jaringan saraf biologis manusia. ANN terdiri atas beberapa komponen utama:

- **Input layer:** Menerima data masukan.
- **Hidden layer:** Melakukan proses komputasi non-linear.
- **Output layer:** Menghasilkan prediksi.

ANN sangat efektif dalam memodelkan hubungan non-linear dalam data, termasuk data finansial yang memiliki pola kompleks.

2.2.5 Deep Learning dan Long Short-Term Memory (LSTM)



Gambar 2.1 Ilustrasi Arsitektur LSTM

Gambar 2.1 menunjukkan ilustrasi arsitektur LSTM yang digunakan untuk mempelajari pola data berurutan. Ilustrasi tersebut membantu menjelaskan

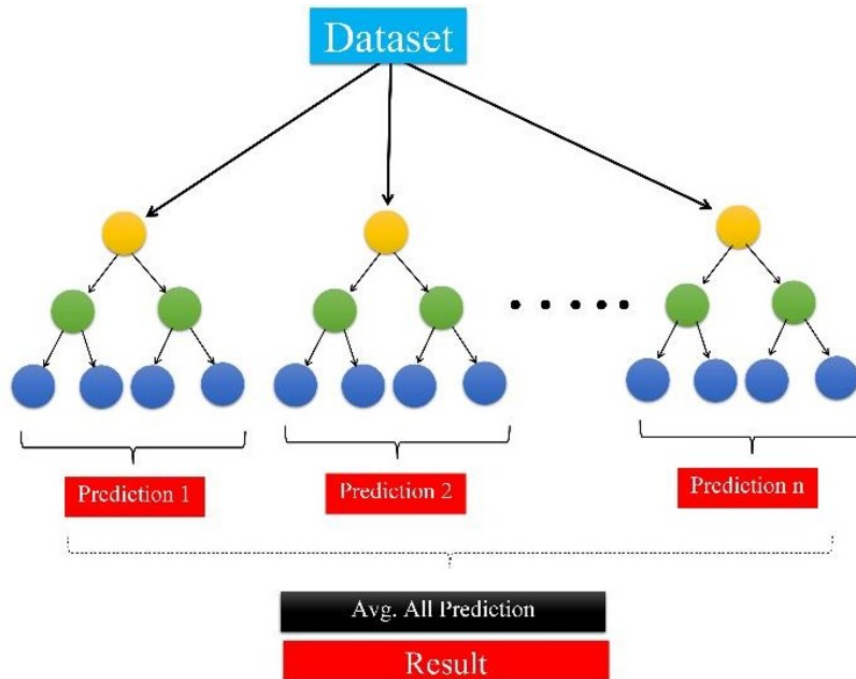
mekanisme utama LSTM dalam menyimpan dan memperbarui informasi melalui struktur gate.

Deep learning merupakan pengembangan dari *Machine Learning* yang memanfaatkan jaringan saraf tiruan dengan banyak lapisan (*deep neural network*) untuk memodelkan hubungan yang kompleks dalam data. Pendekatan ini sangat efektif dalam menangani data berdimensi tinggi serta pola non-linear yang sulit dimodelkan menggunakan metode konvensional. Dalam konteks prediksi data deret waktu (*time series*), salah satu arsitektur yang banyak digunakan adalah *Long Short-Term Memory* (LSTM), yang merupakan varian dari *Recurrent Neural Network* (RNN) [7].

LSTM dirancang untuk mengatasi permasalahan *vanishing gradient* yang sering terjadi pada RNN tradisional, sehingga mampu mempertahankan informasi dalam jangka panjang. Arsitektur LSTM memiliki mekanisme *gating* yang terdiri atas tiga komponen utama, yaitu *forget gate*, *input gate*, dan *output gate*. *Forget gate* berfungsi untuk menentukan informasi yang perlu dipertahankan atau dihapus dari memori, *input gate* mengatur informasi baru yang akan disimpan, sedangkan *output gate* menentukan informasi yang akan diteruskan ke lapisan berikutnya sebagai output model [8].

Kemampuan LSTM dalam mengelola memori jangka panjang menjadikannya efektif dalam menangkap pola temporal dan dependensi antar waktu pada data *time series*. Dalam prediksi harga saham, LSTM mampu mengidentifikasi tren historis serta pola pergerakan harga yang kompleks berdasarkan data sebelumnya. Namun demikian, model ini memiliki keterbatasan dalam merespons perubahan pola yang terjadi secara tiba-tiba, terutama pada kondisi pasar dengan volatilitas tinggi. Hal ini menyebabkan LSTM cenderung menghasilkan prediksi yang bersifat *lagging*, sehingga kurang adaptif terhadap fluktuasi harga yang ekstrem [9].

2.2.6 Ensemble Learning dan Extreme Gradient Boosting (XGBoost)



Gambar 2.2 Ilustrasi Proses XGBoost

Gambar 2.2 menunjukkan ilustrasi proses XGBoost sebagai metode ensemble berbasis pohon keputusan. Proses tersebut menggambarkan bagaimana model membangun beberapa learner secara bertahap untuk memperbaiki kesalahan prediksi sebelumnya.

Ensemble learning merupakan teknik dalam *Machine Learning* yang mengkombinasikan beberapa model untuk meningkatkan performa prediksi dengan cara mengurangi bias dan varians. Salah satu algoritma yang populer dalam kategori ini adalah *Extreme Gradient Boosting* (XGBoost), yang merupakan pengembangan dari metode *gradient boosting* [10].

XGBoost bekerja dengan membangun model secara bertahap (*iterative boosting*) menggunakan serangkaian *decision trees*, di mana setiap model baru berfokus pada kesalahan yang dihasilkan oleh model sebelumnya. Pendekatan ini

memungkinkan model untuk terus memperbaiki performa secara progresif melalui proses optimasi berbasis *gradient descent* terhadap fungsi *loss*. Selain itu, XGBoost dilengkapi dengan mekanisme regularisasi yang efektif dalam mengurangi risiko *overfitting*, serta dirancang untuk memiliki efisiensi komputasi yang tinggi sehingga mampu menangani data dalam jumlah besar [11].

Keunggulan utama XGBoost terletak pada kemampuannya dalam memodelkan hubungan non-linear serta menangani interaksi kompleks antar fitur. Hal ini menjadikan XGBoost sangat cocok digunakan pada permasalahan regresi, termasuk dalam prediksi harga saham yang memiliki karakteristik data yang tidak linear dan dinamis. Dalam konteks prediksi saham, XGBoost memanfaatkan fitur-fitur historis seperti *Open*, *High*, *Low*, *Close*, dan *Volume* untuk menghasilkan estimasi harga yang lebih akurat melalui pendekatan berbasis pohon keputusan yang diperkuat secara bertahap.

Namun demikian, XGBoost tidak memiliki mekanisme eksplisit untuk menangkap dependensi temporal dalam data *time series*. Model ini memperlakukan data sebagai observasi independen, sehingga kurang optimal dalam memahami urutan waktu yang menjadi karakteristik utama data saham. Oleh karena itu, diperlukan kombinasi dengan metode lain, seperti *deep learning*, untuk meningkatkan kemampuan model dalam menangkap pola temporal secara lebih efektif [12].

2.2.7 Perbandingan LSTM dan XGBoost

Secara konseptual, perbedaan kedua model dapat dirangkum sebagai berikut:

Tabel 2.1 Tabel Perbedaan Kedua Model

Aspek	LSTM	XGBoost
Pendekatan	Deep Learning	Ensemble <i>Machine Learning</i>
Fokus	Pola temporal & dependensi jangka panjang	Optimasi prediksi berbasis pohon keputusan
Kelebihan	Menangkap pola sekuensial	Stabil, cepat, dan regularisasi

	kompleks	kuat
Tantangan	Komputasi lebih berat	Tidak eksplisit menangkap urutan waktu

Perbedaan tersebut menjadikan kedua model menarik untuk dibandingkan dalam konteks prediksi harga saham harian pada emiten dengan karakteristik volatilitas dan likuiditas yang berbeda.

2.2.8 Hybrid Stacking LSTM-XGBoost

Hybrid stacking merupakan salah satu pendekatan dalam *ensemble learning* yang mengkombinasikan beberapa model dengan tujuan untuk meningkatkan akurasi prediksi melalui integrasi keunggulan masing-masing model. Berbeda dengan metode ensemble lainnya seperti *bagging* atau *boosting*, teknik *stacking* menggunakan model tambahan yang disebut *meta-learner* untuk menggabungkan output dari beberapa *base learner* menjadi prediksi akhir yang lebih optimal [13].

Dalam konteks prediksi harga saham, pendekatan *stacking* menjadi relevan karena data saham memiliki karakteristik yang kompleks, yaitu mengandung pola *temporal dependency* sekaligus hubungan non-linear antar fitur. Selain itu, data saham juga bersifat dinamis, fluktuatif, dan memiliki tingkat noise yang tinggi, sehingga penggunaan model tunggal seringkali belum mampu menangkap seluruh karakteristik data secara optimal.

Pada penelitian ini, dikembangkan model *Hybrid Stacking LSTM-XGBoost* yang mengkombinasikan keunggulan dari model *deep learning* dan *ensemble learning*. Dalam arsitektur yang diusulkan, *Long Short-Term Memory* (LSTM) berperan sebagai *base learner* sekaligus *feature extractor* yang bertugas untuk mempelajari pola temporal dan dependensi jangka panjang pada data *time series* saham. Model LSTM memproses data historis saham secara sekuensial berdasarkan urutan waktu untuk menghasilkan representasi fitur dan prediksi awal yang mengandung informasi temporal dari pergerakan harga saham.

Dalam implementasinya, output prediksi yang dihasilkan oleh model LSTM digunakan sebagai fitur tambahan (*meta-feature*) yang dikombinasikan dengan fitur historis saham dan indikator teknikal, seperti *Simple Moving Average* (SMA), *Relative Strength Index* (RSI), dan *Bollinger Bands*. Selanjutnya, seluruh fitur tersebut digunakan sebagai input bagi model *Extreme Gradient Boosting* (XGBoost) yang berperan sebagai *meta-learner*.

XGBoost kemudian melakukan proses pembelajaran lanjutan untuk memodelkan hubungan non-linear antar fitur serta mengoptimalkan prediksi akhir melalui mekanisme *gradient boosting*. Dengan demikian, model hybrid ini mampu mengintegrasikan kemampuan LSTM dalam menangkap pola temporal dengan kemampuan XGBoost dalam melakukan optimasi berbasis fitur non-linear secara lebih efektif [14].

Pendekatan hybrid stacking ini diharapkan mampu menghasilkan model prediksi yang lebih akurat, stabil, dan adaptif terhadap perubahan pola pasar, khususnya pada saham dengan karakteristik volatilitas ekstrem seperti emiten Grup Bakrie. Selain itu, kombinasi antara *deep learning*, *ensemble learning*, dan *feature engineering* pada penelitian ini juga menjadi bentuk pengembangan metode prediksi yang lebih komprehensif dalam bidang *financial time series forecasting*.

Secara konseptual, alur kerja model Hybrid Stacking LSTM-XGBoost dalam penelitian ini dapat dijelaskan sebagai berikut:

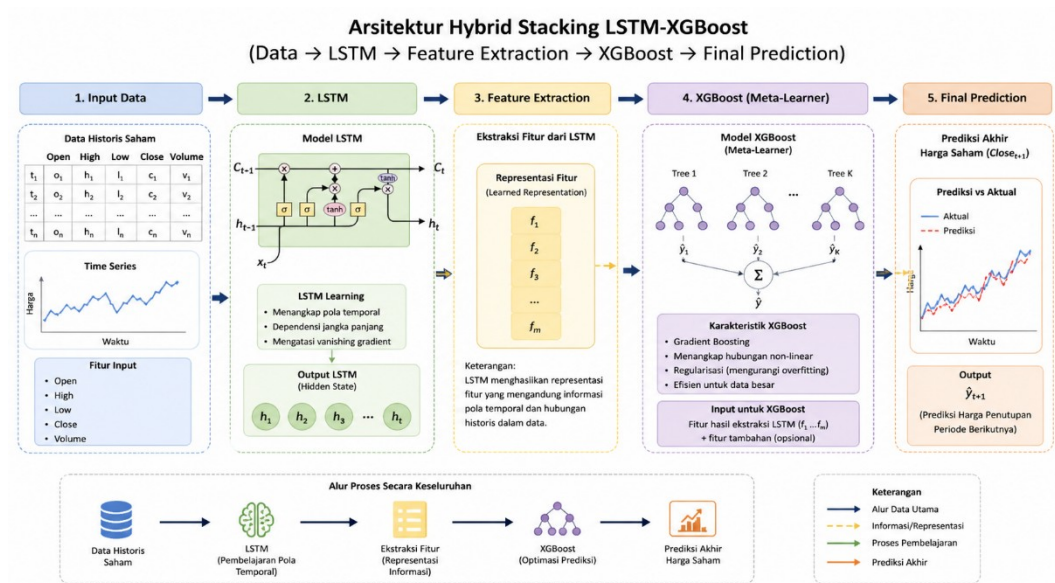
1. Data historis saham dan indikator teknikal (Open, High, Low, Close, Volume, SMA, RSI, Bollinger Bands) diproses dan dibentuk menjadi dataset time series.
2. Model LSTM dilatih untuk mempelajari pola sekuensial dan menghasilkan representasi fitur atau prediksi awal.
3. Output dari LSTM digunakan sebagai input tambahan bagi model XGBoost.
4. Model XGBoost dilatih untuk mengoptimalkan prediksi akhir berdasarkan fitur hasil ekstraksi LSTM dan data input lainnya.

Pendekatan ini memberikan beberapa keunggulan dibandingkan penggunaan model tunggal, antara lain:

- a. Mampu menangkap dependensi temporal melalui LSTM.
- b. Mampu memodelkan hubungan non-linear secara lebih efektif melalui XGBoost.
- c. Mengurangi bias dan varians melalui kombinasi model.
- d. Meningkatkan robustnes model terhadap data dengan volatilitas tinggi.

Dalam konteks saham dengan volatilitas ekstrem seperti emiten Grup Bakrie, pendekatan hybrid stacking menjadi sangat relevan karena mampu menangani fluktuasi harga yang tajam serta pola data yang tidak stabil. Model ini diharapkan dapat menghasilkan prediksi yang lebih akurat dan adaptif dibandingkan model tunggal, khususnya dalam kondisi pasar yang dinamis dan penuh ketidakpastian.

Dengan demikian, penggunaan Hybrid Stacking LSTM-XGBoost dalam penelitian ini tidak hanya bertujuan untuk meningkatkan performa prediksi, tetapi juga sebagai bentuk pengembangan metode analitik yang mengintegrasikan pendekatan deep learning dan ensemble learning dalam satu kerangka yang komprehensif. Pendekatan ini menjadi kontribusi penting dalam bidang time series forecasting, khususnya pada prediksi harga saham dengan karakteristik volatilitas tinggi [15].



2.2.9 Technical Indicators

Technical indicators merupakan indikator matematis yang dibentuk berdasarkan data historis harga dan volume perdagangan saham untuk membantu proses analisis teknikal serta mengidentifikasi pola pergerakan pasar. Dalam konteks prediksi harga saham berbasis *Machine Learning* dan deep learning, *technical indicators* digunakan sebagai bagian dari proses *feature engineering* untuk memperkaya informasi yang dipelajari model. Penggunaan indikator teknikal dapat membantu model dalam menangkap pola tren, momentum, dan volatilitas pasar yang sulit diperoleh hanya dari data harga mentah (*raw historical data*) [16].

Pada penelitian ini, indikator teknikal yang digunakan meliputi *Simple Moving Average* (SMA), *Relative Strength Index* (RSI), dan *Bollinger Bands*. Ketiga indikator tersebut dipilih karena mampu merepresentasikan karakteristik utama pergerakan saham, yaitu tren harga, momentum pasar, dan tingkat volatilitas.

1. Simple Moving Average (SMA)

Simple Moving Average (SMA) merupakan indikator teknikal yang digunakan untuk mengidentifikasi arah tren harga saham berdasarkan rata-rata harga

penutupan dalam periode tertentu. SMA membantu mengurangi noise pada data harga saham sehingga pola tren dapat terlihat lebih jelas.

Secara matematis, SMA dapat dirumuskan sebagai berikut:

$$SMA_n = \frac{1}{n} \sum_{i=1}^n P_i \quad (2.1)$$

Simple Moving Average

Keterangan:

- a. SMA_n : nilai Simple Moving Average pada periode ke- n
- b. P_i : harga penutupan saham
- c. n : jumlah periode perhitungan

Dalam penelitian ini digunakan SMA dengan periode 10 hari (*SMA 10*) dan 20 hari (*SMA 20*) untuk membantu model dalam mengenali pola tren jangka pendek dan menengah pada pergerakan harga saham.

2. Relative Strength Index (RSI)

Relative Strength Index (RSI) merupakan indikator momentum yang digunakan untuk mengukur kekuatan dan kecepatan perubahan harga saham. RSI digunakan untuk mengidentifikasi kondisi *overbought* (jenuh beli) dan *oversold* (jenuh jual) pada pasar saham.

Nilai RSI berada pada rentang 0–100 dan secara matematis dirumuskan sebagai berikut:

$$RSI = 100 - \frac{100}{1 + RS} \quad (2.2)$$

Relative Strength Index

dengan:

$$RS = \frac{\text{Average Gain}}{\text{Average Loss}} \quad (2.3)$$

Relative Strength

Keterangan:

- a. *RSI*: Relative Strength Index
- b. *RS*: Relative Strength
- c. *Average Gain* : rata-rata kenaikan harga
- d. *Average Loss* : rata-rata penurunan harga

Pada penelitian ini digunakan *RSI* periode 14 hari (*RSI 14*) untuk membantu model dalam menangkap momentum perubahan harga saham.

3. Bollinger Bands

Bollinger Bands merupakan indikator teknikal yang digunakan untuk mengukur tingkat volatilitas harga saham. Indikator ini terdiri atas tiga komponen utama, yaitu:

- *Middle Band* : rata-rata bergerak (*moving average*)
- *Upper Band* : batas atas volatilitas
- *Lower Band* : batas bawah volatilitas

Secara matematis, *Bollinger Bands* dirumuskan sebagai berikut:

Middle Band

$$\text{Middle Band} = SM A_n \quad (2.4)$$

Upper Bollinger Band

Upper Band

$$Upper\ Band = SM\ A_n + 2\sigma \quad (2.5)$$

Lower Bollinger Band

Lower Band

$$Lower\ Band = SM\ A_n - 2\sigma \quad (2.6)$$

Middle Bollinger Band

Keterangan:

- a. $SM\ A_n$: Simple Moving Average
- b. σ : standar deviasi harga saham

Bollinger Bands digunakan dalam penelitian ini untuk membantu model dalam mengenali tingkat volatilitas pasar. Semakin lebar jarak antara *upper band* dan *lower band*, maka semakin tinggi tingkat volatilitas saham.

Penggunaan indikator teknikal dalam penelitian ini diharapkan mampu meningkatkan kualitas fitur yang digunakan dalam proses pelatihan model Hybrid Stacking LSTM-XGBoost. Dengan adanya tambahan informasi mengenai tren, momentum, dan volatilitas pasar, model diharapkan dapat menghasilkan prediksi harga saham yang lebih akurat dan adaptif terhadap perubahan kondisi pasar.

2.3 Framework/Algoritma yang digunakan

Framework penelitian ini dirancang untuk membangun model prediksi harga saham harian menggunakan dua pendekatan berbeda, yaitu model berbasis deep learning (LSTM) dan model berbasis ensemble learning (XGBoost). Kedua algoritma tersebut memiliki prinsip kerja matematis yang berbeda dalam memproses data time series.

Dalam konteks penelitian ini, data harga saham harian direpresentasikan sebagai deret waktu:

$$X = \{x_1, x_2, x_3, \dots, x_t\}$$

di mana setiap x_t merepresentasikan fitur saham pada waktu ke- t (Open, High, Low, Close, Volume), dan target prediksi adalah harga penutupan pada periode berikutnya y_{t+1} .

2.3.1 Long Short-Term Memory (LSTM)

LSTM digunakan karena data harga saham memiliki ketergantungan temporal, artinya nilai hari ini dipengaruhi oleh nilai hari sebelumnya. Model linear biasa tidak mampu menangkap ketergantungan jangka panjang tersebut.

1. Konsep Hidden State dan Cell State

Dalam LSTM terdapat dua komponen utama:

- a. **Hidden state** (h_t) $\rightarrow$ output jangka pendek
- b. **Cell state** (C_t) $\rightarrow$ memori jangka panjang

Cell state berfungsi sebagai penyimpan informasi historis yang relevan. Inilah alasan LSTM lebih cocok untuk data time series dibanding ANN biasa.

2. Forget Gate

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.7)$$

Forget Gate LSTM

Penjelasan:

- a. Fungsi sigmoid menghasilkan nilai antara 0 dan 1.
- b. Jika nilai mendekati 0 → informasi dibuang.
- c. Jika mendekati 1 → informasi dipertahankan.

Makna dalam saham:

Forget gate menentukan apakah informasi harga lama masih relevan atau tidak. Misalnya, jika tren lama sudah tidak relevan karena terjadi lonjakan volume besar, maka sebagian informasi historis dapat dikurangi pengaruhnya.

3. Input Gate

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.8)$$

Input Gate LSTM

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.9)$$

Candidate Cell State LSTM

Penjelasan:

- a. i_t menentukan seberapa besar informasi baru dimasukkan.
- b. $\tilde{C}_t$ adalah kandidat informasi baru.

Makna dalam saham:

Model mengevaluasi apakah pergerakan harga dan volume saat ini perlu dimasukkan sebagai memori baru.

4. Update Cell State

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

Rumus ini menunjukkan bahwa memori baru merupakan kombinasi:

- a. Informasi lama yang dipertahankan
- b. Informasi baru yang relevan

Ini sangat penting karena harga saham memiliki pola tren dan momentum.

5. Output Gate

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.10)$$

Cell State Update LSTM

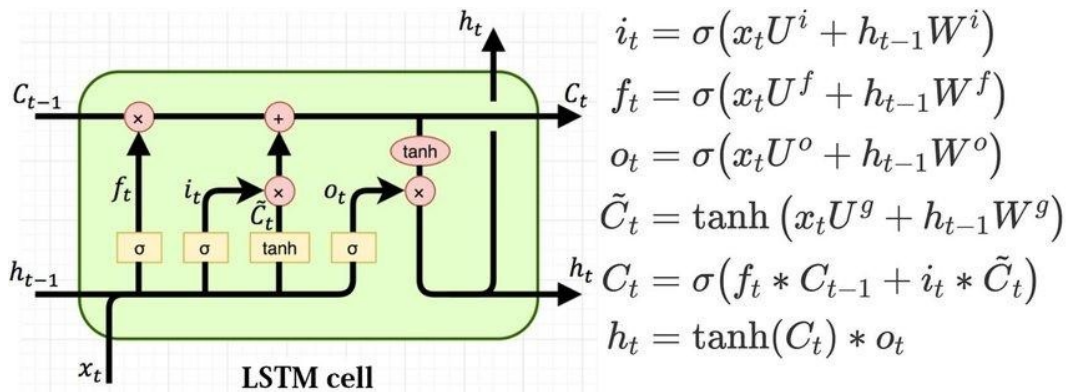
$$h_t = o_t \cdot \tanh(C_t) \quad (2.11)$$

Output Gate LSTM

Hidden state h_t inilah yang digunakan untuk memprediksi harga berikutnya.

Makna matematisnya:

Model melakukan transformasi non-linear untuk menghasilkan estimasi harga berdasarkan seluruh informasi historis yang telah difilter.



Mengapa LSTM Cocok untuk Saham?

Karena:

1. Harga saham bersifat sequential.
2. Ada pola tren jangka panjang.
3. Volatilitas bisa dipengaruhi memori historis.

Secara matematis, LSTM mengimplementasikan sistem dinamika non-linear berbasis fungsi aktivasi sigmoid dan tanh yang mampu menangkap pola kompleks.

2.3.2 Extreme Gradient Boosting (XGBoost)

Berbeda dengan LSTM yang berbasis jaringan saraf, XGBoost berbasis pohon keputusan dan optimasi fungsi loss.

1. Fungsi Objektif

$$Obj = \sum_{i=1}^n \square l(y_i, \hat{y}_i) + \sum_{k=1}^K \square \Omega(f_k) \quad (2.12)$$

Hidden State LSTM

Makna:

- a. Bagian pertama → mengukur kesalahan prediksi.
- b. Bagian kedua → regularisasi (mengontrol kompleksitas model).

Ini penting karena data saham sering noisy dan mudah overfitting.

2. Loss Function (Mean Squared Error)

$$l(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2 \quad (2.13)$$

XGBoost Objective Function

Semakin kecil selisih antara harga aktual dan prediksi, semakin kecil nilai loss.

3. Regularisasi

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda ||w||^2 \quad (2.14)$$

XGBoost Regularization Term

Penjelasan:

- a. T = jumlah daun pohon
- b. w = bobot tiap daun
- c. γ, λ = parameter penalti

Maknanya:

Model dihukum jika terlalu kompleks. Ini membuat XGBoost lebih stabil.

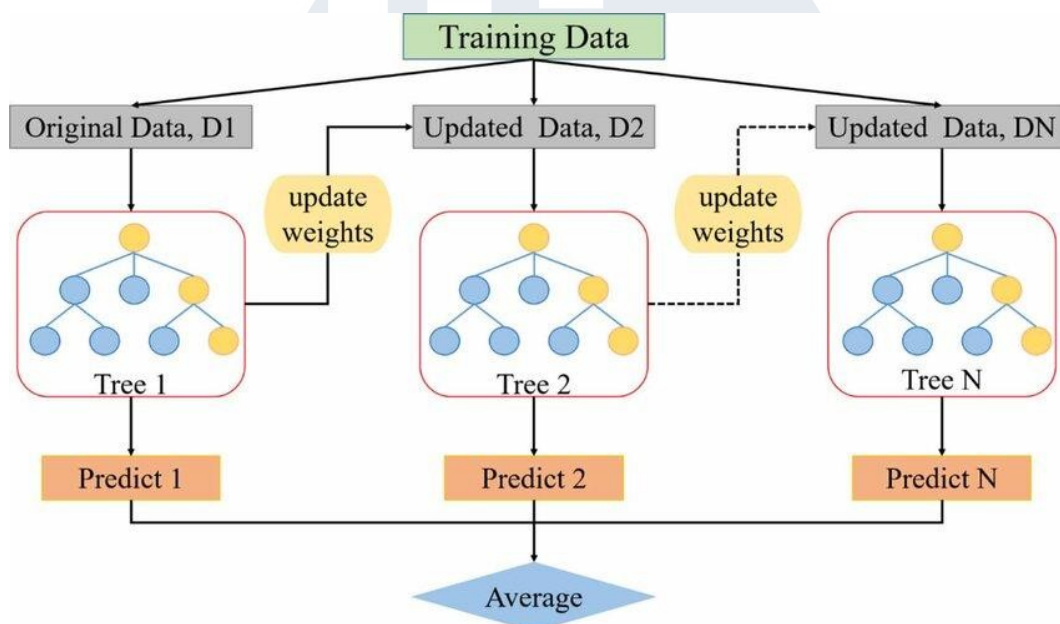
4. Boosting Iteratif

$$\hat{y}^{(t)} = \hat{y}^{(1)} + \eta f_t(x)$$

Artinya:

Prediksi baru = prediksi lama + koreksi error.

Secara matematis, ini adalah proses optimasi bertahap menggunakan gradient descent.



Mengapa XGBoost Cocok untuk Saham?

1. Mampu menangani hubungan non-linear.
2. Tidak memerlukan asumsi distribusi data.
3. Memiliki regularisasi untuk mencegah overfitting.
4. Stabil pada data dengan noise tinggi.

Perbandingan Matematis

Tabel 2.2 Tabel Perbandingan Matematis Kedua Model

Aspek	LSTM	XGBoost
Model	Sistem dinamis non-linear	Ensemble pohon keputusan
Fokus	Dependensi temporal	Mi00000073372asi error global
Optimasi	Backpropagation Through Time	Gradient boosting
Kompleksitas	Tinggi	Sedang

Secara matematis, LSTM bekerja melalui sistem memori berlapis yang mempertahankan informasi temporal, sedangkan XGBoost bekerja melalui optimasi fungsi loss secara iteratif dengan regularisasi. Perbedaan prinsip matematis ini menjadi dasar dilakukannya perbandingan performa kedua model dalam memprediksi harga saham harian pada emiten Grup Bakrie

2.3.3 Evaluation Metrics

Evaluasi model merupakan tahap penting dalam proses pengembangan sistem prediksi untuk mengukur tingkat akurasi dan performa model terhadap data pengujian (*testing data*). Pada penelitian prediksi harga saham, proses evaluasi dilakukan dengan membandingkan nilai prediksi yang dihasilkan model terhadap nilai aktual harga saham. Penggunaan metrik evaluasi yang tepat sangat penting untuk mengetahui kemampuan model dalam meminimalkan kesalahan prediksi serta memahami kualitas hasil prediksi secara lebih komprehensif.

Dalam penelitian ini, evaluasi performa model dilakukan menggunakan beberapa metrik, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan *Directional Accuracy*. Penggunaan beberapa metrik evaluasi dilakukan untuk memperoleh analisis performa model dari berbagai aspek, baik berdasarkan besar kesalahan prediksi maupun kemampuan model dalam memprediksi arah pergerakan harga saham.

1. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) merupakan metrik evaluasi yang digunakan untuk mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi. MAE menunjukkan seberapa besar rata-rata kesalahan prediksi model tanpa memperhatikan arah kesalahan.

Secara matematis, MAE dirumuskan sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.15)$$

Root Mean Square Error

Keterangan:

- a. y_i : nilai aktual
- b. $\hat{y}_i$: nilai prediksi
- c. n : jumlah data pengujian

Nilai MAE yang lebih kecil menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang lebih rendah dan performa yang lebih baik.

2. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) merupakan metrik evaluasi yang digunakan untuk mengukur tingkat kesalahan prediksi dengan memberikan penalti yang lebih besar terhadap error yang ekstrem. RMSE sangat cocok digunakan pada data saham yang memiliki tingkat volatilitas tinggi karena lebih sensitif terhadap kesalahan prediksi yang besar.

Secara matematis, RMSE dirumuskan sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.16)$$

Mean Absolute Error

Keterangan:

- a. y_i : nilai aktual
- b. $\hat{y}_i$: nilai prediksi
- c. n : jumlah data pengujian

Semakin kecil nilai RMSE, maka semakin baik performa model dalam memprediksi harga saham.

3. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) merupakan metrik evaluasi yang digunakan untuk mengukur rata-rata persentase kesalahan prediksi terhadap nilai aktual. MAPE memberikan interpretasi yang lebih mudah dipahami karena dinyatakan dalam bentuk persentase.

Secara matematis, MAPE dirumuskan sebagai berikut:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2.17)$$

Mean Absolute Percentage Error

Keterangan:

- a. y_i : nilai aktual
- b. $\hat{y}_i$: nilai prediksi
- c. n : jumlah data pengujian

Nilai MAPE yang lebih kecil menunjukkan bahwa persentase kesalahan prediksi model semakin rendah.

4. Directional Accuracy

Directional Accuracy merupakan metrik evaluasi yang digunakan untuk mengukur kemampuan model dalam memprediksi arah pergerakan harga saham, apakah mengalami kenaikan atau penurunan dibandingkan periode sebelumnya. Dalam konteks prediksi saham, kemampuan memprediksi arah pergerakan harga menjadi aspek penting karena berkaitan dengan pengambilan keputusan investasi.

Secara matematis, *Directional Accuracy* dapat dirumuskan sebagai berikut:

$$Directional Accuracy = \frac{Jumlah\ Prediksi\ Arah\ Benar}{Jumlah\ Total\ Prediksi} \times 100\% \quad (2.18)$$

Directional
Accuracy

Metrik ini menghitung persentase jumlah prediksi arah harga yang sesuai dengan arah pergerakan harga aktual. Semakin tinggi nilai *Directional Accuracy*, maka semakin baik kemampuan model dalam memprediksi arah pergerakan harga saham.

Dalam penelitian ini, keempat metrik evaluasi tersebut digunakan untuk membandingkan performa model LSTM, XGBoost, dan Hybrid Stacking LSTM-XGBoost. Penggunaan evaluasi multi-metrik diharapkan mampu memberikan analisis performa model secara lebih komprehensif, baik dari aspek besar kesalahan prediksi maupun kemampuan model dalam memahami pola pergerakan harga saham pada kondisi volatilitas ekstrem.

2.4 Tools/software yang digunakan

2.4.1 Python

Python digunakan sebagai bahasa pemrograman utama dalam penelitian ini. *Python* dipilih karena memiliki ekosistem library yang lengkap untuk pengolahan data, *Machine Learning*, dan deep learning. Selain itu, *Python* memiliki sintaks yang sederhana dan banyak digunakan dalam penelitian akademik serta industri data science.

Peran *Python* dalam penelitian ini meliputi:

- a. Pengambilan data saham melalui API.
- b. Preprocessing dan normalisasi data.
- c. Implementasi model LSTM dan XGBoost.
- d. Perhitungan metrik evaluasi.
- e. Visualisasi hasil prediksi.

2.4.2 Library Pandas

Pandas digunakan untuk manipulasi dan pengolahan data berbentuk tabel (dataframe). Library ini berfungsi untuk:

- a. Membersihkan data (handling missing values).

- b. Mengatur struktur time series.
- c. Melakukan slicing dan transformasi data.
- d. Menyiapkan dataset untuk pelatihan model.

Pandas sangat penting karena data harga saham berbentuk deret waktu yang harus dikelola secara sistematis sebelum masuk ke tahap pemodelan.

2.4.3. NumPy

NumPy digunakan untuk operasi numerik dan komputasi matriks. Dalam konteks penelitian ini, NumPy berperan dalam:

- a. Transformasi data menjadi array numerik.
- b. Operasi matematika pada dataset.
- c. Pembentukan dataset time series untuk LSTM.

Karena algoritma LSTM dan XGBoost bekerja dalam bentuk matriks dan tensor, NumPy menjadi komponen penting dalam proses komputasi.

2.4.4 TensorFlow / Keras

TensorFlow dan Keras digunakan untuk mengimplementasikan model Long Short-Term Memory (LSTM). Keras merupakan *high-level API* yang memudahkan pembangunan arsitektur jaringan saraf, sedangkan TensorFlow bertindak sebagai backend komputasi.

Peran TensorFlow/Keras dalam penelitian ini meliputi:

- a. Membangun arsitektur LSTM.
- b. Proses pelatihan model menggunakan backpropagation.
- c. Optimasi parameter menggunakan algoritma seperti Adam optimizer.
- d. Menghasilkan prediksi harga saham berdasarkan model yang telah dilatih.

Penggunaan framework ini memungkinkan implementasi model deep learning secara efisien dan terstruktur.

2.4.5 XGBoost Library

Library XGBoost digunakan untuk mengimplementasikan algoritma Extreme Gradient Boosting. Library ini menyediakan fitur optimasi yang cepat dan stabil dalam menangani permasalahan regresi.

Peran XGBoost dalam penelitian ini adalah:

- a. Membangun model regresi berbasis boosting.
- b. Mengatur parameter seperti learning rate, max depth, dan jumlah estimators.
- c. Melakukan pelatihan model secara iteratif.
- d. Menghasilkan prediksi harga saham berdasarkan fitur historis.

2.4.6 Scikit-Learn

Scikit-Learn digunakan untuk:

- a. Normalisasi data (MinMaxScaler).
- b. Pembagian data training dan testing.
- c. Perhitungan metrik evaluasi seperti MAE, RMSE, MAPE, dan Directional Accuracy.

Library ini mendukung proses validasi model secara sistematis dan terukur.

2.4.7 Matplotlib

Matplotlib digunakan untuk visualisasi hasil prediksi. Grafik yang dihasilkan membantu dalam:

- a. Membandingkan harga aktual dan harga prediksi.

- b. Menganalisis pola kesalahan prediksi.
- c. Menampilkan performa model secara visual.

Visualisasi sangat penting dalam interpretasi hasil penelitian.

2.4.8 Streamlit

Streamlit merupakan framework berbasis *Python* yang digunakan untuk membangun aplikasi web interaktif secara cepat dan sederhana. Dalam penelitian ini, *Streamlit* digunakan untuk mengembangkan prototype sistem prediksi harga saham berbasis web yang menampilkan hasil perbandingan model LSTM, XGBoost, dan Hybrid Stacking LSTM-XGBoost. *Dashboard* juga digunakan untuk menampilkan visualisasi evaluasi model, simulasi prediksi harga saham, feature importance, serta analisis performa model secara interaktif.

Penggunaan *Streamlit* memungkinkan visualisasi hasil prediksi dalam bentuk antarmuka interaktif, sehingga pengguna dapat:

- a. Memilih emiten saham.
- b. Melihat grafik harga aktual dan harga prediksi.
- c. Membandingkan performa model LSTM dan XGBoost.
- d. Menampilkan nilai evaluasi seperti RMSE dan MAE secara langsung.

Implementasi *Streamlit* dalam penelitian ini bertujuan untuk mendemonstrasikan bagaimana model prediksi yang telah dikembangkan dapat diterapkan dalam bentuk sistem pendukung keputusan sederhana. Dengan demikian, penelitian ini tidak hanya berhenti pada tahap analisis model, tetapi juga menghasilkan prototype aplikasi yang dapat digunakan untuk eksplorasi hasil prediksi secara visual.

Streamlit dipilih karena memiliki keunggulan dalam:

- a. Integrasi langsung dengan *Python*.
- b. Kemudahan deployment.

- c. Dukungan visualisasi data.
- d. Tidak memerlukan konfigurasi backend yang kompleks.