



Perbandingan Kinerja Metode Linear Regression, LSTM dan GRU untuk Prediksi Harga Penutupan Saham Coca-Cola

Rosalia Natal Silalahi¹, Muljono^{2*}

^{1,2)} Program Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
Jl. Imam Bonjol No. 207, Pendrikan Kidul, Kec. Semarang Tengah, Kota Semarang, Jawa Tengah 50131

*email: muljono@dsn.dinus.ac.id

(Naskah masuk: 31 Januari 2024; direvisi: 11 September 2024; diterima untuk diterbitkan: 28 September 2024)

ABSTRAK – Dalam pasar saham, membuat prediksi tentang pergerakan harga saham sangat krusial bagi para pelaku transaksi, karena hal ini akan memengaruhi potensi keuntungan atau kerugian mereka. Tingkat keakuratan hasil prediksi sangat bergantung pada metode yang digunakan dan kualitas data yang tersedia. Oleh karena itu, penelitian ini memilih subjek prediksi harga saham Coca-Cola. Penelitian ini akan melakukan perbandingan antara beberapa metode analisis data time series yang berbeda. Metode-metode ini mencakup Linear Regression, LSTM, dan GRU. Perbandingan ke-tiga metode dengan variasi window-width 3, 4, dan 5 memberikan wawasan mendalam terhadap kinerja masing-masing model. Hasil perbandingan menunjukkan bahwa model mencapai kinerja terbaik ketika menggunakan window-width=3 pada metode Linear Regression. Linear Regression dengan MSE sebesar 0.21, RMSE sebesar 0.46, menunjukkan kinerja yang lebih baik dibandingkan dengan LSTM (3.24 & 1.80) dan GRU (0.40 & 0.46). penelitian ini memberikan panduan berharga untuk pengembangan model prediktif di masa depan, dengan fokus meningkatkan akurasi dan ketepatan prediksi harga saham.

Kata Kunci –LSTM; GRU; Linear Regression; Forecasting; Prediksi Harga.

Performance Comparison Of Linear Regression, LSTM & GRU Methods For Coca-Cola Stock Closing Price Prediction

ABSTRACT – In the stock market, making predictions about stock price movements is crucial for traders, as this will affect their potential profits or losses. The accuracy of the prediction results largely depends on the method used and the quality of the data available. Therefore, this research chooses the subject of predicting the stock price of Coca-Cola. This research will conduct a comparison between several different time series data analysis methods. These methods include Linear Regression, LSTM, and GRU. The comparison of the three methods with window-width variations of 3, 4, and 5 provides an in-depth insight into the performance of each model. The comparison results show that the model achieves the best performance when using window-width=3 in the Linear Regression method. Linear Regression with MSE of 0.21, RMSE of 0.46, shows better performance compared to LSTM (3.24 & 1.80) and GRU (0.40 & 0.46). This research provides valuable guidance for future predictive model development, with a focus on improving the accuracy and precision of stock price predictions.

Keywords - LSTM; GRU; Linear Regression; Forecasting; Price Prediction.

1. PENDAHULUAN

Definisi prediksi, yang dapat bervariasi tergantung pada situasi atau konteks tertentu, umumnya diakui sebagai tindakan meramal atau menebak-nebak. Menurut Kamus Besar Bahasa Indonesia, prediksi adalah output dari tindakan memproyeksikan atau menebak nilai di masa depan

berdasarkan data masa lalu[1][2]. Pendekatan ini bertujuan memberikan landasan pemahaman sebelum memasuki konteks khusus prediksi harga saham. Saham adalah instrumen pasar keuangan yang sangat populer saat ini, mewakili kepemilikan dalam sebuah perusahaan. Perusahaan menggunakan penerbitan saham sebagai salah satu

metode untuk memperoleh pendanaan tambahan[1][3].

Prediksi harga saham merupakan tantangan yang signifikan dalam dunia keuangan, dan tingkat keakuratan hasil prediksi sangat bergantung pada metode yang digunakan dan kualitas data yang tersedia[4][5][6]. Saham Coca-Cola dipilih sebagai subjek penelitian ini karena merupakan *blue-chip stock* dengan stabilitas yang tinggi di pasar global. Perusahaan ini memiliki daya tahan yang kuat terhadap volatilitas pasar, menjadikannya dataset yang relevan dan ideal untuk mengevaluasi kinerja model prediksi harga saham. Pemilihan saham ini juga mempertimbangkan cakupan global Coca-Cola, yang dapat merepresentasikan dinamika pasar yang lebih luas, memberikan keandalan dalam hasil prediksi. Oleh karena itu, penelitian ini difokuskan pada prediksi harga saham Coca-Cola, perusahaan minuman terkemuka, dengan membandingkan metode analisis data time series seperti Regresi Linear, LSTM (*Long Short-Term Memory*), dan GRU (*Gated Recurrent Unit*) menggunakan variasi window-width (3, 4, dan 5). Ketiga metode ini dipilih untuk memberikan perbandingan yang lebih komprehensif dalam prediksi harga saham. *Linear Regression* digunakan sebagai *baseline* sederhana untuk menangkap pola linier, sementara LSTM dan GRU dipilih karena kemampuannya dalam memodelkan ketergantungan temporal jangka panjang dalam data time series. Perbandingan ini bertujuan untuk mengidentifikasi metode mana yang paling efektif dalam memprediksi harga saham, khususnya dengan variasi parameter *window-width* [7][8][9][10].

Linear regression dipilih sebagai metode analisis karena mampu menggambarkan hubungan linier antara variabel dependen. Serta Regresi linear memberikan landasan untuk memahami tren langsung secara linier[11][12][13]. LSTM dan GRU digunakan untuk mengatasi kompleksitas data time series dan memanfaatkan dependensi temporal yang penting dalam pergerakan harga saham, atau LSTM dan GRU dipilih karena kemampuannya menangkap pola jangka panjang dan perubahan cepat dalam harga[14][15][16].

Beberapa penelitian untuk prediksi harga saham yang menggunakan algoritma LSTM, GRU dan Linear regression diantaranya seperti Javvaji Venkatarao, DR.V. Deeban Chakravarthy, dkk (2023) dalam penelitiannya yang berjudul "Stock Price Prediction by Normalizing Lstm and Gru Models". Tujuan dari penelitian ini untuk meramalkan nilai saham moneter perusahaan pada g-tahun-tahun berikutnya. Akurasi yang diperoleh RNN 76,9%, LSTM model 83,4% dan yang lebih baik adalah dengan metode GRU model 88,2% dan setelah di normalisasi LSTM-GRU lebih meningkatkan nilai

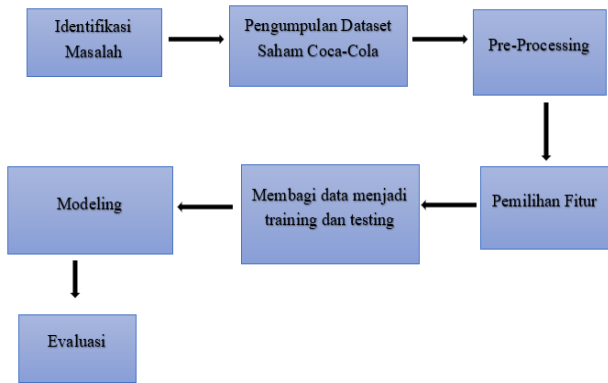
yang baik yaitu 92,4% [14]. Raras Tyasnurita, rifqi Luthfiansyah, dkk (2023) dalam penelitian nya yang berjudul "Gold Price Forecasting using Multiple Linear Regression Method". Hasil penelitian dengan menggunakan metode regresi lineaar berganda menunjukkan bahwa model terbaik ditemukan saat pembagian data pelatihan sebesar 70% dan pengujian sebesar 30%, dengan tingkat kesalahan relatif (MAPE) sebesar 4,7%. Dari hasil ini, disimpulkan bahwa metode regresi linier berganda cocok digunakan unutk peramalan harga emas[17].

Penelitian ini menawarkan pendekatan baru dalam prediksi harga saham dengan membandingkan kinerja model *Linear Regression*, LSTM, dan GRU menggunakan variasi window-width (3, 4, dan 5). Berbeda dari penelitian sebelumnya yang cenderung fokus pada satu model atau membandingkan model tanpa variasi parameter, penelitian ini memberikan kontribusi baru dengan menguji pengaruh variasi window-width terhadap akurasi prediksi harga saham Coca-Cola, yang belum banyak dianalisis dalam konteks serupa. Evaluasi dilakukan dengan ketiga model tersebut menggunakan tiga metrik evaluasi, yaitu MSE, RMSE, dan MAE.

2. METODE DAN BAHAN

Tahapan Penelitian

Alur Penelitian yang dilakukan yaitu, dengan tahapan dimulai dengan identifikasi masalah dilakukan untuk merinci fokus penelitian dan tujuan prediksi. Pengumpulan dataset Coca-Cola Stock Data (1962-2021) dari kaggle, yang kemudian diolah melalui tahapan pre-processing, termasuk *Exploratory Data Analysis* (EDA), *scaling* data. Seleksi fitur dilakukan untuk menentukan variabel yang paling signifikan dalam prediksi harga saham. Data kemudian dibagi menjadi training dan testing untuk melatih dan menguji model sebanyak 80:20. Pembuatan model dilakukan dengan memilih teknik atau algoritma yang sesuai untuk memprediksi harga saham. Akhirnya evaluasi model dilakukan menggunakan matriks evaluasi yang telah ditetapkan sebelumnya untuk mengukur sejauh mana model dapat menghasilkan prediksi yang akurat. Dengan langkah-langkah ini, penelitian ini membentuk kerangka kerja sistematis dari identifikasi masalah hingga evaluasi model dalam konteks prediksi harga penutup saham coco-cola Alur penelitian seperti terlihat pada Gambar 1. Alur penelitian



Gambar 1. Alur Penelitian

Indetifikasi Masalah

Penelitian ini menyoroti kompleksitas prediksi harga penutup saham coca-cola dan menambahkan dimensi perbandingan antara metode linear regression, LSTM dan GRU sebagai pendekatan prediktif. Salah satu masalah yang timbul adalah sejauh mana setiap metode dapat mengatasi fluktuasi harga saham yang dipengaruhi oleh variabel kompleks seperti kondisi pasar, kinerja perusahaan dan faktor eksternal. Komparatif antara metode linear, LSTM dan GRU juga menimbulkan pertanyaan tentang kehandalan dan ketepatan setiap pendekata dalam meramalkan pergerakan harga saham. Dengan memfokuskan analisis pada masalah-masalah ini, penelitian ini berusaha untuk memberikan pemahaman yang lebih mendalam tentang efektivitas metode-metode dalam memprediksi harga penutup saham coca-cola.

Pengumpulan Data

Dalam penelitian ini, dataset harga saham Coca-Cola diambil dari Kaggle, sebuah platform komunitas yang menyediakan akses ke berbagai dataset, termasuk data keuangan. Dataset mencakup periode waktu dari tahun 1962 hingga 2022 dan dapat diakses melalui link berikut: <https://www.kaggle.com/datasets/kalilurrahman/coca-cola-stock-live-and-updated>. Proses pengunduhan data sesuai dengan instruksi Kaggle, dan data kemudian diimpor ke lingkungan analisis data menggunakan Python dengan menggunakan pustaka pandas. Dengan 15311 data harga saham dan 7 kolom termasuk tanggal, harga pembukaan, tertinggi, terendah, penutupan, dividends, stock splits dan volume, penelitian ini fokus pada fitur "close" (harga penutupan) untuk memprediksi harga penutupan saham Coca-Cola. Pendekatan ini dipilih untuk menjadikan model lebih sederhana, mudah diinterpretasikan, dan efisien dalam pelatihan, yang secara langsung berpengaruh terhadap akurasi. Dalam beberapa kasus, model yang lebih sederhana seperti Linear Regression memang tidak seakurat model kompleks seperti LSTM atau GRU, namun

kelebihan dari model sederhana adalah kemampuannya untuk menangkap pola yang lebih mendasar dalam data dengan lebih cepat. Efisiensi dalam pelatihan juga memungkinkan pengujian lebih banyak variasi parameter, yang dapat membantu menemukan kombinasi optimal untuk meningkatkan akurasi. Selain itu, dengan dataset yang lebih kecil atau tren jangka pendek, model yang lebih sederhana sering kali memberikan performa akurasi yang cukup kompetitif tanpa *overfitting*. Gambar 2 merupakan isi dari dataset Coca-Cola.

	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
0	1962-01-02	0.050016	0.051378	0.050016	0.050016	806400	0.0	0
1	1962-01-03	0.049273	0.049273	0.048159	0.048902	1574400	0.0	0
2	1962-01-04	0.049026	0.049645	0.049026	0.049273	844800	0.0	0
3	1962-01-05	0.049273	0.049892	0.048035	0.048159	1420800	0.0	0
4	1962-01-08	0.047787	0.047787	0.046735	0.047664	2035200	0.0	0
...	...	...	...	...	...	...	...	...
15306	2022-10-20	55.770000	55.919998	54.959999	55.080002	16905100	0.0	0
15307	2022-10-21	55.000000	56.110001	54.990002	55.959999	15028000	0.0	0
15308	2022-10-24	56.639999	57.730000	56.570000	57.570000	17416700	0.0	0
15309	2022-10-25	59.040001	59.110001	57.750000	58.950001	28829900	0.0	0
15310	2022-10-26	59.009998	59.779999	58.860001	59.389999	15831400	0.0	0

15311 rows x 8 columns

Gambar 2. DataSet Coca-Cola

Pre-Processing

Preprocessing data adalah proses penting untuk mempersiapkan data sebelum melakukan analisis. Dalam penelitian ini, *pre-processing* data digunakan untuk memastikan kebersihan dan kualitas data, serta untuk menghadapi potensi masalah dalam dataset[18]. Preprocessing data adalah tahap krusial dalam penelitian ini, dirancang untuk memastikan kebersihan dan kualitas data sebelum dilakukan analisis[17]. Tahapan awal dalam preprocessing adalah Exploratory Data Analysis (EDA), sebuah proses investigasi yang bertujuan memahami kondisi data dan menyiapkan data untuk model algoritma[19]. EDA melibatkan langkah-langkah seperti memeriksa dimensi dan jumlah data, mengecek tipe data untuk setiap kolom, mendeteksi dan mengatasi data yang kosong, serta menerapkan analisis statistik deskriptif untuk memahami distribusi data[20][21][1]. Visualisasi data, seperti histogram dan scatter plot, juga digunakan untuk mengidentifikasi pola, hubungan, dan distribusi dalam dataset[22][23][24].

Selanjutnya dilakukan pemilihan atau seleksi fitur yang akan digunakan dalam pembangunan model penelitian. Dalam penelitian ini, pendekatan yang diambil untuk seleksi atribut adalah memusatkan perhatian hanya pada fitur harga penutupan (*close*) sebagai satu-satunya variabel dalam model prediksi. Alasan di balik keputusan ini adalah bahwa harga penutupan mencerminkan nilai terakhir saham pada suatu periode perdagangan, dianggap sebagai indikator kunci dalam banyak strategi investasi [10][25]. Fokus pada fitur *close* memungkinkan model untuk secara khusus memodelkan tren harga penutupan dari waktu ke waktu, menangkap pola historis, dan menggunakan informasi historis dengan menerapkan metode *window*. Dengan demikian, pendekatan ini memungkinkan model untuk lebih fokus pada pergerakan harga penutupan yang dianggap sebagai sinyal penting bagi investor dan pedagang, memudahkan deteksi pola kritis dan meningkatkan kualitas prediksi.

Sebagai penambahan, dalam tahap *preprocessing* ini juga dilakukan *scaling* data untuk memastikan keseragaman dan memperbaiki rentang nilai antar fitur numerik, yang menjadi langkah penting untuk meningkatkan performa model analisis yang akan digunakan.

Split Data

Selanjutnya pemisahan dataset, untuk menyiapkan data secara optimal sebelum dilakukan analisis lebih lanjut. Pemisahan data, atau "*Split data*," merupakan tahap krusial dalam *preprocessing* yang melibatkan pembagian dataset menjadi dua subset utama: data pelatihan (80% dari dataset) dan data pengujian (20% dari dataset). Data pelatihan digunakan untuk melatih model, memahami pola data, dan menghasilkan prediksi, sedangkan data pengujian tidak digunakan dalam pelatihan dan bertujuan untuk menguji seberapa baik model dapat memprediksi data baru yang tidak pernah dilihat sebelumnya.

Sliding Window

Metode *sliding Window* adalah suatu strategi pembentukan struktur pada data *time series* [26]. Pendekatan ini memungkinkan pembentukan data dengan memperhitungkan langkah waktu saat ini, di mana nilai target yang akan diprediksi untuk suatu periode tertentu, misalnya jika data Januari adalah nilai target yang akan diprediksi, maka data *n*-dimensi dapat dibuat dari data *n* bulan sebelumnya untuk membuat prediksi[27].

Modeling

Proses pemodelan dalam penelitian ini melibatkan penerapan tiga metode analisis yang berbeda, yaitu Long Short-Term Memory (LSTM),

Linear regression dan *Gated Recurrent Unit* (GRU) dengan memanfaatkan tiga kali pembagian *window-width* yaitu 3, 4, 5 dan dengan rasio data 80: 20 untuk mengetahui performa algoritma terbaik.

Linear Regression

Linear Regression merupakan model yang digunakan untuk memprediksi nilai numerik berdasarkan data historis dalam suatu periode waktu tertentu. Terdapat dua jenis *Linear regression*: *Linear regression* satu variabel (*simple Linear regression*) dan *linear regression multivariable* (*multiple linear regression*). Dalam *linear regression* satu variabel, fokus adalah pada penentuan korelasi antara variabel independen (*x*) dengan variabel dependen (*y*). sementara dalam *linear regression multivariable*, tujuannya adalah untuk menemukan keterkaitan antara beberapa variabel bebas dengan variabel terikat, sehingga memungkinkan analisis lebih kompleks dan melibatkan lebih dari satu variabel independen[19]. Analisis regresi bertujuan untuk memperkirakan nilai variabel dependen berdasarkan nilai variabel independen [16]. Metode regresi linear berdasarkan pola hubungan data masa lalu, dengan asumsi bahwa variabel yang dapat diprediksi, yang diwakili oleh variabel independen (misalnya, persediaan), dipengaruhi oleh besarnya variabel bebas. Hubungan antara variabel independen dan variabel yang ingin diprediksi dapat dijelaskan sebagai sebuah fungsi [17]. Regresi linear memiliki persamaan model standar [4].

$$Y=a+bX \quad (1)$$

Formula linear regression:

Y = variabel dependent

x = variabel independent

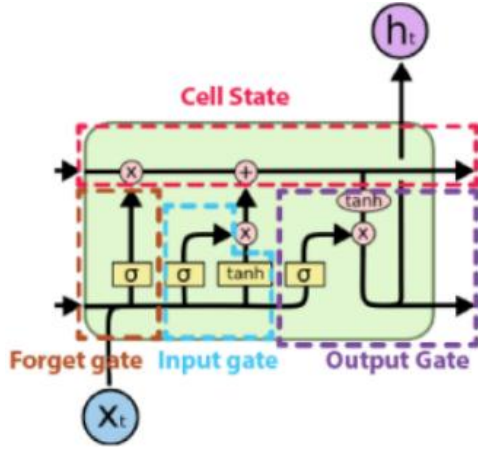
a = konstanta

b = koefisien atau jumlah respon yang dihasilkan

Long Short-Term Memory (LSTM)

Long Short-Term Memory Networks (LSTM) merupakan salah satu jenis dari *Reccurent Neural Network* (RNN). LSTM diusulkan oleh *sepp Hochreiter* dan *jürgen schmidhuber* pada tahun 1997. LSTM dirancang untuk mengatasi masalah ketergantungan jangka panjang yang sering terjadi pada RNN secara umum[21]. LSTM adalah jenis arsitektur jaringan saraf tiruan yang khusus dirancang untuk mengatasi masalah *sequantial* data seperti data harga saham. LSTM memiliki kemampuan untuk mengingat informasi jangka panjang dari data historis dan dapat digunakan untuk membuat prediksi berdasarkan urutan waktu. LSTM sangat cocok untuk analisi harga saham karena mampu menangkap pola-pola yang muncul dari perubahan harga seiring waktu[28]. Model LSTM terdiri dari lapisan input (*input layer*), lapisan tersembunyi (*hidden layers*), dan

lapisan output (*output layers*). Pada tiap blok terdapat beberapa memory cell yang melekat dan tiga unit penggali, yaitu input, output dan forget gates. Dari komponen ini, LSTM berisi tiga gates dan sebuah memory cell yang mampu melupakan atau menghafal informasi yang harus di transfer ke sel berikutnya. Gambar 3 merupakan arsitektur dari model LSTM.



Gambar 3. Arsitektur LSTM

Formula LSTM:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\hat{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

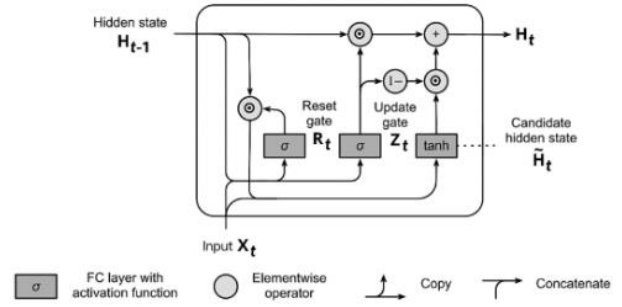
$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (7)$$

Gate Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) adalah jenis arsitektur jaringan syaraf tiruan yang mirip dengan *Long Short-Term Memory (LSTM)*, namun dengan struktur yang lebih sederhana. Tujuan utama dalam pengembangan GRU adalah untuk memungkinkan setiap unit rekurensi untuk secara adaptif menangkap ketergantungan dalam berbagai rentang waktu. GRU digunakan untuk mengatasi masalah data berurutan (*sequential data*) dan dapat digunakan dalam konteks prediksi harga saham. Dalam GRU, terdapat dua gate yang penting, yaitu *reset gate* dan *update gate*. *Reset gate* digunakan untuk mengontrol aliran informasi dalam jaringan, sedangkan *update gate* menentukan sejauh mana informasi dari time step sebelumnya yang dapat disimpan dan digunakan sebagai input untuk time step berikutnya. Keunggulan GRU adalah efisiensi komputasi yang lebih tinggi dibandingkan LSTM, yang membuatnya menjadi pilihan yang baik untuk tugas-tugas yang memerlukan pelatihan model yang lebih cepat[22]. Untuk arsitektur dari GRU dapat dilihat dari Gambar



Gambar 4. Arsitektur GRU

Formula GRU:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (8)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (9)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \quad (10)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (11)$$

Evaluation

Metrik evaluasi digunakan untuk mengukur kinerja dari prediksi model. Pada model time-series, evaluasi dapat digunakan dengan menghitung kesalahan yang dibuat oleh model. Ada beberapa metrik yang umumnya digunakan dalam evaluasi model time-series, diantaranya adalah Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) dan Mean Square Error.

Pada MAE, setiap nilai diambil nilai absolutnya, sehingga nilai negatif diabaikan. Dalam evaluasi model peramalan, MAE memberikan ukuran intuitif rata-rata kesalahan dari keseluruhan data[28].

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y'_i - Y_i| \quad (12)$$

RMSE mengukur besarnya tingkat kesalahan prediksi, di mana semakin kecil (mendekati 0) nilai RMSE, hasil prediksi dianggap semakin akurat[29].

$$RMSE = \sqrt{MSE}$$

MSE sebagai metrik evaluasi membantu mengukur sejauh mana perbedaan antara nilai prediksi model dengan nilai sebenarnya dalam rangkaian waktu. MSE memberikan gambaran lebih mendalam tentang tingkat kesalahan prediksi, yang dapat menjadi informasi penting untuk meningkatkan performa model time-series[28].

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y'_i - Y_i)^2 \quad (13)$$

3. HASIL DAN PEMBAHASAN

Langkah Pre-Processing

Pada tahap ini, data yang sudah diinputkan akan diolah dari data mentah jadi data yang siap untuk digunakan untuk tahap selanjutnya. Disini dilakukan pengecekan data kosong, menangani *missing values*, menghapus kolom yang tidak digunakan. Gambar 5 merupakan hasil dari proses *Exploratory Data Analysis (EDA)*.

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15311 entries, 0 to 15310
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype  
---  --
0   Date             15311 non-null  object  
1   Open             15311 non-null  float64  
2   High             15311 non-null  float64  
3   Low              15311 non-null  float64  
4   Close            15311 non-null  float64  
5   Volume           15311 non-null  int64  
6   Dividends        15311 non-null  float64  
7   Stock Splits     15311 non-null  int64  
dtypes: float64(5), int64(2), object(1)
memory usage: 957.1+ KB

isNull = df.isnull().sum().to_frame(name='isNull').T
isNa = df.isna().sum().to_frame(name='isNa').T
Unique = df.nunique().to_frame(name='Unique').T
summary = pd.concat([Unique, isNa, isNull])
summary
```

	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
Unique	15311	14855	14547	14580	11105	10396	64	3
isNa	0	0	0	0	0	0	0	0
isNull	0	0	0	0	0	0	0	0

Gambar 5. Hasil proses EDA

Langkah pemilihan Fitur dan Window-width

Dalam proses pemilihan fitur, keputusan untuk memfokuskan analisis pada atribut close (harga penutup) didasarkan pada pertimbangan bahwa kolom close memainkan peran kritis sebagai target yang ingin diprediksi dan juga sebagai input dengan penerapan metode window. Fokus pada fitur close terkait dengan tujuan spesifik penelitian ini, yaitu memprediksi harga penutup saham coca-cola. Pendekatan yang digunakan adalah hanya menggunakan fitur close (harga penutupan), pendekatan ini dilakukan dengan mempertimbangkan bahwa fitur Close memiliki peran ganda sebagai target (variabel terikat) dan sebagai input (variabel independen) dengan penggunaan metode *window-width*.

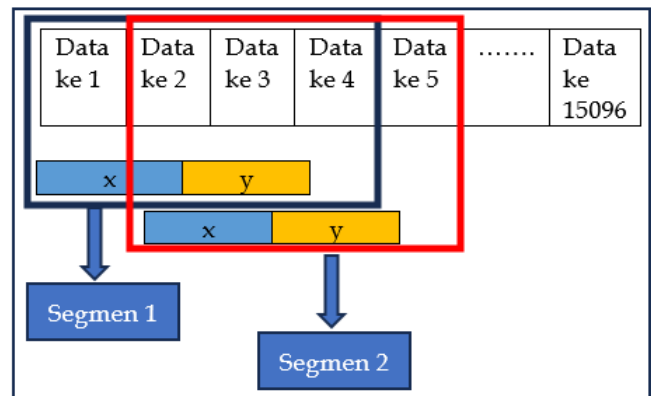
Dalam penggunaan metode *window-width*, kolom close memberikan gambaran yang paling relevan terkait pola dan tren harga penutup saham, yang menjadi fokus utama analisis. Gambar 6 merupakan dataset yang akan digunakan untuk proses pelatihan model selanjutnya.

	Date	Close
0	1962-01-02	0.050016
1	1962-01-03	0.048902
2	1962-01-04	0.049273
3	1962-01-05	0.048159
4	1962-01-08	0.047664
...	...	...
15306	2022-10-20	55.080002
15307	2022-10-21	55.959999
15308	2022-10-24	57.570000
15309	2022-10-25	58.950001
15310	2022-10-26	59.389999

[15311 rows x 2 columns]

Gambar 6. DataSet Close

Setelah ditentukan untuk menetapkan ukuran *window-width* yang akan digunakan dalam penelitian ini yaitu dengan opsi 3, 4 dan 5. Pada dataset yang digunakan data ke 1 sampai dengan 15096 merupakan unit observasi. Dengan menggunakan *window-width* =3, Segmen pertama adalah data ke 1 hingga data ke 3, digunakan pada prediksi data ke 4. Segmen kedua adalah data ke 2 hingga ke 4, digunakan pada prediksi data ke 5. Proses tersebut terus berlanjut hingga data selesai tersegmentasi dengan baik seperti yang tampak pada Gambar 7.



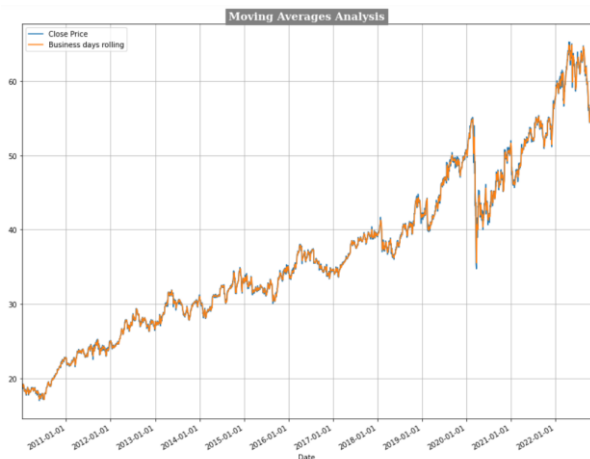
Gambar 7. Sliding Window

Peneliti menentukan percobaan 1 yang digunakan adalah *window-width*=3 yang dimana pengambilan sebagai target ada data close yg ada di index 3 dan data yang berada di index 0,1 dan 2 sebagai fitur atau data latih dan index 3 sebagai target atau data testing.

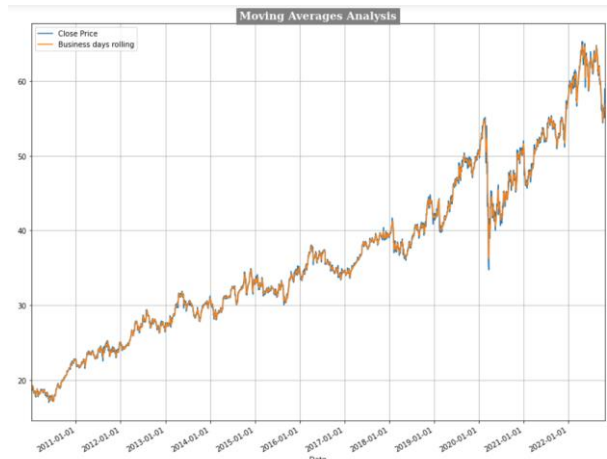
Tabel 1. Contoh Proses *sliding-window* dengan *window-width*=3

X(t-3)	X(t-2)	X(t-1)	y(t)
NaN	NaN	NaN	28,520000
NaN	NaN	28,520000	28,174999
NaN	28,520000	28,174999	28,165001
28,520000	28,174999	28,165001	28,094999
28,174999	28,165001	28,094999	27,575001
.....			
.....			
.....			
56,279999	57,759998	57,799999	58,060001
57,759998	57,799999	58,060001	58,650002
57,799999	58,060001	58,650002	57,730000

Visualisasi grafik pada Gambar 8, 9, dan 10 menunjukkan perbedaan prediksi harga penutupan saham berdasarkan ukuran *window-width* yang digunakan. Grafik ini memperbandingkan nilai asli dengan hasil prediksi yang dipengaruhi oleh pergeseran *window-width*. Pada grafik, garis biru merepresentasikan harga penutupan asli, sedangkan garis oranye menggambarkan nilai prediksi yang dihasilkan dari pergeseran *window-width*.



Gambar 8. Perbandingan Harga Asli dengan *window-width=3*



Gambar 9. Perbandingan Harga Asli dengan *window-width=4*

Peneliti menentukan percobaan 2 yang digunakan adalah *window-width=4* yang dimana pengambilan sebagai target ada data close yg ada di index 5 dan data yang berada di index 1, 2, 3 dan 4 sebagai fitur atau data latih dan index 5 sebagai target atau data testing.

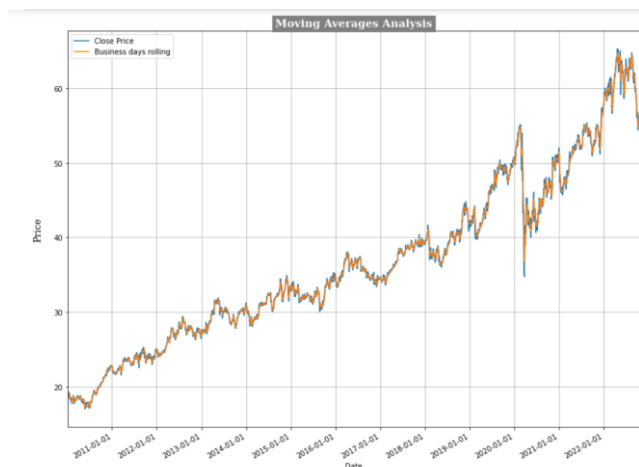
Tabel 2. Contoh Proses *sliding-window* dengan *window-width=4*

$X(t-4)$	$X(t-3)$	$X(t-2)$	$X(t-1)$	$y(t)$
NaN	NaN	NaN	NaN	28,520000
NaN	NaN	NaN	28,520000	28,174999
NaN	NaN	28,520000	28,174999	28,165000
NaN	28,520000	28,174999	28,165000	28,094999
28,520000	28,174999	28,165000	28,094999	27,575000
0	9	1	9	1
.....				
.....				
.....				
54,860000	56,279999	57,759999	57,799999	58,060000
1	9	8	9	1
56,279999	57,759999	57,799999	58,060000	58,650000
9	8	9	1	2
57,759999	57,799999	58,060000	58,650000	57,730000
8	9	1	2	0

Peneliti menentukan percobaan 3 yang digunakan adalah *window-width=5* yang dimana pengambilan sebagai target ada data close yg ada di index 6 dan data yang berada di index 1, 2, 3, 4 dan 5 sebagai fitur atau data latih dan index 6 sebagai target atau data testing.

Tabel 3 Contoh Proses *Sliding-window* dengan *window-width=5*

$X(t-5)$	$X(t-4)$	$X(t-3)$	$X(t-2)$	$X(t-1)$	$y(t)$
NaN	NaN	NaN	NaN	NaN	28,520000
NaN	NaN	NaN	NaN	28,520000	28,174999
NaN	NaN	NaN	28,520000	28,174999	28,165001
NaN	NaN	28,520000	28,174999	28,165001	28,094999
NaN	28,520000	28,174999	28,165001	28,094999	27,575001
.....					
.....					
.....					
55,000000	54,860000	56,279999	57,759999	57,799999	58,060000
000	001	999	998	999	001
54,860000	56,279999	57,759999	57,799999	58,060000	58,650000
001	999	998	999	001	002
56,279999	57,759999	57,799999	58,060000	58,650000	57,730000
999	998	999	001	002	000



Gambar 10. Perbandingan harga Asli dengan *Window-width=5*

Pada Tabel 1, 2 dan 3 merupakan proses pengambilan untuk variabel independen atau fitur (X) dan variabel dependen atau target (Y) sebagai contoh *window-width=3* dimana data yang di awal sebagai target dan $X(t-1)$, $X(t-2)$ dan $X(t-3)$ merupakan fitur pertama, kedua dan ketiga sebelum nilai target dan seterusnya atau nilai NaN dikarenakan tidak ada nilai sebelum nilai target. Untuk proses opsi *window-width* 4 dan 5 sama seperti penyelesaian seperti *window-width=3* diatas hanya ada perubahan di jumlah variabel independen sesuai dengan opsi *window* yang ditentukan.

Langkah Split Data

Tahap selanjutnya adalah split data, Dalam langkah ini data dibagi menjadi dua bagian: data pelatihan (train_Data) yang mencakup 80% dari total dan data uji (test data) yang melibatkan sisa 20%. Dimensi dari kedua dataset dapat diperiksa melalui perintah (train_Data.shape, test_Data.shape). Proses ini mendukung pembentukan model prediksi dengan menggunakan data pelatihan dan menguji performanya dengan data uji. Dapat dilihat pada gambar 11 merupakan data untuk train sebanyak 12077 dan data untuk test 3019 data.

```
#Splitting the Data into two folders: Train data and Test data
training_size = round(len(DF) * 0.80)

train_data = DF.iloc[:training_size]
test_data = DF.iloc[training_size:]

train_data.shape, test_data.shape

((12077, 1), (3019, 1))
```

Gambar 11. Split Data

Selanjutnya, akan menggunakan uji coba model dengan salah satu *window-width* dari yang sudah ditentukan dengan split data menggunakan rasio 80:20 atau 80% data dijadikan data train dan 20% menjadi data tes dapat dilihat pada gambar 12 dibawah ini.

```
x_train, y_train = prepare_time_series_data(train_data,3)
X_test, y_test = prepare_time_series_data(test_data,3)
X_train.shape, y_train.shape, X_test.shape, y_test.shape

((12074, 3, 1), (12074, 1), (3016, 3, 1), (3016, 1))
```

Gambar 12. *prepare_time_series_data*

Langkah Modeling

Hasil dari proses training pada model digunakan untuk prediksi pada model dengan menggunakan data testing. Proses testing pada model dilakukan dengan cara membandingkan data testing (aktual) dan data prediksi untuk mengukur nilai errornya menggunakan matrik evaluasi MAE, RMSE, dan MSE. Hasil dari matriks evaluasi kemudian dianalisis dan direpresentasikan dalam grafik untuk menentukan model dengan batch dan epoch yang memberikan prediksi paling mendekati nilai aktual.

Selain itu, akurasi model juga diukur untuk menilai sejauh mana ketepatan dari nilai prediksinya.

1. Linear Regression

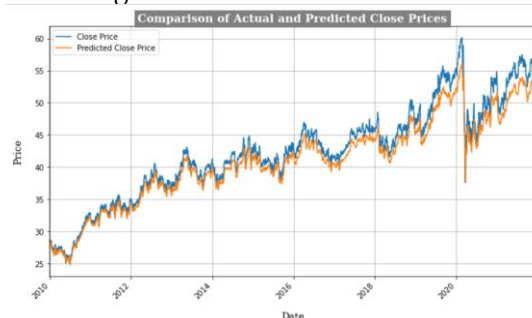
Nilai Aktual vs. Prediksi pada data testing Harga saham coca cola menggunakan *Linear Regression* dengan *window-width=3* dan rasio 80:20 pada gambar 13, nilai aktual ditunjukkan dengan garis berwarna biru, dan nilai prediksi LSTM ditunjukkan dengan garis warna orange. Nilai metrik evaluasi dari model ini menghasilkan MAE 0.29, RMSE 0.46, dan MSE 0.21. Hasil model *Linear Regression* ini nantinya akan dibandingkan dengan model GRU dan LSTM yang memprediksi dataset yang sama. Dapat dilihat pada gambar 13 nilai hasil prediksi dan nilai aktual menggunakan model *linear regression*.



Gambar 13. visualisasi hasil prediksi *Linear Regression* dan nilai aktual

2. LSTM

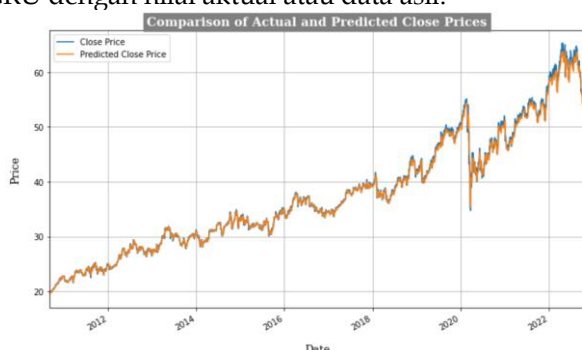
Nilai Aktual vs Prediksi pada data testing Harga saham coca cola menggunakan *LSTM* dengan *window-width=3*, jumlah neuron 64, *batch-size=32*, *epoch=30* dan rasio 80:20, nilai aktual ditunjukkan dengan garis berwarna biru, dan nilai prediksi LSTM ditunjukkan dengan garis warna orange. Model ini dikompilasi menggunakan dengan optimizer 'adam', loss function 'mean_squared_error', dan metric evaluasi 'mean_absolute_error'. Nilai metrik evaluasi dari model ini menghasilkan MAE 1.34, RMSE 1.80, dan MSE 3.24. Hasil model *LSTM* ini nantinya akan dibandingkan dengan model GRU dan *Linear Regression* yang memprediksi dataset yang sama. Gambar 14 merupakan hasil visualisasi dari nilai prediksi yang diperoleh dari data testing pemodelan LSTM dengan nilai aktual atau data asli.



Gambar 14. Visualisasi Hasil prediksi LSTM dan nilai aktual

3. GRU

Nilai Aktual vs Prediksi pada data testing Harga saham coca cola menggunakan GRU dengan *window-width*=3, jumlah neuron 64, *batch-size*=32, *epoch*=30 dan rasio 80:20, nilai aktual ditunjukkan dengan garis berwarna biru, dan nilai prediksi LSTM ditunjukkan dengan garis warna orange. Model ini dikompilasi menggunakan dengan optimizer 'adam', loss function 'mean_squared_error', dan metric evaluasi 'mean_absolute_error'. Nilai metrik evaluasi dari model ini menghasilkan MAE 0.42, RMSE 0.64, dan MSE 0.40. Hasil model GRU ini nantinya akan dibandingkan dengan model LSTM dan *Linear Regression* yang memprediksi dataset yang sama. Gambar 15 merupakan hasil visualisasi dari nilai prediksi yang diperoleh dari data testing pemodelan GRU dengan nilai aktual atau data asli.



Gambar 15. Visualisasi hasil prediksi GRU dan nilai aktual

Evaluasi

Evaluasi model, yang melibatkan pemeriksaan kinerja model melalui perhitungan nilai error menggunakan RMSE (*Root Mean Squared Error*), MSE (*Mean Squared Error*) dan MAE (*Mean Absolute Error*). RMSE dan MSE akan memberikan informasi mengenai tingkat kesalahan prediksi dengan memberikan bobot pada kesalahan besar, sementara MAE memberikan ukuran kesalahan rata-rata tanpa memberikan preferensi berlebihan pada kesalahan tertentu. Dengan menggunakan ketiga metrik ini, dapat menyajikan gambaran yang komprehensif mengenai sejauh mana ketiga model dapat memprediksi harga penutupan saham Coca-Cola. Berikut ini perincian evaluasi antara metode atau model dengan beberapa *window* yang digunakan dan split rasio 80:20.

1. Linear Regression

Pada tabel 4 merupakan hasil metrik evaluasi pada model *linear regression* dengan beberapa *window-width* yang telah di uji coba, dan dapat dilihat bahwa dari ketiga *window-width* tersebut yang memiliki nilai metrik evaluasi yang terbaik adalah *window-width*=3 & 5 dengan nilai RMSE= 0.46, MSE=0.21 dan MAE=0.29.

Tabel 4 Contoh Perbandingan metrik evaluasi dengan beberapa *window-width* pada Model *Linear regression*

<i>window-width</i>	RMSE	MSE	MAE
3	0.46	0.21	0.29
4	0.46	0.22	0.29
5	0.46	0.21	0.29

2. LSTM

Pada tabel 5 merupakan hasil metrik evaluasi pada model LSTM dengan beberapa *window-width* yang telah di uji coba, dan dapat dilihat bahwa dari ketiga *window-width* tersebut yang memiliki nilai metrik evaluasi yang terbaik adalah *window-width*=5 dengan nilai RMSE= 1.39, MSE=1.94 dan MAE=0.94 dibanding dengan nilai metrik evaluasi dari *window-width* 4 dan 3.

Tabel 5 Contoh Perbandingan *metriks evaluation* dengan beberapa *window-width* pada Model LSTM

<i>window-width</i>	RMSE	MSE	MAE
3	1.80	3.24	1.34
4	2.17	4.72	1.59
5	1.39	1.94	0.94

3. GRU

Pada tabel 6 merupakan hasil metrik evaluasi pada model GRU dengan beberapa *window-width* yang telah di uji coba, dan dapat dilihat bahwa dari ketiga *window-width* tersebut yang memiliki nilai metrik evaluasi yang terbaik adalah *window-width*=5 dengan nilai RMSE= 0.48, MSE=0.23 dan MAE=0.31 dibanding dengan nilai metrik evaluasi dari *window-width* 4 dan 3.

Tabel 6. Contoh Perbandingan *metriks evaluation* dengan beberapa *window-width* pada model GRU

<i>window-width</i>	RMSE	MSE	MAE
3	0.64	0.40	0.42
4	0.57	0.32	0.40
5	0.48	0.23	0.31

4. KESIMPULAN

Penelitian ini bertujuan mengevaluasi penerapan metode *Linear Regression*, *Long Short-Term Memory* (LSTM), dan *Gated Recurrent Unit* (GRU) dalam memprediksi harga penutupan saham Coca-Cola, dengan fokus pada memberikan panduan yang berguna bagi individu yang memiliki pengetahuan terbatas dalam investasi saham dengan variasi *window-width*=3, *window-width*=4, dan *window-width*=5 memberikan pemahaman mendalam terhadap kinerja masing-masing model. Hasil

penelitian menunjukkan bahwa *Linear Regression* mencapai kinerja terbaik dengan window-width=3 & 5, mencapai nilai MSE 0.21, RMSE 0.46, dan MAE 0.29, mengungguli LSTM dengan nilai MSE 3.24 & 1.94, RMSE 1.80 & 1.39, MAE 1.34 & 0.94 dan GRU dengan nilai MSE 0.40 & 0.23, RMSE 0.64 & 0.48, dan MAE 0.42 & 0.31 dalam menangkap pola dan tren harga saham Coca-Cola. Berdasarkan temuan ini, dapat disimpulkan bahwa penerapan Linear Regression, khususnya dalam konteks prediksi harga penutupan saham Coca-Cola, dapat memberikan panduan yang lebih akurat dan bermanfaat bagi investor, terutama bagi investor yang mungkin memiliki keterbatasan dalam pemahaman investasi saham. Ini mengindikasikan bahwa model Linear Regression dapat memberikan hasil yang lebih presisi dengan mempertimbangkan data dalam jangka waktu yang lebih singkat, sehingga memungkinkan analisis yang lebih responsif terhadap perubahan harga saham. Model Linear Regression yang sederhana ini, dikombinasikan dengan teknik window width yang efektif, menawarkan alat yang relatif mudah dipahami namun sangat berguna untuk mendukung pengambilan keputusan investasi di pasar saham. Dengan menggunakan pendekatan ini, investor dengan tingkat pengetahuan terbatas dapat memperoleh wawasan yang lebih jelas mengenai pergerakan harga saham, yang pada gilirannya dapat meningkatkan kepercayaan diri dan potensi hasil investasi.

DAFTAR PUSTAKA

- [1] P. Ariyadi, M. M. Effendi, and S. B. Raharjo, "Analisa Prediksi Harga Saham Blue Chip Lq45 Dengan Metode Data Mining Backpropagation Neural Network (Studi Kasus Di Bursa Efek Indonesia)," *Pros. Sains dan Teknol.*, vol. 1, no. 1, pp. 68-76, 2022.
- [2] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock Closing Price Prediction using Machine Learning Techniques," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 599-606, 2020, doi: 10.1016/j.procs.2020.03.326.
- [3] Hilman Winnos, Richashanty Septima, and Husna Gemasih, "Perbandingan Metode Regresi Linier Berganda dan Autoregressive Integrated Moving Average (ARIMA) Untuk Prediksi Saham PT. BSI, Tbk.," *Ocean Eng. J. Ilmu Tek. dan Teknol. Marit.*, vol. 1, no. 4, pp. 15-23, 2022, doi: 10.58192/ocean.v1i4.350.
- [4] D. Priyadi and I. Mardhiyah, "Model Autoregressive Integrated Moving Average (Arima) Dalam Peramalan Nilai Harga Saham Penutup Indeks Lq45," *J. Ilm. Inform. Komput.*, vol. 26, no. 1, pp. 78-94, 2021, doi: 10.35760/ik.2021.v26i1.3695.
- [5] A. Gupta and T. J. Nagalakshmi, "Stock price prediction using linear regression in machine learning," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 1382-1385, 2019, doi: 10.35940/ijitee.L3932.1081219.
- [6] N. Fitriyani and S. Angkasa, "Prediksi Harga Saham Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation Dan Adaptive Neuro Inference System (Anfis)," *J. Ilm. Teknol. Inf. dan Komput. ...*, vol. 15, no. 1, pp. 1-8, 2020, [Online]. Available: <http://stmik.id/jurnal/index.php/pranala/article/view/35>
- [7] T. Hariguna and Ajiono, "Comparison of Three Time Series Forecasting Methods on Linear Regression, Exponential Smoothing and Weighted Moving Average," *Int. J. Informatics Inf. Syst.*, vol. 6, no. 2, pp. 89-102, 2023.
- [8] M. W. Rini and N. Ananda, "Perbandingan Metode Peramalan Menggunakan Model Time Series," *Tekinfo J. Ilm. Tek. Ind. dan Inf.*, vol. 10, no. 2, pp. 88-101, 2022, doi: 10.31001/tekinfo.v10i2.1419.
- [9] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 3285-3292, 2019, doi: 10.1109/BigData47090.2019.9005997.
- [10] S. S. Al-Hasnawi and L. H. Al-Hchemi*, "Closing Price Prediction of Stock Listed on the Iraq Stock Exchange Using Ann-Lstm," *JURISMA J. Ris. Bisnis Manaj.*, vol. 12, no. 2, pp. 173-185, 2022, doi: 10.34010/jurisma.v12i2.8103.
- [11] M. Keren and R. S. Oetama, "Prediksi prospek harga saham perusahaan perbankan Menggunakan Regresi Linear (Study Kasus Bank BCA Tahun 2015-2017)," *JSI J. Sist. Inf.*, vol. 11, no. 1, pp. 1703-1710, 2019, doi: 10.36706/jsi.v11i1.8105.
- [12] J. Leonardo and R. Gantino, "Analisa Teknikal Pembuatan Trading Plan Keputusan Investasi Pada 3 Saham Perbankan Bumh Yang Terdaftar Pada Indeks Lq45 Tahun 2014 - 2019," *J. Ris. Akunt.*, vol. 13, no. 2, pp. 174-185, 2021, doi: 10.34010/jra.v13i2.4651.
- [13] A. I. Caniago, W. Kaswidjanti, and J. Juwairiah, "Recurrent Neural Network With Gate Recurrent Unit For Stock Price Prediction," *Telematika*, vol. 18, no. 3, p. 345, 2021, doi: 10.31315/telematika.v18i3.6650.
- [14] D. R. V. J. Venkatarao, S. Meadi, K. L. Durga, K. Pranay, and N. Purushotham, "Stock Price Prediction by Normalizing LSTM and GRU Models," *Journal of Survey in Fisheries Sciences*,

- vol. 10, pp. 5326–5332, 2023, [Online]. Available: <http://sifisheriessciences.com/journal/index.php/journal/article/view/1875>.
- [15] K. Sofi, A. S. Sunge, S. R. Riady, and A. Z. Kamalia, "Perbandingan Algoritma Linear Regression, LSTM, dan GRU dalam Memprediksi Harga Saham dengan Model Time Series," *Seminastika*, vol. 3, no. 1, pp. 39–46, 2021, doi: 10.47002/seminastika.v3i1.275.
- [16] U. Demirel, H. Cam, and R. Unlu, "Predicting stock prices using machine learning methods and deep learning algorithms: The sample of the istanbul stock exchange," *Journal of Science*, vol. 34, no. 1, pp. 63–82, 2021, doi: 10.35378/gujs.679103.
- [17] T. Raras, L. Rifqi, and B. M. Rayhan, "Gold Price Forecasting using Multiple Linear Regression Method," *J. Teknol. dan Sist. Inf.*, vol. 9, no. 8, pp. 1509–1514, 2023, doi: 10.3844/ajassp.2009.1509.1514.
- [18] J. Senoguchi, "Stock Price Prediction through STL Decomposition using Multivariate Two-way Long Short-term Memory," *J. Comput. Sci. Technol. Stud.*, vol. 4, no. 2, pp. 90–96, 2022, doi: 10.32996/jcsts.2022.4.2.11.
- [19] B. H. Mahendra, L. Chaerani, and G. Gumay, "Analisis Perbandingan Prediksi Harga Saham menggunakan Algoritma Artificial Neural Network dan Linear Regression," *J. Ilm. Komputasi*, vol. 22, no. 2, pp. 303–312, 2023, doi: 10.32409/jikstik.22.2.3357.
- [20] K. Qotimah, L. Kalangi, and C. Korompis, "Pengaruh Analisa Fundamental Terhadap Return Investasi Pada Saham Second Liner Di Sektor Energi Periode 2019-2022 Yang Terdaftar Di Bursa Efek Indonesia," *J. EMBA J. Ris. Ekon. Manajemen, Bisnis dan Akunt.*, vol. 11, no. 3, pp. 12–26, 2023, doi: 10.35794/emba.v11i3.48797.
- [21] A. Fauzi, "Forecasting Saham Syariah Dengan," *Jurnal Lembaga Keuangan dan Perbankan*, vol. 4, no. 1, pp. 65–69, 2019.
- [22] A. M. Priyatno, L. S. Tanjung, W. F. R. Sudirman, P. Cholidhazia, P. Z. Jati, and F. I. Firmananda, "Comparison Random Forest Regression and Linear Regression For Forecasting BBKA Stock Price," *J. Tek. Ind. Terintegrasi*, vol. 6, no. 3, pp. 718–732, 2023, doi: 10.31004/jutin.v6i3.16933.
- [23] Z. Jin, Y. Yang, and Y. Liu, "Stock closing price prediction based on sentiment analysis and LSTM," *Neural Comput. Appl.*, vol. 32, no. 13, pp. 9713–9729, 2020, doi: 10.1007/s00521-019-04504-2.
- [24] S. S. Al Al-Hasnawi and L. H. Al Al-Hchemi, "Stock Closing Price Prediction of ISX-listed Industrial Companies Using Artificial Neural Networks," *Am. J. Bus. Oper. Res.*, vol. 6, no. 2, pp. 47–55, 2022, doi: 10.54216/ajbor.060205.
- [25] P. Chaudhary, "A Deep Learning-based Approach for Stock Price Prediction," *Turkish J. Comput. Math. Educ.*, vol. 11, no. 3, pp. 1983–1989, 2020, doi: 10.17762/turcomat.v11i3.13595.
- [26] C. Chandra and S. Budi, "Analisis Komparatif ARIMA dan Prophet dengan Studi Kasus Dataset Pendaftaran Mahasiswa Baru," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 278–287, 2020, doi: 10.28932/jutisi.v6i2.2676.
- [27] L. Dong, D. Fang, X. Wang, W. Wei, R. Damaševičius, "Prediction of streamflow based on dynamic sliding window LSTM," *Water (Switzerland)*, vol. 12, no. 11, pp. 1–11, 2020, doi: 10.3390/w12113032.
- [28] A. P. Meriani and A. Rahmatulloh, "Perbandingan Gated Recurrent Unit (GRU) Dan Algoritma Long Short Term Memory (LSTM) Linear Refression Dalam Prediksi Harga Emas," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, 2024.
- [29] F. Lin, S. J. Liu, H. C. Chao, and J. S. Pan, "Short-term household load forecasting model based on variational mode decomposition and gated recurrent unit with attention mechanism," *J. Netw. Intell.*, vol. 6, no. 1, pp. 143–153, 2021.