

PERBANDINGAN ALGORITMA LINEAR REGRESSION, LSTM, DAN GRU DALAM MEMPREDIKSI HARGA SAHAM DENGAN MODEL TIME SERIES

STOCK PRICE PREDICTION USING TIME SERIES MODEL BY COMPARING LINEAR REGRESSION, LSTM, AND GRU ALGORITHMS

Khalis Sofi^{1*}, Aswan Supriyadi Sunge², Sasmitoh Rahmad Riady³, Antika Zahrotul Kamalia⁴

^{1,2,4}Universitas Pelita Bangsa, Bekasi

³President University, Bekasi

*E-mail: sofikhalis7@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk memprediksi harga saham dengan membandingkan algoritma Linear Regression, Long Short-Term Memory (LSTM), dan Gated Recurrent Unit (GRU) dengan dataset publik kemudian menentukan performa terbaik dari ketiga algoritma tersebut. Dataset yang diuji bersumber dari Indonesia Stock Exchange (IDX), yaitu dataset harga saham KEJU berbentuk time series dari tanggal 15 November 2019 sampai dengan 08 Juni 2021. Parameter yang digunakan untuk pengukuran perbandingan adalah RMSE (Root Mean Square Error), MSE (Mean Square Error), dan MAE (Mean Absolute Error). Setelah dilakukan proses training dan testing, dihasilkan sebuah analisis bahwa dari hasil perbandingan algoritma yang digunakan, algoritma Gated Recurrent Unit (GRU) memiliki performance paling baik dibandingkan Linear Regression dan Long-Short Term Memory (LSTM) dalam hal memprediksi harga saham, dibuktikan dengan nilai RMSE, MSE, dan MAE dari uji coba GRU paling rendah, yaitu nilai RMSE 0.034, MSE 0.001, dan nilai MAE 0.024.

Kata kunci: linear regression, LSTM, GRU, harga saham

ABSTRACT

This research aims to predict stock price by comparing Linear Regression, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) algorithms with a public dataset and then determine the best performance of these three algorithms. Dataset is sourced from Indonesia Stock Exchange (IDX), that is stock price with code KEJU. Data used in this research is a time series model from November 15, 2019, to June 08, 2021. The parameters used for comparison measurements are RMSE (Root Mean Square Error), MSE (Mean Square Error), and MAE (Mean Absolute Error). After training and testing process, an analysis is produced that Gated Recurrent Unit (GRU) algorithm has the best performance than Linear Regression and Long-Short Term Memory (LSTM) in terms of predicting stock price, as evidenced by the lowest value of RMSE, MSE, and MAE from the GRU trial, were RMSE 0.034, MSE 0.001, and MAE 0.024

Keywords: linear regression, LSTM, GRU, stock price

PENDAHULUAN

Saham adalah secarik kertas yang menunjukkan hak pemodal (pihak yang memiliki kertas tersebut) untuk memperoleh bagian dari prospek atau kekayaan organisasi yang menerbitkan sekuritas tersebut dan berbagai kondisi yang memungkinkan pemodal tersebut menjalankan haknya [1]. Saham merupakan salah satu instrumen pasar keuangan

yang paling populer. Saham merupakan alat investasi yang banyak dipilih para investor karena saham mampu memberikan tingkat keuntungan yang menarik [2]. Akan tetapi banyak investor yang masih ragu dalam berinvestasi, karena fluktuasi indeks harga saham yang sangat cepat dan dikhawatirkan tidak sesuai dengan apa harapan [3]. Indeks saham adalah ukuran statistik yang

mencerminkan keseluruhan pergerakan harga atas sekumpulan saham yang dipilih berdasarkan kriteria dan metodologi tertentu serta dievaluasi secara berkala. Indeks adalah sebagai alat untuk mengukur sentimen pasar atau kepercayaan investor. Perubahan nilai yang tercermin dalam satu indeks dapat dijadikan indikator yang merefleksikan opini kolektif dari seluruh pelaku pasar [4]. Dengan demikian, prediksi atau peramalan sebuah harga saham menjadi alat bantu penting bagi perencanaan yang lebih efektif dan efisien agar para investor dapat dengan mudah mengukur sentimen pasar, baik prediksi menggunakan machine learning [5] maupun deep learning [6]. Machine Learning adalah pendekatan inovatif yang memiliki aplikasi komprehensif dalam prediksi, mencakup berbagai algoritma dan alat pemodelan yang digunakan untuk serangkaian tugas pemrosesan data, yang telah memasuki sebagian besar disiplin ilmu dalam beberapa tahun terakhir [7]. Deep learning adalah salah satu bentuk pembelajaran mesin yang perkembangannya cukup pesat, deep learning juga merupakan bagian dari Machine Learning yang merupakan salah satu dasar dari artificial intelligence [8]. Banyak algoritma yang dapat digunakan untuk menganalisis harga saham. Dalam penelitian ini, peneliti menggunakan algoritma Long Short Term Memory (LSTM) [9], Linear Regression [10], dan Gated Recurrent Unit (GRU) [11]. Regression adalah teknik yang digunakan untuk dua teori. Pertama, analisis regression biasanya digunakan untuk forecasting dan prediction, dimana aplikasi mereka meliputi major dari machine learning. Kedua, analisis refression digunakan dalam beberapa kasus untuk menentukan hubungan sebab akibat antara independent dan dependent variabel. Yang terpenting, regression sendiri hanya menunjukkan hubungan antara variabel dependent dan kumpulan dataset tetap dari variabel yang berbeda. Linear Regression terdiri dari Simple Linear Regression, Multivariate Linear Regression, dan Polynomial Regression [12]. Beberapa penelitian untuk prediksi juga menggunakan algoritma Linear Regression, diantaranya adalah penelitian yang dilakukan oleh Muhammad Iqbal Ramadhan dan Prihandoko untuk menganalisis jumlah bencana di 34 provinsi di Indonesia dengan variable 7 jenis bencana berdasarkan data dari BNPB tahun 2005-2015 yaitu banjir, gempa bumi, kekeringan, putting beliung, gunung meletus,

tanah longsor, dan tsunami. Hasilnya, algoritma Linear Regression mampu melakukan komputasi untuk memprediksi jumlah bencana di Indonesia untuk 4 tahun berikutnya, yaitu tahun 2016 hingga 2020 [13]. Fitriyani dan Rangga Sanjaya juga melakukan analisis terhadap data kebakaran hutan dengan membandingkan algoritma Linear Regression, K-NN, dan Support Vector Machine. Dari penelitian tersebut dihasilkan nilai RMSE terbaik ditunjukkan oleh hasil komputasi algoritma Linear Regression dengan nilai RMSE sebesar 1.496 [14]. Deny Novianti menganalisis data permohonan paten tahun 2014-2018. Dari penelitian ini, didapatkan bahwa metode Linear Regression layak dan efektif untuk memprediksi pendaftaran paten untuk tahun selanjutnya berdasarkan data pendaftaran dari tahun 2014 sampai tahun 2018. Dengan hasil prediksi terbesar didapatkan oleh Taiwan pendaftaran paten selanjutnya sebesar 378 paten, Canada 1.031 paten, dan China 3.150 paten. Hasil dari rumus prediksi yang diperoleh, rumus prediksi diimplementasikan pada data testing sebanyak 20 data dan hasil yang diperoleh dari rumus prediksi memiliki akurasi 35.52% [15]. Berdasarkan beberapa penelitian terdahulu tersebut, metode Linear Regression efektif menganalisis data berbentuk time series untuk menghasilkan performance yang baik.

Long Short-Term Memory (LSTM) merupakan sebuah evolusi dari arsitektur RNN, dimana pertama kali diperkenalkan oleh Hochreiter & Schmidhuber (1997). Unit LSTM terdiri dari sel memori yang menyimpan informasi yang diperbarui oleh tiga gerbang (gate) khusus, yaitu input gate, forget gate, dan output gate [16]. Penelitian terhadap komputasi algoritma LSTM dilakukan oleh Adhib Arfan dan Lussiana ETP, dengan data uji harga saham di Indonesia. Hasilnya, LSTM mampu memprediksi harga saham pada tahun 2017-2019 dengan performa yang baik, dibuktikan dengan tingkat kesalahan yang kecil [17]. Afriansyah dan Ade Irawan melakukan uji coba untuk memprediksi waktu hidup atau matinya sebuah lampu. Dengan memanfaatkan data time series, model LSTM tersebut mampu memprediksi pola pemakaian lampu dengan akurasi 99.28% dan RMSE sebesar 0.091 [18]. LSTM juga digunakan untuk memprediksi penjualan produk, seperti yang dilakukan oleh Laras Wiranda dan Mujiono. Berdasarkan hasil pemodelan data training terhadap data testing bahwa LSTM mampu memprediksi data deret waktu penjualan produk tahun 2017-2019

sebesar 13,762,154.00 untuk RMSE dalam nilai rupiah dan MAPE sebesar 12% [19]. Dengan demikian, dapat ditarik kesimpulan bahwa algoritma LSTM memiliki performance yang baik dan mampu melakukan komputasi dengan baik terhadap data-data berbentuk time series.

Gated Recurrent Unit (GRU) merupakan algoritma dari Deep Learning, performa GRU mirip dengan LSTM, namun pada GRU hanya terdiri dari dua gate (gerbang) saja, yaitu reset gate dan update gate [20]. Wuyan L, et al melakukan penelitian terhadap Dissolved Oxygen (DO) untuk kualitas air aquaculture. Mereka membandingkan tiga algoritma yaitu RNN, LSTM, GRU dengan menggunakan empat indikator yaitu MAE, MSE, MAPE, dan R2. Setelah dilakukan analisis, maka dihasilkan bahwa algoritma GRU memiliki performa paling baik dibandingkan dua algoritma lainnya, yaitu dengan nilai MAE, MSE, MAPE, dan R2 adalah 0.450 mg/L, 0.411, 0.054, dan 0.994 [21]. Yong-gang Zhang et al juga melakukan perbandingan metode GRU dan SVM untuk menganalisis perpindahan tanah longsor Erdahu dengan data time series. Dari penelitian tersebut, dapat diketahui bahwa performa GRU lebih baik dari SVM dengan menunjukkan tingkat akurasi yang lebih tinggi [22]. Prediksi menggunakan algoritma GRU juga pernah dilakukan oleh Rifky Ahmad Saputra, dkk terhadap data Cargo Service Center Tangerang City untuk memprediksi permintaan kargo dengan data Januari 2016-September 2019. Setelah dilakukan uji coba, didapatkan hasil bahwa performa GRU cukup baik dalam memprediksi permintaan kargo dengan model penentuan menggunakan hyperparameter dan dihasilkan nilai RMSE terbaik sebesar 247.3952 [23]. Berdasarkan penelitian-penelitian tersebut, maka dalam penelitian ini prediksi harga saham dengan model data time series akan cocok jika dianalisis menggunakan algoritma GRU, karena algoritma GRU memiliki performance yang baik untuk menghasilkan nilai error yang kecil terbukti dari penelitian-penelitian yang telah dilakukan sebelumnya.

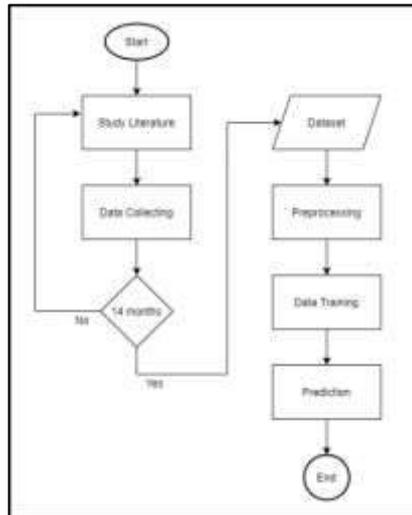
Tujuan dari penelitian ini adalah untuk memprediksi harga saham dengan membandingkan algoritma Linear Regression, Long-Short Term Memory (LSTM), dan Gated Recurrent Unit (GRU) dengan dataset publik berbentuk time series kemudian menentukan performa terbaik dari ketiga algoritma tersebut. Banyak penelitian terdahulu yang membahas

mengenai prediksi harga saham, diantaranya adalah penelitian yang dilakukan oleh R. Hadapiningradja Kusumodestoni, et al. dengan mengelola data saham 20 Juli 2016 pukul 16.26 WIB sampai dengan tanggal 15 September 2016 pukul 17.40 WIB sebanyak 729 dataset, kemudian dilakukan proses training dan testing menggunakan algoritma Support Vector Machine (SVM) dan Neural Network (NN), dan hasilnya adalah algoritma NN memiliki performa yang lebih baik dibandingkan dengan Support Vector Machine, dibuktikan dengan hasil uji model NN akurasi prediksinya sebesar 0.503 +/- 0.009 (mikro 503) sedangkan dengan menggunakan model SVM akurasi prediksinya sebesar 0.477 +/- 0.008 (mikro : 0.477) [24]. Penelitian dengan membandingkan algoritma Support Vector Machine (SVM) dan Multilayer Perceptron (MLP) pernah dilakukan oleh Rusma Insan Nurachim untuk melihat tingkat akurasi dari kedua algoritma tersebut. Setelah dilakukan uji coba, kedua algoritma tersebut mampu melakukan komputasi untuk memprediksi harga saham, dan algoritma MLP memiliki tingkat akurasi lebih tinggi dengan nilai 92,5% dibandingkan dengan SVM yang hanya 90,9% [25]. Andri Pramuntadi melakukan uji coba terhadap dataset saham yang tercatat setiap hari dari tanggal 03 Januari 2000 sampai dengan 13 Juli 2011 dengan membandingkan algoritma Neural Network tanpa Particle Swarm Optimization (PSO) dan dengan PSO. Eksperimen yang dilakukan pada percobaan pertama dilakukan dataset pelatihan dengan NN, dan pada percobaan kedua dengan dataset pelatihan akan dibuat PSO-NN. Dengan menggunakan NN, hasil percobaan yang dilakukan training cycle Neural Network 500, 3 hidden layer, momentum 0 dan learning rate 0,2 mendapatkan 0,466 RMSE. Sedangkan hasil percobaan NN dengan seleksi fitur PSO atau PSO-NN mendapatkan hasil RMSE 0,373. Penelitian yang berbasis PSO-NN ini mampu memprediksi secara lebih akurat [26]. Harga saham JII (Jakarta Islamic Index) dianalisis oleh Sari Setia Ningrum, et al. menggunakan transformasi Wavelet Diskrit Daubechies. Langkah dalam proses memprediksi ini yaitu melakukan estimasi thresholding untuk mendapatkan model terbaik. Hasil penelitian menunjukkan bahwa parameter minimax dengan fungsi hard thresholding maupun soft thresholding memperoleh model terbaik pada level resolusi pertama dan parameter adaptive dengan fungsi soft thresholding memperoleh model terbaik pada level kedua. Namun, model

terbaik untuk memprediksi harga penutupan saham harian JII adalah dengan menggunakan parameter adaptive dengan nilai MAPE (Mean Absolute Percentage Error) sebesar 0,188662% [27]. Prediksi penutupan harga saham juga dilakukan dengan membandingkan metode Backpropagation dan Adaptive Neuro Inference System (ANFIS) oleh Nur Fitriyani dan Seradi Angkasa. Diantara kedua metode tersebut, ANFIS lebih efektif untuk melakukan prediksi harga saham penutupan. Perbandingan prediksi Backpropagation dan ANFIS dengan nilai MAPE 13,44 dan 9,19 [28].

METODOLOGI

1. METODE PENELITIAN



Gambar 1. Flow Metode Penelitian

Berikut adalah metode dalam penelitian ini yang telah diilustrasikan dalam Gambar 1.

- Study literature dari jurnal-jurnal terpercaya yang telah terindeks IEEE, ScienceDirect, SINTA, maupun yang lainnya. Kemudian dilakukan review pada literatur tersebut untuk menemukan topik penelitian.
- Setelah menemukan topik penelitian, kemudian dilakukan pengumpulan data. Data diambil dari Indonesia Stock Exchange (IDX) berupa data penjualan saham dengan kode saham KEJU selama 14 bulan. Dataset tersebut berbentuk time series.
- Kemudian dataset dianalisis tersebut dengan melakukan pelatihan terhadap data-data tersebut. Analisis data menggunakan algoritma Linear Regression, LSTM, dan GRU.
- Setelah dilakukan proses training, maka dihasilkanlah sebuah prediksi, kemudian hasil prediksi dibandingkan untuk menentukan performa terbaik dari ketiga algoritma tersebut.

2. PENGUMPULAN DATA



Gambar 2. Tahapan Pengumpulan Data

Berdasarkan Gambar 2. didapatkan informasi sebagai berikut:

- Objek dalam penelitian ini adalah harga saham pada saham KEJU. Data yang digunakan adalah data publik yang diambil dari <https://www.idx.co.id/> dengan data penjualan saham dari tanggal 25 November 2019 sampai dengan 08 Juni 2021.
- Data diproses untuk ditentukan data training dan data testing.
- Variabel data berisi harga previous saham, harga open, harga high, harga low, harga close, dan volume.

Berikut adalah tabel contoh dataset yang telah diolah.

Table 1. Sample Dataset

date	previous	high	low	close	volume
11/25/2019	750	1125	1125	1125	245800
11/26/2019	1125	1405	1185	1405	5580000
11/27/2019	1405	1755	1405	1505	12136600
11/28/2019	1505	1650	1130	1130	4263000
11/29/2019	1130	1275	1130	1130	1604100
.....
6/2/2021	1255	1270	1245	1265	93600
6/3/2021	1265	1265	1235	1265	48100
6/4/2021	1265	1290	1220	1290	66000
6/7/2021	1290	1310	1245	1270	73200
6/8/2021	1270	1300	1250	1260	196300

3. PREPROCESSING

Bagian ini adalah menjelaskan pengolahan dataset yang telah ada. Pertama, dataset dilakukan proses cleaning guna menghilangkan data yang terdapat missing value. Setelah itu, dilakukanlah proses scaling untuk mempercepat pelatihan data. Kemudian data akan dibagi menjadi dua, yaitu data training dan data testing.

- Data Cleaning, dilakukan untuk menghilangkan data yang terdapat missing value serta menghapus atribut data yang tidak diperlukan. Berikut adalah gambar dataset yang sudah melalui proses cleaning.

	date	previous	high	low	close	volume
0	2019-11-25	750	1125	1125	1125	245800
1	2019-11-26	1125	1405	1185	1405	5580000
2	2019-11-27	1405	1755	1405	1505	12136600
3	2019-11-28	1505	1650	1130	1130	4263000
4	2019-11-29	1130	1275	1130	1130	1604100

Gambar 3. Data Cleaning

- b. Data Scaling, data diubah menggunakan Feature Scaling yang berfungsi untuk membuat data menjadi numerik pada dataset sehingga memiliki rentang nilai yang sama dan tidak ada lagi data yang mendominasi variable data lain, Feature Scaling menggunakan formula sebagai berikut.

Standarisasi:

$$x_{stand} = \frac{x - \text{mean}(x)}{\text{standard deviation}(x)} \quad (1)$$

Normalisasi:

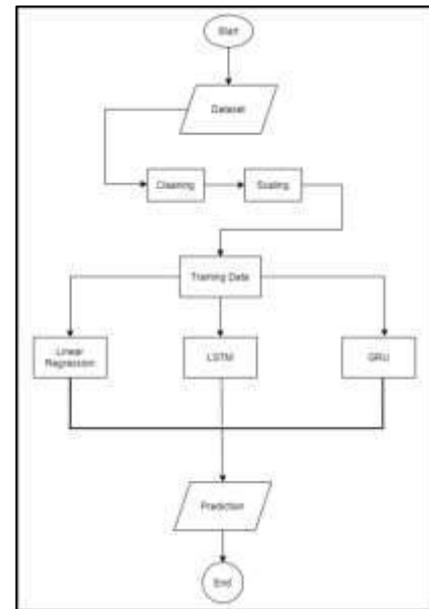
$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

Sehingga data tersebut di-scaler menjadi seperti berikut.

```
[0.06032298 0.0621118 0.02484472 0.04908044 0.02484472 0.04908044
0.04908044 0.0310559 0.03726708 0.0310559 0.02484472 0.02484472
0.02484472 0.02484472 0.02484472 0.0310559 0.01803354 0.02484472
0.02484472 0.01242236 0.05590862 0.07453416 0.07453416 0.11801242
0.19875776 0.19875776 0.1803354 0.18012422 0.19254658 0.19254658
0.19254658 0.1803354 0.1803354 0.18012422 0.1803354 0.1803354
0.19254658 0.1803354 0.1803354 0.19254658 0.19254658 0.20490894
0.19875776 0.19875776 0.19875776 0.22368248 0.31050091 0.32010255
0.29813665 0.27958311 0.27958311 0.29813665 0.31677019 0.34782699
0.39130435 0.41614987 0.45962733 0.47826887 0.48447205 0.51552795
0.52795031 0.43478261 0.42236825 0.39751553 0.39751553 0.39751553
0.41614987 0.40372671 0.36645963 0.37888199 0.40993789 0.39751553
0.39751553 0.45962733 0.42236825 0.45962733 0.45341615 0.45962733
0.47204069 0.47204069 0.48447205 0.49068321 0.51552795 0.52173013
0.53416149 0.55217301 0.68322981 0.68322981 0.68322981 0.68322981
0.68322981 0.67080745 0.05217301 0.33908021 0.68322981 0.67080745
0.67080745 0.65818509 0.04386273]
[0.06032298 0.0621118 0.02484472 0.04908044 0.02484472 0.04908044
0.04908044 0.0310559 0.03726708 0.0310559 0.02484472 0.02484472
0.02484472 0.02484472 0.02484472 0.0310559 0.01803354 0.02484472
0.02484472 0.01242236 0.05590862 0.07453416 0.07453416 0.11801242
0.19875776 0.19875776 0.1803354 0.18012422 0.19254658 0.19254658
0.19254658 0.1803354 0.1803354 0.18012422 0.1803354 0.1803354
0.19254658 0.1803354 0.1803354 0.19254658 0.19254658 0.20490894
0.19875776 0.19875776 0.19875776 0.22368248 0.31050091 0.32010255
0.29813665 0.27958311 0.27958311 0.29813665 0.31677019 0.34782699
0.39130435 0.41614987 0.45962733 0.47826887 0.48447205 0.51552795
0.52795031 0.43478261 0.42236825 0.39751553 0.39751553 0.39751553
0.41614987 0.40372671 0.36645963 0.37888199]
```

Gambar 4. Data Scaling

4. MODEL PELATIHAN



Gambar 5. Model Pelatihan

Dataset yang telah diperoleh berdasarkan penjelasan pada tahap pengumpulan data dilanjutkan dengan preprocessing yang terdiri dari cleaning data dan scaling data. Setelah itu, dilakukanlah proses training data menggunakan algoritma Linear Regression, LSTM, dan GRU. Dari hasil training data, maka dihasilkan prediksi dan kemudian ditentukan performa terbaik dari ketiga algoritma tersebut dengan parameter yang digunakan adalah RMSE, MSE, dan MAE.

5. PENGUKURAN

Pada tahap ini dilakukan pengukuran untuk menentukan performa terbaik dari algoritma Linear Regression, LSTM, dan GRU dengan menentukan nilai Root Mean Square Error (RMSE), Mean Square Error (MSE), dan Mean Absolute Error (MAE).

- a. Root Mean Square Error (RMSE), adalah jumlah dari kesalahan kuadrat atau selisih antara nilai sebenarnya dengan nilai prediksi yang telah ditentukan [29]. Berikut adalah formula untuk menghitung RMSE.

$$RMSE = \sqrt{\frac{\sum (Y' - Y)^2}{n}} \quad (3)$$

Dengan Y' adalah nilai prediksi, Y adalah nilai sejati, dan n merupakan jumlah data.

- b. Mean Square Error (MSE), dalah rata-rata kesalahan kuadrat diantara nilai aktual dan nilai peramalan. Metode MSE secara umum digunakan untuk mengecek estimasi berapa nilai kesalahan pada peramalan. Nilai MSE yang rendah atau mendekati nol menunjukkan bahwa hasil peramalan sesuai dengan data aktual dan bisa dijadikan untuk

perhitungan peramalan di periode mendatang [30]. Berikut adalah formula untuk menghitung MSE.

$$MSE = \sum \frac{(Y' - Y)^2}{n} \quad (4)$$

Dengan Y' adalah nilai prediksi, Y adalah nilai sejati, dan n merupakan jumlah data.

- c. Mean Absolute Error (MAE), menunjukkan nilai kesalahan rata-rata yang error dari nilai sebenarnya dengan nilai prediksi. MAE sendiri secara umum digunakan untuk pengukuran prediksi error pada analisis time series. Berikut adalah formula untuk menghitung MAE [29].

$$MAE = \sum \frac{|Y' - Y|}{n} \quad (5)$$

Dengan Y' adalah nilai prediksi, Y adalah nilai sejati, dan n merupakan jumlah data.

HASIL DAN PEMBAHASAN

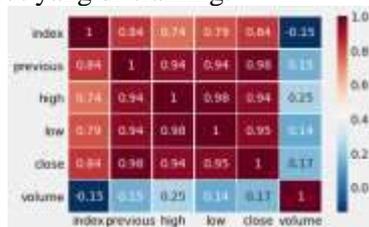
1. TRAINING & TESTING LINEAR REGRESSION

Pada tahap ini, dataset yang sudah diolah akan dilakukan uji coba untuk training pada setiap atribut yang dijadikan inputan. Berikut adalah data input yang akan dilakukan uji coba.

	date	previous	high	low	close	volume	dates
0	2019-11-25	750	1125	1125	1125	245800	0
1	2019-11-26	1125	1405	1185	1405	5580000	1
2	2019-11-27	1405	1755	1405	1505	12136500	2
3	2019-11-28	1505	1650	1130	1130	4263000	3
4	2019-11-29	1130	1275	1130	1130	1604100	4

Gambar 6. Data Input Linear Regression

Dan gambar di bawah ini adalah visualisasi dari data input yang di-training.



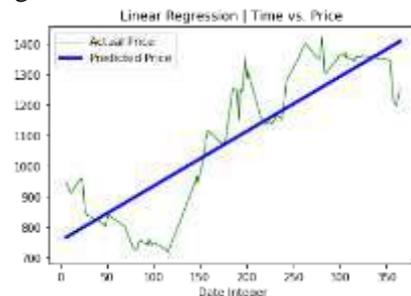
Gambar 7. Visualisasi Data

Dari gambar berikut dapat dilihat bahwa harga saham KEJU sangat fluktuatif dari tahun 2019 sampai dengan 2021. Fluktuasi harga dapat diilustrasikan dengan grafik berikut.



Gambar 8. Histori Harga Saham KEJU 2019-2021

Kemudian dilakukan split data dengan membagi menjadi data training dan data testing. Setelah dilakukan testing, maka dihasilkan nilai prediksi dan nilai aktual yang divisualisasikan dalam grafik berikut.



Gambar 9. Prediksi VS Aktual Linear Regression

Setelah dilakukan proses percobaan, maka dihasilkan nilai RMSE, MSE, dan MAE dari data prediksi dan aktual yang digambarkan dalam tabel berikut.

Table 2. Error Rate Linear Regression

Linear Regression	
Measurement	Value
RMSE	4.621
MSE	2.136
MAE	2.890

2. TRAINING & TESTING LSTM

Percobaan selanjutnya adalah training dataset menggunakan algoritma LSTM. Berikut adalah data input yang akan di-training.

	previous	high	low	close	volume
2019-11-25	750	1125	1125	1125	245800
2019-11-26	1125	1405	1185	1405	5580000
2019-11-27	1405	1755	1405	1505	12136500
2019-11-28	1505	1650	1130	1130	4263000
2019-11-29	1130	1275	1130	1130	1604100

Gambar 10. Data Input LSTM

Kemudian data dibagi menjadi data training dan data testing, dan grafik berikut menunjukkan perbandingan hasil prediksi dengan data aktual.



Gambar 11. Prediksi VS Aktual Training LSTM

Setelah dilakukan proses percobaan, maka dihasilkan nilai RMSE, MSE, dan MAE dari data prediksi dan aktual yang digambarkan dalam tabel berikut.

Table 3. Error Rate LSTM

Long-Short Term Memory	
Measurement	Value
RMSE	0.048
MSE	0.002
MAE	0.038

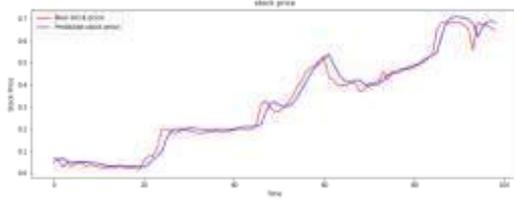
3. TRAINING & TESTING GRU

Percobaan terakhir menggunakan algoritma Gated Recurrent Unit. Berikut adalah data input yang akan di-training.

date	previous	high	low	close	volume
2019-11-25	750	1125	1125	1125	245800
2019-11-26	1125	1405	1185	1405	5580000
2019-11-27	1405	1755	1405	1505	12136500
2019-11-28	1505	1650	1130	1130	4263000
2019-11-29	1130	1275	1130	1130	1604100

Gambar 12. Data Input GRU

Kemudian data dibagi menjadi data training dan data testing, dan grafik berikut menunjukkan perbandingan hasil prediksi dengan data aktual.



Gambar 13. Prediksi VS Aktual Training GRU

Setelah dilakukan proses percobaan, maka dihasilkan nilai RMSE, MSE, dan MAE dari data prediksi dan aktual yang digambarkan dalam tabel berikut.

Table 4. Error Rate GRU

Gated Recurrent Unit	
Measurement	Value
RMSE	0.034
MSE	0.001
MAE	0.024

KESIMPULAN

Model pelatihan dalam penelitian ini adalah menggunakan algoritma Linear

Regression, LSTM, dan GRU dengan tujuan membandingkan performance dari ketiga algoritma tersebut terhadap data harga saham KEJU yang berbentuk time series dari tanggal 15 November 2019 sampai dengan 08 Juni 2021 dengan pengukuran nilai RMSE, MSE, dan MAE. Setelah dilakukan proses training dan testing, hasil yang diperoleh adalah bahwasanya algoritma GRU memiliki performance terbaik dibandingkan dua algoritma lainnya dalam memprediksi harga saham, dibuktikan dengan nilai RMSE, MSE, dan MAE dari uji coba GRU paling rendah, yaitu RMSE 0.034, MSE 0.001, dan MAE 0.024.

SARAN

Berdasarkan hasil penelitian yang telah dijelaskan dalam kesimpulan, makan saran peneliti untuk penelitian selanjutnya adalah untuk selanjutnya penelitian dapat menggunakan dataset yang lain dan menggunakan algoritma deep learning lainnya.

UCAPAN TERIMA KASIH

Terima kasih kepada orang-orang yang telah mendukung terbitnya jurnal ini dan instansi terkait, Universitas Pelita Bangsa dan President University. Kemudian saya ucapkan terima kasih kepada Ngoding Study Club, komunitas yang membuat saya tumbuh.

DAFTAR PUSTAKA

- [1] H. Dalimunthe, "Pengaruh Marjin Laba Bersih, Pengembalian Atas Ekuitas, Dan Inflasi Terhadap Harga Saham," *J. Akunt. DAN BISNIS J. Progr. Stud. Akunt.*, vol. 4, no. 2, p. 62, 2018.
- [2] A. Novita, "Prediksi Pergerakan Harga Saham Pada Bank Terbesar Di Indonesia Dengan Metode Backpropagation Neural Network," *Jutisi*, vol. 05, no. 01, pp. 965–972, 2016.
- [3] A. Izzah and R. Widyastuti, "Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression untuk Pencegahan Data Outlier," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 2, no. 3, pp. 141–150, 2017.
- [4] Indonesia Stock Exchange, "IDX Stock Index Handbook V1.2," p. 52, 2019.
- [5] R. Chowdhury, M. R. C. Mahdy, T. N. Alam, G. D. Al Quaderi, and M. Arifur Rahman, "Predicting the stock price of frontier markets using machine learning

- and modified Black–Scholes Option pricing model,” *Phys. A Stat. Mech. its Appl.*, vol. 555, p. 124444, 2020.
- [6] H. Liu and Z. Long, “An improved deep learning model for predicting stock market price time series,” *Digit. Signal Process. A Rev. J.*, vol. 102, p. 102741, 2020.
- [7] G. Carleo *et al.*, “Machine learning and the physical sciences,” *Rev. Mod. Phys.*, vol. 91, no. 4, p. 45002, 2019.
- [8] V. K. Ayyadevara, “Recurrent Neural Network,” *Pro Mach. Learn. Algorithms*, pp. 217–257, 2018.
- [9] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, “Deep Learning with Long Short-Term Memory for Time Series Prediction,” *IEEE Commun. Mag.*, vol. 57, no. 6, pp. 114–119, 2019.
- [10] S. Rath, A. Tripathy, and A. R. Tripathy, “Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model,” *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 5, pp. 1467–1474, 2020.
- [11] D. Lien Minh, A. Sadeghi-Niaraki, H. D. Huy, K. Min, and H. Moon, “Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network,” *IEEE Access*, vol. 6, no. c, pp. 55392–55404, 2018.
- [12] D. Maulud and A. M. Abdulazeez, “A Review on Linear Regression Comprehensive in Machine Learning,” *J. Appl. Sci. Technol. Trends*, vol. 1, no. 4, pp. 140–147, 2020.
- [13] M. I. Ramadhan, “Penerapan Data Mining untuk Analisis Data Bencana Milik BNPB Menggunakan Algoritma K-Means dan Linear Regression,” *J. Inform. dan Komput.*, vol. 22, no. 1, pp. 57–65, 2017.
- [14] Fitriyani and R. Sanjaya, “Komparasi Algoritma Lr , K-Nn Dan Svm Untuk Estimasi Area Kebakaran Hutan,” *Infotronik*, vol. 3, no. 2, pp. 103–110, 2018.
- [15] D. Novianti, N. Palasara, and M. Qomaruddin, “Algoritma Regresi Linear pada Prediksi Permohonan Paten yang Terdaftar di Indonesia Linear Regression Algorithm for Prediction of Patent Applications Registered in Indonesia,” vol. 09, no. 2, pp. 81–85, 2021.
- [16] J. Cao, Z. Li, and J. Li, “Financial time series forecasting model based on CEEMDAN and LSTM,” *Phys. A Stat. Mech. its Appl.*, vol. 519, pp. 127–139, 2019.
- [17] A. Arfan and L. ETP, “Prediksi Harga Saham Di Indonesia Menggunakan Algoritma Long Short-Term Memory,” *SeNTIK*, vol. 3, no. 1, pp. 225–230, 2019.
- [18] A. Afriansyah and A. Irawan, “Otomatisasi Pengoperasian Alat Elektronik Berdasarkan Hasil Prediksi Algoritma Long Short Term Memory,” *JITCE (Journal Inf. Technol. Comput. Eng.)*, vol. 4, no. 02, pp. 83–89, 2020.
- [19] L. Wiranda and M. Sadikin, “Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma,” *J. Nas. Pendidik. Tek. Inform.*, vol. 8, no. 3, pp. 184–196, 2019.
- [20] N. Giarsyani, “Komparasi Algoritma Machine Learning dan Deep Learning untuk Named Entity Recognition : Studi Kasus Data Kebencanaan,” *Indones. J. Appl. Informatics*, vol. 4, no. 2, p. 138, 2020.
- [21] W. Li, H. Wu, N. Zhu, Y. Jiang, J. Tan, and Y. Guo, “Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU),” *Inf. Process. Agric.*, vol. 8, no. 1, pp. 185–193, 2021.
- [22] Y. gang Zhang, J. Tang, Z. ying He, J. Tan, and C. Li, *A novel displacement prediction method using gated recurrent unit model with time series analysis in the Erdaohe landslide*, vol. 105, no. 1. Springer Netherlands, 2021.
- [23] R. A. Saputra, “Prediksi Permintaan Kargo pada Cargo Service Center Tangerang City Menggunakan Metode Gated Recurrent Unit,” *J. Repos.*, vol. 2, no. 8, pp. 1113–1122, 2020.
- [24] R. H. Kusumodestoni and S. Sarwido, “Komparasi Model Support Vector Machines (Svm) Dan Neural Network Untuk Mengetahui Tingkat Akurasi Prediksi Tertinggi Harga Saham,” *J. Inform. Upgris*, vol. 3, no. 1, 2017.
- [25] R. I. Nurachim, “Pemilihan Model Prediksi Indeks Harga Saham Yang Dikembangkan Berdasarkan Algoritma Support Vector Machine (Svm) Atau Multilayer Perceptron (Mlp) Studi

- Kasus : Saham Pt Telekomunikasi Indonesia Tbk,” vol. 5, no. 1, pp. 29–35, 2019.
- [26] A. Pramuntadi, “Model Prediksi Rentet Waktu Neural Network Berbasis Particle Swarm Optimization Untuk Prediksi Harga Saham,” *Telematika*, vol. 14, no. 2, pp. 100–106, 2017.
- [27] S. S. N. H. Fransiskus Fran, “Prediksi Harga Saham Jii Menggunakan Transformasi Wavelet Diskrit Daubechies,” *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 8, no. 4, pp. 659–666, 2019.
- [28] N. Fitriadini and S. Angkasa, “Jaringan Syaraf Tiruan Backpropagation Dan Adaptive Neuro Inference System (Anfis),” *Prediksi Harga Saham Menggunakan Metod. Jar. Syaraf Tiruan Backpropagation Dan Adapt. Neuro Inference Syst.*, vol. 15, pp. 1–58, 2020.
- [29] T. Chai and R. R. Draxler, “Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature,” *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014.
- [30] M. S. Salman, O. Kukrer, and A. Hocanin, “Recursive inverse algorithm: Mean-square-error analysis,” *Digit. Signal Process. A Rev. J.*, vol. 66, pp. 10–17, 2017.