

2.27%

SIMILARITY OVERALL

SCANNED ON: 21 JUL 2025, 10:18 AM

Similarity report

Your text is highlighted according to the matched content in the results above.

● CHANGED TEXT 2.27% ● QUOTES 0.49%

Report #27589781

1 BAB I PENDAHULUAN 1.1 Latar Belakang Masalah Emas telah dianggap sebagai aset tempat berlindung yang aman, memberikan stabilitas selama masa ketidakstabilan ekonomi. Emas memiliki berbagai fungsi dalam pasar keuangan yang membangun hubungan kuat dengan berbagai faktor keuangan dan makroekonomi, menjadikan emas aset yang kompleks untuk diprediksi (Qian et al., 2019). Sepanjang sejarah, emas telah terbukti tahan banting, terutama pada masa krisis; misalnya, selama penurunan ekonomi global pada 2008, harga emas terus meningkat, sementara aset lainnya mengalami kerugian yang signifikan. Ketahanan ini, yang dipadukan dengan volatilitasnya, menjadikan emas area studi yang penting bagi para ahli keuangan, karena fluktuasi harganya menciptakan peluang keuntungan yang signifikan sekaligus tantangan dalam manajemen risiko. Oleh karena itu, pengembangan model prediksi yang akurat untuk harga emas sangat penting, karena model ini dapat memberikan wawasan krusial yang membantu investor membuat keputusan yang lebih informasi dan strategis (Livieris et al., 2020). Selain itu, peningkatan akurasi prediksi melalui metodologi canggih seperti analisis deret waktu dan pembelajaran mendalam semakin vital untuk memahami lebih baik kompleksitas perilaku harga emas (Zhang & Ci, 2020). Hubungan antara aset lainnya dengan harga emas ini bersifat tidak statis dan berubah seiring waktu terutama setelah krisis finansial pada tahun 2008 (Beckmann et al., 2019). Berdasarkan studi

Sheikh et al. (2020) bahwa terdapat tantangan dalam menyesuaikan dengan perubahan harga pada nilai tukar la innya sehingga variabel- variabel ini menjadi semakin kompleks karena in vestor m emiliki berbagai reaksi yang berbed a terhadap fluktuasi ekon omi. Oleh karena itu, prediksi harga emas menjadi tantangan unik baik dari akademis maupun pemerintahan dan aktivis ekonomi. Seiring penelitia n in i dilaksanakan, harga emas sendiri sudah melambung tinggi. Harga emas global diperkirakan akan menyentu h di le vel harga USD 3,50 2 Amani (2025). Hal ini diperkuat oleh faktor geopolitik sehingga membuat ha rga emas mengalami penguatan (Amani, 2025). Harga komoditas seperti perak dan emas ikut mengalami kenaikan signifikan karena permintaan yang semakin banyak (Adam, 2025). Faktor seperti indeks dolar Am erika Serikat (AS) yang melemah serta kenaikan harga minyak menjadi sentimen positif bagi pasar untuk logam mulia (Adam, 2025). Penelitian mengenai dinamika harga emas dan peramalan fluktuasi harga emas telah muncul sebagai area studi yang penting, dengan berbagai metodologi yang diusulkan untuk meningkatkan akurasi prediksi. Terdapat tiga kategori utama metode prediksi yaitu pendekatan deret waktu klasik, kedua adalah metodologi kecerdasan buatan, dan ketiga adalah m etodologi hibrida (Jabeur et al., 2024). 9 Pertama, dalam ranah metode deret waktu klasik, pendekatan seperti pemodelan matematis sederhana mencakup model Autoregressive Integrated Moving Average (ARIMA) (Dong et al., 2020; Li et al., 2021; Yang, 2019), serta teknik regresi multi- linear (Manoj & K, 2019), yang keduanya telah banya k diterapkan dalam peramalan harga emas. Kedua, beberapa teknik pembelajar an mesin telah digunakan un tuk prediksi harga emas, termasuk Support Vector Machine (SVM) (Li et al., 2021), Extreme Gradient Boosting (XGBoost)(Jabeur et al., 2024), dan pendekatan regresi ensemble (Kilim ci, 2022). Mengingat kompleksitas dan non-linearitas yang melekat pada karakteristik harga emas , metode cerdas, khu susnya Jaringan Saraf Tiruan (Artificial Neural Networks/ ANN), sering kali lebih efektif dalam memberikan prediksi yang akurat tentang pergerakan harga emas (Alameer et al., 2019; He et al., 2019;

Livieris et al., 2020; Vidal & Kristjanpoller, 2020). Ketiga, metode hibrida mengaplikasikan kombinasi metode kecerdasan buatan dengan berbagai metode lainnya untuk meramalkan harga emas, seperti metode hibrida Convolutional Neural Networks - Long Short Term Memory (CNN-LSTM) (Livieris et al., 2020; Vidal & Kristjanpoller, 2020). LSTM, sebagai salah satu metode cerdas, merupakan varian dari metode Recurrent Neural Network (RNN), yang dapat mengatasi kekurangan RNN dalam menangani masalah ketergantungan jangka panjang maupun hilang atau meledaknya gradien (Emmert-Streib et al., 2020). Metode LSTM digunakan sebagai salah satu metode yang dapat secara efisien menangkap pola-pola yang berderet 3 berdasarkan arsitekturnya yang dikhususkan (Livieris et al., 2020). Sebuah studi yang dilakukan oleh Makalesi & Yurtsever, (2021), menerapkan tiga buah metode untuk meramalkan harga emas yaitu LSTM, Bi-Directional LSTM, dan Gated Recurrent Unit (GRU). Studi lainnya yang dilakukan oleh Song et al, (2020) menerapkan LSTM sebagai metode pemodelan jaringan syaraf tiruan untuk memprediksi deretan waktu dan menerapkan modifikasi Attention. Diharapkan bahwa metode LSTM dapat memberikan prediksi yang lebih akurat mengenai pergerakan harga emas di masa depan, yang dapat membantu investor dan pengambil kebijakan dalam merumuskan keputusan yang lebih terinformasi dan berbasis data. Dengan kemampuannya dalam menangani data berurutan dan menangkap pola temporal yang kompleks, LSTM berpotensi mengurangi ketidakpastian dalam prediksi harga emas yang sering dipengaruhi oleh berbagai faktor eksternal, seperti perubahan ekonomi global.

1.2 Identifikasi Masalah 1.2 **22** 1. Rumusan Masalah Penulis merangkum dan menjabarkan rumusan masalah dari latar belakang sebagai berikut : 1. Bagaimana proses harga emas dapat diprediksi menggunakan metode LSTM berdasarkan analisis time series? 2. Model apa yang memiliki persentase nilai akurasi tertinggi yang dihasilkan oleh metode LSTM dalam proses prediksi harga emas? **1 3 4 5 7** 3. Model apa yang memiliki besar nilai error atau kesalahan terkecil dalam, dalam bentuk Mean Squared Error (MSE), Root Mean Squared Error (RMSE), dan

Mean Absolute Error (MAE), dan Mean Absolute Error Percentage Error (MAPE)? 1.2.2.

Batasan Masalah Penulis merangkum beberapa batasan masalah dalam penelitian ini sebagai berikut :

1. Penelitian ini akan mencakup analisis data harga emas spots (XAU/USD) dalam rentang waktu 1 Januari 2017 sampai 28 Februari 2025.
2. Penelitian ini menggunakan data yang didapatkan pada situs informasi harga yaitu Yahoo Finance atau TradingView.
3. Penelitian ini mencakup fitur-fitur eksternal seperti harga minyak mentah, harga perak, harga indeks saham S&P 500, dan indeks dolar AS.

1.3 Tujuan Penelitian Penulis meneliti penelitian ini bertujuan untuk :

1. Mengevaluasi dan melakukan analisa terhadap akurasi model prediksi harga emas harian dengan LSTM dalam deretan waktu dari 1 Januari 2020 sampai 28 Februari 2025.
2. Mengevaluasi dan melakukan analisa faktor eksternal yang mempengaruhi pergerakan harga emas.
3. Mengembangkan model dari metode LSTM yang optimal dalam memprediksi harga emas yang memberikan error atau kesalahan terkecil.

1.4. Manfaat Penelitian Manfaat penelitian ini dibagi menjadi tiga bagian, yaitu manfaat bagi masyarakat, peneliti, dan ilmu pengetahuan. Manfaat itu dirangkum sebagai berikut.

1.4.1. Manfaat bagi Masyarakat Penelitian ini dapat membantu meningkatkan akurasi prediksi harga emas, sehingga investor yang menggunakan instrumen emas maupun analisis keuangan dapat membuat keputusan investasi yang lebih tepat.

8 Selain itu, penelitian ini membantu investor dalam pengambilan keputusan dengan memiliki model prediksi harga emas yang akurat, investor dapat membuat keputusan investasi yang lebih tepat dan mengurangi risiko kerugian. Penelitian ini juga dapat membantu meningkatkan efisiensi pasar emas dengan menyediakan informasi yang lebih akurat tentang harga emas di masa depan. Hasil penelitian ini dapat dijadikan sebagai usulan penerapan pemodelan analisis prediktif harga emas pada aplikasi analisis pergerakan harga emas.

5 1.4.2. Manfaat bagi Peneliti Penelitian ini dapat membantu mengembangkan metode prediksi harga emas yang lebih baik, sehingga dapat digunakan sebagai referensi bagi penelitian lainnya. Penelitian ini dapat membantu

meningkatkan pemahaman tentang pola harga emas dan faktor-faktor yang mempengaruhi harga emas. 1.4.3. Manfaat bagi Ilmu Pengetahuan Penelitian ini dapat menjadi kontribusi bagi pengembangan ilmu di bidang data science dan ekonomi dengan menciptakan model prediksi yang lebih akurat. 1.5 Kebaruan Pada subbab kebaruan, akan dijadikan sebuah poin-poin untuk menjabarkan kebaruan dari penelitian sebagai berikut :

1. Penelitian ini menggunakan metode LSTM yang relatif baru dalam prediksi harga emas, sehingga dapat memberikan kontribusi pada pengembangan metode prediksi yang lebih baik.
2. Penelitian ini menggunakan analisis time series dengan deep learning untuk memprediksi harga emas, sehingga dapat memberikan wawasan baru tentang pola harga emas.
3. Pengembangan model prediksi yang lebih akurat: Penelitian ini bertujuan untuk mengembangkan model prediksi harga emas yang lebih akurat dengan menggunakan metode LSTM, sehingga dapat memberikan kontribusi pada peningkatan akurasi prediksi harga emas.
4. Penggunaan data harga emas historis yang terkini: Penelitian ini menggunakan data harga emas historis yang terkini untuk melatih model LSTM, sehingga dapat memberikan hasil yang lebih relevan dengan kondisi pasar saat ini.

1.6 Kerangka

Penulisan Laporan penelitian ini disusun berdasarkan pedoman yang sudah diterbitkan oleh Universitas Pembangunan Jaya sesuai dengan sistematika penulisan dalam Program Studi Informatika yaitu tiga bab utama. BAB I PENDAHULUAN Bab ini terdiri dari subbab latar belakang penelitian, identifikasi masalah penelitian atau perumusan masalah, tujuan penelitian, manfaat penelitian terbagi jadi tiga bagian yaitu masyarakat, peneliti, dan ilmu pengetahuan, kebaruan, dan kerangka penulisan. BAB II TINJAUAN PUSTAKA Bab ini terdiri dari subbab pencapaian terdahulu dan tinjauan teoritis. BAB III Metode Penelitian Bab ini terdiri dari paradigma penelitian, metode penelitian, metode pengumpulan data, metode analisis data, dan metode pengujian data. BAB IV Rancangan Penelitian Bab ini terdiri dari subbab langkah-langkah penelitian dan rancangan pengujian. BAB V Hasil dan Pembahasan Bab ini terdiri dari subbab hasil dan subbab pembahasan.

VI Penutup Bab ini terdiri dari subbab kesimpulan dan saran. 7 BAB II TINJAUAN PUSTAKA Bab ini terdiri dari penelusuran literatur, pencapaian terdahulu, dan tinjauan teoritis. Isi bab ini untuk memperkuat landasan dari penelitian Peneliti. 2.1. Pencapaian Terdahulu Penelitian ini berfokus dalam pengembangan model prediksi harga emas dengan pendekatan time series forecasting. Penelusuran literatur ini menggunakan beberapa kata kunci untuk membantu penulis dalam mencari literatur. Beberapa di antaranya adalah : 1. "Gold price prediction with time series method 2. "Gold price prediction with technical analysis comparison 3. "Systematic review time series forecasting Penelusuran literatur ini memiliki tujuan untuk menjawab pertanyaan penelitian di antaranya adalah " Apa saja metode atau pendekatan time series untuk memprediksi harga? , " Apa saja variabel sebagai parameter yang relevan untuk memprediksi harga emas? ". Penelusuran literatur ini dilakukan dengan menelusuri beberapa database literatur yaitu Google Scholar. Penulis menggunakan mesin penelusuran yaitu Google dan sebuah software bernama Harzing Publish and Perish untuk mencari literatur. Tabel 2.1 Hasil penelitian terdahulu

No	Nama (Tahun)	Judul
1	Ioannis E. Livieris, Emmanuel Pintelas, Panagiotis Pintelas (2020)	A CNN-LSTM model for gold price time-series forecasting
2	Andrés Vidal, Werner Kristjanpoller (2019)	Gold volatility prediction using a CNN-LSTM approach

Penelitian ini mengembangkan sebuah model dengan metode hibrida yaitu CNN-LSTM yang menerapkan lapisan konvolusi dan lapisan pooling sebagai ekstraksi fitur serta LSTM unit untuk menganalisa data harga emas sehingga dapat diprediksi. Model ini dievaluasi dengan metrik MAE dan RMSE yang menunjukkan bahwa model ini menghasilkan MAE dan RMSE yang kecil. Hasil prediksi harga akan naik dan turun juga mendapatkan akurasi sebesar 55,53% (Livieris et al., 2020). Penelitian ini mengembangkan sebuah model dengan metode CNN-LSTM untuk memprediksi volatilitas harga emas dengan mengubah data deretan waktu menjadi citra lalu diekstraksi datanya sehingga dapat dianalisa oleh LSTM. Model CNN-LSTM menghasilkan error atau selisih paling kecil

dibandingkan pengujian dengan model lainnya. Perbedaan antara model LSTM dengan hibrida CNN-LSTM sebesar -18,1% atau model CNN-LSTM lebih unggul dalam memprediksi (Vidal & Kristjanpoller, 2020).

3 Junling Luo, Zhongliang Zhang, Yao Fu, Feng Rao (2021) Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms Sebuah penelitian mengkaji penerapan metode LSTM dan XGBoost untuk memprediksi kasus harian terkonfirmasi COVID-19 di Amerika. Studi ini menggunakan data serial waktu dari 1 April hingga 30 September 2020, yang dibagi menjadi 90% data latih dan 10% data uji. Fitur masukan untuk model mencakup riwayat kasus harian dari 14 hari sebelumnya, serta nilai rata-rata dan standar deviasi dari dua minggu terakhir. Hasil pengujian menunjukkan bahwa model LSTM memiliki kinerja prediksi yang lebih superior, mencapai MAPE sebesar 2,32%. Sementara itu, model XGBoost mencatatkan MAPE 7,21%, namun unggul dalam mengidentifikasi fitur paling berpengaruh, yaitu rata-rata 9 kasus dan hari dalam seminggu (Luo et al., 2021).

2 4 Sami Ben Jabeur, Salma Mefteh-Wali, Jean-Laurent Viviani (2021) Forecasting gold price with the XGBoost algorithm and SHAP interaction values Penelitian ini menerapkan algoritma XGBoost dan menerapkan pendekatan SHapley Additive exPlanation (SHAP) untuk menerjemahkan hasil dari model ML. Nilai dari SHAP dapat membantu dalam memahami berbagai model termasuk kepentingan dari fitur, ketergantungan fitur, kesimpulan lokal maupun plot kesimpulan. Penelitian ini menggunakan dataset dari berbagai harga yang dapat mempengaruhi harga emas, dalam harga Dolar AS, dari tahun 1986 sampai 2019. Berdasarkan uji dengan 5 model selain XGBoost, XGBoost mendapatkan RMSE terendah dengan nilai 34.92 (Jabeur et al., 2024).

5 Manjula, K. A., Karthikeyan, P. Gold Price Prediction using Ensemble based Machine Learning Techniques Penelitian ini menerapkan teknik ensemble yang menggabungkan beberapa model lalu dibentuk menjadi kesatuan model bertujuan menghasilkan performa yang lebih kuat. Penelitian ini menggunakan metode pohon keputusan yaitu Random Forest Regressor serta metode pohon regresi-gradien, lalu

dibandingkan dengan metode regresi linear. Penelitian ini menerapkan fitur seperti harga minyak mentah, harga saham, nilai tukar rupee terhadap Dolar AS, inflasi, serta suku bunga. Hasil 10 yang didapatkan adalah metode dengan teknik ensemble lebih unggul dengan MAE 3007,8 dan 2804,44 dibandingkan metode Regresi Linear dengan MAE 4384,43 (Manjula & Karthikeyan, 2019).

1 6 Zakaria Alameer, Mohamed Abd Elaziz, Ahmed A. Ewees, Haiwang Ye, Zhang Jianhua (2019) Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm Penelitian ini menerapkan pemodelan jaringan syaraf tiruan yang melibatkan lapisan-lapisan syaraf. Penelitian ini menerapkan algoritma optimasi Whale Optimization Algorithm (WOA) pada model (WOA-NN) untuk mengatasi masalah regresi non-linear seperti peramalan fluktuasi harga emas. Penelitian ini menerapkan fitur yaitu harga komoditas, nilai tukar, dan inflasi, dengan periode bulanan, dengan kurun waktu September 2008 sampai Agustus 2017. Hasil yang didapatkan adalah model WOA-NN mendapatkan hasil yang lebih baik (MAE = ,021) dibandingkan syaraf tiruan lainnya (MAE = 0,027) dan pemodelan matematika seperti ARIMA (MAE = 0,057) (Alameer et al., 2019).

7 Mustafa Yurtsever (2021) Gold Price Forecasting Using LSTM, Bi-LSTM and GRU Penelitian ini menerapkan metode LSTM serta Bidirectional-LSTM (Bi-LSTM) dan Gated Recurrent Unit (GRU). Data yang digunakan selama penelitian ini melibatkan variabel yang mewakili indikator ekonomi seperti harga minyak mentah, Consumer Price Index (CPI), nilai tukar efektif, suku bunga, dan indeks S&P 500. Hasil penelitian adalah model LSTM lebih unggul dibandingkan model Bi-LSTM dan GRU. Model LSTM menghasilkan nilai evaluasi (MAE = 48,85), mengalahkan model Bi-LSTM dengan (MAE = 61,53) dan GRU dengan 11 (MAE = 71,24) (Makalasi & Yurtsever, 2021).

8 Pinyi Zhang, Bicong Ci (2020) Deep belief network for gold price forecasting Penelitian ini menerapkan metode Deep Belief Network (DBN) yang merupakan penerapan RBM dan model Backpropagation Neural Network (BPNN) untuk mempelajari data

dari menerapkan bobot dan bias lalu dikoreksi dan diubah agar lebih menyesuaikan dengan target. Penelitian ini menggunakan fitur CPI, minyak mentah, suku bunga, indeks saham Dow Jones, dan nilai tukar Dolar AS. Hasil dari penelitian ini bahwa model dengan sepuluh neuron menghasilkan performa terbaik dengan nilai evaluasi (MAE = 0,04) dan waktu yang diperlukan untuk melatih selama 179,4 detik (Zhang & Ci, 2020).

9 Zhanhong He, Junhao Zhou, Hong-Ning Dai, Hao Wang (2019) Gold Price Forecast based on LSTM-CNN Model Penelitian ini menggunakan metode hibrida yaitu LSTM-CNN serta menerapkan teknik Attention Mechanism untuk memberikan bobot yang lebih penting pada data yang lebih berharga dan mengurangi bobot pada data yang kurang relevan. Penelitian ini menerapkan fitur harga seperti harga emas harian dalam harga Dolar AS, mata uang euro, RMB, Dolar Hongkong, per ounce. Berdasarkan hasil penelitian, model LSTM-Attention-CNN (MAE = 21,71) lebih baik dibandingkan model LSTM (MAE = 60,40) dan LSTM-CNN (MAE = 42,50) (He et al., 2019).

12 10 Xuanyi Song, Yuetian Liu, Liang Xue, Jun Wang, Jingzhe Zhang, Junqiang Wang, Long Jiang, Ziyang Cheng (2020) Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model Penelitian ini menerapkan metode LSTM untuk memprediksi serta menguji performa dari LSTM dibandingkan dengan model RNN maupun ANN.

4 Penelitian ini menggunakan teknik optimisasi seperti Particle Swarm Optimization (PSO). Penelitian ini menerapkan model untuk memprediksi laju produksi minyak harian dari reservoir vulkanik. Penelitian ini menghasilkan penemuan bahwa model LSTM (MAE = 1.60) mendapatkan performa lebih baik dibandingkan model pengujian lainnya seperti RNN (3,18) (Song et al., 2020).

Penelitian-penelitian sebelumnya menunjukkan bahwa metode hybrid deep learning dan ensemble sangat efektif untuk peramalan data deret waktu, khususnya harga emas. Beberapa penelitian menyoroti keunggulan model hibrida CNN-LSTM. Model ini terbukti menghasilkan error prediksi yang lebih kecil dibandingkan model tunggal. Sebagai contoh, sebuah studi mengembangkan model CNN-LSTM untuk prediksi harga emas dan menemukan bahwa model ini

menghasilkan nilai Mean Absolute Error (MAE) dan Root Mean Square Error (RMSE) yang paling kecil (Livieris et al., 2020). Penelitian lain juga mengonfirmasi bahwa model CNN-LSTM 18,1% lebih unggul dalam memprediksi volatilitas harga emas dibandingkan model LSTM saja (Vidal & Kristjanpoller, 2019). Keunggulan model hibrida juga terlihat pada penelitian yang menggabungkan LSTM-CNN dengan Attention Mechanism, yang menghasilkan MAE jauh lebih rendah (MAE=21,71) dibandingkan dengan model LSTM standar (MAE=60,40) atau LSTM-CNN (MAE=42,50) (He et al., 2019). Metode LSTM secara konsisten menunjukkan kinerja yang superior dalam berbagai aplikasi peramalan. Dalam prediksi penularan COVID-19, LSTM (MAPE 2,32%) terbukti lebih akurat daripada XGBoost (MAPE 7,21%) (Luo et al., 2021). Dalam peramalan harga emas, model LSTM (MAE=48,85) juga mengungguli model lain seperti Bi-LSTM (MAE=61,53) dan GRU (MAE=71,24) (Yurtsever, 13 2021). Selain itu, pada prediksi performa sumur minyak, LSTM (MAE=1.60) juga lebih baik daripada model RNN konvensional (Song et al., 2020). Metode berbasis pohon keputusan seperti XGBoost dan ensemble learning juga menunjukkan hasil yang kuat. Sebuah studi yang menggunakan XGBoost untuk peramalan harga emas dari tahun 1986 hingga 2019 menemukan bahwa model ini memiliki RMSE terendah (34.92) dibandingkan lima model lainnya (Jabeur et al., 2024). Sementara itu, penelitian lain yang menerapkan teknik ensemble seperti Random Forest dan gradient-boosted trees juga membuktikan bahwa teknik ini lebih unggul (MAE 2804,44) dibandingkan regresi linear (MAE 4384,43) dalam memprediksi harga emas (Manjula & Karthikeyan, 2019). Penelitian lain mengeksplorasi berbagai arsitektur jaringan saraf dan algoritma optimisasi. Misalnya, model Deep Belief Network (DBN) menunjukkan performa yang sangat baik untuk peramalan harga emas dengan nilai MAE serendah 0.04 (Zhang & Ci, 2020). Selain itu, penggunaan Whale Optimization Algorithm (WOA) untuk mengoptimalkan jaringan saraf tiruan (WOA-NN) juga terbukti lebih efektif (MAE=0.021) untuk memprediksi fluktuasi harga emas dibandingkan dengan model jaringan saraf lainnya (Alameer et al., 2019).

2.2. Tinjauan Teoritis Tinjauan Teoritis adalah subbab dari bab 2.

Subbab ini adalah landasan dari teori-teori yang digunakan peneliti untuk memperkuat penyelesaian masalah dalam penelitian. Teori-teori ini secara detail akan dijelaskan sebagai berikut.

2.2.1 Prediksi Harga Emas Harga emas merupakan salah satu topik atau objek penelitian yang relevan karena harganya yang dapat dikatakan cenderung naik dalam jangka panjang. Penelitian mengenai harga emas cukup beragam, baik dari analisa faktor yang mempengaruhi harga emas (Qian et al., 2019), dinamika harga emas dalam perannya di tengah ketidakpastian (Beckmann et al., 2019), maupun penelitian mengenai prediksi harga emas dengan berbagai metode Machine Learning seperti XGBoost (Jabeur et al., 2024), jaringan syaraf tiruan (Alameer et al., 2019), hingga 14 pemodelan hibrida atau gabungan dua metode seperti CNN-LSTM (Livieris et al., 2020; Vidal & Kr istjanpoller, 2020) maupun LSTM-CNN (He et al., 2019). Harga emas sendiri terbilang kompleks karena emas sendiri memiliki hubungan dengan variabel finansial maupun makroekonomi (Qian et al., 2019). Berdasarkan studi dari Qian et al., (2019) mengenai analisa yang mempengaruhi harga emas, penelitian ini menunjukkan dalam kurun waktu Januari 2020 hingga Desember 2018 bahwa CPI atau tingkat inflasi konsumen berkorelasi positif dengan harga emas, sedangkan indeks Dolar AS, suku bunga, nilai tukar, harga minyak, dan S&P 500 memiliki korelasi negatif dengan harga emas. Penelitian lainnya yang dilakukan oleh Alameer et al., (2019) mengenai peramalan harga emas dengan model jaringan syaraf tiruan, menguji variabel-variabel seperti harga komoditas, nilai tukar, dan inflasi dari harga dolar AS maupun inflasi mata uang RMB terhadap harga emas, dalam kurun waktu September 1987 hingga Agustus 2017. Hasil dari pengujian dengan uji korelasi bahwa harga komoditas khususnya harga perak, merupakan nilai korelasi yang tinggi sebesar 0,956 terhadap harga emas. Ini menandakan bahwa harga perak naik memiliki kecenderungan yang tinggi bahwa harga emas akan naik. Variabel lainnya seperti nilai tukar mata uang cina memiliki korelasi negatif terhadap

harga emas sebesar -0,23. Meskipun demikian, nilai tukar mata uang india dan nilai tukar mata uang afrika selatan memiliki korelasi positif terhadap harga emas sebesar 0,619 dan 0,605. Variabel lainnya yaitu tingkat inflasi mata uang memiliki korelasi negatif terhadap harga emas. Ini menandakan bahwa harga emas memiliki kecenderungan naik jika tingkat inflasi mengecil, meskipun nilai korelasinya tidak begitu kuat.

2.2.2 Machine Learning (ML)

ML merupakan sebuah terminologi bahwa sebuah mesin dapat mempelajari berbagai hal sehingga dapat melakukan berbagai tugas berdasarkan parameter yang telah ditentukan. ML ini dikembangkan untuk menghasilkan nilai aproksimasi berdasarkan pola-pola dalam data yang berguna dalam melaksanakan tugasnya. Pola-pola ini dapat membantu mesin untuk memahami proses, seperti berdasarkan pola-pola dapat meramal atau memprediksi maupun berasumsi dalam masa depan. ML menggunakan teori dari statistika dalam pengembangan model matematika berdasarkan sampel-sampel lalu menjadi sebuah kesimpulan (Alpaydin, 2020). ML memiliki berbagai tugas, salah satunya adalah regresi. Tugas regresi merupakan salah satu dari Supervised Learning atau pembelajaran yang diawasi, yang terdapat nilai input sebagai X dan Y sebagai output, serta dapat mempelajari pemetaan dari X ke Y. Menurut Alpaydin (2020), pendekatan ini mengoptimalkan parameter, dengan itu error atau selisih aproksimasi juga diperkecil, yang membuat estimasi output sedekat mungkin dengan output yang sesuai. (Introduction to Machine Learning, Fourth Edition - Ethem Alpaydin - Google Books, n.d.)

2.2.2.1 Deep Learning (DL)

Metode pembelajaran mesin dengan menggunakan jaringan saraf tiruan sudah digunakan dari tahun 1957. Meskipun demikian, jaringan saraf tiruan yang lebih mendalam baru digunakan sejak 2006. Karakteristik dari berbagai variasi model DL merupakan banyaknya lapisan dari neuron tersembunyi. Komponen dasar dari setiap jaringan saraf merupakan sebuah neuron di mana ide dasar dari pengembangan neuron adalah sebuah input x , bersama dengan bias b , diberikan sebuah bobot oleh w , lalu dijumlahkan bersama. Penjumlahan ini dinyatakan sebagai,

$$z = \sum w_i x_i + b$$

ada sebuah argumen untuk sebuah fungsi aktivasi, menghasilkan luaran untuk sebuah neuron. Terdapat berbagai fungsi aktivasi pada neuron, salah satunya merupakan Rectified Linear Unit (ReLU) yang merupakan fungsi aktivasi terpopuler untuk jaringan saraf mendalam. Fungsi aktivasi lainnya yaitu fungsi softmax yang mengubah vektor n -dimensi x menjadi vektor n -dimensi y , di mana seluruh komponen y adalah 1, yang berarti nilai-nilai dari y mewakili probabilitas dari masing-masing elemen (Emmert-Streib et al., 2020). Untuk mengembangkan jaringan saraf, neuron-neuron diperlukan untuk saling terhubung hingga membentuk sebuah struktur jaringan saraf. Struktur paling sederhana merupakan Feedforward Neural Networks, yang sering disebutkan Multilayer Perceptron (MLP), dapat menggunakan fungsi aktivasi yang non-linear. Perlu diperhatikan bahwa MLP tidak memiliki siklus yang memungkinkan adanya umpan balik. Itu sebabnya, diperlukan sebuah metode optimal untuk dapat fungsi loss dengan algoritma yang optimal untuk mencari parameter dengan mengurangi error selama pembelajaran (Emmert-Streib et al., 2020). 2.2 3 2.2

Long Short Term Memory (LSTM) LSTM adalah suatu versi atau varian dari Recurrent Neural Network (RNN). Sherstinsky (2018) membuat literatur mengenai dasar atau fundamental dari RNN dan LSTM. Penjelasan dimulai dengan menguraikan formulasi RNN karonik berasal dari persamaan diferensial atau turunan dan teknik dengan istilah “Unfolding” atau “Unrolling” dalam RNN. Pembahasan dilanjutkan dengan membahas kesulitan dalam melatih RNN standar dan menangani kesulitan tersebut dengan mengubah RNN ini menjadi istilah “Vanilla LSTM”. 3 12 LSTM sebagai varian atau versi yang dikembangkan dari RNN, dirancang untuk mengatasi masalah dependensi jangka panjang yang sering terjadi jika menggunakan RNN standar. Literatur paper ini memberikan sebuah identifikasi peluang baru untuk memperkaya sistem LSTM dan dipadukan eksistensi peluang tersebut ke dalam jaringan Vanilla LSTM. Paduan ini menghasilkan varian LSTM yang paling umum hingga saat ini. LSTM ini merupakan salah satu Jaringan Saraf Tiruan (JST) yang mampu untuk mempelajari ketergantungan jangka panjang

dalam data. Ini membuat model LSTM cocok untuk penafsiran bahasa, pengenalan bahasa, dan prediksi berderet waktu. Prediksi berderet waktu ini cocok untuk prediksi harga yang memiliki data historis. 20 Hal ini disebabkan oleh LSTM memiliki tiga buah gate yaitu Forget Gate, Input Gate, dan Output Gate. 17

Gambar 2. 1. Rangkaian Kerja LSTM Forget Gate menghasilkan nilai baru untuk perhitungan nilai memori jangka panjang dari perkalian antar bobot dari setiap input dan nilai memori jangka pendek awal lalu ditambahkan bias. Hasil pertama bahan ini akan menjadi nilai sumbu x pada fungsi aktivasi yaitu persamaan logistik yang memiliki range nilai (0,1). Hasil ini akan dikalikan dengan nilai memori jangka panjang awal dan disimpan untuk menghitung nilai memori jangka panjang yang baru. Input Gate menghasilkan nilai baru untuk menghitung nilai memori jangka panjang yang baru. Input Gate dibagi menjadi dua proses yaitu proses di mana proses pertama itu menggunakan fungsi aktivasi persamaan logistik sehingga menghasilkan range nilai (0,1), sedangkan proses kedua menggunakan fungsi aktivasi persamaan hiperbolik tangen atau istilah persamaan tanh sehingga menghasilkan range nilai (-1,1). Masing-masing proses ini menggunakan memori jangka pendek dan nilai input yang telah dikalikan dengan bobot lalu hasilnya akan ditambahkan juga dengan bias pada setiap proses. Hasil dari setiap proses akan menjadi nilai sumbu x pada fungsi aktivasi proses masing-masing lalu range nilai pada kedua proses akan dikalikan. Hasil perkalian kedua range nilai ini akan ditambahkan pada hasil perhitungan pada Forget Gate dan hasil pertambahan ini akan menjadi nilai memori jangka panjang yang baru. Output Gate menghasilkan nilai baru untuk menghitung nilai memori jangka pendek yang baru. Output Gate memiliki dua proses berbeda yaitu salah satu prosesnya adalah menghasilkan range nilai (0,1) dengan fungsi aktivasi persamaan logistik dan proses satunya lagi menghasilkan range nilai (-1,1) dengan fungsi aktivasi persamaan tanh. Proses pertama adalah menghitung nilai memori jangka pendek awal dikalikan bobotnya, ditambah dengan nilai input yang dikalikan bobotnya, lalu

penambahan ini ditambahkan lagi dengan bias. Hasil perhitungan akan menjadi nilai sumbu x pada fungsi aktivasi persamaan logistik yang menghasilkan range nilai (0,1). Proses kedua yaitu nilai memori jangka panjang baru akan menjadi nilai sumbu x pada persamaan tanh yang memiliki range nilai (- 1,1) lalu hasilnya akan dikalikan dengan nilai yang didapatkan pada proses menggunakan fungsi aktivasi persamaan logistik. Hasil perkalian ini akan menjadi nilai memori jangka pendek yang baru. Kumbure et al. (2022) memberikan kesimpulan dari pengembangan tinjauan literatur teknik mesin belajar dan data untuk harga saham bahwa penelitian berbasis Deep Learning telah menerapkan LSTM untuk memprediksi pasar saham. Kesimpulan dari tinjauan literatur adalah LSTM ini lebih efektif dibandingkan Deep Learning lainnya dengan hasil perbandingan yaitu prediksi yang lebih kuat.

2.2.2.3 Mean Absolute Error (MAE)

Mean Absolute Error adalah salah satu metrik yang digunakan untuk mengukur performa model prediktif, terutama dalam konteks regresi. MAE mengukur rata-rata kesalahan absolut antara nilai yang diprediksi oleh model dan nilai aktual. Kesalahan absolut adalah selisih antara nilai yang diprediksi dan nilai sebenarnya tanpa memperhitungkan arah kesalahan (positif atau negatif).

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

Keterangan : n adalah jumlah observasi (data points). y_i adalah nilai aktual untuk observasi ke- i . \hat{y}_i adalah nilai prediksi untuk observasi ke- i . $|y_i - \hat{y}_i|$ adalah nilai absolut dari selisih antara nilai aktual dan nilai prediksi untuk observasi ke- i .

19 Penilaian MAE ini akan diinterpretasikan dengan penilaian yang lebih kecil menandakan bahwa model memiliki performa lebih baik karena kesalahan yang lebih kecil. MAE yang bernilai relatif lebih besar menandakan bahwa model memiliki performa kurang baik karena kesalahan yang lebih besar. Penelitian yang dilakukan oleh Manjula & Karthikeyan, (2019) juga menerapkan MAE dalam mengukur kesalahan hasil prediksi terhadap harga emas.

2.2.3.4 Mean Squared Error (MSE)

Mean Squared Error (MSE) adalah metrik yang sering digunakan untuk mengukur performa model

prediktif, termasuk model Long Short-Term Memory (LSTM). MSE mengukur rata-rata dari kuadrat kesalahan antara nilai yang diprediksi oleh model dan nilai aktual.

Berbeda dengan (MAE) yang hanya menjumlahkan kesalahan prediksi dalam bentuk absolut, MSE memberikan penalti yang lebih besar terhadap kesalahan besar karena kesalahan tersebut dikuadratkan. Ini berarti bahwa MSE lebih

sensitif terhadap outlier, yang dapat menyebabkan nilai MSE menjadi sangat tinggi jika model tidak dapat menggeneralisasi dengan baik terhadap data yang menyimpang. Sebagai contoh dalam Studi dari Vidal & Kristjanpoller, (2020), MSE divalidasi sebagai pengukuran kesalahan yang paling sesuai untuk mengukur volatilitas harga emas. Ini mengindikasikan bahwa MSE efektif digunakan dalam situasi di mana perbedaan besar antara nilai yang diprediksi dan nilai aktual dianggap penting. MSE dirumuskan secara matematis sebagai berikut:

Di mana: A. n adalah jumlah observasi (data points). B. y_i adalah nilai aktual untuk observasi ke-i. C. y'_i

adalah nilai prediksi untuk observasi ke-i. D. adalah kuadrat dari selisih antara nilai aktual dan nilai prediksi untuk observasi ke-i. Penilaian

MSE ini akan diinterpretasikan dengan penilaian yang lebih kecil menandakan bahwa model memiliki performa lebih baik karena rata-rata kesalahan yang lebih kecil. MSE yang bernilai relatif lebih besar menandakan bahwa model memiliki performa kurang baik karena rata-rata kesalahan yang lebih besar. Penelitian yang dilakukan oleh Manjula & Karthikeyan, (2019) juga menerapkan MSE dalam mengukur kesalahan hasil

prediksi terhadap harga emas. 2.2.3.5 Root Mean Squared Error (RMSE) RMSE adalah salah satu metrik kesalahan untuk mengevaluasi kinerja

model dalam memprediksi deretan waktu. Metrik RMSE sering digunakan untuk mengukur kinerja model dari metode prediktif, khususnya LSTM. RMSE merupakan nilai akar kuadrat dari MSE, sehingga RMSE dapat memberikan gambaran mengenai tingkat kesalahan dalam satuan yang sama dengan data yang diprediksi. RMSE dirumuskan secara matematis sebagai berikut:

Di mana: A. n adalah

jumlah observasi (data points). B_i adalah nilai aktual untuk observasi ke- i . C_i adalah nilai prediksi untuk observasi ke- i . $(B_i - C_i)^2$ adalah kuadrat dari selisih antara nilai aktual dan nilai prediksi untuk observasi ke- i . Sama seperti metrik evaluasi pengukuran kinerja lainnya yaitu MAE dan MSE, semakin rendah nilai RMSE maka semakin akurat model dalam memprediksi data yang mengartikan bahwa hasil prediksi semakin mendekati dengan nilai asli. Kebalikannya, semakin besar nilai RMSE maka semakin besar selisih atau error yang dihasilkan model untuk memprediksi harga. Sebagai contoh, penelitian yang dilakukan oleh Nguyen & Walther, (2020) menerapkan RMSE sebagai salah satu metrik pengujian dalam mengukur kesalahan hasil peramalan volatilitas komoditas terhadap nilai asli. Penelitian yang dilakukan oleh Manjula & Karthikeyan, (2019) juga menerapkan RMSE dalam mengukur kesalahan hasil prediksi terhadap harga emas.

2.2.3.6 Mean Absolute Percentage (MAPE) MAPE adalah salah satu metrik kesalahan yang umum digunakan untuk mengevaluasi kinerja model prediktif, terutama dalam prediksi deret waktu. Metrik ini mengukur rata-rata persentase kesalahan antara nilai yang diprediksi oleh model dan nilai aktual, memberikan gambaran tentang seberapa besar kesalahan model dalam bentuk persentase. Salah satu kelebihan MAPE adalah hasilnya yang lebih mudah dipahami, karena kesalahan diungkapkan dalam persentase dari nilai aktual.

Secara matematis, MAPE dirumuskan sebagai berikut :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{B_i - C_i}{B_i} \right| \times 100$$

Di mana: n adalah jumlah observasi (data points). B_i adalah nilai aktual untuk observasi ke- i . C_i adalah nilai prediksi untuk observasi ke- i . $(B_i - C_i)$ adalah kuadrat dari selisih antara nilai aktual dan nilai prediksi untuk observasi ke- i . Sama seperti metrik evaluasi pengukuran kinerja lainnya yaitu MAE dan MSE, semakin rendah nilai RMSE maka semakin akurat model dalam memprediksi data yang mengartikan bahwa hasil prediksi semakin mendekati dengan nilai asli. Kebalikannya, semakin besar nilai RMSE maka semakin besar selisih atau error yang dihasilkan model untuk memprediksi harga. MAPE mengukur kesalahan relatif antara

nilai yang diprediksi dan nilai aktual. Semakin rendah nilai MAPE, semakin akurat model dalam memprediksi data, yang berarti bahwa hasil prediksi semakin mendekati nilai aktual. Sebaliknya, semakin besar nilai MAPE, semakin besar kesalahan relatif yang dihasilkan model dalam memprediksi data (Jeong, 2024). Namun, MAPE memiliki beberapa kelemahan, terutama ketika nilai aktual mendekati nol, karena dapat menghasilkan kesalahan yang sangat besar atau bahkan tidak terdefinisi. Oleh karena itu, meskipun MAPE mudah dipahami, penggunaannya perlu hati-hati dalam situasi tertentu. Beberapa penelitian terkini menunjukkan bahwa MAPE tetap menjadi metrik yang populer dan efektif untuk mengukur kinerja model prediktif dalam memprediksi deretan waktu (He et al., 2019; Luo et al., 2021; Makalesi & Yurtsever, 2021; Zhang & Ci, 2020).

2.2.3.7 Jaringan Syaraf Tiruan (JST) Jaringan Syaraf Tiruan (JST) atau Artificial Neural Network (ANN) merupakan sistem komputasi yang terinspirasi cara kerja otak manusia. Secara umum, ANN tersusun atas lapisan input, satu atau beberapa lapisan tersembunyi, dan lapisan output. Setiap neuron melakukan kombinasi linear terhadap input yang dibobot dan diberi bias, lalu menerapkan fungsi aktivasi seperti ReLU, Sigmoid, atau Tanh untuk menghasilkan sinyal keluaran (Kurniasari et al., 2023). Latihan model ANN umumnya menggunakan algoritma backpropagation. Prosesnya melibatkan perhitungan kesalahan (loss), kemudian menggunakan turunan derivatif kesalahan tersebut untuk memperbarui bobot secara bertahap melalui metode seperti gradient descent. Untuk pemodelan data berurutan (seperti time-series), varian ANN seperti RNN dan LSTM digunakan. RNN memiliki hidden state yang mempertahankan informasi dari langkah sebelumnya, memungkinkan penangkapan ketergantungan temporal. Namun, RNN pada dasarnya rentan terhadap permasalahan vanishing atau exploding gradient ketika menangani urutan panjang. LSTM dan GRU mengatasi hal tersebut dengan mekanisme gerbang (gating) untuk mengatur aliran informasi (Kim et al., 2025). Model ANN dan variannya banyak digunakan dalam berbagai aplikasi, mulai dari peramalan cuaca, ekonomi,

hingga konsumsi energi. CNN 1-D, meskipun pertama kali dikembangkan untuk data citra, juga sering diaplikasikan untuk mengekstraksi fitur dari data time-series.

24 BAB III METODE PENELITIAN

Pada bab tiga ini akan menjelaskan pendekatan penelitian yang diterapkan pada penelitian ini, dimulai dari pemilihan metode kuantitatif sebagai rancangan penelitian, metode penelitian, metode pengumpulan data, metode analisis data, dan metode pengujian data.

3.1. Paradigma Penelitian

Paradigma penelitian ini menggunakan sebuah diagram yaitu diagram tulang ikan atau Fishbone Diagram. Fishbone Diagram atau dikenal sebagai Diagram Sebab-Akibat adalah sebuah pendekatan terstruktur dalam menganalisis dan menggambarkan hubungan sebab-akibat dari suatu permasalahan (Putu Widnyana et al., 2022.). Diagram Fishbone digambarkan dalam bentuk tulang-tulang ikan yang membentuk penyebab permasalahan lalu dihubungkan ke kepala ikan sebagai akibat dari permasalahan. Berikut diagram Fish bone yang digunakan dalam penelitian analisis time series atau deretan waktu.

Gambar 3. 1. Diagram Fishbone Paradigma Penelitian Berdasarkan gambar Diagram Fishbone di atas, terdapat empat buah penyebab dari suatu akibat yaitu sulitnya memprediksi pergerakan harga emas 25 dengan metode LSTM. Empat buah penyebabnya adalah faktor data, manusia, metode, dan juga mesin. Faktor atau penyebab pertama adalah data yang kurang representatif atau kurang relevan sehingga kurang menggambarkan pergerakan harga emas secara nyata. Data yang kurang cukup juga menjadi faktor dari kurangnya efektif metode LSTM dalam memprediksi harga emas. Faktor atau penyebab kedua adalah manusia yang kurang dalam pengetahuan untuk memilih model tanpa menganalisis karakteristik dari objek data, khususnya harga emas. Penyebab dari segi manusia juga berasal dari manusia yang terlalu mengandalkan intuisi atau pengalaman dari praktik kerja untuk menerapkan metode LSTM dalam memprediksi harga emas. Faktor atau penyebab ketiga adalah metode di mana pemilihan metode LSTM yang kurang ideal terlebih dengan faktor pertama yaitu data, membuat metode LSTM tidak optimal dalam memprediksi pergerakan harga emas. Selain itu, terdapat faktor

lainnya seperti metode LSTM memiliki keterbatasan dalam menangkap pergerakan harga emas yang sangat dipengaruhi oleh faktor eksternal. Faktor atau penyebab keempat yaitu mesin, di mana kurangnya sumber daya komputasi yang memadai untuk menunjang pelatihan model LSTM. Faktor lainnya mengenai mesin juga dari kurangnya tingkat akurasi yang cukup tinggi untuk memprediksi harga emas. Diharapkan dengan adanya Diagram Fishbone pada Gambar 3.1. Diagram Fishbone Paradigma Penelitian, penelitian dalam memprediksi harga emas dapat dilakukan secara sistematis dan menggambarkan permasalahan utama sehingga menjadi acuan untuk memecahkan masalah yang diangkat dalam kasus ini. Penelitian mengenai analisa time series harga emas dengan menerapkan prediksi pada pendekatan Deep Learning : LSTM ini dilakukan perbandingan pada akurasi dari masing-masing model yang telah dilatih untuk menghasilkan hasil yang optimal dalam memprediksi harga emas. Oleh karena itu, diharapkan penelitian ini dapat memberikan hasil akurasi yang tinggi serta nilai error atau kesalahan atau selisih antara nilai asli dengan hasil prediksi terkecil.

3.2. Metode Penelitian

Metode penelitian dapat dipahami sebagai sebuah teknik maupun pendekatan yang sistematis untuk digunakan dalam suatu penelitian. Dengan kata lain, metode penelitian merupakan sebuah langkah-langkah sistematis untuk digunakan selama penelitian.

17 Metode penelitian yang umum digunakan merupakan metode kuantitatif, kualitatif, dan gabungan baik kuantitatif dan kualitatif.

23 Dalam penelitian ini, penulis menggunakan metode kuantitatif. Tujuan dari penggunaan metode kuantitatif untuk menghasilkan nilai numerik yang dapat dianalisis secara statistik dalam penekanan hasil yang objektif (Fidella, 2024). Penulis membuat sub bab Metode Penelitian dipecah menjadi tiga subsub bab yaitu metode pengumpulan data, metode analisis data, dan metode pengujian data.

3.2.1 Metode Pengumpulan Data

Dalam melakukan penelitian dengan menggunakan metode penelitian kuantitatif, di mana pengumpulan data dilakukan dengan cara studi literatur terlebih dahulu. Cara ini melibatkan pencarian sumber-sumber yang relevan untuk dijadikan acuan dalam menerapkan data penelitian yang terdiri dari

variabel-variabel independen dan variabel dependen. Studi literatur ini dilakukan dengan mencari artikel ilmiah dalam menerapkan prediksi mengenai harga emas, melibatkan berbagai faktor-faktor eksternal dari harga emas seperti variabel finansial maupun makroekonomi. Dalam penelitian ini, penulis menggunakan data historis harga emas yang diperoleh dari sumber laman keuangan yaitu yahoo finance. Penulis memperoleh data-data historis ini dengan frekuensi harian dari 1 Januari 2017 sampai 28 Februari 2025 atau sebanyak 2051 baris data. Data yang digunakan selama penelitian ini mencakup data harga emas sebagai variabel dependen, lalu data sebagai prediktor atau variabel independen mencakup harga minyak mentah, perak, indeks harga S&P 500, dan indeks harga Dolar Amerika Serikat (USD Index / USDX). Masing-masing harga memiliki harga terendah, tertinggi, pembuka, penutup, dan volume yang diukur secara frekuensi harian.

3.2.2 Metode Analisis Data

Analisis Data adalah tahapan memeriksa, menguraikan, dan mempelajari masalah lebih mendalam sehingga mendapatkan informasi dan kesimpulan yang berguna. Sebuah artikel yang ditulis oleh Rian Tineges (2021) pada situs web 27 DQLab menjelaskan bahwa analisis data merupakan proses penting dalam memahami data dan mengambil keputusan berdasarkan informasi yang diperoleh.

14 Dalam kasus penelitian ini, analisis pada variabel harga untuk dapat memahami pola, tren, perubahan harga sehingga dapat memprediksi harga emas pada masa depan. Secara umum, analisis dibagi menjadi empat jenis, yaitu: Analisis deskriptif yang meliputi penjelasan data; Analisis diagnostik yang melibatkan penyebab suatu fenomena; Analisis prediktif yang melibatkan analisis prediksi atau peramalan masa depan pada data historis; Analisis preskriptif yang melibatkan pemberian rekomendasi tindakan ke depannya (Rian Tineges, 2021).

18 Dalam penelitian ini, metode analisis data yang digunakan dengan pendekatan analisis deskriptif dan analisis prediktif. Pendekatan analisis deskriptif meliputi identifikasi tren, pola, maupun outlier yang terdapat dalam data sehingga dapat memberikan wawasan dasar untuk menggambarkan karakteristik harga emas maupun harga sebagai variabel

independen lainnya yang digunakan untuk meramalkan harga emas. Pendekatan analisis prediktif meliputi penggunaan data-data yang memadai untuk membangun model prediksi. Menurut Informati Digital, (2024), analisis prediktif digunakan untuk memperkirakan tren atau perilaku di masa mendatang berdasarkan data historis, seperti dalam kasus prediksi harga. Salah satu metode yang digunakan dalam analisis prediktif adalah Deep Learning (DL) dengan menggunakan metode LSTM, yang dapat menangkap pola temporal dan ketergantungan jangka panjang dalam data. Model dengan metode LSTM ini dilatih menggunakan data harga emas dan variabel independen lainnya, sehingga mampu mengidentifikasi pola-pola temporal dan hubungan jangka panjang yang mempengaruhi fluktuasi harga emas. Dengan demikian, model LSTM dapat memberikan prediksi yang lebih akurat dan relevan berdasarkan data historis yang tersedia.

3.2.3 Metode Pengujian Data

Dalam penelitian ini, pengujian model dilakukan dengan membandingkan kinerja beberapa model yang telah dilatih untuk memprediksi harga emas menggunakan metrik evaluasi seperti MSE, RMSE, MAE, dan MAPE. Metrik-metrik ini digunakan untuk menilai sejauh mana model dapat menghasilkan prediksi yang akurat berdasarkan data historis. Pengujian ini berfokus pada pengukuran kesalahan prediksi untuk menilai performa model dalam memprediksi harga emas. Model yang diukur pada kesalahan atau selisih terkecil merupakan model yang terbaik. Melalui perbandingan ini, diharapkan dapat ditemukan model yang memberikan hasil terbaik dalam memprediksi fluktuasi harga emas dengan mempertimbangkan faktor-faktor yang ada pada data.

29 BAB IV PERENCANAAN

Pada bab tiga ini akan menjelaskan langkah-langkah penelitian hingga perancangan pengujian yang memiliki spesifikasi tertentu.

4.1. Langkah-Langkah Penelitian

Penelitian ini dilakukan dengan tujuan untuk mengembangkan model prediksi harga emas menggunakan pendekatan DL. Langkah-langkah yang akan diambil mencakup berbagai tahapan mulai dari pengumpulan data, pra-pemrosesan data, pelatihan model, hingga evaluasi kinerja model. Proses ini dilakukan untuk memperoleh hasil yang optimal dalam meramalkan fluktuasi harga emas

berdasarkan data historis yang ada. Setiap tahap penelitian di rancang untuk memastikan bahwa model yang dibangun dapat memberikan prediksi yang akurat dan relevan. Berikut alur mengembangkan sebuah ML dengan pendekatan DL metode LSTM yang akan dirancang untuk sebagai model pengujian. Gambar 4. 1 Diagram cara kerja mengembangkan Deep Learning (DL) Berdasarkan diagram cara kerja mengembangkan DL dapat diuraikan sebagai sub- subbab berikut.

4.1.1. Koleksi Data Pada tahapan penelitian koleksi data, penelitian ini menerapkan pemuatan data dari situs Yahoo Finance untuk memuatkan data historis harian harga emas dan 30 harga sebagai variabel independen eksternal lainnya. Tahapan ini melibatkan komputasi untuk pemuatan data yang dimulai dengan memasukkan tanggal mulai dan tanggal akhir yang diinginkan untuk dijadikan data penelitian. Setelah memasukkan tanggal, data akan dimuat dan disimpan ke file comma-separated values (.csv).

4.1.2. Seleksi Fitur Pada tahapan penelitian seleksi fitur, penelitian ini menerapkan metode - metode analisis data sebelum masuk ke pemrosesan data melalui pelatihan model. Tahapan ini diperlukan agar model tidak mempelajari data yang kurang relevan maupun data yang tidak terlalu berkorelasi dengan variabel dependen atau harga emas sehingga model mempelajari data lebih baik dan menghasilkan performa yang lebih unggul dibandingkan tanpa seleksi fitur. Penelitian ini dalam menerapkan seleksi fitur dimulai dengan memuat file data dalam format .csv untuk kemudian menerapkan statistik deskriptif guna mengidentifikasi tren dan pola dalam dataset. Setelah itu, dilakukan visualisasi data dengan plotting untuk setiap variabel, termasuk harga emas, guna memperoleh gambaran yang lebih jelas mengenai distribusi dan hubungan antar variabel. Selanjutnya, uji korelasi diterapkan untuk mengukur sejauh mana hubungan antar fitur yang ada, dengan fokus pada identifikasi variabel yang memiliki korelasi signifikan terhadap variabel target, yaitu harga emas. Variabel-variabel yang relevan dan memiliki pengaruh signifikan dipilih sebagai fitur yang akan digunakan dalam model, sementara fitur yang tidak relevan atau memiliki korelasi rendah akan

dieliminasi untuk meningkatkan akurasi dan efisiensi model. 4.1.3.

Normalisasi Data Pada tahapan penelitian normalisasi data, penelitian ini menerapkan agar data-data yang digunakan ini dalam skala yang sama atau seragam, penting untuk metode seperti LSTM yang sensitif terhadap skala. Salah satu teknik yang diterapkan dalam penelitian ini adalah Min-Max Scaling, di mana setiap nilai dari fitur akan dipetakan ke dalam rentang nilai [0,1]. Berikut adalah rumus matematis dalam penerapan penskalaan Min-Max Scaling.

$$x' = \frac{x - \min}{\max - \min}$$

Berdasarkan rumus matematis di atas, nilai x' merupakan nilai hasil penskalaan Min-Max Scaling dan nilai x merupakan nilai asli dari sebuah data pada fitur. Nilai \max merupakan nilai tertinggi atau maksimum dari fitur tersebut dan nilai \min merupakan nilai terendah atau minimum dari fitur tersebut. Teknik ini memastikan agar semua fitur memiliki kontribusi yang setara terhadap model, menghindari dominasi fitur dengan skala yang lebih besar dan mempermudah berbagai metode DL, khususnya metode LSTM lebih baik untuk mengenali pola dalam data deretan waktu. Berikut diagram dari normalisasi data khususnya dengan teknik Min-Max Scaling. Gambar 4.2 Gambar diagram tahapan penelitian normalisasi data dengan teknik Min-Max Scaling

4.1.4. Pembuatan Dataset atau Kumpulan Data Pada tahapan atau langkah penelitian pembuatan kumpulan data, penelitian ini dimulai dengan menerapkan teknik sliding window. Teknik ini digunakan untuk menyusun data dalam bentuk yang dapat diterima oleh beberapa pemodelan berbasis DL, khususnya LSTM. Dalam hal ini, setiap jendela yang bergerak akan mengumpulkan sejumlah pengamatan dari deret waktu sebelumnya dan menyusunnya menjadi fitur input yang akan digunakan oleh model untuk memprediksi nilai masa depan. Sebagai contoh, jika model menggunakan data dari Januari 2021, jendela ini akan menangkap pola dari periode tersebut untuk meramalkan nilai pada Februari 2021 (Dong et al., 2020), dengan tujuan memungkinkan model untuk mempelajari hubungan temporal dan dinamika dalam data. Setelah menerapkan teknik sliding window, penelitian ini menerapkan pembuatan

dataset untuk menghasilkan dua variabel utama yaitu X dan y . X merupakan fitur-fitur untuk prediksi, yaitu data-data yang terkait dengan harga emas berdasarkan sliding window yang diterapkan, sedangkan y adalah target atau harga emas pada hari berikutnya. Setelah dataset atau kumpulan data dibuat, penelitian ini menerapkan pembagian kumpulan data dengan ukuran yang ingin ditetapkan. Pembagian ini meliputi data untuk pelatihan dan juga data untuk pengujian. Berikut adalah diagram tahapan penelitian pembuatan dataset yang meliputi teknik sliding-window, pembuatan dataset, dan pembagian dataset. Gambar 4.3

Gambar diagram tahapan pembuatan dataset 4.1.5. Penerapan Metode LSTM Pada tahapan penelitian penerapan metode LSTM, penelitian ini menerapkan metode LSTM sebagai model untuk memprediksi harga emas berdasarkan data time-series yang telah disiapkan. Berdasarkan Gambar 2.1 yang merupakan rancangan kerja sebuah Unit LSTM, bahwa rancangan kerja dipilih karena kemampuannya untuk mengenali hingga memproses data yang memiliki ketergantungan waktu, seperti yang terdapat pada data-data berderet waktu. Dalam penelitian ini, fitur-fitur harga emas yang telah dijadikan sebuah data pelatihan maupun pengujian, akan menjadi masukkan atau input deretan waktu. Input ini akan diterima di lapisan pertama LSTM yang memiliki LSTM Unit, setiap LSTM Unit dalam lapisan pertama akan mengolah input deretan waktu tersebut dengan mempertimbangkan hubungan temporal antar fitur. Setiap unit LSTM terdiri dari tiga gerbang utama, yaitu gerbang input, gerbang lupa, dan gerbang output. Ketiga gerbang ini bekerja secara bersamaan untuk mengontrol aliran informasi mengenai pola-pola pergerakan harga emas yang masuk, lalu diolah untuk dijadikan memori serta memutuskan baik dipertahankan atau dilupakan dari memori jangka pendek maupun jangka panjang. Gerbang input berperan dalam menentukan informasi mengenai pergerakan harga emas baru apa yang akan disimpan dalam memori. Gerbang lupa akan memutuskan bagian mana dari informasi mengenai pola-pola harga emas lama yang tidak lagi relevan dan harus dilupakan. Sedangkan gerbang output akan menentukan bagian

mana dari memori yang akan dikeluarkan sebagai output pada waktu tertentu. Proses ini membuat LSTM sangat efektif dalam mengingat pola-pola penting pada data historis, seperti pergerakan harga emas dalam periode waktu tertentu. Berdasarkan novel mengenai mekanisme efisiensi memori Backpropagation Through Time (BPTT) yang dituliskan oleh Gruslys et al., (2016), dengan menggunakan mekanisme tersebut bahwa model penerapan LSTM akan memperbarui bobot dan bias dari setiap gerbang selama proses pelatihan. Penerapan mekanisme ini agar model dapat belajar dari kesalahan, mengurangi biaya komputasi, serta meningkatkan akurasi prediksi. Penelitian ini menerapkan lapisan kedua dari lapisan LSTM sehingga output atau luaran atau hasil pembelajaran lapisan pertama akan menjadi masukan untuk lapisan kedua dari LSTM. Sama halnya dengan lapisan pertama, informasi-informasi yang didapatkan dari pola-pola pergerakan harga emas akan dipelajari serta diolah untuk menjadikan memori atau luaran serta mempertahankannya atau tidak melupakan informasi tersebut pada memori. Penelitian ini menerapkan jaringan syaraf tiruan (JST) sebagai lapisan jaringan untuk memproses informasi yang didapatkan dari lapisan kedua LSTM dan memetakan harga emas secara linear ke lapisan output berupa analisis numerik pergerakan harga emas. Berikut diagram tahapan atau langkah penelitian perancangan arsitektur model. Gambar 4.4

Gambar diagram tahapan penelitian penerapan metode LSTM Berdasarkan gambar diagram di atas, tahapan penelitian ini dimulai dengan memasukkan data pelatihan ke lapisan pertama model LSTM. Setelah itu, masukkan nilai untuk jumlah LSTM unit yang dirancang sebagai lapisan yang memproses data berurutan dengan panjang sliding window yang ditetapkan pada data pelatihan. LSTM unit dalam lapisan ini bertanggung jawab untuk menangkap serta mengenali pola-pola serta ketergantungan dalam data historis. Selanjutnya, masukkan nilai untuk jumlah LSTM unit lebih kecil dari lapisan pertama yang dirancang sebagai lapisan yang memperkecil model agar tidak terlalu kompleks namun tetap mempertahankan kemampuan dalam menangkap hubungan lebih mendalam antara data-data sebelumnya.

Selanjutnya, model dilengkapi dengan lapisan jaringan syaraf tiruan untuk memproses memori-memori berupa informasi pergerakan harga emas yang telah diekstraksi oleh lapisan LSTM. Lapisan ini menggunakan neuron yang relatif lebih kecil jumlahnya dibanding LSTM unit untuk memperhalus proses pemetaan memori dari hasil ekstraksi lapisan LSTM ke lapisan output.

Berikut adalah gambar penerapan metode LSTM dalam penelitian. Gambar 4.

5 Gambar penerapan metode LSTM yang digunakan selama penelitian 4.1 **15** 6. Pelatihan

Model Pada tahapan pelatihan model, model LSTM yang telah dirancang akan dilatih menggunakan data pelatihan yang telah disiapkan sebelumnya. Proses pelatihan

ini akan melibatkan optimasi model menggunakan algoritma Adam dan metrik MSE dan MAE untuk mengukur seberapa besar perbedaan antara hasil prediksi model dan nilai sebenarnya pada data pelatihan selama pelatihan berlangsung. Dalam pelatihan ini, batch size digunakan untuk menentukan berapa banyak sampel data yang diproses sekaligus dalam satu iterasi pembaruan bobot. Penggunaan batch size yang tepat penting untuk mengontrol seberapa cepat model belajar dan untuk menghindari pembaruan bobot yang terlalu besar atau terlalu kecil. Di sisi lain, jumlah epoch merujuk pada berapa kali seluruh dataset pelatihan digunakan untuk melatih model. Semakin banyak jumlah epoch, semakin banyak kesempatan bagi model untuk belajar dari data, namun juga harus dipertimbangkan agar tidak terjadi overfitting. Pemilihan batch size dan jumlah epoch ini bertujuan untuk mencapai keseimbangan antara kecepatan pelatihan dan kemampuan model dalam mempelajari pola yang relevan dari data. Berikut adalah diagram tahapan atau langkah penelitian untuk pelatihan model.

36 Gambar 4.6 Gambar diagram tahapan penelitian pelatihan model LSTM

4.1.7. Evaluasi Model Pada tahapan penelitian evaluasi model, penelitian ini menerapkan metrik MAE, MSE, RMSE, dan MAPE untuk menguji kinerja model.

Model yang telah dilatih akan dimuat lalu diprediksi dengan data baru. Hasil dari prediksi dengan data pengujian akan dievaluasi lalu

dicetak berdasarkan metrik evaluasi. 4.2. Rancangan Pengujian Pada subbab

Rancangan Ujian ini akan dipaparkan beberapa spesifikasi dalam perancangan

untuk dapat melaksanakan penelitian sehingga pengujian hipotesa dapat berjalan dengan baik. Beberapa spesifikasi ini meliputi dari tahapan atau langkah-langkah penelitian yang memiliki spesifikasi khusus. Berikut adalah beberapa spesifikasi yang diusulkan untuk melaksanakan penelitian.

4.2.1. Spesifikasi Seleksi Fitur Pada spesifikasi seleksi fitur, terdapat berbagai metode dalam menyeleksi variabel-variabel yang akan diseleksi sehingga fitur-fitur yang relevan akan menjadi efektif pada model dan fitur yang kurang relevan akan dibuang. Beberapa seleksi fitur ini meliputi analisis deskriptif dengan statistik deskriptif, visualisasi untuk plotting pada setiap variabel, dan uji korelasi antar hubungan variabel independen terhadap variabel dependen (harga emas).

37 4.2.1.1 Spesifikasi Statistik Deskriptif Analisis statistik deskriptif digunakan untuk memberikan gambaran yang jelas mengenai distribusi dan karakteristik utama dari variabel-variabel numerik yang ada dalam dataset. Dengan pendekatan ini, kita dapat memperoleh informasi yang lebih mendalam tentang data, seperti kecenderungan pusat, variasi, dan bentuk distribusi. Analisis ini tidak hanya diterapkan pada data harga emas, tetapi juga pada variabel eksternal lainnya, seperti harga minyak mentah, harga perak, indeks S&P 500, dan indeks harga dolar AS (USDIX). Dengan memperhitungkan faktor-faktor eksternal ini, kita dapat memahami hubungan antar variabel dan dinamika pasar yang lebih luas. Beberapa ukuran statistik deskriptif yang akan digunakan dalam analisis ini meliputi:

- A. Mean (Rata-rata): Mengukur nilai rata-rata dari setiap variabel dalam dataset.
- B. Median (Nilai Tengah): Menunjukkan nilai tengah yang membagi data menjadi dua bagian yang sama besar, berguna untuk mengatasi adanya outlier.
- C. Nilai Maksimum dan Minimum: Memberikan informasi mengenai batas atas dan batas bawah dari data yang dianalisis.
- D. Standar Deviasi: Mengukur sejauh mana variasi atau penyebaran data dari rata-rata.
- E. Skewness (Kemencengan Distribusi): Menggambarkan sejauh mana distribusi data miring ke kiri atau ke kanan, yang memberikan informasi tentang simetri distribusi.
- F. Kurtosis:

Mengukur sejauh mana distribusi data cenderung memiliki puncak yang tajam atau datar, yang menunjukkan ketajaman atau kelembutan distribusi data.

4.2.1.2 Spesifikasi Visualisasi Plotting Untuk mendukung analisis statistik deskriptif pada harga emas dan variabel eksternal lainnya (harga minyak mentah, harga perak, indeks S&P 500, dan indeks harga dolar AS), digunakan beberapa jenis visualisasi sebagai berikut:

4.2.1.2.1 Plot Distribusi Distribution plot digunakan untuk menggambarkan bentuk distribusi dari variabel numerik dalam dataset, sehingga dapat memberikan pemahaman visual mengenai kecenderungan pusat, penyebaran data, dan potensi adanya outlier. Plot ini menggabungkan histogram dengan kurva Kernel Density Estimation (KDE) yang memperhalus visualisasi distribusi, sehingga kita dapat melihat apakah data cenderung berdistribusi normal, simetris, atau miring (skewed). Dalam konteks analisis ini, plot distribusi diterapkan pada variabel harga penutupan dari berbagai instrumen, yaitu emas_close, minyak_close, perak_close, sp500_close, dan usdx_close. Ini bertujuan untuk melihat pola distribusi masing-masing harga instrumen. Secara teknis, sumbu horizontal (X) menunjukkan nilai dari masing-masing variabel, sementara sumbu vertikal (Y) menggambarkan frekuensi atau kepadatan data. Untuk memperkaya informasi visual, ditambahkan garis vertikal yang menandai nilai mean dan median, yang masing-masing dibedakan dengan warna khusus. Dengan visualisasi ini, peneliti dapat lebih mudah mengenali struktur distribusi data dan perbedaan karakteristik antar variabel secara intuitif.

4.2.1.2.2  2 Boxplot Plotting Boxplot, atau yang dikenal sebagai grafik kotak garis, merupakan visualisasi statistik yang digunakan untuk menampilkan persebaran data berdasarkan lima ukuran utama: nilai minimum, kuartil pertama (Q1), median (Q2), kuartil ketiga (Q3), dan nilai maksimum. Visualisasi ini sangat berguna untuk mengidentifikasi sebaran data, melihat simetri atau ketidaksimetrian distribusi, serta mendeteksi keberadaan outlier secara visual. Dalam analisis ini, boxplot digunakan untuk menggambarkan karakteristik data dari variabel harga penutupan, yaitu emas_close, minyak_

close, perak_close, sp500_close, dan usdx_close. Setiap boxplot menunjukkan Rentang Interkuartil atau Interquartile Range (IQR) melalui bentuk kotak, dengan garis horizontal di dalamnya sebagai penanda median. Garis vertikal (whiskers) merepresentasikan rentang data di luar kuartil hingga batas minimum dan maksimum, sementara titik-titik di luar whiskers menunjukkan nilai-nilai outlier. **6** Sumbu horizontal mewakili kategori variabel, sedangkan sumbu vertikal menunjukkan nilai numerik dari masing-masing variabel. Dengan 39 visualisasi ini, pengguna dapat dengan cepat membandingkan persebaran, variabilitas, dan potensi outlier antar variabel secara efisien dan intuitif. 4.2.1.2 **2 11** **3** Scatter Plot atau Grafik Sebar Scatter plot, atau grafik sebar, merupakan jenis visualisasi yang digunakan untuk menunjukkan hubungan antara dua variabel numerik. Dengan menempatkan satu variabel pada sumbu X dan variabel lainnya pada sumbu Y, scatter plot memungkinkan pengguna untuk mengamati pola hubungan, apakah linier, non-linier, positif, negatif, atau bahkan tidak ada hubungan sama sekali. Dalam analisis ini, scatter plot digunakan untuk menggambarkan hubungan antara variabel harga penutupan dari emas dengan variabel harga penutupan instrumen lainnya seperti harga minyak mentah, perak, indeks S&P 500, dan indeks dolar AS. Visualisasi ini berperan sebagai langkah eksplorasi awal dalam mengidentifikasi potensi korelasi sebelum dilakukan pengujian statistik formal. Titik-titik pada plot merepresentasikan pasangan nilai pada waktu yang sama, sedangkan pola distribusi titik memberikan indikasi arah dan kekuatan hubungan. **13** Scatter plot juga dapat dilengkapi dengan garis regresi linier untuk memperjelas kecenderungan hubungan. Oleh karena itu, grafik ini tetap ditempatkan dalam spesifikasi visualisasi, sebagai pendukung pemahaman awal sebelum uji korelasi dilakukan. 4.2 **6 19** **1.3** Spesifikasi Uji Korelasi Uji korelasi digunakan untuk mengukur kekuatan dan arah hubungan linier antara dua variabel numerik. Dalam analisis ini, digunakan metode Pearson Correlation Coefficient, yang mengukur hubungan linear antar variabel dengan nilai koefisien berkisar antara -1 hingga 1. **2** Nilai mendekati 1 menunjukkan

korelasi positif kuat, mendekati -1 menunjukkan korelasi negatif kuat, sedan gkan nilai mendekati menunjukkan tidak adanya hubungan linear yang signifikan. Uji ini dilakukan dengan tingkat signifikansi se besar 5% ($\alpha = 0,05$) untu k menentukan apakah korelasi yang ditem ukan secara statistik signifikan atau tidak. Hasil korelasi kemudian divisualisasikan menggunakan heatmap, yaitu peta warna dua dimensi yang menunjukkan nilai koefisien korelasi antar pasangan variabel dalam bentuk grid. Warna dalam heatmap me representasikan kekuatan hu bungan, dengan gradasi dari biru (negatif), putih (netral), hingga merah (positif). Dalam konteks ini, uji 40 korelasi Pearson dilakukan pada variabel emas_close, mi nyak_close, perak_close, sp500_close, dan usdx_close sebagai bagian dari analisis hubungan antar aset global. Heatmap ini tidak hanya memberikan gambaran numerik, tetapi juga menyajikan pola hubungan yang mudah diinterpretasikan secara visual, s ehingga sangat berguna sebagai dasar dalam pemodelan prediktif maupun analisis lanjutan.

4.2.2. Spesifikasi Pembuatan Dataset

Pada spesifikasi pembuatan dataset, di tetapkan bahwa ukuran jendela data berderet waktu ini 30. Uku ran sliding window sebesar tiga puluh, yang di artikan bahwa data untuk me mprediksi data contoh waktu pa da h ari t (misalnya harga emas pada hari ke-31) akan menggunakan data d ari 30 hari sebelumnya, yaitu dari hari t- 30 hingga hari t-1. Dengan demikian, setiap window mencakup informasi historis dalam periode waktu tertentu yang digunakan untuk memprediksi nilai target di masa depan. Dataset ini memanfaatkan ke tergantungan temporal dalam data, di mana pola h arg a pada masa lalu diharapkan memiliki ki pengaruh terhadap harga pada masa yang akan datang. Spesifikasi lainnya adalah penelit ian ini menerapka n tiga buah ukuran dataset dengan pe rbandingan ukuran data pelatihan dan data pengujian (pelatihan : pengujian) yaitu ukuran 70:30, 80:20, dan 90:10. Pembagian dataset ini bertujuan untuk mengevaluasi kinerja model dengan berbagai proporsi data pelatihan dan pengujian, se hingga dapat mengidentifikasi perbandingan terbaik yang m emberikan hasil prediksi yang optimal. Pada pembagian

70:30, sebagian besar data digunakan untuk pelatihan, sementara 30% sisanya digunakan untuk menguji kemampuan model dalam melakukan generalisasi. Demikian pula pada pembagian 80:20 dan 90:10, di mana proporsi data pelatihan semakin besar, memungkinkan model untuk belajar lebih banyak dari data yang ada, tetapi dengan pengujian yang lebih kecil untuk menilai kemampuannya pada data yang belum pernah dilihat sebelumnya. Dengan variasi perbandingan ini, diharapkan dapat ditemukan keseimbangan yang tepat antara jumlah data yang digunakan untuk pelatihan dan seberapa baik model mampu menguji kinerjanya pada data yang tidak terlihat.

4.2.3. Spesifikasi Penerapan Metode LSTM

Pada spesifikasi penerapan metode LSTM, penelitian ini menggunakan beberapa konfigurasi arsitektur untuk mengoptimalkan kinerja model LSTM dalam memprediksi harga emas. Setiap konfigurasi berbeda pada jumlah LSTM unit di lapisan pertama dan kedua, serta jumlah neuron pada lapisan jaringan syaraf tiruan. Berikut adalah tabel spesifikasi konfigurasi arsitektur model dengan penerapan metode LSTM.

Tabel	Spesifikasi konfigurasi arsitektur pada penerapan metode LSTM
Model	Jumlah LSTM Units Lapisan Pertama
Jumlah LSTM Units Lapisan Kedua	Jumlah neuron dalam lapisan JST
1	64 32 25 2 96 64 30 3 128 64 30 4 128 64 40 5 256 128

Berdasarkan tabel spesifikasi di atas, penelitian ini menguji beberapa konfigurasi model LSTM dengan variasi jumlah unit pada lapisan pertama LSTM dan Lapisan kedua LSTM, serta neuron pada lapisan jaringan syaraf tiruan, yang bertujuan untuk menemukan arsitektur terbaik dalam memprediksi harga emas. Lapisan pertama LSTM adalah lapisan pertama yang bertanggung jawab untuk memproses data berurutan dan menangkap pola-pola dalam data historis. Pada eksperimen ini, jumlah unit pada lapisan pertama bervariasi mulai dari 64 hingga 256 untuk mengukur dampaknya terhadap kinerja model. Lapisan kedua LSTM digunakan untuk memperdalam pemahaman model terhadap hubungan temporal dalam data, dengan jumlah unit yang lebih kecil dari lapisan pertama untuk mencegah kompleksitas model yang berlebihan. Lapisan jaringan syaraf tiruan berfungsi untuk memetakan

hasil ekstraksi fitur dari lapisan LSTM ke prediksi akhir, dengan jumlah neuron yang lebih kecil dibandingkan unit LSTM untuk menjaga kesederhanaan model dan mencegah overfitting.

4.2.4. Spesifikasi Pelatihan Model

Pada spesifikasi pelatihan model, penelitian ini menguji berbagai kombinasi epoch dan batch size untuk mengoptimalkan proses pelatihan model LSTM.

Epoch mengacu pada jumlah iterasi di mana seluruh dataset pelatihan digunakan untuk memperbarui bobot model. Pengaturan jumlah epoch yang bervariasi bertujuan untuk memberikan model cukup kesempatan untuk mempelajari pola-pola yang relevan dari data tanpa menyebabkan overfitting.

Sementara itu, batch size menentukan jumlah sampel yang diproses dalam satu iterasi pembaruan bobot.

Pemilihan batch size yang tepat diharapkan dapat membantu model belajar dengan lebih stabil dan efisien.

Berikut adalah tabel dari berbagai kombinasi epoch dan batch size yang diuji pada pelatihan model:

Tabel 4.2 Tabel spesifikasi jumlah epoch dan ukuran batch data

Epochs	Batch-Size
120	8
230	8
308	4
308	16
40	16

Berdasarkan tabel spesifikasi jumlah epochs dan ukuran batch data, penelitian ini bertujuan untuk menemukan pengaturan yang optimal antara jumlah epoch dan ukuran batch yang dapat meningkatkan akurasi prediksi harga emas tanpa mengorbankan efisiensi pelatihan. Setiap kombinasi akan dievaluasi berdasarkan performa model pada data pelatihan dan pengujian yang telah disiapkan sebelumnya.

4.3 BAB V HASIL DAN PEMBAHASAN

Pada bab tiga ini akan menjelaskan hasil dari perancangan serta pembahasan mengenai hasil yang telah dicapai selama penelitian ini.

5.1 Hasil

Pada subbab ini, dipaparkan hasil dari penelitian yang dimulai dari pra-pemrosesan data seperti seleksi fitur, normalisasi, pembuatan dataset. Setelah pra-pemrosesan data telah dicapai, hasil analisa dengan menerapkan pemodelan dari metode LSTM akan dipaparkan.

5.1.1 Pra-Pemrosesan Data

Pada subbab ini akan memaparkan hasil dari seleksi fitur dengan melibatkan statistik deskriptif, visualisasi, uji korelasi hubungan antar variabel atau fitur terhadap target, kesimpulan dari seleksi fitur, normalisasi data, dan pembuatan dataset. Tabel 5.

1 Tabel statistik deskriptif emas 5.1.1.1. Statistik Deskriptif 5.1.1.1.1.

Statistik deskriptif harga emas Tabel 5.1 Tabel statistik deskriptif emas Harga Pembuka Harga Terendah Harga Tertinggi Harga Close Volume Harga Valid 2051,000 2051,000 2051,000 2051,000 2051,000 Median 1766,2 1759,400 1776,000 1767,40 187,00 Rerat a 1724,617 1716,623 1732,328 1724,359 5270,296 Std. Dev 403,046 400,091 405,624 402,747 29550,255 Skewness 0.693 0.695 0.686 0.692 7447,000 Kurtosis 0.139 0.136 0.129 0.138 61066,000 Minimum 1160,400 1146,500 1160,800 1155,200 0,000 Maksimum 2947,900 2933,100 2957,900 2949,100 386334,000 44

Berdasarkan tabel statistik deskriptif harga emas, ter lihat bahwa jumlah data valid untuk setiap variabel mencapai 2.051 dan tidak terdapat data yang hilang. Nilai median dari harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran 1.759 hingga 1.776, yang mencerminkan harga tengah dari data dalam rentang tersebut. Mean dari keempat harga tersebut juga berada pada kisaran yang relatif dekat, yaitu sekitar 1.716 hingga 1.732. Jika diperhatikan lebih lanjut, nilai mean sedikit lebih rendah dibandingkan median, yang secara umum dapat mengindikasikan distribusi miring ke kiri (negatively skewed). Namun, nilai skewness yang tercatat justru berada pada kisaran positif (antara 0,686 hingga 0,695), yang menandakan adanya kemiringan ke kanan. Ketidakesesuaian ini dapat terjadi karena bentuk distribusi yang tidak sepenuhnya simetris atau karena pengaruh nilai-nilai ekstrem atau outlier di sisi atas distribusi. Nilai standar deviasi yang berada di kisaran 400 untuk seluruh harga mengindikasikan adanya variasi harga yang cukup signifikan dari rata-ratanya. Sementara itu, nilai kurtosis yang sangat rendah (sekitar 0,13) menunjukkan bahwa distribusi data cenderung lebih mendatar, yang berarti kemungkinan terjadinya harga ekstrem atau outlier relatif rendah. Untuk volume harga, terjadi perbedaan yang sangat mencolok antara nilai median (187.000) dan rata-rata (5.270.296), serta nilai skewness dan kurtosis yang sangat tinggi masing-masing sebesar

7.447 dan 61.066. Hal ini menunjukkan bahwa distribusi volume harga sangat tidak simetris dan memiliki ekor panjang di sebelah kanan (positively skewed), mengindikasikan adanya beberapa hari perdagangan dengan volume transaksi yang sangat besar dibandingkan hari-hari lainnya.

Nilai minimum volume adalah 0, yang mengindikasikan adanya hari tanpa transaksi, sementara maksimum mencapai 386.334.000, menunjukkan lonjakan volume yang sangat besar pada titik tertentu.

5.1.1.1.2. Statistik

Deskriptif Harga Minyak Mentah Tabel 5. 2 Tabel statistik deskriptif minyak mentah

Volume	Volume	Volume	Volume	Volume	Volume
2051,000	2051,000	2051,000	2051,000	2051,000	2051,000

45	0,000	0,000	0,000	0,000	0,000	Media	67,540	66,410
----	-------	-------	-------	-------	-------	-------	--------	--------

68,520	67,500	405596,000	Rerat	66,109	64,956	67,248	66,146	457327,722
--------	--------	------------	-------	--------	--------	--------	--------	------------

S td.	Dev	17,702	17,419	17,931	17,661	225152,781	Skewness	0.073
-------	-----	--------	--------	--------	--------	------------	----------	-------

-0.060	0.254	0.135	0.943	Kurtosis	0.964	1,159	0.536	0.638	2611,000
--------	-------	-------	-------	----------	-------	-------	-------	-------	----------

Minimum	-37,630	-40,320	13,690	-14,000	0,000	Maksimum	123,700
---------	---------	---------	--------	---------	-------	----------	---------

120,790	130,500	124,660	2288,000
---------	---------	---------	----------

Berdasarkan tabel statistik deskriptif harga minyak mentah, diketahui bahwa seluruh variabel memiliki 2.051 data

valid tanpa adanya nilai yang hilang. Median harga pembuka, harga

terendah, harga tertinggi, dan harga penutupan berkisar antara 66,41

hingga 68,52, sedangkan nilai rata-rata dari keempat harga tersebut

berada sedikit di bawah median, yaitu antara 64,96 hingga 67,25. Hal

ini menunjukkan bahwa distribusi data secara umum cenderung sedikit miring

ke kiri (negatively skewed), meskipun nilai skewness yang tercatat sangat

kecil (berkisar antara -0,060 hingga 0,254) menunjukkan bahwa distribusi

harga relatif simetris atau hanya memiliki kemiringan yang sangat ringan.

Nilai standar deviasi sekitar 17 untuk seluruh harga menunjukkan

adanya variasi yang signifikan, meskipun tidak sebesar variabilitas pada

harga emas. Seluruh nilai kurtosis berada di bawah 3, termasuk harga

terendah yang tercatat sebesar 1,159, yang menunjukkan distribusi datar

dengan kemungkinan outlier ekstrem yang rendah. Namun, perhatian khusus

perlu diberikan pada nilai minimum negatif, seperti -40,32 untuk harga

terendah dan -37,63 untuk harga pembuka, yang secara logis tidak mungkin terjadi pada harga minyak mentah. Fenomena harga minyak mentah yang turun hingga ke nilai negatif ini tidak selalu mencerminkan harga fisik negatif, melainkan lebih berkaitan dengan mekanisme pasar berjangka dan kondisi ekstrem mendasar. Berdasarkan artikel Energy Information Administration (EIA) dari AS yang ditulis oleh (French, 2021), pada April 2020 terjadi kondisi pasar yang luar biasa: permintaan minyak turun drastis akibat pembatasan mobilitas selama pandemi COVID-19, sementara kapasitas penyimpanan sudah penuh. Akibatnya, kontrak berjangka minyak seperti West Texas Intermediate (WTI) sempat diperdagangkan di bawah nol bahkan sempat mencapai -37,63 USD per barel. Hal ini terjadi karena para pemegang kontrak yang hampir jatuh tempo tidak ingin mengambil pengiriman fisik jika tidak ada tempat untuk menyimpannya, sehingga mereka bersedia membayar agar kontrak tersebut diambil alih orang lain. Pada titik itu, biaya penyimpanan dan potensi biaya shutdown produksi menjadi lebih mahal dibandingkan menjual kontrak berjangkanya dengan harga negatif. Dengan demikian, penurunan harga hingga negatif dalam data deskriptif kemungkinan bersumber dari perilaku pasar kontrak berjangka yang terjadi dalam periode ekstrem tersebut, bukan perwakilan harga riil minyak fisik. Oleh karena itu, data ini harus ditandai sebagai anomali namun valid. Untuk volume harga, median tercatat sebesar 405.596, sementara rata-ratanya lebih tinggi yaitu 457.327, mengindikasikan sedikit kemiringan ke kanan (positively skewed), yang juga didukung oleh nilai skewness sebesar 0,943. Nilai kurtosis volume yang sangat tinggi (2611) menunjukkan adanya outlier signifikan, yaitu hari-hari dengan volume transaksi yang jauh lebih besar dibandingkan hari-hari lainnya. Nilai minimum volume adalah 0, dan maksimum mencapai 2.288.000, menunjukkan rentang yang sangat lebar.

5.1.1.1.3. Statistik Deskriptif Indeks Harga S&P 500

Statistik	Nilai
Pembuka	2051,000
Terendah	-37,63
Tertinggi	2051,000
Close	2051,000
Volume	405.596

2051,000 2 051,000 2051,000 M issin g 0,000 0,000 0,000 ,000 0,000
M edia n 3727,750 3686,250 3774,750 3 725,000 1475,000 R erat a
3727,765 3698,823 3751,914 3 725,848 1553,000 S td. De v 1028,804
1023,365 1033,926 1 029,119 559545,450 S kewnes s 0.513 0.521 0.503
.512 1195,000 K urtosi s -0.678 -0.671 -0.689 -0.678 4037,000 M inim u
m 2220,500 2174,000 2259,500 2 220,250 0,000 47 M aksimu m 6163,000
6129,250 6166,500 6 153,750 5715,000 Berdasarkan tabel statis tik
deskriptif indeks S&P 500, terlihat bahwa jumlah data valid untuk setiap
variabel mencapai 2.051 da n tidak terdapat data yang hilang (missing).
Nilai median dari harga pembuka, harga terendah, harga tertinggi, dan
harga penutupan berada pada kisaran 3.686 hingga 3.775 , yang
mencerminkan posisi tengah dari distribusi data. Rata-rata (mean) dari
keem pat harga tersebut juga berada dalam kisaran yang sangat dekat,
yaitu sekitar 3.698 hingga 3.751. Jik a diperhatikan lebih lanjut, nilai
mean pada semua ka tegori harga cenderung sedikit lebih tinggi
dibandingk an med ian. Hal ini secara umum mengindikasikan adanya
distribusi miring ke kanan (positively skewed). Kondisi ini juga
dikonfirmasi oleh nilai skewness yang seluruhnya berada di atas nol
(sekitar 0,503 hingga 0,521), yang men unjukkan adanya kemiringan ke
kanan, artinya terdapat sebagian kecil nilai-nilai harga yang lebih
tinggi dari rata- rata, sehingga menarik ekor distribusi ke arah kanan.
Nilai st andar deviasi yang berada di kisaran 1.023 hingga 1.033 me
nunjukkan adanya variabilita s harga yang cukup besar, mencerminkan
fluktuasi harga indeks S&P 500 dalam periode pengamatan yang signi
fikan. Meskipun besar, variasi ini dapat dianggap wajar mengingat tingkat
harga yang memang berada dalam ribuan poin. Nilai kurtosis untuk
seluruh kategori harga adalah negatif (sekitar -0 ,67 hingga -0,68), ya
ng menunjukkan bahwa distribusi data cenderung lebih mendatar dibandingkan
distribusi norm al. Artinya, data harga tidak memiliki puncak yang tajam
dan probabilitas terjadinya nilai ekstrem atau outlier cenderung rendah.
Untuk volume harga, median tercatat sebesar 1.475 dan mean sed ikit

lebih tinggi sebesar 1.553, yang menunjukkan kemiringan positif yang ringan, sesuai dengan nilai skewness sebesar 1.195. Sementara itu, kurtosis volume sangat tinggi, yakni mencapai 4.037, yang mengindikasikan adanya outlier atau lonjakan volume ekstrem dalam beberapa hari perdagangan. Hal ini diperkuat dengan perbedaan besar antara nilai minimum (0) dan maksimum (5.715.000), menunjukkan adanya hari-hari tanpa transaksi sama sekali, serta hari-hari dengan volume transaksi yang sangat tinggi.

4.8.5.1.1.1.4. Statistik Deskriptif Harga Perak

Tabel	5	4	Tabel
statistik deskriptif harga perak	Harga Pembuka	Harga Terendah	Harga Tertinggi
	Volume	Harga	V ali d
2051,000	2051,000	2051,000	2051,000
2051,000	2051,000	M issin g	0,000 0,000 0,000 0,000 0,000
21,533	21,355	21,675	21,510 45,000
R erat a	21,220	21,087	21,364
21,231	1821,461	S td. De v	5,020 4 ,961 5,078 5,021 10554,935
S kewnes s	0.417 .418 0.418 0.421	7596,000	K urtosi s -0.844 -0.834
-0.849 -0.838	63080,000	M inimu m	11,735 11,735 12,205 12,070 ,000
M aksimu m	34,831 34,425 34,835 34,555	131415,000	Berdasarkan tabel statistik deskriptif harga perak, seluruh variabel memiliki jumlah data valid sebanyak 2.051 dan tidak terdapat data yang hilang.

Median dari harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran 21,355 hingga 21,675, yang mencerminkan nilai tengah dari distribusi harga. Sementara itu, nilai rata-rata (mean) sedikit lebih rendah dibandingkan median, yaitu antara 21,087 hingga 21,364. Pola ini biasanya mengindikasikan adanya kemiringan ke kiri (negatively skewed), namun dalam tabel terlihat bahwa nilai skewness untuk keempat harga berada di kisaran positif (sekitar 0,417 hingga 0,421), yang berarti secara statistik, distribusi memiliki kemiringan ke kanan (positively skewed). Ini menunjukkan bahwa meskipun secara visual tampak seimbang antara mean dan median, masih terdapat beberapa nilai tinggi yang mendorong ekor distribusi ke kanan. Nilai standar deviasi berada di kisaran 4,961 hingga 5,078, yang menunjukkan adanya tingkat variasi harga yang cukup moderat jika dibandingkan dengan nilai

rata-ratanya . Ini berarti harga perak relatif stabil dalam kisaran yang 49 tidak terlalu ekstrem, setidaknya dalam periode pengamatan. Nilai kurtosis seluruhnya bernilai negatif (sekitar $-0,83$ hingga $-0,84$), menunjukkan bahwa distribusi harga cenderung lebih mendatar dibandingkan distribusi normal. Artinya, kemungkinan terjadinya nilai ekstrem atau outlier relatif rendah dan distribusinya memiliki puncak yang lebar. Sementara itu, untuk volume harga, terdapat perbedaan yang sangat mencolok antara median (45.000) dan rata-rata (1.821.461), serta nilai skewness dan kurtosis yang sangat tinggi masing-masing sebesar 7.596 dan 63.080. Hal ini menunjukkan bahwa distribusi volume perdagangan sangat tidak simetris dan memiliki ekor panjang ke kanan (positively skewed), yang menunjukkan adanya beberapa hari dengan volume transaksi yang jauh lebih besar dibandingkan hari-hari lainnya. Nilai minimum volume adalah 0, menunjukkan kemungkinan hari tanpa transaksi, sedangkan nilai maksimum mencapai 131.415.000, yang memperkuat indikasi adanya lonjakan besar dalam volume di waktu tertentu.

5.1.1.1.5. Statistik deskriptif indeks dolar AS

Tabel 5.5 Tabel statistik deskriptif indeks harga dolar AS

Pembuka	Harga Terendah	Harga Tertinggi	Harga Close	Volume	Harga Valid
2051,000	2051,000	2051,000	2051,000	2051,000	Missing 0,000
0,000	0,000	0,000	0,000	0,000	0,000
Median	97,469	97,230	97,720	97,490	19691,000
Rerata	98,509	98,207	98,804	98,503	21945,578
Std. Dev	5,405	5,338	5,457	5,400	10879,227
Skewness	0.308	0.286	0.330	0.308	1850,000
Kurtosis	-0.723	-0.753	-0.683	-0.716	7384,000
Minimum	88,505	88,150	88,890	88,490	0,000
Maksimum	114,047	113,260	114,745	114,105	113489,000

Berdasarkan tabel statistik deskriptif indeks dolar AS (USD_X), seluruh variabel memiliki jumlah data valid sebanyak 2.051 dan tidak terdapat data yang hilang, menunjukkan kelengkapan data yang baik. Nilai median untuk harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran 50 97,230 hingga 97,720, sementara rata-rata (mean) dari keempat harga tersebut berada pada kisaran 98,207 hingga 98,804. Karena rata-rata lebih tinggi daripada median di seluruh

variabel harga, hal ini menunjukkan bahwa distribusi data cenderung miring ke kanan (positively skewed). Informasi ini diperkuat oleh nilai skewness positif di semua harga (sekitar 0,286 hingga 0,330), yang menandakan adanya ekor distribusi yang lebih panjang di sisi kanan. Ini berarti ada sejumlah nilai tinggi yang mungkin menjadi outlier atau hari dengan pengamatan nilai indeks secara signifikan. Nilai standar deviasi berada di kisaran 5,338 hingga 5,457, yang menunjukkan tingkat variasi harga yang relatif moderat dalam konteks nilai indeks dolar. Artinya, meskipun terdapat fluktuasi, indeks ini tergolong cukup stabil dalam periode yang dianalisis. Nilai kurtosis berada pada kisaran negatif (sekitar -0,683 hingga -0,753), menunjukkan bahwa distribusi data lebih mendatar dari distribusi normal. Ini mengindikasikan bahwa data harga memiliki kemungkinan lebih kecil terhadap outlier dan lebih tersebar merata di sekitar rata-ratanya. Untuk volume harga, terdapat perbedaan mencolok antara median (19.691) dan rata-rata (21.945,578), disertai nilai skewness dan kurtosis yang sangat tinggi masing-masing sebesar 1.850 dan 7.384. Hal ini menunjukkan bahwa distribusi volume perdagangan sangat tidak simetris dan memiliki ekor panjang di kanan (positively skewed), mengindikasikan adanya hari-hari tertentu dengan volume transaksi yang jauh lebih tinggi dibandingkan rata-rata. Nilai minimum volume adalah 0, yang bisa menunjukkan tidak adanya transaksi pada hari tertentu, sementara nilai maksimumnya mencapai 113.489, mengindikasikan lonjakan volume yang sangat tinggi.

5.1.1.2. Visualisasi Plotting

Untuk memahami karakteristik distribusi data secara visual, dilakukan eksplorasi menggunakan tiga jenis metode visualisasi, yaitu distribution plot, boxplot, dan scatter plot. Ketiga metode ini dipilih karena mampu memberikan gambaran menyeluruh mengenai pola penyebaran, potensi outlier, dan hubungan antar variabel secara intuitif. 51 Dalam visualisasi ini, digunakan harga penutupan sebagai representasi utama dari masing-masing komoditas dan indeks, yaitu harga emas, harga perak, harga minyak mentah, indeks S&P 500, dan indeks

dolar AS (USD). 5.1.1.2.1. Distribution Plot atau Grafik Distribusi

Gambar 5. 1 Gambar visualisasi distribution plot harga penutupan emas Berdasarkan grafik distribusi harga emas_Close menunjukkan pola distribusi yang multimodal, dengan beberapa puncak yang terlihat pada kisaran harga sekitar 1250, 1850, dan indikasi lainnya di atas 2500. Pola ini selaras dengan hasil statistik deskriptif, yang menunjukkan bahwa nilai median harga emas berada di kisaran 1767,400, sementara mean sedikit lebih rendah, yaitu sekitar 1724,359. Perbedaan antara mean dan median ini umumnya mengarah pada interpretasi distribusi yang miring ke kiri (negatively skewed), namun nilai skewness positif yang tercatat 0.692 justru mengonfirmasi bahwa distribusi miring ke kanan (positively skewed). Hal ini juga tampak jelas dalam plot distribusi, di mana terdapat ekor panjang di sisi kanan (harga di atas 2000), yang mengindikasikan adanya nilai-nilai ekstrem atau outlier atau fase lonjakan harga yang signifikan. Rentang harga yang tercakup dalam distribusi, yakni dari 1155,2 dan 2949,10, mencerminkan tingkat volatilitas yang tinggi, yang juga diperkuat oleh nilai standar deviasi yang cukup besar 402,75 dalam analisis deskriptif. Selain itu, kepadatan data tertinggi yang muncul di kisaran 1250 sampai 1350 dan 1800 sampai 1900 mencerminkan dua fase harga emas yang paling sering terjadi, yang kemungkinan besar berkaitan dengan dua periode historis dengan karakteristik pasar berbeda. Sementara itu, nilai kurtosis yang rendah sekitar 0,13, menandakan bahwa distribusi data cenderung lebih mendatar dibanding distribusi normal, sehingga kemungkinan terjadinya lonjakan ekstrem sebenarnya relatif rendah, meskipun tetap terlihat pada sebagian kecil data.

Gambar 5. 2 Gambar visualisasi distribution plot harga penutupan minyak mentah Berdasarkan grafik distribusi harga penutupan minyak mentah atau minyak_close yang mencakup histogram dengan tumpang-tindih kepadatan (density plot), terlihat bahwa distribusi memiliki pola unimodal dengan satu puncak utama di sekitar harga 75,00. Kepadatan harga tertinggi terjadi pada kisaran ini, yang menunjukkan bahwa harga minyak paling se

ring berada dalam rentang 53 tersebut selama periode analisis. Ini selaras dengan statistika deskriptif, di mana median harga tercatat sebesar 67,50, sementara rata-rata sedikit lebih rendah di 66,15, menandakan bahwa sebagian besar nilai harga berpusat pada rentang tersebut dengan beberapa ekstrem yang menarik distribusi. Kepadatan distribusi juga memperkuat temuan mengenai kemiringan data yang ringan ke kiri (negatively skewed), meskipun nilai skewness yang sangat kecil (0,135) menunjukkan bahwa distribusi relatif simetris. Dalam density plot, terdapat ekor panjang di sisi kiri, yang menunjukkan adanya sejumlah harga ekstrem di bawah 0, yang juga tercermin dalam nilai minimum yang tercatat -14,00. Keberadaan ekor ini menunjukkan bahwa walaupun sebagian besar data terkonsentrasi di sekitar harga 75, beberapa nilai rendah tetap berkontribusi pada persebaran distribusi. Volatilitas harga minyak juga terlihat dalam kepadatan distribusi, di mana rentang harga -14 hingga 124,6 memperlihatkan variasi yang cukup signifikan, diperkuat oleh nilai standar deviasi sebesar 17,66. Grafik distribusi menunjukkan bentuk distribusi yang lebih mendatar dibandingkan distribusi normal, sesuai dengan nilai kurtosis rendah di nilai 1,16, yang mengindikasikan bahwa probabilitas munculnya lonjakan harga ekstrem tetap ada, tetapi tidak dominan. Secara keseluruhan, kepadatan distribusi memberikan gambaran tambahan bahwa harga minyak memiliki pola persebaran yang cukup stabil tetapi tetap menunjukkan beberapa titik ekstrem yang memengaruhi pergerakan harga secara keseluruhan.

54 Gambar 5.3 Gambar visualisasi distribution plot harga penutupan indeks S&P 500 Berdasarkan grafik distribusi harga penutupan indeks S&P 500, distribusi menunjukkan pola unimodal dengan satu puncak utama yang berada di kisaran harga sekitar 2500 sampai 4250, di mana kepadatan tertinggi terjadi. Kepadatan yang tumpang-tindih pada grafik distribusi memperlihatkan bahwa mayoritas nilai harga terkonsentrasi di sekitar puncak ini, sedangkan ekor panjang di sisi kanan menunjukkan adanya sejumlah harga tinggi yang lebih jarang terjadi namun tetap memengaruhi

persebaran distribusi. Kepadatan distribusi memperkuat interpretasi bahwa indeks S&P 500 memiliki kemiringan ke kanan (positively skewed), dengan nilai skewness positif antara 0.512 yang menunjukkan bahwa sebagian kecil harga yang lebih tinggi dari rata-rata menarik ekor distribusi ke arah kanan. Nilai median harga penutupan 3725,00, yang sedikit lebih rendah dibandingkan rata-rata 3725,85, juga menjadi indikasi bahwa distribusi harga lebih berat di sisi kiri dibandingkan kanan. Selain itu, nilai standar deviasi 1029,12 mengonfirmasi bahwa indeks ini mengalami variabilitas harga yang cukup besar. Namun, tingkat variabilitas ini tetap dalam batas wajar untuk indeks pasar yang memang bernilai tinggi. Kepadatan distribusi juga menunjukkan bahwa probabilitas terjadinya lonjakan harga ekstrem relatif rendah, sejalan dengan nilai kurtosis negatif $-0,678$, yang mengindikasikan bahwa bentuk distribusi lebih mendatar dibandingkan distribusi normal. Secara keseluruhan, kepadatan distribusi mempertegas pola pergerakan indeks S&P 500 yang stabil namun tetap mengalami beberapa harga tinggi yang memengaruhi kemiringan distribusi ke kanan.

Gambar 5.4 Gambar visualisasi distribution plot harga perak Berdasarkan grafik distribusi harga penutupan perak, pola distribusi menunjukkan karakteristik bimodal, dengan dua puncak utama yang terlihat di kisaran 17 dan 24, yang mencerminkan adanya dua kelompok harga dominan dalam periode pengamatan. Kepadatan tertinggi terjadi di sekitar kedua puncak ini, menandakan bahwa harga perak paling sering berada dalam dua rentang yang berbeda, yang mungkin dipengaruhi oleh kondisi pasar tertentu dalam waktu yang berbeda. Distribusi ini selaras dengan hasil statistik deskriptif, di mana median harga penutupan tercatat di 21,51, sementara nilai rata-rata sedikit lebih rendah, yaitu sekitar 21,23. Perbedaan kecil antara mean dan median ini biasanya mengarah pada interpretasi distribusi miring ke kiri (negatively skewed), tetapi nilai skewness positif (sekitar 0,417 hingga 0,421) mengonfirmasi bahwa distribusi sebenarnya miring ke kanan (positively skewed).

56 Kepadatan yang tumpang-tindih pada grafik distribusi

memperjelas bahwa meskipun mayoritas harga terkonsentrasi di sekitar dua puncak utama, terdapat ekor panjang di sisi kanan, yang menunjukkan beberapa harga ekstrem yang lebih tinggi dari rata-rata. Ini mengindikasikan adanya peristiwa tertentu yang menyebabkan lonjakan harga di luar rentang normalnya. Selain itu, nilai standar deviasi sekitar 5,02 menunjukkan bahwa variasi harga cukup moderat, mengisyaratkan stabilitas relatif dalam pergerakan harga perak, meskipun tetap menunjukkan fluktuasi dalam kisaran yang signifikan. Nilai kurtosis negatif -0,838 mempertegas bahwa distribusi data relatif mendatar dibandingkan distribusi normal, sehingga probabilitas terjadinya lonjakan harga ekstrem lebih rendah dan harga perak cenderung memiliki distribusi yang lebih merata tanpa puncak tajam. Secara keseluruhan, kepadatan distribusi semakin memperjelas pola bimodal dengan kecenderungan miring ke kanan, serta menunjukkan dua kelompok harga yang paling sering muncul. Fluktuasi harga cukup stabil tetapi tetap menunjukkan adanya periode tertentu di mana harga perak mengalami pergerakan signifikan, terutama di sisi harga yang lebih tinggi.

57 Gambar 5.5 Gambar visualisasi distribution plot harga indeks dolar AS Berdasarkan grafik distribusi harga penutupan indeks dolar AS, pola distribusi menunjukkan kemiringan ke kanan (positively skewed), dengan ekor panjang di sisi kanan yang mengindikasikan adanya beberapa nilai tinggi yang lebih jarang terjadi tetapi signifikan. Grafik kepadatan memperlihatkan pola distribusi yang lebih halus dibandingkan histogram, dengan puncak utama yang menunjukkan frekuensi tertinggi dari pergerakan harga indeks. Kepadatan tertinggi tampak pada kisaran 95–100. Distribusi ini sejalan dengan hasil statistik deskriptif, di mana median harga penutupan berada di 97,49, sedangkan rata-rata lebih tinggi, yakni 98,50, yang mengonfirmasi adanya kemiringan ke kanan. Grafik kepadatan juga memperjelas keberadaan ekor panjang di sisi kanan, yang menunjukkan beberapa nilai harga ekstrem yang lebih tinggi dari rata-rata dan dapat dianggap sebagai outlier, yang mendorong distribusi ke arah kanan. Nilai skewness positif sebesar 0,30

8 semakin mempertegas bahwa distribusi lebih berat di sisi kiri, dengan sejumlah harga tinggi menarik ekor distribusi ke kanan. Dalam grafik kepadatan, bentuk yang tidak sepenuhnya simetris semakin menguatkan karakteristik ini. Sementara itu, nilai standar deviasi 5,40 menunjukkan 58 bahwa variabilitas harga cukup moderat dibandingkan dengan tingkat harga indeks, menandakan bahwa fluktuasi harga masih dalam rentang yang wajar. Selain itu, nilai kurtosis negatif (-0,716) mengindikasikan bahwa distribusi data lebih mendatar dibandingkan distribusi normal, sehingga probabilitas terjadinya lonjakan harga yang sangat ekstrem tetap rendah. Secara keseluruhan, grafik kepadatan semakin menegaskan pola distribusi indeks dolar AS yang menunjukkan kemiringan ke kanan dengan stabilitas relatif, di mana beberapa nilai harga tinggi menjadi faktor utama dalam pergerakan indeks. Bentuk distribusi yang terstruktur ini mencerminkan fluktuasi yang terjadi dalam periode pengamatan tanpa indikasi perubahan yang sangat tajam.

5.1.1.2.2. Boxplot Atau Grafik Kotak-Garis Gambar 5.6

Gambar visualisasi boxplot pada harga penutupan emas memberikan wawasan visual mengenai distribusi harga yang telah dianalisis sebelumnya. Median harga emas tercatat berada di kisaran 1767,4, dengan rata-rata sedikit lebih rendah sekitar 1724,359, yang biasanya menunjukkan distribusi miring ke kiri. Namun, dalam analisis statistik, nilai skewness positif di 0.692 justru mengonfirmasi bahwa distribusi ini sebenarnya miring ke kanan (positively skewed), dengan ekor panjang yang terlihat di sisi kanan. Keberadaan outlier di atas 2000 dalam boxplot semakin memperkuat karakteristik skewness positif yang telah diidentifikasi dalam distribusi harga. Outlier ini mencerminkan fase lonjakan harga yang signifikan, di mana beberapa periode mengalami kenaikan harga yang lebih tinggi dibandingkan mayoritas data lainnya. Selain itu, whisker yang cukup panjang terutama ke arah atas menegaskan adanya fluktuasi harga yang cukup tinggi. Dengan rentang harga yang luas, yaitu antara 1155,20 dan 2949,10, volatilitas harga emas tampak jelas dan diperkuat dengan nilai standar deviasi

yang cukup besar sekitar 402,75. Ini menunjukkan bahwa pergerakan harga dalam periode pengamatan cukup dinamis dan mengalami perubahan signifikan dari waktu ke waktu. Kepadatan data tertinggi yang ditemukan di kisaran 1250 sampai 1350 dan 1800 sampai 1900 dalam distribusi menandakan bahwa terdapat dua fase harga yang paling sering terjadi, yang kemungkinan besar berhubungan dengan peristiwa historis tertentu yang mempengaruhi pasar emas. Sementara itu, nilai kurtosis yang rendah (sekitar 0,13) menunjukkan bahwa distribusi data lebih mendatar dibandingkan distribusi normal, sehingga probabilitas lonjakan harga yang sangat ekstrem sebenarnya relatif rendah, meskipun tetap terlihat dalam beberapa outlier yang berada di sisi kanan distribusi. Secara keseluruhan, boxplot ini mengonfirmasi hasil analisis statistik sebelumnya dengan menunjukkan pola distribusi yang luas, volatilitas tinggi, dan kecenderungan harga ekstrem yang berada di sisi kanan distribusi. Gambar 5.7 Gambar visualisasi boxplot pada harga penutupan minyak mentah. Boxplot harga penutupan minyak mentah memberikan gambaran visual yang memperkuat analisis distribusi sebelumnya, dengan menunjukkan persebaran harga yang relatif luas serta beberapa nilai ekstrem yang signifikan. Median harga berada di sekitar 70, yang selaras dengan temuan sebelumnya bahwa harga minyak paling sering berada dalam kisaran sekitar 67,50–75,00. Posisi median dalam boxplot terlihat lebih dekat ke bagian bawah kotak, yang menandakan bahwa sebagian besar harga cenderung berada di rentang bawah, meskipun ada beberapa nilai tinggi yang memperpanjang distribusi ke sisi kanan. Rentang IQR dalam boxplot menunjukkan bahwa mayoritas harga minyak terkonsentrasi antara 60 hingga 80, mencerminkan variabilitas yang cukup stabil dalam periode analisis. Namun, garis whisker yang panjang ke arah bawah serta keberadaan outlier di bawah semakin mempertegas bahwa terdapat beberapa harga ekstrem yang mendorong distribusi ke sisi kiri. Hal ini konsisten dengan distribusi sebelumnya, di mana ekor panjang ke kiri terlihat dalam plot kepadatan, menunjukkan bahwa walaupun sebagian besar harga berada di kisaran stabil, terdapat beberapa

titik harga yang berada dalam rentang sangat rendah. Di sisi lain, outlier di atas 120 memperkuat identifikasi bahwa meskipun distribusi secara umum cukup stabil, terdapat sejumlah lonjakan harga ekstrem yang terjadi pada periode tertentu. Ini selaras dengan nilai standar deviasi sebesar 17,66, yang menunjukkan bahwa harga minyak mengalami fluktuasi cukup besar, tetapi tetap dalam batas yang wajar. Selain itu, boxplot mengonfirmasi bahwa distribusi harga lebih mendasar dibandingkan distribusi normal, sebagaimana tercermin dalam nilai kurtosis yang rendah (sekitar 1,16), menandakan bahwa probabilitas munculnya lonjakan harga ekstrem tetap ada, tetapi tidak dominan. Secara keseluruhan, boxplot harga minyak mentah menegaskan hasil analisis distribusi sebelumnya, dengan menunjukkan skewness ringan ke kiri, volatilitas harga yang cukup tinggi, dan keberadaan beberapa outlier ekstrem. Bentuk distribusi ini mencerminkan pola harga minyak yang cenderung stabil dalam rentang tertentu, tetapi tetap menunjukkan beberapa fluktuasi signifikan yang dapat mempengaruhi tren pasar. Gambar 5.8 Gambar visualisasi boxplot pada harga penutupan indeks S&P 500 Berdasarkan boxplot harga penutupan indeks S&P 500, semakin memperkuat analisis distribusi sebelumnya dengan memberikan visualisasi yang jelas mengenai persebaran harga. Median harga berada di sekitar 4000, yang sesuai dengan pola distribusi sebelumnya yang menunjukkan bahwa mayoritas harga 62 indeks berkisar antara 2500 hingga 4250, di mana kepadatan tertinggi terjadi. Kotak dalam boxplot menggambarkan rentang IQR, yakni antara 3000 dan 5000, menunjukkan bahwa sebagian besar harga indeks terkonsentrasi di dalam rentang ini, meskipun ada beberapa nilai ekstrem yang memperpanjang distribusi ke arah luar. Keberadaan outlier di bawah 2000 semakin menegaskan adanya harga ekstrem yang lebih rendah dari batas normal. Outlier ini sesuai dengan ekor panjang yang terlihat dalam distribusi, yang mencerminkan periode volatilitas yang lebih tinggi dalam harga indeks S&P 500. Namun, di sisi lain, boxplot juga menunjukkan bahwa harga tertinggi yang bukan outlier mencapai sekitar 6000, yang mengonfirmasi bahwa meskipun ada

fluktuasi harga yang signifikan, sebagian besar pergerakan tetap dalam kisaran yang terstruktur. Struktur boxplot juga mendukung temuan mengenai kemiringan ke kanan (positively skewed), seperti yang telah dianalisis dalam distribusi sebelumnya, dengan nilai skewness positif sebesar 0.512, yang menunjukkan adanya sebagian kecil harga yang lebih tinggi dari rata-rata yang menarik ekor distribusi ke arah kanan. Standar deviasi sebesar 1029,12 semakin mempertegas adanya variabilitas harga yang cukup besar, namun tetap dalam batas wajar untuk indeks saham bernilai tinggi seperti S&P 500. Selain itu, kurtosis negatif sebesar $-0,67$

8 mengindikasikan bahwa distribusi harga lebih mendatar dibandingkan distribusi normal, menandakan bahwa probabilitas terjadinya lonjakan harga ekstrem lebih kecil. Secara keseluruhan, boxplot memberikan perspektif tambahan terhadap distribusi harga penutupan indeks S&P 500, dengan mengonfirmasi bahwa meskipun terdapat beberapa outlier yang mendorong distribusi, indeks ini tetap menunjukkan pola yang stabil dengan persebaran harga yang cukup terstruktur dalam IQR utama. Hal ini semakin mempertegas bahwa pergerakan harga indeks S&P 50 memiliki karakteristik volatilitas yang tinggi tetapi tetap dalam pola distribusi yang dapat diprediksi. 63 Gambar 5. 9 Gambar visualisasi boxplot pada harga penutupan perak Berdasarkan boxplot, harga penutupan perak memberikan gambaran visual yang memperkuat hasil analisis distribusi sebelumnya. Hal ini menegaskan pola bimodal dengan kecenderungan miring ke kanan (positively skewed) serta keberadaan beberapa harga ekstrem. Median harga berada di sekitar 21,51, yang sesuai dengan kepadatan distribusi sebelumnya, di mana sebagian besar harga terkonsentrasi di dua puncak utama di kisaran 17 dan 24. Posisi median yang lebih dekat ke bagian bawah kotak dalam boxplot menunjukkan bahwa sebagian besar harga berada dalam rentang yang lebih rendah, meskipun terdapat sejumlah harga tinggi yang memperpanjang distribusi ke sisi kanan. Dari Boxplot diperlihatkan IQR yang cukup lebar, mengindikasikan bahwa mayoritas harga perak berada dalam kisaran 20 hingga 25, namun dengan beberapa

harga ekstrem yang muncul di luar batas ini. Keberadaan whisker yang panjang di sisi atas mengonfirmasi adanya harga tinggi yang terjadi dalam beberapa periode tertentu, yang juga tercermin dalam distribusi kepadatan yang menunjukkan ekor panjang di sisi kanan. Di sisi lain, terdapat outlier di bawah 10, yang memperkuat bukti bahwa terdapat nilai-nilai harga yang jauh lebih rendah dari sebaran utama, meskipun kejadian ini tergolong jarang. Selain itu, nilai skewness positif 0.421 yang terlihat dalam analisis deskriptif juga tercermin dalam boxplot, dengan persebaran harga yang lebih berat di sisi bawah namun tetap memiliki beberapa harga tinggi yang menarik distribusi ke kanan. Standar deviasi sekitar 5,02 menunjukkan variasi harga yang cukup moderat tetapi tetap memberikan fluktuasi signifikan. Kurtosis negatif -0,838 semakin mempertegas bahwa distribusi lebih mendatar dibandingkan distribusi normal, sehingga probabilitas terjadinya lonjakan harga yang sangat ekstrem tetap rendah. Gambar 5. 10 Gambar visualisasi boxplot pada harga penutupan indeks dolar AS Berdasarkan boxplot, harga penutupan indeks dolar AS memberikan gambaran visual yang mempertegas hasil analisis distribusi sebelumnya, menunjukkan kemiringan ke kanan (positively skewed) dengan persebaran data yang cukup stabil tetapi tetap memperlihatkan beberapa outlier di sisi kanan distribusi. Median harga berada di sekitar 100, yang sesuai dengan hasil distribusi sebelumnya yang menunjukkan kepadatan tertinggi pada kisaran 95–100, mengindikasikan bahwa harga indeks dolar paling sering muncul dalam rentang ini. Rentang IQR dalam boxplot menunjukkan bahwa mayoritas harga indeks terkonsentrasi di antara 95 dan 105, mencerminkan variabilitas yang moderat dalam pergerakan harga. Sementara itu, whisker yang lebih panjang di sisi atas menunjukkan adanya harga ekstrem yang lebih tinggi dari batas normal distribusi, yang sesuai dengan karakteristik ekor panjang di sisi kanan yang terlihat dalam grafik kepadatan. Keberadaan outlier di bawah 90 semakin memperjelas bahwa meskipun sebagian besar harga terkonsentrasi di sekitar median, terdapat

beberapa nilai rendah yang relatif jarang terjadi tetapi tetap berpengaruh pada persebaran data. Selain itu, nilai skewness positif sebesar 0,308 yang terlihat dalam analisis distribusi juga tercermin dalam boxplot, dengan persebaran harga yang lebih berat di sisi bawah tetapi tetap memiliki beberapa harga tinggi yang menarik distribusi ke kanan. Standar deviasi sekitar 5,40 semakin memperjelas adanya variabilitas harga yang cukup moderat, mencerminkan pergerakan indeks dolar yang cukup stabil tetapi tetap mengalami beberapa lonjakan harga dalam periode tertentu. Kurtosis negatif (-0,716) menunjukkan bahwa distribusi lebih mendatar dibandingkan distribusi normal, yang mengindikasikan bahwa probabilitas munculnya lonjakan harga yang sangat ekstrem tetap rendah.

5.1.1.2.3. Scatter Plot Atau Grafik Sebar 66 Gambar 5. 11 Gambar visualisasi scatter plot persebaran harga emas terhadap harga minyak mentah. Scatter plot yang ditampilkan menunjukkan hubungan antara harga penutupan minyak mentah dengan harga penutupan emas. Pada grafik ini, harga penutupan emas berada pada sumbu X, sementara harga penutupan minyak ditempatkan pada sumbu Y. Titik-titik data yang tersebar mengindikasikan bagaimana pergerakan harga kedua komoditas saling berkaitan. Dari distribusi titik-titik, terlihat bahwa terdapat kecenderungan positif, di mana peningkatan harga emas sering kali diikuti oleh kenaikan harga minyak. Meskipun pola ini tidak sepenuhnya linier, keberadaan garis regresi linier berwarna biru memberikan indikasi bahwa ada hubungan proporsional antara kedua variabel. Selain itu, grafik juga menyertakan density plot di bagian atas dan kanan, yang menggambarkan penyebaran data masing-masing variabel. Ini memperjelas pola distribusi harga, yang mungkin menunjukkan adanya konsentrasi data di rentang tertentu. Hal ini terbukti dari analisis distribusi dan analisis dengan boxplot yang memiliki kepadatan pada rentang-rentang masing-masing harga. Dengan demikian, scatter plot ini berfungsi sebagai alat eksplorasi awal untuk memahami korelasi harga emas dan minyak, serta memberikan wawasan awal sebelum dilakukan uji korelasi. emas_Close - sp500_Close Gambar 5. 12

Gambar visualisasi scatter plot persebaran harga emas terhadap indeks S&P 500 Scatter plot yang ditampilkan menggambarkan hubungan antara indeks penutupan S&P 500 dengan harga penutupan emas. Harga emas ditampilkan pada sumbu X, sedangkan indeks S&P 500 berada pada sumbu Y. Distribusi titik-titik data menunjukkan pola yang cenderung membentuk hubungan positif, di mana 68 kenaikan harga emas sering kali sejalan dengan peningkatan indeks S&P 500. Garis regresi linier berwarna biru memberikan indikasi bahwa kedua variabel memiliki tren yang bergerak bersama dalam rentang tertentu. Selain itu, density plot di bagian atas dan kanan scatter plot menampilkan distribusi data masing-masing variabel, membantu dalam memahami pola penyebaran harga dan indeks. Pola ini dapat menjadi dasar analisis lebih lanjut untuk memahami bagaimana pergerakan pasar ekuitas dan harga emas berinteraksi dalam konteks ekonomi dan investasi.

emas_Close - perak_Close Gambar 5. 13 Gambar visualisasi scatter plot persebaran harga emas terhadap harga perak Scatter plot yang ditampilkan menunjukkan hubungan antara harga penutupan perak dengan harga penutupan emas. Harga emas berada pada sumbu X, sementara harga perak ditempatkan pada sumbu Y, dengan titik-titik data yang 69 mewakili pasangan nilai dari kedua variabel pada waktu yang sama. Dari pola distribusi titik-titik tersebut, terlihat adanya hubungan positif, di mana kenaikan harga emas cenderung diikuti oleh kenaikan harga perak. Keberadaan garis regresi linier berwarna biru dalam scatter plot memberikan indikasi bahwa hubungan antara harga emas dan perak memiliki tren yang relatif proporsional dalam rentang tertentu. Meskipun titik-titik data menunjukkan sedikit penyebaran yang mungkin mengindikasikan fluktuasi harga, secara umum pola ini mencerminkan hubungan yang erat antara kedua komoditas. Selain itu, histogram yang berada di sisi atas dan kanan scatter plot memberikan gambaran distribusi harga masing-masing variabel, memperlihatkan bagaimana harga perak dan emas tersebar dalam rentang tertentu. Visualisasi ini dapat digunakan sebagai alat eksplorasi awal dalam memahami ke

terkait pergerakan harga perak terhadap emas, sebelum dilakukan analisis statistik formal seperti uji korelasi. 70 Gambar 5. 14 Gambar visualisasi scatter plot persebaran harga emas terhadap indeks dolar AS

Scatter plot yang ditampilkan menggambarkan hubungan antara indeks penutupan dolar AS dengan harga penutupan emas. Pada sumbu X, variabel yang ditampilkan adalah harga penutupan emas, sedangkan pada sumbu Y adalah indeks penutupan dolar AS. Titik-titik data pada grafik merepresentasikan pasangan nilai dari kedua variabel pada waktu yang sama, memungkinkan analisis terhadap pola hubungan antara harga emas dan pergerakan indeks dolar AS. Dari distribusi titik-titik pada scatter plot, terlihat bahwa hubungan antara kedua variabel ini cenderung negatif, di mana peningkatan harga emas sering kali disertai dengan penurunan nilai indeks dolar AS. Garis regresi linier yang ditambahkan dalam grafik memberikan indikasi kecenderungan hubungan positif yang mengisyaratkan adanya pergerakan kenaikan yang bersamaan antara harga emas dan indeks dolar AS. Selain itu, density 71 plot yang berada di sisi atas dan kanan scatter plot menunjukkan distribusi masing-masing variabel, membantu dalam memahami pola penyebaran nilai indeks dan harga emas dalam rentang tertentu. Visualisasi ini berfungsi sebagai langkah eksplorasi awal untuk memahami bagaimana pergerakan indeks dolar AS berkorelasi dengan harga emas sebelum dilakukan analisis statistik lebih mendalam, seperti uji korelasi. 5.1.1.3.

Uji Korelasi Uji korelasi dilakukan untuk mengevaluasi kekuatan dan arah hubungan linear antar harga penutupan dari lima instrumen global, yaitu emas, minyak mentah, perak, indeks S&P 500, dan indeks dolar AS (USD). Dengan menggunakan metode Pearson Correlation Coefficient pada tingkat signifikansi 5% ($\alpha = 0,05$), analisis ini bertujuan mengidentifikasi apakah terdapat keterkaitan signifikan secara statistik di antara pasangan variabel tersebut. Nilai koefisien korelasi yang diperoleh berkisar antara -1 hingga 1, yang merepresentasikan arah dan kekuatan hubungan, mulai dari negatif kuat, tidak berkorelasi, hingga positif kuat. Untuk mempermudah interpretasi, hasil uji ini divisualisasikan dalam bentuk

heatmap, yang menggambarkan nilai korelasi antar variabel secara visual menggunakan gradasi warna. Visualisasi ini membantu dalam mengidentifikasi pola keterkaitan antar aset yang dapat digunakan sebagai dasar dalam strategi analisis hubungan antar pasar maupun untuk seleksi fitur dalam analisis prediktif sehingga dapat membuang fitur yang tidak memiliki korelasi.

72 Gambar 5. 15 Gambar heatmap korelasi Pearson antar variabel harga penutupan masing-masing harga Berdasarkan heatmap hasil uji korelasi Pearson menunjukkan hubungan linier antara harga penutupan emas (emas_Close) dengan berbagai variabel lainnya, termasuk minyak_Close, perak_Close, S&P 500 (sp500_Close), dan indeks dolar AS (usdx_Close). Visualisasi ini memberikan wawasan tentang keterkaitan antar aset global, di mana intensitas hubungan direpresentasikan melalui gradasi warna yang mencerminkan kekuatan korelasi. Dari analisis ini, terlihat bahwa emas_Close memiliki korelasi yang sangat kuat dengan dirinya sendiri (1.0000), seperti yang diharapkan. Korelasi yang tinggi juga ditemukan antara emas_Close dan S&P 500 (0.9352) serta perak_Close (0.9162), menunjukkan bahwa pergerakan harga emas memiliki hubungan erat dengan pasar saham AS serta harga perak, yang sering dianggap sebagai aset komplementer dalam dunia investasi. Sebaliknya, korelasi dengan indeks dolar AS (0.5037) menunjukkan hubungan yang lebih moderat, mengindikasikan bahwa walaupun emas sering digunakan sebagai lindung nilai terhadap dolar, pergerakannya tidak selalu berbanding terbalik dalam pola yang konsisten. Korelasi terlemah dalam heatmap ini terlihat antara harga penutupan emas dan harga penutupan minyak mentah (0.3716). Walaupun terdapat hubungan positif, nilainya relatif rendah dibandingkan dengan aset lainnya, menunjukkan bahwa harga minyak tidak selalu bergerak sejalan dengan harga emas. Sebagai tambahan untuk uji korelasi terhadap semua variabel dari semua instrumen harga yang didapatkan, berikut adalah visualisasi dari heatmap uji korelasi Pearson untuk semua variabel terhadap harga penutupan emas. Gambar 5. 16 Gambar heatmap korelasi Pearson semua variabel harga Berdasarkan dari heatmap

semua variabel dari semua instrumen, dapat disimpulkan bahwa volume perdagangan dari berbagai aset menunjukkan hubungan 74 yang lemah atau bahkan negatif terhadap harga penutupan emas. Korelasi negatif ini mengindikasikan bahwa peningkatan volume transaksi tidak selalu sejalan dengan perubahan harga emas, dan dalam beberapa kasus, bahkan bergerak berlawanan. Secara keseluruhan, hubungan negatif atau lemah antara volume perdagangan dan harga penutupan emas menunjukkan bahwa aktivitas transaksi tidak memiliki keterkaitan linear yang kuat dengan pergerakan harga. Volume sering kali mengalami fluktuasi yang tidak sejalan dengan perubahan harga emas, mencerminkan bahwa faktor-faktor harga lainnya memiliki pengaruh yang lebih signifikan dalam menentukan tren harga. Korelasi negatif pada beberapa aset, seperti volume minyak terhadap harga emas, menandakan bahwa peningkatan jumlah transaksi tidak selalu beriringan dengan kenaikan harga emas, melainkan dapat berfluktuasi secara independen. Selain itu, korelasi yang sangat kecil antara volume perdagangan emas sendiri terhadap harga penutupan emas mengindikasikan bahwa jumlah transaksi dalam satu periode tidak dapat digunakan sebagai fitur prediksi untuk pergerakan harga emas. Meskipun emas dan perak sering bergerak dalam pola yang serupa, volume perak juga menunjukkan korelasi yang rendah terhadap harga emas, yang menegaskan bahwa meskipun harga kedua logam mulia dapat saling terkait, aktivitas perdagangan tidak selalu merefleksikan hubungan harga yang kuat. Dengan demikian, volume perdagangan dalam berbagai aset lebih bersifat sebagai indikator aktivitas pasar daripada penentu langsung pergerakan harga emas, sehingga dalam analisis harga, faktor-faktor lain yang lebih mencerminkan pergerakan harga perlu mendapatkan prioritas lebih tinggi dibandingkan volume transaksi.

5.1.1.4. Kesimpulan Seleksi Fitur

Kesimpulan seleksi fitur dalam penelitian ini didasarkan pada analisis statistik deskriptif, analisa berdasarkan visualisasi, dan uji korelasi antar variabel terhadap harga penutupan emas (emas_Close). Dari hasil eksplorasi data, beberapa 75 aspek utama yang diperhatikan meliputi distribusi data, hubungan antar fitur,

serta signifikansi dalam representasi karakteristik aset. Berdasarkan statistik deskriptif, seluruh variabel memiliki jumlah data valid yang lengkap tanpa adanya nilai yang hilang. Distribusi harga emas menunjukkan pola positively skewed, dengan median yang sedikit lebih tinggi dibandingkan rata-rata, serta standar deviasi yang mencerminkan tingkat volatilitas yang cukup signifikan. Harga perak dan indeks S&P 50 menunjukkan pola distribusi yang relatif stabil, meskipun tetap menunjukkan variasi harga yang cukup besar. Sebaliknya, harga minyak mentah memiliki beberapa nilai ekstrem, termasuk harga negatif yang terjadi akibat fenomena pasar berjangka saat krisis pandemi. Indeks dolar AS menunjukkan pola distribusi yang cenderung simetris dengan sedikit kemiringan ke kanan. Uji korelasi Pearson mengonfirmasi hubungan yang kuat antara harga emas dengan harga perak (0.9162) dan indeks S&P 50 (0.9352), menandakan keterkaitan erat antara logam mulia dan pasar ekuitas. Korelasi dengan indeks dolar AS (0.5037) bersifat moderat, menunjukkan bahwa meskipun emas sering digunakan sebagai lindung nilai terhadap dolar, pergerakannya tidak selalu berkebalikan dalam pola yang konsisten. Sementara itu, harga minyak mentah memiliki korelasi yang rendah terhadap harga emas (0.3716), mengindikasikan bahwa fluktuasi harga minyak tidak selalu bergerak selaras dengan pergerakan emas. Dalam penelitian ini, volume perdagangan dari berbagai instrumen menunjukkan korelasi yang lemah atau negatif terhadap harga penutupan emas, sehingga tidak digunakan sebagai fitur dalam analisis time-series dengan pendekatan DL metode LSTM. Korelasi negatif antara volume minyak dan harga emas, misalnya, menunjukkan bahwa peningkatan jumlah transaksi tidak selalu beriringan dengan kenaikan harga emas, melainkan dapat berfluktuasi secara independen. Selain itu, volume perdagangan emas sendiri memiliki korelasi yang sangat kecil terhadap harga penutupannya, sehingga tidak dapat dijadikan fitur prediksi dalam pemodelan harga berbasis time-series. 76 Dengan mempertimbangkan berbagai analisis deskriptif seperti statistik deskriptif, distribusi data, hasil visualisasi, khususnya uji korelasi

Pearson, penelitian ini memutuskan untuk mengecualikan variabel volume dalam pemodelan prediktif menggunakan LSTM. Fokus analisis diarahkan pada variabel harga yang memiliki keterkaitan lebih erat dengan harga emas, sehingga model dapat menangkap pola pergerakan harga dengan lebih akurat tanpa dipengaruhi oleh fluktuasi volume yang tidak memiliki hubungan linear yang konsisten terhadap harga emas. Keputusan ini didasarkan pada evaluasi statistik yang menunjukkan bahwa volume lebih bersifat sebagai indikator aktivitas pasar daripada sebagai penentu langsung pergerakan harga emas. Dengan demikian, pemodelan prediktif dengan metode LSTM dapat lebih optimal dalam mengidentifikasi tren harga berdasarkan variabel yang memiliki hubungan lebih kuat dan signifikan.

5.1.1.5. Normalisasi Data

Setelah menerapkan seleksi fitur, semua fitur yang digunakan dalam analisis telah dipilih berdasarkan relevansi terhadap harga penutupan emas (emas_Close). Untuk memastikan bahwa data berada dalam skala yang seragam dan mengurangi pengaruh perbedaan rentang nilai antar fitur, dilakukan proses normalisasi menggunakan teknik Min-Max Scaling. Dengan pendekatan ini, setiap fitur akan memiliki nilai yang dipetakan dalam rentang hingga 1, sehingga skala variabel menjadi lebih seragam tanpa mengubah pola distribusi data secara signifikan. Berdasarkan hasil statistik deskriptif, fitur harga yang digunakan dalam penelitian ini, seperti harga penutupan emas, perak, minyak, indeks S&P 500, dan indeks dolar AS, memiliki rentang nilai yang cukup beragam. Oleh karena itu, normalisasi Min-Max dilakukan untuk setiap fitur harga guna memastikan bahwa model pembelajaran tidak dipengaruhi oleh perbedaan skala antar variabel. Implementasi normalisasi ini bertujuan untuk meningkatkan stabilitas dan konvergensi model saat proses pelatihan dengan metode LSTM, sehingga model dapat lebih akurat dalam menangkap pola pergerakan harga tanpa terpengaruh oleh skala awal dari fitur yang berbeda-beda. Setelah proses normalisasi selesai, dataset yang telah diubah ke skala [0,1] siap digunakan dalam tahap pemodelan berikutnya.

77 Berikut tabel lima baris terakhir dari semua fitur yang telah dinormalisasi dengan Min-Max

REPORT #27589781

Scaling yang dimulai dengan harga emas, minyak mentah, indeks S&P 500, harga perak, dan indeks dolar AS. Tabel 5. 6 Tabel hasil normalisasi 5 baris terakhir pada harga emas Date emas_Open emas_High emas_Low emas_Close 2025-02- 24 1 1 1 0,991081 2025-02- 25 0,97572 0,99182 0,975876 0,994202 2025-02- 26 0,982601 0,977241 0,98875 0,980155 2025-02- 27 0,963804 0,980469 0,964905 0,982942 2025-02- 28 0,937846 0,955039 0,944587 0,959864

Tabel 5. 7 Tabel hasil normalisasi 5 baris terakhir pada harga minyak mentah Date minyak_Open minyak_High minyak_Low minyak_Close 2025-02- 24 0,671481 0,490112 0,683508 0,604356 2025-02- 25 0,66051 0,492852 0,676556 0,612433 2025-02- 26 0,658588 0,475901 0,67457 0,59938 2025-02- 27 0,669311 0,486688 0,676122 0,597288 2025-02- 28 0,665654 0,484548 0,679412 0,607024

Tabel 5. 8 Tabel hasil normalisasi 5 baris terakhir pada indeks S&P 500 Date sp500_Open sp500_High sp500_Low sp500_Close 2025-02- 24 0,958846 0,974661 0,965931 0,971272 2025-02- 25 0,951046 0,961479 0,948107 0,962565 2025-02- 26 0,951237 0,963463 0,953543 0,956337 2025-02- 27 0,927267 0,961095 0,935213 0,955828 2025-02- 28 0,949334 0,949962 0,928892 0,931168

Tabel 5. 9 Tabel hasil normalisasi 5 baris terakhir pada harga perak Date perak_Open perak_High perak_Low perak_Close 2025-02- 24 0,902321 0,912506 0,918466 0,923505 2025-02- 25 0,868808 0,894167 0,884354 0,891928 2025-02- 26 0,888379 0,885904 0,904275 0,897621 2025-02- 27 0,868808 0,887539 0,869766 0,89682 2025-02- 28 0,843609 0,859037 0,852799 0,867467

Tabel 5. 10 Tabel hasil normalisasi 5 baris terakhir pada indeks dolar AS Date usdx_Open usdx_High usdx_Low usdx_Close 2025-02- 24 0,704878 0,687488 0,712266 0,701542 2025-02- 25 0,693877 0,691549 0,714855 0,70974 2025-02- 26 0,697831 0,684007 0,71366 0,688073 2025-02- 27 0,730992 0,709534 0,724811 0,703299 2025-02- 28 0,745948 0,723458 0,755476 0,727894

5.1.1.6. Pembuatan Dataset Setelah seluruh variabel dinormalisasi dengan teknik Min-Max Scaling, dataset diorganisasikan menggunakan pendekatan sliding window dengan ukuran jendela waktu 30 hari. Artinya, dalam proses prediksi harga emas pada hari t, data historis dari 30 hari sebelumnya

digunakan sebagai input. Penerapan teknik ini memungkinkan model untuk menangkap pola tren harga berdasarkan informasi dari periode sebelumnya, yang berperan penting dalam analisis time-series. Selain itu, dataset atau kumpulan data dipartisi menjadi tiga ukuran berdasarkan rasio data pelatihan dan data pengujian, yaitu 70:30, 80:20, dan 90:10. Pembagian ini dilakukan untuk mengevaluasi dampak jumlah data pelatihan terhadap performa prediksi model. Berikut adalah perbandingan jumlah baris data pelatihan dan data pengujian pada masing-masing ukuran dataset yang dijabarkan dalam bentuk tabel.

Ukuran Dataset	Jumlah Baris Data Pelatihan	Jumlah Baris Data Pengujian
Ukuran Dataset 70:30	1414	606
Ukuran Dataset 80:20	1616	404
Ukuran Dataset 90:10	1818	202

Dengan struktur dataset ini, model LSTM dapat memanfaatkan data berderet waktu untuk membangun prediksi berdasarkan pola historis yang relevan. Variasi dalam pembagian data set memungkinkan evaluasi efektivitas model dengan berbagai tingkat pelatihan dan pengujian, sehingga memberikan gambaran optimal tentang bagaimana pendekatan time series dapat digunakan dalam analisis harga emas.

5.1.2 Analisa Data Dengan Metode LSTM

Pada subsubbab ini akan memaparkan hasil pelatihan dan pengujian model LSTM dengan berbagai kumpulan data atau dataset yang telah diusulkan dalam perancangan. Hasil dari prediksi untuk model dalam memprediksi data tahunan akan dipaparkan dalam sub subbab ini. Hasil pelatihan ini mencakup nilai evaluasi yaitu MSE, MAE, RMSE, dan MAPE.

5.1.2.1. Hasil Pelatihan dan Pengujian Model LSTM

A. Hasil pelatihan model dengan kumpulan data ukuran 70:30

Mode	MSE pelatihan	RMSE pelatihan	MAE pelatihan	MAPE pelatihan	MSE pengujian	RMSE pengujian	MAE pengujian	MAPE pengujian
Mode 1	431,49	20,77	17,18	1,10%	5510,31	74,23	50,87	2,12%
Mode 2	97,31	9,86	7,34	0,47%	4981,73	7,58	4,6	,92
Mode 3	316,07	17,78	15,94	1,05%	2233,96	47,26	32,43	1,35%
Mode 4	233,47	15,28						

12,520,82% 2018,3944,9331,231,30% Model 5 104,8
10,27,520,48% 2212,7247,434,
1,41% Tabel 5. 13 Tabel kenaikan dari pelatihan ke pengujian ukuran
dataset 70:30 Model MSE pelatihan → MSE pengujian RMSE
pelatihan → RMSE pengujian MAE pelatihan → MAE peng
ujian MAPE pelatihan → MAPE pengujian Model 1
1177,04% 257,39% 196,10% 92,73% Model 2 5019,44% 615,82% 539,24% 302,13%
Model 3 606,79% 165,80% 103,45% 28,57% Model 4 764,52% 194,04% 149,44%
58,54% Model 5 2025,98% 361,18% 352,13% 193,75% Model 1 menunjukkan
adanya overfitting yang jelas, di mana kinerjanya pada data pelatihan
jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat
dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE, terjadi
kenaikan sebesar 1.177,04% antara MSE pelatihan dan MSE pengujian, yang
menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat
baik, tetapi kesulitan saat diterapkan pada data pengujian. Begitu pula
dengan MAE, yang mengalami kenaikan sebesar 196,10%, mengindikasikan bahwa
model lebih akurat pada data pelatihan namun prediksinya jauh lebih
buruk pada data pengujian. Kenaikan juga terjadi pada RMSE, dengan
peningkatan sebesar 257,39%, dan MAPE yang meningkat sebesar 92,73%,
semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan
tidak dapat menggeneralisasi dengan baik pada data pengujian. Model
2 menunjukkan adanya overfitting yang lebih parah dibandingkan Model 1,
di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan
dengan data pengujian. Hal ini terlihat dari kenaikan yang sangat
signifikan pada metrik evaluasi. Untuk MSE, terjadi kenaikan sebesar
5.019,44% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa
model mampu mempelajari data pelatihan dengan sangat baik, tetapi
kesulitan ketika diterapkan pada data pengujian. Begitu pula dengan
MAE, yang mengalami kenaikan sebesar 81539,24%, mengindikasikan bahwa
model lebih akurat pada data pelatihan namun prediksinya jauh lebih
buruk pada data pengujian. Kenaikan juga terjadi pada RMSE, dengan

peningkatan sebesar 615,82%, dan MAPE yang meningkat sebesar 302,13%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Model 3 menunjukkan adanya overfitting yang cukup jelas, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE, terjadi kenaikan sebesar 606,79% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan baik, namun kesulitan saat diterapkan pada data pengujian. Begitu pula dengan MAE, yang mengalami kenaikan sebesar 103,45%, mengindikasikan bahwa model lebih akurat pada data pelatihan namun prediksinya kurang memadai pada data pengujian. Kenaikan juga terjadi pada RMSE, dengan peningkatan sebesar 165,80%, dan MAPE yang meningkat sebesar 28,57%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Meskipun demikian, model 3 merupakan model yang paling stabil dibanding kelima model lainnya karena kenaikan yang tidak setajam model lainnya untuk ukuran dataset 70:30. Gambar 5. 17

Prediksi model 3 pada dataset pengujian dengan ukuran 70:30 Model 4 menunjukkan hasil yang sangat baik dan merupakan model yang paling optimal dibandingkan dengan model-model sebelumnya. Meskipun ada sedikit indikasi overfitting, kenaikan pada metrik evaluasi tetap terjadi dalam batas yang wajar. Untuk MSE, terjadi kenaikan sebesar 764,52% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model ini berhasil mempelajari data pelatihan dengan sangat baik dan tetap mempertahankan kinerja yang cukup baik pada data pengujian. MAE mengalami kenaikan sebesar 149,44%, yang menunjukkan bahwa prediksi pada data pelatihan jauh lebih akurat, namun tetap menunjukkan performa yang relatif baik pada data pengujian. Kenaikan pada RMSE sebesar 194,04% dan MAPE yang meningkat sebesar 58,54% semakin memperlihatkan bahwa meskipun terjadi peningkatan kesalahan saat diuji pada data pengujian, model ini

tetap lebih baik daripada model-model sebelumnya dalam menjaga keseimbangan antara akurasi pelat ihan dan generalisasi pada data pengujian. Dengan hasil ini, Model 4 dapat dianggap sebagai model yang paling efektif dalam menghasilkan prediksi yang stabil dan lebih dapat diandalkan. Gambar 5. 18 Prediksi model 4 pada dataset peng ujian dengan ukuran 70:30 Model 5 menunjukkan adanya overfitting yang cukup signifikan, meskipun kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Kenaikan yang terjadi pada metrik evaluasi sangat tajam. Untuk MSE, terjadi kenaikan sebesar 2.025,98% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mempelajari data pelatihan dengan sangat baik, namun kesulitan dalam menerapkan pengetahuan tersebut pada data pengujian. Begitu pula dengan MAE, yang mengalami kenaikan sebesar 352,13%, menandakan bahwa model jauh lebih akurat pada data pelatihan, tetapi prediksinya sangat buruk pada data pengujian. RMSE meningkat sebesar 361,18%, dan MAPE yang meningkat sebesar 193,75%, semakin memperlihatkan bahwa model ini terlalu "menghafal" data pelatihan, namun tidak dapat menggeneralisasi dengan baik pada data pengujian. Meskipun demikian, model 5 memiliki performa yang lebih buruk dibandingkan model-model sebelumnya dalam hal generalisasi, menandakan bahwa meskipun kinerja pada pelatihan sangat baik, model ini terlalu sesuai dan kurang efektif untuk prediksi pada data yang belum dilihat sebelumnya.

B. Hasil pelatihan model dengan kumpulan data ukuran 80:20

Model	MSE pelatihan	RMSE pelatihan	MAE pelatihan	MAPE pelatihan	MSE pengujian	RMSE pengujian	MAE pengujian	MAPE pengujian
Model 1	240,87	15,52	13,42	0,87%	3864,16	62,16	45,67	1,79%
Model 2	134,51	11,68	9,99	0,55%	3666,76	60,55	49,62	2,00%
Model 3	615,41	24,81	21,39	1,31%	770,34	27,76	24,46	1,11%
Model 4	98,40	9,92	7,28	0,45%	1543,11	39,28	28,95	1,15%
Model 5	167,81	12,95	10,32	0,66%				

1264,0435,5526,851,07% Tabel 5. 15 Tabel kenaikan dari pelatihan ke pengujian ukuran dataset 80:20 Model MSE pelatihan \rightarrow MSE pengujian RMSE pelatihan \rightarrow RMSE pengujian MAE pelatihan \rightarrow MAE pengujian MAPE pelatihan \rightarrow MAPE pengujian Model 1 1504,23% 300,52% 240,31% 105,75% Model 2 2626,01% 421,98% 451,95% 263,64% Model 3 25,18% 11,89% 14,35% -15,27% Model 4 1468,20% 295,97% 297,66% 155,56% Model 5 653,26% 174,52% 160,17% 62,12%

Model 1 dengan dataset 80:20 menunjukkan adanya overfitting yang cukup jelas, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE, terjadi kenaikan sebesar 1.504,23% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat baik, tetapi kesulitan saat diterapkan pada data pengujian. Begitu pula dengan MAE, yang mengalami kenaikan sebesar 240,31%, mengindikasikan bahwa model lebih akurat pada data pelatihan, namun prediksi jauh lebih buruk pada data pengujian. Kenaikan juga terjadi pada RMSE, dengan peningkatan sebesar 300,52%, dan MAPE yang meningkat sebesar 105,75%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Kenaikan yang tajam pada metrik ini menandakan bahwa meskipun model sangat baik dalam mempelajari data pelatihan, performanya menurun secara signifikan ketika diuji pada data yang tidak terlihat sebelumnya. Model 2 menunjukkan adanya overfitting yang lebih parah dibandingkan dengan Model 1, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang sangat signifikan pada metrik evaluasi. Untuk MSE, terjadi kenaikan sebesar 2.626,01% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat baik, namun kesulitan saat diterapkan pada data pengujian. Begitu pula dengan MAE, yang mengalami kenaikan sebesar 451,95%, mengindikasikan bahwa model lebih akurat pada

data pelatihan, namun prediksinya sangat buruk pada data pengujian. Kenaikan juga terjadi pada RMSE, dengan peningkatan sebesar 421,98%, dan MAPE yang meningkat sebesar 263,64%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Kenaikan yang sangat tajam pada metrik ini menandakan bahwa meskipun model mempelajari data pelatihan dengan sangat baik, kinerjanya menurun sangat drastis ketika diterapkan pada data yang tidak terlihat sebelumnya. Model 3 menunjukkan hasil yang lebih optimal dan stabil dibandingkan dengan model-model lainnya, meskipun masih terdapat indikasi overfitting. Kenaikan pada metrik evaluasi antara data pelatihan dan data pengujian terjaga dalam batas yang wajar. Untuk MSE, terjadi kenaikan sebesar 25,18%, yang relatif kecil dibandingkan dengan model-model lainnya, menunjukkan bahwa meskipun model mampu mempelajari data pelatihan dengan baik, kinerjanya pada data pengujian tetap cukup stabil. MAE mengalami kenaikan sebesar 14,35%, yang juga menunjukkan stabilitas prediksi pada data pengujian. RMSE mengalami 85 peningkatan sebesar 11,89%, sementara MAPE justru menurun sebesar -15,27%, yang menunjukkan bahwa meskipun ada sedikit peningkatan kesalahan pada data pengujian, model ini masih dapat memperbaiki akurasi dan generalisasinya, dengan MAPE yang lebih rendah pada data pengujian dibandingkan data pelatihan. Penurunan MAPE ini mengindikasikan bahwa meskipun model menghadapi beberapa peningkatan kesalahan, ia tetap lebih mampu dalam menjaga tingkat ketepatan yang lebih baik dan konsisten pada data pengujian. Dengan kenaikan yang relatif moderat pada metrik evaluasi dan penurunan MAPE, Model 3 menjadi model yang paling stabil dan efektif dalam menjaga keseimbangan antara akurasi pada data pelatihan dan kemampuannya untuk menggeneralisasi pada data pengujian, menjadikannya pilihan terbaik di antara model lainnya.

Gambar 5. 19 Prediksi model 3 pada dataset pengujian dengan ukuran 80:20 Model 4 menunjukkan adanya overfitting yang cukup jelas, meskipun masih lebih stabil dibandingkan dengan Model 1 dan Model 2.

Kinerja model pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian, yang terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE, terjadi kenaikan sebesar 1.468,20%, menunjukkan bahwa meskipun model mampu mempelajari data pelatihan dengan baik, performanya menurun saat diterapkan pada data pengujian. Begitu pula dengan MAE, yang mengalami kenaikan sebesar 297,66%, mengindikasikan bahwa prediksi model pada data pengujian lebih buruk dibandingkan dengan data pelatihan. Kenaikan juga terjadi pada RMSE, dengan peningkatan sebesar 295,97%, dan MAPE yang meningkat sebesar 155,56%, semakin memperlihatkan bahwa meskipun model belajar dengan baik pada data pelatihan, ia kesulitan dalam menggeneralisasi pada data pengujian. Meskipun demikian, Model 4 menunjukkan peningkatan yang lebih moderat pada metrik ini dibandingkan dengan Model 1 dan Model 2, menjadikannya sedikit lebih stabil meskipun masih ada tanda-tanda overfitting yang cukup jelas. Model 5, meskipun masih menunjukkan indikasi overfitting, menunjukkan performa yang lebih baik dibandingkan dengan Model 1, Model 2, dan Model 4. Kenaikan pada metrik evaluasi antara data pelatihan dan data pengujian cukup signifikan, tetapi tidak sebesar pada model-model sebelumnya. Untuk MSE, terjadi kenaikan sebesar 653,26%, yang menunjukkan bahwa meskipun model mampu mempelajari data pelatihan dengan baik, performanya menurun pada data pengujian. MAE mengalami kenaikan sebesar 160,17%, yang mengindikasikan bahwa prediksi model pada data pengujian lebih buruk dibandingkan dengan data pelatihan. RMSE mengalami peningkatan sebesar 174,52%, dan MAPE meningkat sebesar 62,12%, yang menunjukkan bahwa model masih "menghafal" data pelatihan namun kesulitan saat diuji pada data pengujian. Namun, kenaikan pada metrik ini lebih moderat dibandingkan model-model lainnya selain Model 3, dan dengan MAPE yang relatif rendah pada data pengujian, Model 5 menunjukkan kinerja yang cukup baik dalam menjaga akurasi pada data pengujian meskipun tetap ada sedikit penurunan dibandingkan data pelatihan. Gambar 5. 20 Prediksi Model 5 pada dataset pengujian dengan ukuran 80:20 C. Hasil pelatihan

model dengan kumpulan data ukuran 90:10 87 Tabel 5. 16 Tabel hasil
 pelatihan dan pengujian dengan kumpulan data 90:10 Model MSE pelatihan
 n RMSE pelatihan MAE pelatihan MAPE pelatihan MSE pengujian
 RMSE pengujian MAE pengujian MAPE pengujian Model 1
 359,31 18,96 16,45 1,03% 389,71 15,30 19,74 0,60% Model 2 74,45
 8,63 6,03 0,36% 165,77 9,36 12,88 0,37% Model 3 204,09
 14,29 11,67 0,70% 265,81 12,85 16,30 0,51% Model 4 108,81 10,43
 7,51 0,45% 205,09 10,65 14,32 0,42% Model 5 79,2
 48,96 41,039% 232,15 15,24 11,95 0,46%

Tabel 5. 17 Tabel kenaikan dari pelatihan ke pengujian ukuran dataset
 90:10 Model MSE pelatihan \rightarrow MSE pengujian RMSE pelatihan
 \rightarrow RMSE pengujian MAE pelatihan \rightarrow MAE pengujian
 n MAPE pelatihan \rightarrow MAPE pengujian Model 1 8,46% -19,30%
 20,00% -41,75% Model 2 122,66% 8,46% 113,60% 2,78% Model 3 30,24%
 -10,08% 39,67% -27,14% Model 4 88,48% 2,11% 90,68% -6,67% Model 5
 192,97% 71,24% 86,43% 17,95% Model 1 dengan dataset 90:10 menunjukkan
 adanya peningkatan yang signifikan pada metrik evaluasi, meskipun tidak
 separah pada dataset 80:20. Kinerja model pada data pelatihan lebih baik
 dibandingkan dengan data pengujian, dengan kenaikan MSE sebesar 8,46%,
 yang menunjukkan bahwa model dapat mempelajari data pelatihan dengan baik
 tetapi kesulitan saat diterapkan pada data pengujian. Begitu pula dengan
 MAE yang mengalami kenaikan sebesar 20,00%, yang menandakan bahwa model
 lebih akurat pada data pelatihan, namun prediksinya agak buruk pada data
 pengujian. RMSE mengalami penurunan sebesar -19,30%, dan MAPE justru
 menurun sebesar -41,75%, menunjukkan bahwa meskipun ada sedikit penurunan
 kesalahan, model ini lebih mampu menggeneralisasi dengan lebih baik
 pada data pengujian dibandingkan dengan data pelatihan. Dengan penurunan
 pada RMSE dan MAPE, Model 1 menunjukkan 88 perbaikan dalam kemampuan
 generalisasi meskipun masih ada tanda-tanda overfitting yang harus di
 perhatikan. Model 2 menunjukkan adanya overfitting yang lebih moderat
 dibandingkan dengan Model 1, meskipun masih terdapat perbedaan signifikan

antara kinerja pada data pelatihan dan data pengujian. Kenaikan MSE sebesar 122,66% menunjukkan bahwa meskipun model belajar dengan sangat baik pada data pelatihan, performanya menurun pada data pengujian. Begitu pula dengan MAE yang mengalami kenaikan sebesar 113,60%, yang mengindikasikan bahwa model lebih akurat pada data pelatihan, namun hasil prediksinya pada data pengujian lebih buruk. RMSE mengalami kenaikan sebesar 8,46%, dan MAPE yang sedikit meningkat sebesar 2,78%. Hal ini menunjukkan bahwa meskipun ada sedikit peningkatan pada kesalahan pada data pengujian, Model 2 mampu mempertahankan akurasi yang baik, dengan perbedaan antara data pelatihan dan data pengujian yang tidak terlalu tajam, menjadikannya model yang cukup stabil meskipun masih menunjukkan sedikit overfitting. Meskipun demikian, Model 2 merupakan model yang menghasilkan nilai terkecil dibanding model lainnya. Gambar 5. 21

Prediksi Model 2 pada dataset pengujian dengan ukuran 90:10 Model 3 menunjukkan hasil yang lebih stabil dibandingkan dengan Model 1 dan Model 2, meskipun masih ada indikasi overfitting. Untuk MSE, terjadi kenaikan sebesar 30,24%, yang masih relatif rendah dibandingkan dengan model lainnya, menunjukkan bahwa model dapat mempertahankan kinerjanya pada data pengujian meskipun tidak sebaik pada data pelatihan. MAE mengalami kenaikan sebesar 39,67%, yang mengindikasikan sedikit penurunan akurasi pada data pengujian dibandingkan data pelatihan. RMSE mengalami penurunan sebesar -8910,08%, dan MAPE juga menurun sebesar -27,14%, yang menunjukkan bahwa meskipun terjadi sedikit peningkatan kesalahan pada data pengujian, model ini lebih baik dalam menggeneralisasi dan mempertahankan akurasi yang lebih konsisten. Penurunan MAPE ini mengindikasikan bahwa Model 3 mampu menjaga akurasi yang lebih baik pada data pengujian, menjadikannya pilihan yang lebih stabil di antara model lainnya. Gambar 5. 22

Prediksi Model 3 pada dataset pengujian dengan ukuran 90:10 Model 4 menunjukkan adanya peningkatan yang moderat pada metrik evaluasi antara data pelatihan dan data pengujian, dengan MSE meningkat sebesar 88,48%, yang menunjukkan penurunan kinerja

model pada data pengujian meskipun masih cukup stabil pada data pelatihan. MAE mengalami kenaikan sebesar 90,68%, yang menandakan bahwa prediksi model lebih buruk pada data pengujian, meskipun ada peningkatan yang lebih moderat dibandingkan dengan model lainnya. RMSE mengalami kenaikan sebesar 2,11%, dan MAPE mengalami penurunan sebesar -6,67%, yang menunjukkan bahwa meskipun ada sedikit peningkatan kesalahan pada data pengujian, model ini cukup baik dalam mempertahankan akurasi pada data pengujian. Dengan kenaikan yang lebih terkendali pada metrik evaluasi dan penurunan MAPE, Model 4 menunjukkan performa yang cukup stabil meskipun masih ada sedikit overfitting. Model 5 menunjukkan adanya overfitting yang lebih jelas dibandingkan Model 4, meskipun tetap cukup stabil dalam hal kinerja. MSE mengalami kenaikan yang signifikan sebesar 192,97%, menunjukkan penurunan performa pada data pengujian. MAE mengalami kenaikan sebesar 86,43%, yang menunjukkan bahwa model lebih akurat pada data pelatihan, namun hasil prediksinya lebih buruk pada data pengujian. RMSE mengalami kenaikan sebesar 71,24%, dan MAPE mengalami kenaikan sebesar 17,95%, yang mengindikasikan bahwa model mengalami kesulitan saat menggeneralisasi pada data pengujian meskipun tetap dapat mempertahankan akurasi yang cukup baik. Dengan kenaikan pada metrik evaluasi yang lebih besar dibandingkan Model 4, Model 5 masih menunjukkan adanya peningkatan pada prediksi data pengujian, tetapi kurang stabil dibandingkan dengan model lainnya.

5.1.2.2. Hasil Prediksi Model LSTM

Pada sub subbab ini akan memaparkan hasil prediksi model LSTM dengan data tahunan berfrekuensi harian selama lima tahun terakhir, dengan kata lain terdapat lima kali pengujian. Lima kali pengujian ini terdiri dari awal tahun 2020 hingga akhir tahun 2020 sampai awal tahun 2024 hingga akhir tahun 2024. Model-model yang digunakan merupakan model terbaik dari ketiga ukuran dataset, sehingga jumlah model yang diprediksi berjumlah tiga buah. Perwakilan model terbaik dari ukuran dataset 70:30 adalah Model 4, dalam pengujian analisis time-series ini Model 4 dinantikan Model 1. Model 2 dalam

pengujian ini merupakan perwakilan model terbaik dari ukuran dataset 80:20, yaitu Model 3. Model 3 atau model ketiga dalam pengujian ini merupakan perwakilan model terbaik dari ukuran dataset 90:10, yaitu Model 2. Proses analisis yang dilakukan selama pengujian analisis time-series ini dengan membandingkan kenaikan tingkat kesalahan atau selisih serta akurasi. Hasil prediksi model ini mencakup nilai evaluasi yaitu MSE, RMSE, MAE, dan MAPE, serta persentase kenaikan dari setiap nilai evaluasi per pengujian dengan tahun yang berbeda.

Tabel 5. 18 Tabel hasil prediksi dengan model 1 pada pengujian tahunan Metrik/ Tahun awal 202 - akhir 2020 awal 2021 - akhir 2021 awal 2022 - akhir 2022 awal 2023 - akhir 2023 awal 2024 - akhir 2024

Metric	2020	2021	2022	2023	2024
MSE	446,91	155,75	346,46	154,99	3193,08
MAE	17,33	9,84	16,07	9,73	49,07
RMSE	21,14	12,48	18,61	12,45	56,51
MAPE	0,99%	0,55%	0,90%	0,50%	1,95%

Tabel 5. 19 Tabel persentase kenaikan relatif pada prediksi Model 1 Metrik/ Tahun 2020→2021 2021→2022 2022→2023 2023→2024

Metric	2020→2021	2021→2022	2022→2023	2023→2024
MSE	-34,76%	33,10%	-51,77%	1552,68%
MAE	-15,35%	14,50%	-36,21%	316,37%
RMSE	-19,23%	15,37%	-30,55%	306,53%
MAPE	-16,04%	14,61%	-41,18%	215,00%

Secara keseluruhan, Model 1 menunjukkan performa yang cukup baik pada tahun 2021 dan 2023, namun mengalami penurunan yang signifikan pada tahun 2024. Pada tahun 2021, sebagian besar metrik, termasuk MSE yang turun sebesar 34,76%, MAE yang turun -15,35%, RMSE yang turun -19,23%, dan MAPE yang turun -16,04%, menunjukkan peningkatan akurasi Model 1. Namun, pada 2022, terjadi penurunan performa, yang tercermin dari peningkatan MSE sebesar 33,10%, MAE yang meningkat 14,50%, RMSE yang meningkat 15,37%, dan MAPE yang meningkat 14,61%, meskipun perbaikan kembali terlihat pada 2023 dengan penurunan MSE sebesar -51,77%, MAE -36,21%, RMSE -30,55%, dan MAPE -41,18%. Lonjakan besar pada 2024, dengan kenaikan yang sangat drastis di semua metrik, seperti MSE yang meningkat sebesar 1552,68%, MAE yang meningkat 316,37%, RMSE yang meningkat 306,53%, dan MAPE yang meningkat 215,00%, menunjukkan adanya masalah yang sangat serius pada Model

1. Tabel 5. 20 Tabel hasil prediksi dengan Model 2 Metrik/ Tahun

awal 202 - akhir 2020 awal 2021 - akhir 2021 awal 2022 - akhir
2022 awal 2023 - akhir 2023 awal 2024 - akhir 2024 MSE 840,49
377,38 1475,15 1236,87 465,19 MAE 26,19 17,73 36,85 33,82 18
,45 RMSE 28,99 19,43 38,41 35,17 21,57 MAPE

1,48% 0,99% 2,07% 1,74% 0,77% Tabel 5. 21 Tabel persentase kenaikan
relatif pada prediksi Model 2 Metrik/ Tahun 202 →2021 2021 →2

022 2022→2023 2023→2024 MSE -55,10% 290,89% -16,15% -62,39% MAE
-32,30% 107,80% -8,22% -45,45% 92 MAE -32,99% 97,71% -8,43% -38,67%

MAPE -33,11% 109,09% -15,94% -55,75% Model 2 menunjukkan fluktuasi
performa yang cukup signifikan sepanjang periode yang diamati, dengan
beberapa penurunan dan peningkatan yang cukup mencolok pada masing-masing
metrik. Pada tahun 2021, model menunjukkan perbaikan yang jelas
dibandingkan tahun 2020, terlihat dari penurunan MSE sebesar -55,10%,
MAE yang turun -32,30%, RMSE yang menurun -32,99%, dan MAPE yang turun
-33,11%. Namun, pada tahun 2022, terjadi lonjakan besar pada MSE yang
meningkat 290,89%, serta peningkatan MAE sebesar 107,80%, RMSE yang
meningkat 97,71%, dan MAPE yang meningkat 109,09%, menunjukkan
penurunan performa yang sangat signifikan. Tahun 2023 memberikan sedikit
perbaikan, dengan MSE yang turun -16,15%, MAE yang turun -8,22%,
RMSE yang turun -8,43%, dan MAPE yang turun -15,94%, menunjukkan bahwa
model mulai kembali menunjukkan hasil yang lebih baik. Namun, pada
tahun 2024, model mengalami penurunan yang signifikan lagi, dengan MSE
turun -62,39%, MAE turun -45,45%, RMSE turun -38,67%, dan MAPE turun
-55,75%, meskipun secara keseluruhan tetap lebih baik dibandingkan dengan
2022.

Tabel 5. 22 Tabel hasil prediksi dengan Model 3 Metrik/ Tahun

awal 202 - akhir 2020 awal 2021 - akhir 2021 awal 2022 - akhir
2022 awal 2023 - akhir 2023 awal 2024 - akhir 2024 MSE 149,24
64,43 105,75 86,66 160,98 MAE 8,79 6,03 7,71 6,71 9,7
RMSE 12,22 8,03 10,28 9,31 12,69 MAPE 0,50% 0,34% 0,43%

0,34% 0,37% Tabel 5. 23 Tabel persentase kenaikan relatif pada

prediksi Model 3 Metrik/ Tahun 2020→2021 2021→2022 2022→2023 2023→2024 MSE -56,83% 64,12% -18,05% 85,76% MAE -31,43% 27,95% -12,98% 35,08% MAE -34,29% 28,11% -9,47% 36,29% MAPE -32,00% 26,47% -20,93% 8,82%

93 Model 3 m enunjukkan kinerja yang relatif stabil dengan beberapa fluktuasi yang terlihat jelas pada metrik yang digunakan, meskipun ada perbaikan yang cukup baik di beberapa tahun. Pada tahun 2021, model menunjukkan penurunan yang signifikan dalam MSE sebesar -56,83%, MAE turun -31,43% , RMSE turun - 34,29%, dan MAPE turun -32,00%, yang menandakan perbaikan akurasi model dibandingkan dengan tahun 2020. Namun, pada tahun 2022, model mengalami peningkatan MSE sebesar 64,12%, serta kenaikan pada MAE 27,95% , RMSE 28,11%, dan MAPE 26,47%, yang mengindikasikan penurunan performa yang signifikan dibandingkan tahun sebelumnya. Setelah lonjakan di 2022, model kembali menunjukkan perbaikan pada 2023, dengan penurunan MSE sebesar - 18,05%, MAE turun -12,98%, RMSE turun -9,47%, dan MAPE turun -20,93%, yang menunjukkan pemulihan kinerja. Namun, pada tahun 2024, meskipun metrik-metrik tersebut kembali menunjukkan penurunan, ada lonjakan besar pada MSE yang meningkat sebesar 85,76%, MAE meningkat 35,08%, RMSE meningkat 36,29%, dan MAPE meningkat 8,82%. Hal ini menunjukkan adanya penurunan performa model yang signifikan, yang perlu dianalisis lebih lanjut untuk mengetahui penyebabnya, apakah terkait dengan perubahan data atau potensi kesalahan lainnya. Secara keseluruhan, Model 3 menunjukkan fluktuasi yang cukup besar dalam performanya, dengan perbaikan yang terlihat pada 2021 dan 2023, namun penurunan signifikan terjadi pada 2022 dan 2024. Fluktuasi ini mengindikasikan bahwa meskipun model mampu beradaptasi dengan baik di beberapa titik, ada kemungkinan bahwa faktor eksternal seperti perubahan dalam data atau pengaturan model mempengaruhi kinerjanya secara tidak konsisten.

Tabel 5. 24 Tabel hasil prediksi tahunan pada tiga buah model Metrik / Tahun Model 1 (2020-2024) Model 2 (2020-2024) Model 3 (2020-2024) 2020 MSE 446,91 840,49 149,24 MAE 17,33 26,19 8,79 RMSE 21,14 28,99 12,22 MAPE 0,99% 1,48% 0,50%

REPORT #27589781

2020→2021 94 MSE -65,15% -55,10% -56,83% MAE -43,26% -32,30% -31,43%
RMSE -40,97% -32,99% -34,29% MAPE -44,44% -33,11% -32,00% 2021→2022
MSE 122,44% 290,89% 64,12% MAE 63,39% 107,80% 27,95% RMSE 49,14%
97,71% 28,11% MAPE 63,64% 109,09% 26,47% 2022→2023 MSE -55,26%
-16,15% -18,05% MAE -39,48% -8,22% -12,98% RMSE -33,12% -8,43% -9,47%
MAPE -44,44% -15,94% -20,93% 2023→2024 MSE 1960,21% -62,39% 85,76%
MAE 404,55% -45,45% 35,08% RMSE 353,90% -38,67% 36,29% MAPE 290,00%
-55,75% 8,82% Secara keseluruhan, Model 3 menunjukkan performa yang paling
konsisten dan stabil dibandingkan dengan Model 1 dan Model 2, meskipun
ada beberapa fluktuasi dalam hasilnya. Pada tahun 2020, Model 3
memiliki nilai MSE, MAE, RMSE, dan MAPE yang jauh lebih rendah
daripada kedua model lainnya, menunjukkan akurasi yang lebih baik di
awal periode. Model 1 dan Model 2 memiliki nilai MSE yang jauh
lebih tinggi pada tahun 2020, dengan Model 2 bahkan mencapai MSE
sebesar 840,49, jauh lebih besar dibandingkan Model 3 yang hanya 149,24.
Model 1 juga menunjukkan performa yang lebih baik dibandingkan Model
2 pada 2020, tetapi masih jauh di belakang Model 3 dalam hal
akurasi. Pada tahun 2021, ketiga model mengalami perbaikan
signifikan, dengan Model 1 menunjukkan penurunan yang paling tajam pada
MSE -65,15% dan MAE -43,26%. Meskipun demikian, Model 2 dan Model
3 juga mengalami perbaikan, meskipun tidak sebesar Model 1. Namun,
ketika memasuki tahun 2022, terjadi lonjakan besar dalam MSE, MAE, RM
SE, dan MAPE untuk semua model. Model 2 mengalami lonjakan terbesar,
dengan MSE meningkat sebesar 290,89%, sementara Model 1 juga
mengalami peningkatan MSE sebesar 122,44%. Model 3 mengalami lonjakan
lebih kecil dengan MSE meningkat 64,12%, tetapi tetap menunjukkan
peningkatan yang lebih moderat dibandingkan dua model lainnya. Pada tahun
2023, ketiga model menunjukkan pemulihan yang lebih baik dibandingkan
tahun 2022. Model 1 mengalami penurunan MSE yang signifikan -
55,26%, MAE -39,48%, dan RMSE -33,12%, yang menunjukkan perbaikan
performa yang cukup solid. Model 2 dan Model 3 juga menunjukkan

penurunan, meskipun lebih kecil daripada Model 1. Namun, pada tahun 2024, ada lonjakan besar yang sangat signifikan pada Model 1, dengan MSE meningkat sebesar 1960,21%, MAE meningkat 404,55%, RMSE meningkat 353,90%, dan MAPE meningkat 290,00%. Hal ini menunjukkan penurunan performa yang sangat besar dan menunjukkan bahwa Model 1 mengalami masalah serius di tahun 2024. Sementara itu, Model 2 dan Model 3 menunjukkan penurunan yang lebih moderat pada tahun 2024. Model 2 mengalami penurunan yang sangat besar pada MSE -62,39%, MAE -45,45%, dan MAPE -55,75%, yang menunjukkan pemulihan yang baik setelah lonjakan besar di tahun 2022. Model 3, meskipun ada lonjakan kecil pada MSE 85,76% dan MAE 35,08%, tetap menunjukkan performa yang lebih stabil dibandingkan dengan Model 1, dengan MAPE yang tetap rendah 8,82%.

5.2. Pembahasan

Dalam penelitian analisis time-series untuk prediksi harga emas menggunakan pendekatan DL metode LSTM, dilakukan berbagai proses untuk dapat mengoptimalkan analisis prediktif dengan pemodelan LSTM. Proses-proses seperti pra-pemrosesan data yang meliputi seleksi fitur, normalisasi, pembuatan dataset hingga pelatihan model LSTM serta analisa prediksi. Dimulai dengan seleksi fitur, seleksi fitur merupakan salah satu tahap yang sangat penting dalam penelitian ini karena menentukan fitur-fitur yang paling relevan dengan pergerakan harga emas, sehingga model prediktif dapat bekerja dengan lebih akurat. Proses seleksi fitur diawali dengan analisis statistik deskriptif, visualisasi distribusi data, serta uji korelasi Pearson untuk mengevaluasi hubungan linear antar variabel dengan harga penutupan emas. Hasil analisis menunjukkan bahwa variabel volume perdagangan dari berbagai aset memiliki variabilitas yang sangat tinggi dan distribusi yang tidak konsisten. Selain itu, uji korelasi Pearson mengungkapkan bahwa volume memiliki hubungan yang sangat lemah atau bahkan negatif terhadap harga penutupan emas, yang menandakan bahwa peningkatan jumlah transaksi tidak selalu beriringan dengan perubahan harga emas. Oleh karena itu, penelitian ini memutuskan untuk mengecualikan variabel volume dari analisis prediktif guna memastikan bahwa

model hanya menggunakan fitur yang memiliki keterkaitan erat dengan pergerakan harga. Setelah seleksi fitur, dilakukan normalisasi data dengan teknik Min-Max Scaling untuk memastikan bahwa seluruh fitur berada dalam rentang yang seragam [0,1]. Normalisasi ini berfungsi untuk menghilangkan perbedaan skala antar variabel agar model LSTM dapat lebih efektif dalam menangkap pola pergerakan harga tanpa dipengaruhi oleh perbedaan rentang nilai yang terlalu besar. Selanjutnya, pembuatan dataset dilakukan dengan pendekatan sliding window berukuran 30 hari, yang berarti setiap observasi untuk prediksi harga emas menggunakan informasi dari 30 hari sebelumnya. Hal ini bertujuan agar model dapat memahami pola historis pergerakan harga dan memanfaatkannya dalam analisis time-series. Selain itu, dataset dibagi menjadi tiga skenario berdasarkan proporsi data pelatihan dan pengujian, yaitu 70:30, 80:20, dan 90:10. Dengan kombinasi seleksi fitur yang tepat, normalisasi data, dan strategi pembuatan dataset yang mempertimbangkan ketergantungan temporal dalam data time-series, penelitian ini memastikan bahwa model LSTM memiliki fondasi yang kuat untuk menangkap pola pergerakan harga emas dengan lebih optimal. Langkah-langkah ini bertujuan untuk meningkatkan akurasi prediksi dan memberikan pemahaman yang lebih mendalam terhadap faktor-faktor yang memengaruhi pergerakan harga emas dalam periode waktu tertentu. Pembahasan selanjutnya membahas mengenai hasil pelatihan serta pengujian dari model LSTM dengan ukuran dataset 70:30. Berdasarkan dari lima buah model dengan spesifikasi berbeda, Model 4 dan Model 3 merupakan model yang terbaik untuk ukuran dataset 70:30. Meskipun kedua model ini merupakan 97 model terbaik, hasil MSE menunjukkan bahwa kedua model ini mengalami overfitting dengan nilai MSE pada model 4 sebesar 2018,39. Ini menandakan bahwa model terlalu menghafal data pelatihan sehingga kinerja menurun pada data pengujian. Namun, berdasarkan Gambar 5.17 dan Gambar 5.18, dapat disimpulkan bahwa data pengujian ini memiliki outlier pada harga emas di atas yang terbukti oleh Gambar 5.6 bahwa adanya outlier pada harga tinggi emas. Ini memperkuat

bahwa baik Model 3 dan Model 4 kurang mempelajari di namika harga emas yang terdapat outlier sehingga performanya kurang baik. Selanjutnya, berdasarkan dari lima buah model dengan spesifikasi berbeda dan ukuran dataset 80:20, Model 3 dan Model 5 merupakan model terbaik. Meskipun kedua model ini merupakan model terbaik, hasil MSE menunjukkan bahwa model masih mengalami overfitting atau kead aan model terlalu menghafal data pelatihan sehingga kurang baik performanya dalam memprediksi data pada data pengujian. Namun, Model 3 merupakan model terbaik yang memiliki MSE pengujian sebesar 770,34 dan model yang terstabil dibandingkan model lainnya dalam ukuran dataset 80:20. **13** Jika dilihat dari hasil prediksi data pengujian pada Gambar 5. 19 dan Gambar 5. 20, model masih kurang baik dalam menganalisis harga emas yang mengalami outlier. Hal ini dapat disimpulkan bahwa dalam data pelatihan ini kurang menggambarkan dinamika harga emas yang cukup untuk menangkap pergerakan harga emas yang fluktuatif. Berdasarkan hasil pelatihan lima buah model dengan spesifikasi yang berbeda pada ukuran dataset 90:10, Model 2 dan Model 4 merupakan model yang terbaik bahkan untuk semua model pada ukuran dataset yang berbeda. Model 2 merupakan model terbaik dengan MSE pengujian sebesar 165,77 dan estimasi akurasi 99,63%. Jika dilihat hasil prediksi kedua model ini pada data pengujian yaitu Gambar 5. 21 dan Gambar 5. 22, dapat disimpulkan bahwa kedua model dapat menangkap pola-pola pergerakan harga emas yang fluktuatif. Dapat disimpulkan bahwa pada data pelatihan ini memiliki referensi yang cukup untuk menggeneralisasi serta menangkap pola-pola pergerakan harga emas yang fluktuatif.

98 BAB VI PENUTUP Bab keenam atau bab penutup merupakan bab terakhir untuk mengakhiri laporan penelitian dengan menyajikan kesimpulan penelitian serta saran yang berkaitan dengan penelitian. Kesimpulan diambil berdasarkan rumusan masalah yang dijawab selama penelitian serta saran berdasarkan masukan untuk keberlanjutan penelitian selanjutnya.

6.1. Kesimpulan Kesimpulan dalam penelitian ini dapat dituliskan dalam bentuk poin-poin untuk menjawab rumusan masalah

sebagai berikut. 1. Proses harga emas dapat diprediksi dengan menerapkan pemodelan dengan analisis prediktif metode LSTM. Proses dimulai dengan variabel-variabel penentu pergerakan harga emas dipilih untuk dijadikan fitur dalam penentuan pergerakan harga emas. Fitur-fitur ini dijadikan masukkan dan diterapkan pengurutan waktu yang berderet untuk mempelajari pergerakan harga emas berdasarkan pola deretan waktu. Setelah pemodelan analisis prediktif metode LSTM mempelajari fitur penentu pergerakan harga emas, pemodelan analisis prediktif dapat memprediksi harga emas. Hasil prediksi harga emas akan dievaluasi untuk menilai kinerja dalam menganalisis secara prediktif. Evaluasi yang dilakukan dengan melakukan kalkulasi nilai hasil prediksi harga emas terhadap harga emas asli, semakin kecil kesalahan atau selisih hasil prediksi harga emas terhadap harga asli maka semakin akurat. Hasil evaluasi kinerja pemodelan analisis prediktif dengan MSE, RMSE, MAE, dan MAPE. 2. Berdasarkan hasil yang didapat, pemodelan analisis prediktif mendapatkan estimasi akurasi tertinggi sebesar 99,63% dalam evaluasi terhadap data pengujian, yang termasuk kategori dalam akurasi tinggi. 3. Berdasarkan evaluasi yang didapatkan, dapat diperkirakan bahwa kesalahan hasil analisis prediktif terhadap harga emas asli adalah sekitar 12,88 poin, 99 dengan MSE sebesar 165,77 dan RMSE sebesar 9,36. MAPE yang tercatat sebesar 0,37% juga menunjukkan tingkat kesalahan yang sangat rendah dalam prediksi harga emas tersebut. 6.2. Saran Berdasarkan hasil penelitian yang telah didapatkan, terdapat beberapa saran yang dapat diberikan untuk keberlanjutan penelitian selanjutnya sehingga kinerja dalam analisis prediktif metode LSTM dapat lebih optimal dalam memprediksi harga emas. Saran-saran untuk keberlanjutan penelitian selanjutnya meliputi penggunaan data lebih mendalam, analisis deskriptif harga emas lebih lanjut, serta beberapa optimasi parameter dalam perancangan model prediksi. Saran pertama yaitu penggunaan data yang lebih mendalam dan berkorelasi tinggi (baik positif dan negatif) sehingga analisis prediktif dapat menganalisis pergerakan harga emas lebih baik dan nyata,

REPORT #27589781

membuat analisis prediktif lebih akurat dan optimal. Saran kedua adalah analisis deskriptif harga emas lebih lanjut serta mendalam untuk mempelajari karakteristik harga emas dan berbagai fitur-fitur lainnya sehingga berbagai fitur-fitur dari variabel yang kurang relevan dapat dieliminasi guna mengoptimalkan hasil analisis secara nyata dan akurat. Saran ketiga merupakan optimasi parameter untuk menunjang pembelajaran model dengan metode LSTM lebih mempelajari pergerakan harga emas serta mengekstraksi pergerakan harga pada instrumen lainnya yang membantu dalam penentu pergerakan harga emas. Melalui penerapan saran-saran ini, diharapkan keberlanjutan penelitian dapat menghasilkan analisis prediktif baik dengan metode LSTM maupun metode lainnya lebih optimal baik tingkat akurasi yang lebih tinggi maupun kesalahan yang lebih kecil. Saran-saran ini dapat mengatasi berbagai hasil penelitian dalam model yang memiliki masalah overfitting.

REPORT #27589781

Results

Sources that matched your submitted document.

● IDENTICAL ● CHANGED TEXT

INTERNET SOURCE		
1.	0.4% staff.universitaspahlawan.ac.id https://staff.universitaspahlawan.ac.id/upload/riset/537-lampiran.pdf	●
INTERNET SOURCE		
2.	0.38% umsupress.umsu.ac.id https://umsupress.umsu.ac.id/wp-content/uploads/2025/05/File-Isi-Panduan-Le..	●
INTERNET SOURCE		
3.	0.24% ojs.uma.ac.id https://ojs.uma.ac.id/index.php/jite/article/download/10714/5438/48581	●
INTERNET SOURCE		
4.	0.2% www.trivusi.web.id https://www.trivusi.web.id/2023/03/perbedaan-mae-mse-rmse-dan-mape.html	●
INTERNET SOURCE		
5.	0.2% publikasi.uyelindo.ac.id https://publikasi.uyelindo.ac.id/index.php/hoaq/article/download/634/327/2106	●
INTERNET SOURCE		
6.	0.15% repository.mediapenerbitindonesia.com http://repository.mediapenerbitindonesia.com/301/1/%28Revisi-margin%29%2...	●
INTERNET SOURCE		
7.	0.14% pdfs.semanticscholar.org https://pdfs.semanticscholar.org/31f3/f348b12fa6a88245cc227d3c653f28165a4b..	●
INTERNET SOURCE		
8.	0.13% www.cloudeka.id https://www.cloudeka.id/id/berita/cloud/metode-pengolahan-data/	●
INTERNET SOURCE		
9.	0.11% repository.unhas.ac.id http://repository.unhas.ac.id/id/eprint/43869/	●

REPORT #27589781

INTERNET SOURCE		
10.	0.11% ejurnal.seminar-id.com https://ejurnal.seminar-id.com/index.php/bits/article/download/7399/3809/	●
INTERNET SOURCE		
11.	0.1% bookdown.org https://bookdown.org/dsciencelabs/statistika_dasar/_book/Ukuran_Penyebara...	●
INTERNET SOURCE		
12.	0.1% pdfs.semanticscholar.org https://pdfs.semanticscholar.org/33bf/1d506924c8cea532c134491ab9eccf9a753...	●
INTERNET SOURCE		
13.	0.1% publikasi.teknokrat.ac.id https://publikasi.teknokrat.ac.id/index.php/teknokompak/article/download/54/...	●
INTERNET SOURCE		
14.	0.09% www.fanruan.com https://www.fanruan.com/id/blog/jenis-jenis-grafik-yang-umum-digunakan	●
INTERNET SOURCE		
15.	0.09% jurnal.umt.ac.id https://jurnal.umt.ac.id/index.php/jt/article/download/9099/4575	●
INTERNET SOURCE		
16.	0.08% eprints.umg.ac.id http://eprints.umg.ac.id/12814/7/BAB%20II.pdf	●
INTERNET SOURCE		
17.	0.07% jsi.stikom-bali.ac.id https://jsi.stikom-bali.ac.id/index.php/jsi/article/download/193/166	●
INTERNET SOURCE		
18.	0.07% library.umy.ac.id https://library.umy.ac.id/penelitian-kuantitatif-vs-kualitatif-apa-bedanya/	●
INTERNET SOURCE		
19.	0.07% perpustakaan.bappenas.go.id https://perpustakaan.bappenas.go.id/e-library/file_upload/koleksi/migrasi-data..	●
INTERNET SOURCE		
20.	0.06% www.trivusi.web.id https://www.trivusi.web.id/2022/07/ algoritma-lstm.html	●



REPORT #27589781

INTERNET SOURCE

21. **0.06%** ojs.unikom.ac.id

<https://ojs.unikom.ac.id/index.php/jamika/article/download/12828/4372>



INTERNET SOURCE

22. **0.06%** repository.syekhnurjati.ac.id

<http://repository.syekhnurjati.ac.id/6847/2/bab%20i.pdf>



INTERNET SOURCE

23. **0.03%** eprints.umm.ac.id

<https://eprints.umm.ac.id/18230/1/TESIS%20.pdf>



INTERNET SOURCE

24. **0.03%** repository.ubb.ac.id

<https://repository.ubb.ac.id/id/eprint/3193/4/BAB%20I.pdf>



● QUOTES

INTERNET SOURCE

1. **0.18%** pdfs.semanticscholar.org

<https://pdfs.semanticscholar.org/76ac/9e24b39e2ac1599c7684fdc6a287d2067a...>

INTERNET SOURCE

2. **0.17%** journal.eng.unila.ac.id

<https://journal.eng.unila.ac.id/index.php/jitet/article/download/4827/1970>

INTERNET SOURCE

3. **0.06%** ojs.bsi.ac.id

<https://ojs.bsi.ac.id/index.php/justian/article/download/8107/2095>

INTERNET SOURCE

4. **0.05%** journal.unipdu.ac.id

<https://journal.unipdu.ac.id/index.php/jmpm/article/view/1071>