BAB II
TINJAUAN PUSTAKA

2.1 Investasi dan Emas sebagai Instrumen Keuangan

Investasi adalah kegiatan menanamkan modal pada suatu aset dengan harapan memperoleh keuntungan di masa mendatang. Dalam ilmu keuangan, instrumen investasi dikategorikan ke dalam dua kelompok utama: aset keuangan (saham, obligasi, reksa dana) dan aset riil (properti, komoditas, logam mulia). Emas termasuk dalam kategori aset riil sekaligus berfungsi sebagai aset keuangan karena dapat diperdagangkan secara aktif di pasar global.

Emas memiliki karakteristik unik yang membedakannya dari instrumen investasi lain. Pertama, emas bersifat tangible memiliki nilai intrinsik fisik yang tidak dapat dihapus oleh kebangkrutan perusahaan atau gagal bayar pemerintah. Kedua, emas berfungsi sebagai safe haven asset, yaitu aset yang nilainya cenderung meningkat saat kondisi ekonomi memburuk, inflasi meningkat, atau ketidakpastian geopolitik meningkat. Ketiga, emas memiliki likuiditas tinggi karena dapat diperjualbelikan kapan saja di pasar global yang beroperasi 24 jam sehari, 5 hari seminggu. Keempat, emas diakui sebagai alat lindung nilai (hedging) terhadap inflasi karena secara historis harganya bergerak berlawanan arah dengan nilai Dolar AS (WGC, 2026).

Dalam pasar keuangan global, emas diperdagangkan dalam satuan XAU/USD, di mana XAU adalah kode ISO 4217 untuk emas dan USD adalah Dolar Amerika Serikat. Harga XAU/USD merepresentasikan berapa Dolar AS yang diperlukan untuk membeli satu troy ons emas (setara dengan 31,1035 gram). Perdagangan emas dunia dapat dilakukan melalui berbagai instrumen: spot trading (pembelian langsung), futures contract (kontrak berjangka), Exchange-Traded Fund (ETF) emas, dan Contract for Difference (CFD) pada platform forex.

2.2 Harga Emas Dunia dan Faktor-Faktor yang Mempengaruhinya

Pergerakan harga emas dunia dipengaruhi oleh banyak variabel yang saling berinteraksi secara kompleks dan non-linear. (Rahman et al., 2025) menegaskan bahwa kompleksitas hubungan antar variabel penentu harga emas bersifat tidak linear dan berubah sepanjang waktu, sehingga tidak dapat dimodelkan secara optimal oleh metode statistik konvensional. Secara umum, faktor-faktor yang mempengaruhi harga emas dapat dikelompokkan ke dalam empat kategori utama.

2.2.1 Faktor Makroekonomi

Kebijakan suku bunga Federal Reserve AS merupakan salah satu penggerak terkuat harga emas. Secara historis, kenaikan suku bunga The Fed cenderung menekan harga emas karena aset berbunga (seperti obligasi) menjadi lebih menarik dibandingkan emas yang tidak menghasilkan yield. Sebaliknya, penurunan suku bunga mendorong kenaikan harga emas. Indeks Dolar AS (DXY) juga memiliki hubungan invers yang kuat dengan harga emas: ketika Dolar melemah, harga emas dalam USD cenderung naik karena emas menjadi lebih murah bagi pemegang mata uang lain. Tingkat inflasi (diukur melalui CPI dan PCE) juga berperan penting, saat inflasi meningkat, investor beralih ke emas sebagai pelindung nilai.

2.2.2 Faktor Pasar Keuangan

Yield obligasi pemerintah AS tenor 10 tahun (US Treasury 10-year yield) memiliki korelasi negatif yang konsisten dengan harga emas. Ketika yield obligasi naik, biaya kesempatan (opportunity cost) memegang emas meningkat sehingga harga emas tertekan. Performa pasar saham global juga mempengaruhi harga emas: saat pasar saham jatuh, investor mengalihkan aset ke emas sebagai safe haven, yang mendorong harga emas naik. Harga minyak mentah dunia memiliki korelasi positif dengan harga emas karena keduanya sama-sama sensitif terhadap tekanan inflasi dan ketidakpastian geopolitik.

2.2.3 Faktor Geopolitik dan Permintaan Fisik

Ketidakpastian geopolitik seperti konflik bersenjata, sanksi ekonomi, dan ketegangan perdagangan secara konsisten mendorong kenaikan harga emas karena memicu peningkatan permintaan safe haven. (WGC, 2026) mencatat bahwa bank sentral global membeli 1.045 metrik ton emas sepanjang tahun 2024, menjadikannya tahun ketiga berturut-turut dengan pembelian melampaui 1.000 ton sebuah laju yang tidak pernah terjadi sejak era 1960-an. Selain itu, permintaan emas perhiasan dari negara-negara berkembang seperti India dan Tiongkok, serta permintaan industri untuk keperluan elektronik dan medis, turut menjadi komponen permintaan fisik yang mempengaruhi harga.

2.2.4 Faktor Psikologi Pasar dan Sentimen

Sentimen pasar dan psikologi investor memainkan peran yang tidak kalah penting. Indeks ketakutan pasar (VIX Volatility Index) yang tinggi biasanya berkorelasi positif dengan kenaikan harga emas. Pernyataan pejabat bank sentral, data ekonomi yang di luar ekspektasi, atau peristiwa tak terduga (black swan events) dapat menggerakkan harga emas secara tiba-tiba dalam waktu sangat singkat. Kombinasi dari seluruh faktor di atas itulah yang menjadikan data harga XAU/USD bersifat sangat non-linear, tidak stasioner, dan multi-dimensional.

2.3 Data Time Series dalam Konteks Harga Emas

Data time series adalah rangkaian observasi yang dikumpulkan secara berurutan dalam interval waktu tertentu. Data harga XAU/USD harian merupakan salah satu contoh data time series dalam bidang keuangan. Terdapat empat komponen utama dalam data time series:

1. Trend yaitu kecenderungan data bergerak naik atau turun dalam jangka panjang.
2. Seasonality yaitu pola berulang yang terjadi dalam periode waktu tertentu.
3. Cyclical yaitu fluktuasi jangka panjang yang tidak memiliki periode tetap.

4. Irregular/Noise yaitu variasi acak yang tidak dapat dijelaskan oleh ketiga komponen lainnya.

Data harga emas bersifat tidak stasioner (non-stationary), artinya nilai rata-rata, variansi, dan struktur korelasinya berubah seiring waktu. Kondisi ini merupakan tantangan utama dalam pemodelan prediksi karena sebagian besar metode statistik klasik mengasumsikan data stasioner. Selain itu, data harga emas memiliki sifat long memory dependency, yaitu nilai di masa lalu yang sangat jauh masih memiliki pengaruh terhadap nilai saat ini karakteristik inilah yang menjadi alasan utama mengapa LSTM, dengan kemampuannya mengingat dependensi jangka panjang, sangat cocok untuk memodelkan data harga emas (Agusmawati, 2023).

2.4 Machine Learning untuk Prediksi Data Keuangan

Machine learning adalah cabang kecerdasan buatan yang memungkinkan sistem untuk belajar dan meningkatkan kinerjanya secara otomatis dari data tanpa diprogram secara eksplisit. Dalam konteks prediksi harga keuangan, machine learning menawarkan keunggulan signifikan dibandingkan metode statistik konvensional karena tidak bergantung pada asumsi distribusi data tertentu dan mampu menangkap pola non-linear yang kompleks secara langsung dari data historis. (Wahyuningsih, 2023) menyimpulkan bahwa dari seluruh metode yang dibandingkan untuk prediksi harga emas mencakup SVR, Random Forest, ANN, dan LSTM metode berbasis Neural Network menghasilkan nilai RMSE, MAE, dan MAPE yang paling rendah secara konsisten.

Secara umum, algoritma machine learning yang digunakan untuk prediksi harga keuangan dapat dibagi menjadi dua kategori. Pertama, algoritma berbasis ensemble seperti Random Forest dan XGBoost yang bekerja dengan membangun banyak model sederhana (decision tree) dan menggabungkan hasilnya untuk menghasilkan prediksi yang lebih robust. Kedua, algoritma berbasis deep learning seperti RNN, LSTM, dan GRU yang mampu memodelkan ketergantungan temporal dalam data sekuensial. Masing-masing kategori memiliki kekuatan dan keterbatasan

yang berbeda, yang menjadi dasar pertimbangan pengembangan model hybrid dalam penelitian ini.

2.5 Long Short-Term Memory (LSTM)

2.5.1 Konsep Dasar dan Arsitektur

Long Short-Term Memory (LSTM) merupakan varian dari Recurrent Neural Network (RNN) yang dikembangkan oleh Hochreiter dan Schmidhuber pada tahun 1997. LSTM dirancang khusus untuk mengatasi kelemahan mendasar RNN standar, yaitu masalah vanishing gradient kondisi di mana sinyal gradien menjadi sangat kecil selama proses backpropagation pada urutan data yang panjang, sehingga jaringan tidak dapat mempelajari ketergantungan jangka panjang secara efektif. LSTM mengatasi masalah ini melalui unit memori khusus yang disebut cell state dan mekanisme tiga gerbang (gate) yang mengatur aliran informasi secara selektif.

2.5.2 Mekanisme Gerbang LSTM

LSTM bekerja melalui tiga gerbang yang masing-masing memiliki fungsi spesifik dalam mengatur informasi. Pertama, Forget Gate berfungsi memutuskan informasi mana dari cell state sebelumnya yang harus dibuang atau dipertahankan. Gate ini menggunakan fungsi sigmoid yang menghasilkan nilai antara 0 (lupakan sepenuhnya) dan 1 (pertahankan sepenuhnya). Kedua, Input Gate berfungsi memutuskan informasi baru apa yang akan disimpan ke dalam cell state. Gate ini terdiri dari dua bagian: sigmoid layer yang menentukan nilai mana yang akan diperbarui, dan tanh layer yang membuat vektor nilai baru yang akan ditambahkan ke cell state. Ketiga, Output Gate berfungsi menentukan bagian mana dari cell state yang akan menjadi output. Output gate menggunakan fungsi sigmoid untuk menentukan bagian cell state yang akan dikeluarkan, kemudian cell state diproses melalui fungsi tanh untuk menghasilkan nilai output akhir.

2.5.3 Keunggulan dan Keterbatasan LSTM

Keunggulan utama LSTM dalam konteks prediksi harga emas adalah kemampuannya menangkap dependensi temporal jangka panjang yang tidak dapat ditangkap oleh metode konvensional. Kemampuan ini telah terbukti secara empiris (Agusmawati, 2023) membuktikan LSTM menghasilkan MAE 0,0389 untuk prediksi harga emas, mengungguli GRU secara konsisten. (Rahman et al., 2025) pada RIGGS Journal melaporkan akurasi 99,55% pada prediksi harga emas ANTAM menggunakan 4.545 data harian. Meski demikian, LSTM memiliki keterbatasan yang perlu diperhatikan: membutuhkan volume data yang besar untuk pelatihan optimal, rentan terhadap overfitting jika arsitektur tidak dikonfigurasi dengan tepat, dan kurang efektif dalam menyeleksi serta membobot banyak fitur eksternal secara bersamaan.

2.6 Extreme Gradient Boosting (XGBoost)

2.6.1 Konsep Dasar dan Algoritma

Extreme Gradient Boosting (XGBoost) adalah algoritma ensemble learning berbasis gradient boosting yang dikembangkan oleh Chen dan Guestrin dan dipublikasikan pada tahun 2016. XGBoost merupakan implementasi yang dioptimalkan dari gradient boosting decision tree (GBDT), dirancang untuk efisiensi komputasi tinggi, fleksibilitas, dan portabilitas. XGBoost bekerja dengan membangun serangkaian pohon keputusan (decision tree) secara berurutan, di mana setiap pohon baru dibangun untuk memperbaiki kesalahan prediksi (residual) dari gabungan pohon-pohon sebelumnya menggunakan teknik gradient descent pada fungsi loss.

2.6.2 Komponen Utama XGBoost

XGBoost memiliki beberapa komponen teknis yang membedakannya dari implementasi gradient boosting standar. Pertama, Regularization: XGBoost menambahkan term regularisasi L1 (Lasso) dan L2 (Ridge) langsung ke dalam fungsi objektif, yang secara efektif mencegah overfitting dengan menghukum

kompleksitas model yang berlebihan. Kedua, Tree Pruning: berbeda dengan GBDT konvensional yang menggunakan pendekatan greedy depth-first, XGBoost menggunakan pendekatan max_depth dan memangkas pohon ke belakang (backward pruning), menghasilkan pohon yang lebih efisien. Ketiga, Handling Missing Values: XGBoost mampu menangani nilai yang hilang secara otomatis dengan mempelajari arah default terbaik untuk data yang missing selama proses pelatihan. Keempat, Feature Importance: XGBoost menghasilkan skor kepentingan fitur (feature importance score) yang memungkinkan identifikasi variabel mana yang paling berpengaruh terhadap prediksi.

2.6.3 Keunggulan dan Keterbatasan XGBoost

Keunggulan XGBoost yang paling relevan untuk penelitian ini adalah kemampuannya mengolah banyak fitur input secara bersamaan, kecepatan pelatihan yang tinggi karena mendukung komputasi paralel, dan ketahanannya terhadap overfitting. XGBoost mencapai R^2 89,09% dan MAPE 3,21% untuk prediksi harga saham. XGBoost kompetitif dengan LSTM dalam memprediksi harga Bitcoin menggunakan fitur sentimen dan Google Trends Index. Namun, keterbatasan mendasar XGBoost adalah ketidakmampuannya memproses urutan temporal secara alami setiap observasi diperlakukan secara independen tanpa memori terhadap urutan data sebelumnya, sehingga tidak dapat menangkap long-range temporal dependency yang esensial dalam data time series harga emas.

2.7 Model Hybrid dalam Machine Learning

2.7.1 Konsep Model Hybrid

Model hybrid dalam machine learning adalah pendekatan yang mengombinasikan dua atau lebih algoritma berbeda dengan tujuan memanfaatkan kekuatan masing-masing algoritma sekaligus menutupi kelemahannya secara mutual. Konsep ini didasarkan pada prinsip bahwa tidak ada satu algoritma tunggal yang optimal untuk semua jenis data dan semua kondisi sebuah prinsip yang dalam komunitas machine learning dikenal sebagai No Free Lunch Theorem

(Wolpert, 1996). Model hybrid umumnya menghasilkan performa yang lebih baik dibandingkan model individual karena mampu menangkap dimensi yang berbeda dari kompleksitas data secara bersamaan.

2.7.2 Skema Hybrid XGBoost-LSTM

Dalam penelitian ini, model hybrid dirancang dengan mekanisme dua tahap yang saling melengkapi. Pada tahap pertama, XGBoost digunakan untuk mengekstraksi fitur-fitur paling relevan dari data multi-variabel harga XAU/USD. XGBoost akan secara otomatis mengidentifikasi dan membobot variabel-variabel yang paling berpengaruh terhadap harga emas melalui mekanisme feature importance, serta menghasilkan prediksi awal berbasis fitur terstruktur. Pada tahap kedua, output dari XGBoost beserta data sekuensial historis digunakan sebagai input bagi LSTM. LSTM kemudian memodelkan ketergantungan temporal jangka panjang dari data tersebut untuk menghasilkan prediksi harga emas final. Dengan skema ini, kelemahan XGBoost dalam memproses urutan temporal ditutupi oleh LSTM, sementara kelemahan LSTM dalam menyeleksi banyak fitur eksternal ditutupi oleh XGBoost.

2.7.3 Bukti Empiris Efektivitas Hybrid

Efektivitas model hybrid XGBoost-LSTM telah dikonfirmasi oleh beberapa penelitian yang terpublikasi di jurnal terakreditasi SINTA di Indonesia. Jurnal Analysis of Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost) Algorithms to Predict the Number of Airplane Passengers at Makassar Sultan Hasanuddin International Airport : Systematic Literature Review menyimpulkan bahwa pendekatan hybrid LSTM-XGBoost menghasilkan performa lebih baik dibandingkan model tunggal untuk prediksi data time series (Idham, 2025). Penelitian pada Informatika: Jurnal Teknik Informatika dan Multimedia secara langsung membuktikan bahwa hybrid LSTM-XGBoost menghasilkan RMSE 117,89 dan MAE 92,45 pada prediksi harga saham BBRI,

lebih unggul dari kedua model secara individual. Temuan-temuan ini memperkuat justifikasi ilmiah penggunaan model hybrid XGBoost-LSTM pada penelitian ini.

2.8 Metrik Evaluasi Model Prediksi

Untuk mengukur dan membandingkan performa model prediksi secara objektif, penelitian ini menggunakan tiga metrik evaluasi yang umum digunakan dalam literatur prediksi time series keuangan.

2.8.1 Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) adalah akar kuadrat dari rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual. RMSE dihitung menggunakan rumus:

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

di mana yi adalah nilai aktual, ŷi adalah nilai prediksi, dan n adalah jumlah data. RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar (outlier) karena menggunakan operasi kuadrat. Semakin kecil nilai RMSE, semakin akurat model prediksi yang dihasilkan. RMSE memiliki satuan yang sama dengan data asli sehingga mudah diinterpretasikan secara langsung.

2.8.2 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) adalah rata-rata nilai absolut dari selisih antara nilai prediksi dan nilai aktual. MAE dihitung dengan rumus:

$$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}|\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}| \qquad (2)$$

Berbeda dengan RMSE, MAE memberikan bobot yang sama untuk semua kesalahan prediksi tanpa memberikan penalti lebih besar pada outlier. MAE lebih robust terhadap nilai ekstrem dan mudah diinterpretasikan karena menyatakan rata-rata besarnya kesalahan prediksi dalam satuan yang sama dengan data aslinya. (Agusmawati, 2023) menggunakan MAE sebagai salah satu metrik utama dalam membandingkan performa LSTM dan GRU untuk prediksi harga emas.

2.8.3 Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) adalah rata-rata persentase absolut dari selisih antara nilai prediksi dan nilai aktual terhadap nilai aktual. MAPE dihitung dengan rumus:

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \qquad (3)$$

Keunggulan MAPE adalah tidak bergantung pada skala data sehingga dapat digunakan untuk membandingkan akurasi model pada dataset yang berbeda skala. Secara umum, nilai MAPE kurang dari 10% menunjukkan prediksi yang sangat akurat, 10%–20% akurat, 20%–50% cukup akurat, dan di atas 50% tidak akurat. Penelitian pada Jurnal Perbandingan Metode Extreme Gradient Boosting (XGBOOST) dengan Long Short-Term Memory (LSTM) Untuk Prediksi Saham Pt. Bank Mandiri Tbk. (BMRI)

Tabel 2. 1 Ringkasan Metrik Evaluasi Model

Metrik	Rumus	Interpretasi	Kelebihan	Kekurangan
RMSE	$\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n}}$	Semakin kecil semakin baik	Sensitif terhadap outlier (penalti besar)	Sulit dibandingkan antar dataset berskala berbeda
MAE	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i \rvert$	Semakin kecil semakin baik	Robust terhadap outlier, mudah diinterpretasi	Tidak memberikan penalti lebih pada kesalahan besar
MAPE	$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i - \hat{y}_i}{y_i}\right\rvert$	<10%: sangat akurat	Bebas skala, mudah dipahami dalam persentase	Tidak stabil jika nilai aktual mendekati nol

2.9 Preprocessing Data Time Series

Preprocessing atau pra-pemrosesan data merupakan tahapan kritis dalam pipeline machine learning yang bertujuan mempersiapkan data mentah agar dapat diproses secara optimal oleh algoritma. Untuk data time series harga emas, terdapat empat tahapan preprocessing utama yang perlu dilakukan.

2.9.1 Penanganan Missing Value

Data harga XAU/USD dapat memiliki nilai yang hilang (missing value) akibat hari libur bursa, gangguan teknis, atau inkonsistensi sumber data. Penanganan missing value dapat dilakukan dengan beberapa metode: forward fill

(mengisi dengan nilai terakhir yang tersedia), interpolasi linear (mengisi dengan nilai tengah antara dua titik data terdekat), atau penghapusan baris yang mengandung missing value jika persentasenya sangat kecil. Pemilihan metode harus mempertimbangkan karakteristik data agar tidak menimbulkan bias pada model.

2.9.2 Normalisasi Data

Normalisasi data bertujuan mengubah skala nilai fitur ke dalam rentang tertentu agar proses pelatihan algoritma deep learning seperti LSTM menjadi lebih stabil dan konvergen lebih cepat. Metode normalisasi yang umum digunakan adalah Min-Max Normalization dengan rumus: x' = (x - x_min) / (x_max - x_min), yang menghasilkan nilai dalam rentang [0, 1]. Setelah prediksi selesai dilakukan, nilai harus didenormalisasi kembali ke skala aslinya menggunakan transformasi invers. Perlu diperhatikan bahwa nilai x_min dan x_max yang digunakan untuk denormalisasi harus berasal dari data training, bukan data keseluruhan, untuk menghindari data leakage.

2.9.3 Feature Engineering

Feature engineering adalah proses menciptakan fitur-fitur baru yang bermakna dari data mentah untuk meningkatkan kemampuan prediksi model. Untuk data harga emas, fitur teknikal yang umum digunakan meliputi: Moving Average (MA-7, MA-14, MA-30) yang menghaluskan fluktuasi jangka pendek, Relative Strength Index (RSI) yang mengukur momentum pergerakan harga, Bollinger Bands yang mengukur volatilitas harga, MACD (Moving Average Convergence Divergence) sebagai indikator tren, dan fitur lag (nilai harga pada t-1, t-2, t-n) yang merepresentasikan nilai historis sebagai input model.

2.9.4 Pembagian Data Training dan Testing

Untuk memvalidasi performa model secara objektif, data dibagi menjadi dua subset: data training (data latih) dan data testing (data uji). Untuk data time

series, pembagian harus dilakukan secara kronologis data awal digunakan untuk training dan data terbaru digunakan untuk testing untuk menghindari data leakage dan memastikan evaluasi yang realistis. Rasio pembagian yang umum digunakan adalah 80:20 atau 70:30 (training:testing). Pendekatan walk-forward validation atau time series cross-validation juga dapat digunakan untuk validasi yang lebih robust pada data time series.

2.10 Kerangka Teori Penelitian

Berdasarkan seluruh kajian teori dan penelitian terdahulu yang telah diuraikan, kerangka teori penelitian ini dapat disintesiskan sebagai berikut. Data harga emas dunia (XAU/USD) adalah data time series yang bersifat non-linear, tidak stasioner, dan dipengaruhi oleh banyak variabel secara bersamaan karakteristik yang tidak dapat dimodelkan secara optimal oleh metode statistik konvensional maupun oleh satu algoritma machine learning tunggal. LSTM unggul dalam menangkap dependensi temporal jangka panjang namun lemah dalam seleksi fitur multi-variabel. XGBoost unggul dalam pengolahan fitur terstruktur namun tidak mampu memproses urutan temporal. Model hybrid XGBoost-LSTM yang dirancang dengan mekanisme dua tahap XGBoost sebagai feature extractor dan LSTM sebagai temporal modeler secara teoritis mampu mengatasi keterbatasan kedua algoritma secara bersamaan. Prediksi yang dihasilkan dievaluasi menggunakan metrik RMSE, MAE, dan MAPE, serta dibandingkan dengan model tunggal XGBoost dan LSTM sebagai baseline.