SNESTIK

Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika

https://ejurnal.itats.ac.id/snestik dan https://snestik.itats.ac.id

Informasi Pelaksanaan :
SNESTIK V - Surabaya, 26 April 2025
Fakultas Teknik Elektro dan Teknologi Informasi, Institut Teknologi Adhi Tama Surabaya

Informasi Artikel:
DOI : 10.31284/p.snestik.2025.7034

Prosiding ISSN 2775-5126

Fakultas Teknik Elektro dan Teknologi Informasi-Institut Teknologi Adhi Tama Surabaya
Gedung A-ITATS, Jl. Arief Rachman Hakim 100 Surabaya 60117 Telp. (031) 5945043
Email : snestik@itats.ac.id

Implementasi Algoritma Pengklasifikasi Long Short–Term Memory (LSTM) untuk Data Time Series

Arum I. Sari, Handi F. R. Wibowo, Chatarina N. Putri, Gamaliel V.F. Lande,

Exacta B. Aldero, Dian P. Hapsari

Teknik Informatika, Institut Teknologi Adhi Tama Surabaya
e-mail: arumindahsari03@gmail.com

ABSTRACT

Forecasting gold prices is essential for making informed financial decisions, offering valuable information for investors and stakeholders in the gold market. Gold prices tend to fluctuate due to several influencing factors, such as global economic conditions, inflation, and demand and supply. The instability causes investors to face higher risks in determining the optimal investment strategy. Thus, gold price prediction becomes an important requirement to help investors reduce losses. In recent years, deep learning methods have made significant progress in various fields, such as image recognition, sentiment analysis, and time series data prediction. This paper implements a Long and Short Term Memory (LSTM) network for data classification and then its performance is compared with a linear regression model to predict gold price fluctuations. Analysis of daily gold price prediction shows that the LSTM model achieves an accuracy rate of 88%, while the linear regression model performs slightly better with an accuracy rate of 98%. By utilising the strengths of both models, this research provides important insights for investors to improve investment strategies as well as understand future gold price trends.

Keywords: *machine learning; supervised; LSTM; linear regression.*

ABSTRAK

Meramalkan harga emas sangat penting untuk membuat keputusan keuangan yang tepat, menawarkan informasi berharga bagi investor dan pemangku kepentingan di pasar emas. Harga emas memiliki pola cenderung fluktuatif hal ini disebabkan karena beberapa faktor yang mempengaruhi, seperti kondisi ekonomi global, inflasi, dan permintaan serta penawaran. Adanya ketidakstabilan tersebut menyebabkan investor menghadapi resiko yang lebih tinggi dalam menentukan strategi investasi yang optimal. Sehingga, prediksi harga emas menjadi kebutuhan penting untuk membantu investor dalam mengurangi kerugian. Beberapa tahun terakhir, metode pembelajaran mendalam telah

membuat kemajuan signifikan di berbagai bidang, seperti pengenalan gambar, analisis sentiment, dan prediksi ddata deret waktu. Makalah ini mengimplementasikan jaringan Memori Jangka Panjang dan Pendek (LSTM) untuk klasifikasi data kemudian kinerjanya dibandingkan dengan model regresi linier untuk memprediksi fluktuasi harga emas. Analisis prediksi harga emas harian menunjukkan bahwa model LSTM mencapai tingkat akurasi 88%, sedangkan model regresi linier berkinerja sedikit lebih baik dengan tingkat akurasi 98%. Dengan memanfaatkan kekuatan kedua model, penelitian ini memberikan wawasan penting bagi investor untuk meningkatkan strategi investasi serta memahami tren harga emasi di masa depan.

Kata kunci: pembelajaran mesin; pembelajaran mesin; LSTM; regresi linier.

PENDAHULUAN

Investasi merupakan bentuk penanaman dana atau aset dengan tujuan untuk memperoleh keuntungan di masa yang akan datang. Dengan kata lain, individu melakukan aktivitas ini untuk mengalokasikan sebagian dari kekayaan mereka di masa sekarang sebagai salah satu upaya perencanaan keuangan jangka panjang. Investasi yang sering dilakukan oleh sebagian besar masyarakat yaitu investasi dalam bentuk emas [1]. Emas, sebagai salah satu jenis logam mulia selalu dianggap sebagai lambang kemakmuran dan kemewahan. Dengan berkembangnya teknologi pertambangan emas, menjadikan posisinya semakin kuat untuk dijadikan sebagai aset berharga dan sebagai cadangan mata uang di tingkat nasional dan internasional. Kestabilan harga dari emas memberikan nilai tambah sehingga harga emas dianggap sebagai instrumen proteksi terhadap krisis ekonomi, kondisi geopolitik, inflasi, dan depresiasi mata uang Rupiah. Pada bulan Januari 2020, pada saat implementasi *lockdown* akibat pandemi COVID-19, harga emas dalam hitungan USD mencapai 1,525/troy ounce dan pada bulan Desember 2021 harga emas mencapai USD 1.810/troy ounce setelah mengalami peningkatan 18,68% [2]. Guna memaksimalkan investasi emas, para investor harus membuat keputusan investasi yang tepat waktu. Oleh karena itu, upaya memprediksi pergerakan harga emas memegangi peranan penting dalam mengevaluasi strategi investasi[3].

Para ekonom dan peneliti telah secara efektif memanfaatkan metode matematika, seperti algoritma Naive Bayes, sebagai alat statistik penting untuk memprediksi tren masa depan dalam produk keuangan. Namun, sektor keuangan menimbulkan tantangan penting, termasuk tuntutan komputasi yang substansial dan sifat rumit data keuangan. Meskipun berharga, pengklasifikasi Naive Bayes menghadapi keterbatasan dalam menangani kumpulan data skala besar dan fitur berkelanjutan. Selain itu, model faktor kuantitatif tradisional sering kali kesulitan untuk menangkap hubungan yang kompleks dan nonlinier yang melekat dalam domain keuangan. Banyak faktor yang mempengaruhi estimasi hubungan antara prediktor potensial dan pengembalian yang diharapkan. Oleh karena itu, melangkah maju melampaui metodologi saat ini sangat penting untuk mencapai akurasi dan keandalan yang lebih tinggi dalam pemodelan keuangan.

Kemajuan terkini dalam bidang pembelajaran mendalam telah menarik perhatian yang signifikan, khususnya karena kinerjanya yang luar biasa dalam domain seperti visi komputer dan pemrosesan bahasa alami. Diskusi seputar penggunaan pembelajaran mendalam untuk peramalan keuangan tetap bersemangat dan terkadang kontroversial, prinsip yang mendasarinya melibatkan penciptaan hubungan antara data masukan dan hasil keuntungan finansial. Kemampuan inheren ini memberdayakan pembelajaran mendalam untuk mengungkap hubungan kompleks yang terkait dengan fluktuasi harga emas, bahkan saat ia menavigasi karakteristik yang tidak dapat diprediksi dan nonlinier yang melekat dalam data tersebut. Lebih khusus lagi, sejumlah penelitian penting telah dilakukan oleh para peneliti yang menyelidiki topik ini.

Pada penelitian studi empiris telah dilakukan dengan tujuan guna mengidentifikasi faktor-faktor yang memengaruhi harga emas untuk memperkirakannya. Analisis statistik dan ekonometrika menunjukkan berbagai variabel penjelas. Artikel ini memberikan beberapa informasi baru untuk memperkirakan harga emas berdasarkan indikator ekonomi. Metode LSTM, BiLSTM, dan GRU digunakan untuk memperkirakan harga emas [4]. Meskipun harga emas telah diperkirakan dengan berbagai metode dalam literatur, harga tersebut masih asli dalam hal variabel yang digunakan dalam penelitian. Zhang menyelidiki model jaringan saraf LSTM-P untuk memprediksi nilai Bitcoin dan emas. Pertama-tama kami Diawali dengan menggunakan pendekatan pengurangan derau berdasarkan

transformasi wavelet untuk menghaluskan fluktuasi data harga, yang telah terbukti meningkatkan akurasi prediksi [5][6].

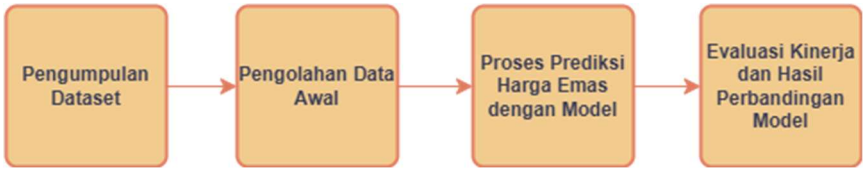
Penelitian lain juga telah dilakukan [7], penelitian tersebut menggunakan metode LSTM (*Long Short-Term Memory*) untuk memprediksi harga emas berdasarkan data dari gold.org. Hasil yang didapatkan dari penelitian yaitu metode LSTM dapat memprediksi harga emas untuk rentang waktu 1-3 hari kedepannya. Serta nilai akurasi yang diperoleh dari percobaan yang telah dilakukan dengan mengimplementasikan metode LSTM yaitu sebesar 88,84%, dengan jumlah data testing sebanyak 941 data.

Terdapat penelitian lain yang bertujuan untuk memprediksi harga emas, penelitian tersebut membandingkan antara algoritma *Support Vector Regression* (SVr) dan *Linear Regression.* Data yang digunakan pada penelitian tersebut ialah data emas harian yang bersumber dari Investing.com, dengan jumlah data sebanyak 739 dataset. Dari penelitian tersebut didapatkan hasil pengujian menggunakan MSE (*Mean Squared Error*) terkecil yang dihasilkan oleh algoritma Regresi Linier sebesar 4.04444791059, sedangkan MSE untuk *Support Vector Regression* sebesar 7.524505784357. Dengan nilai error terkecil yang dihasilkan oleh algoritma Regresi Linier, dapat dikatakan algoritma tersebut lebih optimal dibandingkan algoritma *Support Vector Regression* [8].

Berdasarkan ulasan serta penelitian-penelitian yang telah dilakukan sebelumnya, pada penelitian ini penulis mencoba untuk membandingkan akurasi dari model LSTM dan *Linear Regression*. Dengan dilakukan penelitian ini, diharapkan dapat mengetahui model mana yang menghasilkan akurasi paling optimal dalam memprediksi data time series yaitu prediksi harga emas.

METODE PENELITIAN

Pada penelitian ini terdapat beberapa tahapan yang disusun secara sistematis, dimulai dari pengumpulan dataset, pengolaha data awal, proses prediksi harga emas dengan model, dan yang terakhir yaitu evaluasi kinerja dan hasil perbandingan model. Adapun tahapan dalam penelitian ini digambarkan pada gambar 1.



Gambar 1. Tahapan Penelitian

1. Pengumpulan Dataset
   Data yang digunakan dalam penelitian ini merupakan dataset Harga Emas (*Gold Prices*) dengan rentang waktu 5 tahun terakhir, yaitu dari 2019 sampai 2024, dengan jumlah data sebanyak 1258 data yang bersumber dari Kaggle.
2. Pengolahan Data Awal
   Tahap ini bertujuan untuk memastikan bahwa data yang digunakan bersih dari *outlier, noise*, dan ketidakkonsistenan [9]. Proses pengolahan data awal adalah sebagai berikut.
   a. Validasi Data
      Proses validasi data digunakan untuk mengidentifikasi dan menangani data yang hilang (*missing value), outlier, noise*, dan ketidaksesuaian format pada data.
   b. Normalisasi Data
      Untuk menormalisasi data dapat menggunakan metode *Min-Max Normalization*, metode ini digunakan dengan tujuan untuk menyamakan skala atribut ke dalam rentang 0 hingga 1. Hal ini dapat mempercepat proses pelatihan dan meningkatkan performa model.

3. Proses Prediksi Harga Emas dengan Model
   Pada penelitian ini, model yang diusulkan atau yang digunakan yaitu *Long Short-Term Memory* (LSTM) dan Regresi Linier untuk memprediksi harga emas.
4. Evaluasi Kinerja dan Hasil Perbandingan Model
   Evaluasi digunakan untuk mengukur performa model dalam melakukan prediksi. Pada penelitian ini, pengukuran akurasi model menggunakan metrik *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE).

Prediksi

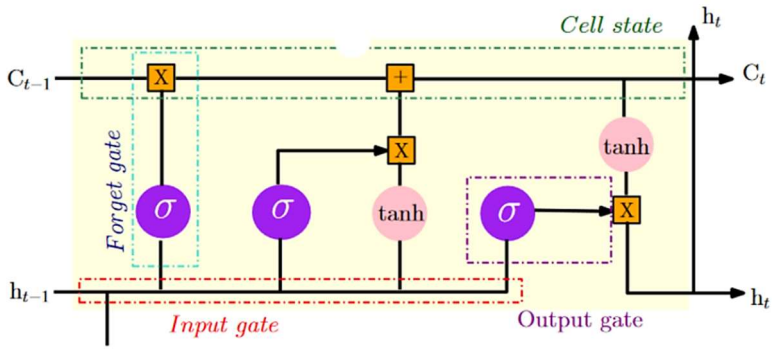
Prediksi atau peramalan merupakan proses estimasi atau memperkirakan kondisi di masa mendatang yang berlandaskan pada data historis dan informasi aktual. Secara teknis, prediksi dapat diartikan sebagai proses yang menghasilkan data dalam bentuk informasi prediksi atau perkiraan di masa mendatang. Prediksi digunakan untuk memberikan gambaran atau informasi tren atau perkembangan yang dapat dipakai sebagai dasar pengambilan keputusan di berbagai bidang, seperti ekonomi, bisnis, dan keuangan [10].

Prediksi umumnya dibagi menjadi dua, yaitu prediksi yang bersifat kualitatif dan prediksi yang bersifat kuantitatif. Prediksi kualitatif sifatnya subjektif karena biasanya berdasarkan opini atau intuisi dari ahli. Kualitas hasil dari prediksi tersebut lebih bervariasi dan tidak dapat diukur secara objektif. Sedangkan prediksi yang bersifat objektif atau kuantitatif menggunakan pendekatan matematis atau statistika untuk menghasilkan sebuah keluaran dalam bentuk angka atau bilangan. Hasil dari prediksi kuantitatif sangat dipengaruhi oleh metode yang digunakan, karena perbedaan metode dapat menghasilkan tingkat akurasi yang berbeda-beda [11].

***Long Short-Term Memory* (LSTM)**

Memori jangka pendek atau Long Short-Term Memory (LSTM) adalah metode untuk mengembangkan jaringan saraf tipe jaringan saraf berulang atau Recurrent Neural Network (RNN), pertama kali diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997 [12]. Pengembangan RNN di LSTM terletak pada mekanisme gerbang khusus untuk mengendalikan akses sel memori. Sel memori sebagai area penyimpanan informasi akan mendapatkan informasi yang dipilih oleh mekanisme gerbang dalam LSTM. Menurut Kalchbrenner et al., mekanisme gerbang dalam LSTM terdiri dari tiga unit gerbang vektor, gerbang input, gerbang lupa (forget gate), dan gerbang output [13]. Setiap unit gerbang vektor memiliki fungsi yang berbeda. Gerbang vektor input berfungsi untuk mengontrol jumlah vektor input yang akan mempengaruhi memori. Gerbang vektor lupa mengontrol jumlah memori lama yang akan dihapus. Gerbang vektor output mengontrol jumlah memori yang disimpan dalam keadaan tersembunyi [14].

Saat menerapkan model LSTM untuk tugas prediksi, beberapa faktor harus dipertimbangkan antara lain, pertama persiapan data, data input harus disusun dengan tepat, biasanya sebagai urutan di mana setiap sampel terdiri dari beberapa langkah waktu. Ini memungkinkan LSTM untuk belajar dari pengamatan masa lalu secara efektif. Tahap kedua proses pelatihan, model ini dilatih menggunakan data historis dengan fokus pada meminimalkan perbedaan antara nilai – nilai yang diprediksi dan aktual. Pelatihan ini melibatkan penyesuaian parameter seperti bobot dan bias melalui backpropagation. Dan tahap evaluasi yang dilakukan setelah pelatihan, penting untuk memvalidasi kinerja model menggunakan data uji terpisah untuk memastikannya menggeneralisasi dengan baik ke data yang tidak terlihat. Singkatnya, LSTM mewakili alat yang kuat untuk tugas prediksi urutan karena arsitektur canggih mereka yang mengatasi tantangan umum yang dihadapi oleh RNN tradisional. Fleksibilitas mereka di berbagai domain menjadikan mereka pilihan populer dalam aplikasi pembelajaran mesin.



Gambar 2. Tahapan LSTM

LSTM telah berhasil diterapkan di berbagai domain karena kemampuannya untuk menangani data berurutan seperti peramalan seri waktu atau Time Series Forcasting, LSTM dapat memprediksi nilai -nilai di masa depan berdasarkan data historis, membuatnya ideal untuk aplikasi seperti prediksi harga saham, peramalan cuaca, dan peramalan konsumsi energi. Pemrosesan bahasa alami atau Natural Language Processing, mereka banyak digunakan dalam tugas – tugas seperti terjemahan mesin, ringkasan teks, dan analisis sentimen dengan menangkap ketergantungan antara kata -kata dalam kalimat. LSTM juga diimplementasikan pada pengenalan pembicaraan atau Speech Recognition, LSTM membantu menyalin bahasa lisan menjadi teks dengan pola belajar dalam sinyal audio.

Pada LSTM terdapat beberapa komponen penting sebagai berikut; Keadaan - 657 -egative cell state ©: keadaan sel bertindak sebagai memori jangka - 657 -egativ dari sel LSTM, yang memungkinkannya untuk membawa informasi yang relevan di seluruh - 657 -egativ waktu. Ini diperbarui secara selektif berdasarkan input dan keadaan sel sebelumnya, memungkinkan jaringan untuk mempertahankan dependensi jangka - 657 -egativ dalam data. Berikutnya gerbang input, gerbang input mengontrol berapa banyak informasi baru dari input saat ini harus ditambahkan ke keadaan sel. Ini memproses input saat ini dan keadaan tersembunyi sebelumnya melalui fungsi aktivasi sigmoid, menghasilkan nilai antara 0 dan 1 yang menunjukkan berapa banyak informasi untuk masuk ke keadaan sel. Gerbang ini dapat ditulis dalam bentuk model seperti pada persamaan (1) berikut;

$$i_t = \sigma(W_i \cdot [h_t - 1, x_t] + bi) \quad (1)$$

Gerbang Lupakan atau forget gate merupakan gerbang yang menentukan informasi mana dari keadaan sel yang harus dibuang. Dibutuhkan keadaan tersembunyi sebelumnya dan input saat ini, menerapkan matriks bobot dan bias, dan melewati hasilnya melalui fungsi aktivasi sigmoid untuk menghasilkan output biner yang menunjukkan apakah akan menyimpan atau melupakan informasinya. Gerbang ini dapat ditulis dalam bentuk model seperti pada persamaan (2) berikut;

$$f_t = \sigma(W_f \cdot [h_t - 1, x_t] + b_f) \quad (2)$$

Gerbang Output: Gerbang ini mengontrol informasi apa yang dikirim ke lapisan atau output berikutnya. Ini menggunakan keadaan sel yang diperbarui untuk menentukan output. Gerbang output memutuskan berapa banyak keadaan sel yang harus diekspos sebagai output. Ini menggabungkan input saat ini dan keadaan tersembunyi sebelumnya, memprosesnya melalui fungsi sigmoid, dan melipatgandakan output ini dengan tanh dari keadaan sel untuk menghasilkan output akhir dari sel LSTM. Gerbang ini dapat ditulis dalam bentuk model seperti pada persamaan (3) berikut;

$$o_t = \sigma(W_o \cdot [h_t - 1, x_t] + b_o) \quad (3)$$

Bagian akhir merupakan komponen Kandidat status sel, komponen ini sering dianggap sebagai bagian dari gerbang input, komponen ini menghasilkan nilai kandidat baru yang dapat ditambahkan ke keadaan sel. Ini menerapkan fungsi aktivasi TANH untuk membuat nilai yang kemudian ditingkatkan oleh gerbang input sebelum ditambahkan ke keadaan sel yang ada [15]. Komponen – komponen ini bekerja - 658 -egativ dalam sel LSTM untuk memungkinkan retensi memori selektif dan pemrosesan informasi, membuat LSTM sangat efektif untuk tugas -tugas yang melibatkan data berurutan seperti peramalan deret waktu, pemrosesan - 658 -egati alami, dan banyak lagi. Fungsi aktivasi TANH memainkan peran penting dalam sel memori jangka pendek (LSTM), terutama dalam mengelola keadaan sel dan mengatur aliran informasi.

Peran kunci dari fungsi aktivasi TANH dalam sel LSTM antara lain, Mengatur Nilai Negara Sel: Fungsi TANH digunakan untuk membuat nilai kandidat untuk memperbarui keadaan sel. Ini mengubah input menjadi kisaran antara -1 dan 1, yang membantu mempertahankan nilai -nilai dalam kisaran yang dapat dikelola. Ini sangat penting karena mencegah keadaan sel menjadi terlalu besar atau kecil, memfasilitasi pembelajaran yang lebih baik selama pelatihan. Transformasi non-linear: Tanh memperkenalkan non-linearitas ke dalam model, memungkinkan LSTM untuk mempelajari pola kompleks dalam data berurutan. Dengan memberikan gradien yang mulus, ia membantu dalam backpropagation yang efisien melalui waktu (BPTT), yang sangat penting untuk melatih jaringan berulang. Menggabungkan dengan output sigmoid: Sehubungan dengan fungsi sigmoid yang digunakan dalam gerbang input dan output, TANH memungkinkan LSTM untuk mengontrol berapa banyak informasi yang ditambahkan atau dihapus dari keadaan sel.

Sementara nilai output sigmoid antara 0 dan 1 (bertindak sebagai gerbang), Tanh memungkinkan nilai positif dan - 658 -egative, memungkinkan meningkat dan berkurang dalam keadaan sel. Memfasilitasi pembaruan memori: Kombinasi fungsi tanh dan sigmoid memungkinkan LSTM untuk secara efektif mengelola memori dengan menentukan informasi apa yang harus disimpan atau dilupakan. Kemampuan fungsi TANH untuk mengeluarkan nilai - 658 -egative membantu dalam menyesuaikan keadaan sel ke bawah bila perlu, sedangkan gerbang sigmoid mengontrol berapa banyak penyesuaian ini terjadi. Singkatnya, fungsi aktivasi TANH merupakan bagian integral dari sel LSTM karena memastikan bahwa nilai-nilai keadaan sel tetap dalam kisaran yang wajar, memperkenalkan non-linearitas untuk pembelajaran yang efektif, dan bekerja di samping fungsi sigmoid untuk mengelola aliran informasi dalam jaringan.

Dalam penelitian ini akan menggunakan kelompok data time series, kumpulan data deret waktu memiliki karakteristik berbeda yang membedakannya dari jenis data lainnya. Karakteristik Utama Kumpulan Data Deret Waktu antara lain, Stempel Waktu: Setiap titik data dalam kumpulan data deret waktu dikaitkan dengan stempel waktu, yang menunjukkan kapan pengamatan direkam. Urutan kronologis ini penting untuk analisis karena memungkinkan pelacakan perubahan dari waktu ke waktu dan digunakan untuk mengindeks data. Struktur: Data deret waktu biasanya terstruktur, sering kali dengan tipe data yang telah ditentukan sebelumnya atau panjang yang tetap. Sifat terstruktur ini memudahkan pemrosesan dan analisis. Kemudian Sifat seperti Aliran: Data deret waktu dapat dilihat sebagai aliran informasi berkelanjutan yang dikumpulkan dari waktu ke waktu, sering kali dari berbagai sumber seperti sensor atau log. Data ini independen dan terus mengalir ke dalam basis data.

Fokus Tren: Menganalisis tren dari waktu ke waktu lebih penting daripada berfokus pada nilai individual pada stempel waktu tertentu. Analisis deret waktu sering kali bertujuan untuk mengidentifikasi pola atau tren yang bertahan meskipun terjadi fluktuasi pada data. Rasio Tulis atau Baca Tinggi: Data deret waktu dicirikan oleh frekuensi tulis (data yang direkam) yang tinggi dibandingkan dengan baca (data yang dianalisis). Ini berarti bahwa meskipun data terus diperbarui, data tersebut lebih jarang dibaca untuk tujuan analisis atau pelaporan. Kekekalan:Data deret waktu umumnya tidak dapat diubah, artinya setelah direkam, data tersebut tidak berubah. Entri baru ditambahkan ke kumpulan data yang ada daripada mengubah entri sebelumnya, memastikan catatan historis yang andal.

Jenis Data: Kumpulan data deret waktu dapat diklasifikasikan menjadi univariat (variabel tunggal yang diukur dari waktu ke waktu), multivariat (beberapa variabel terkait yang diukur dari waktu ke waktu), dan beberapa deret waktu (pengukuran beberapa entitas independen). Pola dan Komponen: Data deret

waktu sering kali menunjukkan pola seperti tren (pergerakan jangka panjang), musiman (fluktuasi reguler), dan residual (gangguan acak). Memahami komponen-komponen ini sangat penting untuk peramalan dan analisis yang efektif. Interval Reguler vs. Tidak Reguler: Data deret waktu dapat dikumpulkan pada interval reguler (misalnya, pembacaan suhu per jam) atau interval tidak teratur (misalnya, peristiwa yang dicatat saat terjadi). Perbedaan ini memengaruhi bagaimana data dapat dimodelkan dan dianalisis.

Ketergantungan Serial: Nilai-nilai dalam deret waktu sering kali bergantung secara statistik pada nilai-nilai sebelumnya, yang menunjukkan autokorelasi. Properti ini memungkinkan prediksi berdasarkan pengamatan historis, menjadikan analisis deret waktu berharga untuk meramalkan hasil di masa mendatang. Singkatnya, karakteristik unik dari kumpulan data deret waktu—seperti sifatnya yang terstruktur, fokus pada tren, rasio tulis/baca yang tinggi, dan ketergantungan serial—membuatnya penting untuk berbagai aplikasi termasuk peramalan, deteksi anomali, dan pemahaman dinamika temporal dalam bidang-bidang yang digerakkan oleh data.

Regresi Linier

Regresi linier adalah salah satu metode dalam statistika untuk memprediksi nilai suatu variabel dengan cara membuat hubungan matematis antara variabel dependen (Y) dan variabel independen (X) [16]. Variabel dependen adalah variabel yang nilainya dipengaruhi oleh variabel lain. Sedangkan variabel independen adalah variabel yang memengaruhi. Apabila nilai dari variabel independen sudah diketahui, maka kita bisa memperkirakan atau memprediksi nilai dari variabel dependen [17].

Regresi linier menjadi salah satu metode yang populer digunakan di industri praktisi dalam memprediksi karakteristik baik kualitas ataupun kuantitas dari produk [18]. Hal ini dikarenakan dengan menggunakan regresi linier, perusahaan dapat memperkirakan jumlah kombinasi yang optimal untuk memproduksi sebuah produk, sehingga perusahaan dapat memaksimalkan profit dan juga dapat memperkirakan jumlah produksi dengan tepat [19].

Adapun rumus untuk Regresi Linier dengan metode kuadrat terkecil atau sederhana seperti pada persamaan (4)-(6) berikut:

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \quad (4)$$

$$b = \frac{\mathrm{n}(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (5)$$

$$y = \mathrm{a} + \mathrm{b.x} \quad (6)$$

Rumus di atas menjelaskan bahwa variabel y merupakan variabel dependen atau output yang diprediksi, variabel x merupakan variabel independe atau sebagai input (dapat berupa waktu), a merupakan konstanta, dan b merupakan koefisien regresi.

RMSE (*Root Mean Square Error*)

Ketepatan dari suatu hasil peramalan atau prediksi perlu dilakukan evaluasi, hal ini bertujuan guna mengetahui seberapa akurat atau seberapa besar kesalahan prediksi antara data aktual atau data yang sebenarnya dengan data yang diprediksi. Salah satu cara yang paling umum untuk menghitung tingkat kesalahan prediksi adalah dengan menggunakan *Root Mean Square Error* (RMSE). *Root Mean Square Error* merupakan cara yang paling umum untuk menerapkan ukuran statistik yang memperlihatkan betapa dekat hasil prediksi model dengan kebenaran atau sesuai dengan data aktual [20]. Adapun rumus dari RMSE seperti pada persamaan (7) berikut.

$$RMSE = \sqrt{(MSE)} \quad (7)$$

Keterangan:

MSE adalah *Mean Squared Error*

√ adalah fungsi akar kuadrat

Metode ini adalah akar rata-rata dari selisih kuadrat antara nilai yang diprediksi oleh model dan dari nilai pengamatan yang sebenarnya. RMSE sering digunakan dalam pemodelan data numerik karena mampu memberikan nilai kesalahan dengan tepat, dapat berupa nilai kecenderungan atau tren kesalahan dan dapat menunjukkan sejarah variabilitas kesalahan yang terjadi. Jika nilai RMSE semakin kecil, maka tingkat akurasi model dalam melakukan prediksi semakin baik. Oleh karena itu, RMSE dengan nilai kecil menjadi salah satu pertimbangan untuk penentuan model peramalan yang akan digunakan [21].

MAE (*Mean Absolute Error*)

Mean Absolute Error (MAE) merupakan salah satu metrik evaluasi yang sering digunakan dalam bidang pembelajaran mesin, dengan tujuan yaitu untuk menilai seberapa akurat atau tepat hasil prediksi dari model yang digunakan terhadap data sebenarnya (aktual). MAE menghitung rata-rata dari nilai mutlak (*absolute*) selisih antara hasil prediksi dengan nilai sebenarnya, tanpa memperhitungkan arah kesalahan, baik itu bernilai positif maupun negatif [22]. Adapun rumus dari MAE seperti pada persamaan (8) berikut.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - F(X_i)| \quad (8)$$

Keterangan:

n : jumlah total observasi.

Y_i : nilai aktual (sebenarnya).

$F(X_i)$: nilai prediksi dari model untuk observasi ke-i.

$Y_i - F(X_i)$: selisih *absolute* antara nilai aktual dan nilai prediksi.

$\frac{1}{n}$: kebalikan dari jumlah total observasi yang merupakan rata-rata dari selisih *absolute*.

$\sum_{i=1}^{n} |Y_i - F(X_i)|$: menunjukkan operasi penjumlahan dari semua selisih *absolute*.

MAE mengukur rata-rata besarnya kesalahan prediksi dalam satuan yang sama dengan variabel yang diamati. Semakin kecil nilai MAE yang dihasilkan, maka semakin tinggi tingkat akurasi model dalam melakukan prediksi atau dapat dikatakan bahwa model tersebut dapat memprediksi dengan baik.

HASIL DAN PEMBAHASAN

Dataset

Kelompok data yang digunakan pada penelitian itu berupa dataset numerik dari harga Emas (*Gold Prices*), yang bersumber dari Kaggle. Dataset tersebut berjumlah 1258 data numerik, dengan rentang waktu 5 tahun terakhir (2019 – 2024). Dataset yang digunakan terdiri dari beberapa atribut antara lain "Date", "Open", "High", "Low", "Close", "Dividends", "Stock Split", dan "Capital Gains" dari harga Emas. Kelompok data yang digunakan dapat diakses pada Alamat berikut ini https://www.kaggle.com/datasets/kusumakar/gold-prices-for-5-years-financial-predictions/data. Pada kelompok data yang digunakan dalam penelitian ini, harga emas yang digunakan dalam satuan US Dollar. Berikut adalah data harga emas dari tahun 2019 – 2024 yang dapat dilihat pada **Tabel 1**.

Tabel 1. Dataset Harga Emas (*Gold Prices*)

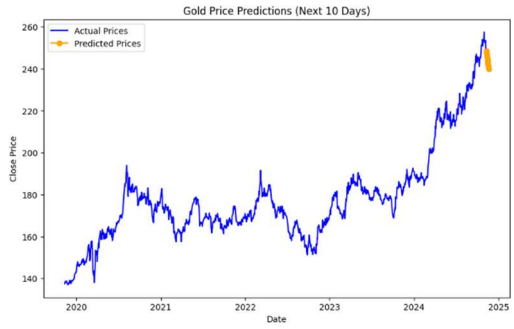
No.	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits	Capital Gains

1	2019-11-11 00:00:00-05:00	Rp13.761.000.061.035.100	Rp1.377.899.932.861.320,00	Rp13.644.000.244.140.600,00	Rp13.705.999.755.859.300,00	7037300	0.0	0.0	0.0
2	2019-11-12 00:00:00-05:00	13.702.999.877.929.600	Rp13.758.999.633.789.000,00	Rp13.619.000.244.140.600,00	Rp13.742.999.267.578.100,00	6448600	0.0	0.0	0.0
3	2019-11-13 00:00:00-05:00	1.378.000.030.517.570	Rp13.822.000.122.070.300,00	Rp13.763.999.938.964.800,00	Rp13.797.999.572.753.900,00	8776000	0.0	0.0	0.0
4	2019-11-14 00:00:00-05:00	13.838.999.938.964.800	Rp13.894.000.244.140.600,00	Rp1.378.699.951.171.870,00	Rp13.855.999.755.859.300,00	5220500	0.0	0.0	0.0
…	……..	…….	…….	…….	……..	…….	…….	…….	…….
1256	2024-11-05 00:00:00-05:00	25.377.999.877.929.600	25.399.000.549.316.400	25.241.000.366.210.900	25.339.999.389.648.400	6029900	0.0	0.0	0.0
1257	2024-11-06 00:00:00-05:00	24.563.999.938.964.800	24.735.000.610.351.500	2.449.499.969.482.420	2.456.999.969.482.420	14600200	0.0	0.0	0.0
1258	2024-11-07 00:00:00-05:00	2.482.899.932.861.320	25.024.000.549.316.400	2.481.199.951.171.870	24.964.999.389.648.400	8821400	0.0	0.0	0.0
1259	2024-11-08 00:00:00-05:00	24.907.000.732.421.800	24.935.000.610.351.500	24.777.999.877.929.600	2.479.600.067.138.670	6166900	0.0	0.0	0.0

Pengujian Model

Dalam penelitian ini tahapan ujicoba diawali dengan kegiatan inisialisasi hiperparameter LSTM yang akan diproses. Diawali dengan menetukan nilai learning rate yang akan digunakan dalam penelitian ini adalah 0,01; 0,001; dan 0,0001; serta epoch dengan nilai 100, 500, dan 1000. Selanjutnya menghitung nilai rata-rata RMSE dari setiap kombinasi hiperparameter untuk setiap skenario dan mengambil hiperparameter terbaik. Menangani masalah *underfitting* dan *overfitting* yang sering ditemui dalam penggunaan LSTM dengan penghentian awal. Penghentian awal akan menghentikan proses lebih cepat ketika terjadi penurunan validasi dan mempertahankan kondisi optimal tanpa menjalankan seluruh epoch.

Regresi Linier

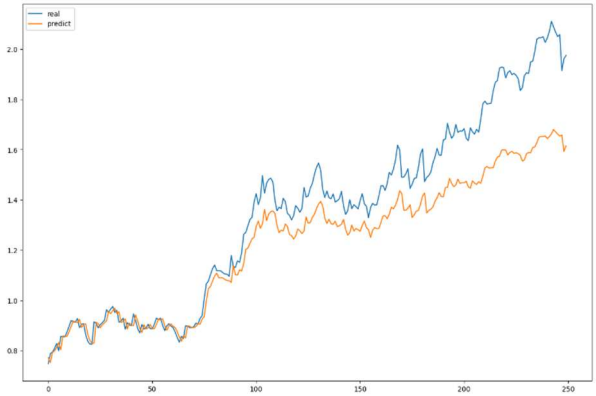


Gambar 3. Grafik Prediksi Harga Emas Regresi Linier

Sebagai pembanding dilakukan ujicoba menggunakan algoritma Regresi Linier dengan menentukan nilai Intercept 0.00390555, Training Loss 0.0002, dan Validation Loss 0.0001.

Gambar 3 merupakan grafik hasil dari pengujian prediksi harga emas menggunakan algoritma Regresi Linier. Pada grafik, garis biru menggambarkan data harga emas yang sebenarnya dari tahun 2020 hingga awal 2025, sedangkan titik oranye menunjukkan hasil prediksi untuk sepuluh hari ke depan. Model regresi linier ini mampu menghasilkan prediksi yang sejalan dengan pola tren dari data aktual, terutama di periode akhir grafik. Dengan demikian, prediksi ini memberikan estimasi harga emas jangka pendek yang sejalan dengan tren kenaikan harga yang terlihat dalam data historis.

Long Short-Term Memory (LSTM)



Gambar 4. Grafik Prediksi Harga Emas LSTM

Gambar 4 merupakan Grafik prediksi harga emas dengan model LSTM, grafik tersebut memperlihatkan perbandingan antara nilai aktual dan prediksi harga emas yang dihasilkan melalui model LSTM. Garis biru merepresentasikan harga emas yang sebenarnya, sedangkan garis oranye menunjukkan hasil prediksi. Pada tahap awal, prediksi yang dihasilkan model cukup akurat dan mampu mengikuti pola harga aktual dengan baik. Namun, seiring berjalannya waktu, terlihat adanya peningkatan deviasi, terutama pada periode di mana harga mengalami lonjakan yang lebih signifikan. Secara keseluruhan, model LSTM dapat mencapai tingkat akurasi sebesar 88% dalam memprediksi harga emas harian, menunjukkan kemampuan model ini dalam menangkap pola tren penting meskipun terdapat sedikit perbedaan pada beberapa titik.

Evaluasi Hasil Pengujian

Pengukuran kinerja model yang dihasilkan oleh pengklasifikasi dapat menggunakan beberapa jenis ukuran. Dalam analisis regresi linier, baik Mean Squared Error (MSE) maupun Root Mean Squared Error (RMSE) berfungsi sebagai ukuran untuk mengevaluasi kinerja model, khususnya terkait seberapa baik model memprediksi variabel dependen. Akan tetapi, RMSE sering kali lebih disukai daripada MSE karena beberapa alasan, terutama karena sifatnya yang mudah ditafsirkan. Istilah "Mean Squared Error" mengacu pada ukuran statistik yang digunakan untuk menilai perbedaan kuadrat rata-rata antara nilai prediksi dan nilai aktual. Ini adalah metrik umum untuk mengevaluasi keakuratan model dalam analisis regresi, yang menunjukkan seberapa dekat prediksi dengan hasil sebenarnya. Dengan mengkuadratkan kesalahan, ini menekankan perbedaan yang lebih besar, yang dapat berguna dalam mengidentifikasi ketidak-akuratan prakiraan yang signifikan.

Tabel 2. Perbandingan Kinerja Model

MODEL	DATA	RMSE	MAE

Linear Regression	Training	1.64	1.19
	Testing	2.17	1.69
LSTM	Training	1.72	1.25
	Testing	9.61	7.06

Hasil prediksi model pengklasifikasi, seperti hasil yang ditunjukkan pada **Tabel 2** diatas. Analisis akurasi prediksi mengenai kenaikan dan penurunan harga emas harian menunjukkan bahwa model LSTM mencapai tingkat akurasi sebesar 88%. Hal ini menunjukkan bahwa model LSTM memiliki tingkat keberhasilan sedang dalam memprediksi dengan tepat perubahan arah harga emas. Model LSTM memiliki RMSE sebesar 1,72 pada kegiatan training dan 9.61 pada kegiatan testing. Nilai MAE model LSTM sebesar 1.25 saat kegiatan training dan 7.06 pada kegiatan testing. MAE merupakan metode yang digunakan untuk mengukur tingkat keakuratan model peramalan. Nilai MAE yang kecil mengindikasikan kinerja model yang dihasilkan pengklasifikasi cukup baik.

Model Regresi linier memiliki RMSE 1.64 pada kegiatan training dan 2.17 pada kegiatan testing. RMSE adalah deviasi standar dari residual, yang merupakan jarak antara garis regresi dan titik data. Saat sebuah model memiliki nilai RSME yang kecil, hal tersebut mengindikasikan kinerja model yang baik. Nilai RMSE dari model yang dihasilkan pengklasifikasi regresi linier lebih kecil dibandingkan dengan model yang dihasilkan pengklasifikasi LSTM. Nilai MAE model Regresi linier sebesar 1.19 saat kegiatan training dan 1.69 pada kegiatan testing. Terlihat bahwa akurasi model Regresi Linier mencapai sedikit lebih tinggi. Hal ini menunjukkan bahwa model Regresi Linier memiliki tingkat keberhasilan yang sedikit lebih tinggi dalam memprediksi perubahan arah harga emas dengan tepat dibandingkan dengan model LSTM.

KESIMPULAN

Sebagai kesimpulan, penelitian yang dilakukan dalam studi ini menggunakan pengklasifikasi Regresi Linier dan model LSTM untuk peramalan atau memprediksi tren harga emas secara akurat. Berdasarkan hasil dari penelitian yang telah dilakukan, tingkat akurasi dari model LSTM sebesar 88%, sedangkan pada tingkat akurasi Regresi Linier jauh lebih tinggi yaitu sebesar 98%. Dari tingkat akurasi yang dihasilkan mengindikasikan bahwa, untuk memprediksi tren jangka panjang dari harga emas secara akurat, dapat menggunakan model LSTM. Kemampuannya untuk menangkap ketergantungan temporal dan pola rumit dapat menguntungkan untuk tujuan peramalan tersebut. Di sisi lain, ketika fokusnya adalah pada penilaian detail fluktuasi harga emas yang lebih halus, seperti memprediksi perubahan jangka pendek atau menganalisis pergerakan harga tertentu, model Regresi Linier mungkin lebih cocok.

Secara umum, penelitian ini memberikan wawasan berharga tentang kekuatan dan keterbatasan model LSTM dan Regresi Linier untuk peramalan harga emas. Hasil penelitian mengungkap beberapa keterbatasan yang terkait dengan kedua model. Salah satu keterbatasan yang menonjol adalah keterbatasannya dalam memprediksi harga dalam jangka waktu pendek. Untuk mengatasi ketidaksempurnaan ini, penelitian di masa mendatang harus memprioritaskan pengembangan model yang mengatasi kekurangan yang teridentifikasi. Hal ini dapat melibatkan peningkatan arsitektur dan proses pelatihan model LSTM untuk meningkatkan kinerjanya dalam menangkap pola temporal yang rumit dalam data harga emas.

DAFTAR PUSTAKA

[1] N. Dwi and S. B. Pengembangan, “Penerapan Algoritma Support Vector Machine untuk Prediksi Harga Emas,” 2015.

[2] B. Pengelola and K. Haji, *Investasi Emas BPKH*. 2022. [Online]. Available: www.bpkh.go.id

[3] C. Tanudy, T. Handhayani, and J. Hendryli, “Prediksi Harga Emas di Indonesia Menggunakan Gated Recurrent Unit,” *JURNAL FASILKOM*, 2023.

[4] M. Yurtsever, “Gold Price Forecasting Using LSTM , Bi-LSTM and GRU LSTM , Bi-LSTM ve GRU ile Altın Fiyatı Tahmini,” vol. 31, no. 31, pp. 341–347, 2021, doi: 10.31590/ejosat.959405.

[5] X. Zhang, L. Zhang, Q. Zhou, and X. Jin, “A Novel Bitcoin and Gold Prices Prediction Method Using an LSTM-P Neural Network Model,” *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/1643413.

[6] J. Lei and Q. Lin, “Analysis of gold and bitcoin price prediction based on LSTM model,” *Academic Journal of Computing & Information Science*, vol. 5, no. 6, pp. 95–100, 2022, doi: 10.25236/ajcis.2022.050614.

[7] V. Riandaru Prasetyo *et al.*, “SISTEMASI: Jurnal Sistem Informasi Prediksi Harga Emas Berdasarkan Data gold.org menggunakan Metode Long Short Term Memory Gold Price Prediction Based on Gold.org Data using the Long Short Term Memory Method,” 2022. [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[8] A. Aulia, B. Aprianti, Y. Supriyanto, and C. Rozikin, “Prediksi Harga Emas dengan Menggunakan Algoritma Support Vector Regression (Svr) dan Linear Regression (LR),” *Jurnal Ilmiah Wahana Pendidikan*, Apr. 2022.

[9] S. Handayani, Taslim, and D. Toresa, “Convolutional Neural Network – Long Short Term Memory Untuk Prediksi Harga Emas Indonesia,” *Indonesian Journal of Computer Science*, 2022.

[10] L. Harianti Hasibuan, S. Musthofa, P. Studi Matematika, and U. Imam Bonjol Padang, “Journal of Science and Technology Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang,” 2022.

[11] Y. Hakimah, “ANALISIS KEBUTUHAN ENERGI LISTRIK DANPREDIKSI PENAMBAHAN PEMBANGKIT LISTRIK DI SUMATERA SELATAN,” 2019.

[12] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, 1997, doi: 10.1162/neco.1997.9.8.1735.

[13] N. Kalchbrenner, I. Danihelka, and A. Graves, “Grid long short-term memory,” *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, pp. 1–15, 2016.

[14] J. Lasenby, “T HE UNREASONABLE EFFECTIVENESS OF THE FORGET,” pp. 1–15, 2015.

[15] N. G. Timmons, “Approximating Activation Functions”.

[16] P. Juanta, S. Tamba Parsaoran, W. Purba, Y. Ferdinand Zai, and E. Ghozali, “PENERAPAN METODE REGRESI LINEAR MEMPREDIKSI TINGKAT PENJUALAN SEPEDA MOTOR HONDA PADA PT. PLATINA MULIA ABADI,” *Jurnal TEKINKOM*, vol. 6, no. 2, p. 2023, 2023, doi: 10.37600/tekinkom.v6i2.941.

[17] T. Indarwati, T. Irawati, and E. Rimawati, “PENGGUNAAN METODE LINEAR REGRESSION UNTUK PREDIKSI PENJUALAN SMARTPHONE,” *Jurnal Teknologi Informasi dan Komunikasi (TIKomSiN)*, vol. 6, no. 2, Jan. 2019, doi: 10.30646/tikomsin.v6i2.369.

[18] G. N. Ayuni and D. Fitrianah, "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ," *Jurnal Telematika*, vol. 14, no. 2, 2019.

[19] Princessa Sissy, Maharani Aisadira, Dyta Khairani, Fadhila Syahda, and Nur Aini Rakhmawati, "Pengaruh Marketing Digital Shopee Affiliate pada Minat Beli Mahasiswa Sistem Informasi ITS," *Jurnal Manuhara : Pusat Penelitian Ilmu Manajemen dan Bisnis*, vol. 2, no. 1, pp. 182–197, Nov. 2023, doi: 10.61132/manuhara.v2i1.447.

[20] M. Muharrom, "Analisis Penggunaan Orange Data Mining untuk Prediksi Harga USDT/BIDR Binance," *Bulletin of Information Technology (BIT)*, vol. 4, no. 2, pp. 178–184, 2023, doi: 10.47065/bit.v3i1.

[21] E. Fitri, "ANALISA PERBANDINGAN MODEL PREDICTION DALAM PREDIKSI HARGA SAHAM PADA TIGA SEKTOR INDUSTRI DI INDONESIA`TESIS," 2020.

[22] I. Amansyah, J. Indra, E. Nurlaelasari, and A. R. Juwita, "Prediksi Penjualan Kendaraan Menggunakan Regresi Linear: Studi Kasus pada Industri Otomotif di Indonesia," *INNOVATIVE: Journal Of Social Science Research*, vol. 4, pp. 1199–1216, 2024.

Halaman ini sengaja dikosongkan