

Algoritma *Gated Recurrent Unit* untuk Prediksi Harga Indeks Penutupan Saham LQ45

Venia R. Danestiara¹, Elia Setiana², Imannudin Akbar³, Taufik Hidayah⁴

^{1,2}Informatika, Universitas Informatika dan Bisnis Indonesia, Indonesia

^{3,4}Sistem Informasi, Universitas Informatika dan Bisnis Indonesia, Indonesia

veniarestreva@unibi.ac.id

Info Artikel

Sejarah artikel :

Diterima Februari 2024

Direvisi Maret 2024

Disetujui Maret 2024

Diterbitkan Maret 2024

ABSTRACT

The Indonesia Stock Exchange (IDX) states that stocks, including LQ45 stocks, which constitute the stock market index for the IDX, have become one of the preferred investment options for the public. Investors need to have accurate analysis and information to gain significant profits as stock prices fluctuate due to company performance, industry factors, changes in interest rates, liquidity, global market conditions, market sentiment, and investor psychology. The *Gated Recurrent Unit* algorithm is suitable for application on historical stock data sets because they are time series data, can be computed and compared on a numerical scale. This algorithm is a variant of the *Long Short-Term Memory* algorithm or other types of processing modules for *Recurrent Neural Networks*. The data set used consists of closing price data or close features, comprising a training data set of 4,406 data and a test data set of 1,889 data that have undergone data preparation using various techniques, including data cleansing, data scrubbing, data splitting, data normalization, and data reshaping. The results showed that the *Gated Recurrent Unit* algorithm is the right strategy because it obtains a good evaluation of model performance, namely MSE of 0.0009; RMSE of 0.17325 and MAE of 0.0207.

Keywords : *Gated Recurrent Unit*; LQ45 Stock; Stock Price Prediction.

ABSTRAK

The Indonesia Stock Exchange (IDX) atau PT Bursa Efek Indonesia menyatakan bahwa saham, termasuk saham LQ45 yang merupakan indeks pasar saham untuk IDX, telah menjadi salah satu pilihan investasi yang diminati oleh masyarakat. Para investor perlu memiliki analisis dan informasi akurat agar mendapatkan banyak keuntungan karena harga saham mengalami fluktuasi yang disebabkan oleh kinerja perusahaan, faktor industri, perubahan dalam tingkat bunga, likuiditas, kondisi pasar global, sentimen pasar dan psikologi investor. Algoritma *Gated Recurrent Unit* tepat untuk diterapkan pada himpunan data histori saham karena bersifat *time series*, dapat dihitung dan dibandingkan pada skala numerik. Algoritma ini merupakan varian dari algoritma *Long Short-Term Memory* atau jenis modul pemrosesan lain untuk algoritma *Recurrent Neural Network*. Himpunan data yang digunakan berupa data harga penutupan atau fitur *close* yang terdiri dari data latih berisi 4.406 data dan data uji berisi 1.889 data yang telah melalui data *preparation* dengan berbagai teknik, diantaranya data *cleansing*, data *scrubbing*, data *splitting*, data *normalization* dan data *reshaping*. Hasil penelitian menunjukkan bahwa algoritma *Gated Recurrent Unit* merupakan strategi yang tepat karena memperoleh evaluasi performa model yang baik, yakni MSE sebesar 0,0009; RMSE sebesar 0,17325 dan MAE sebesar 0,0207.

Kata Kunci : *Gated Recurrent Unit*; Prediksi Harga Saham; Saham LQ45.

PENDAHULUAN

Menurut *The Indonesia Stock Exchange (IDX)* atau PT Bursa Efek Indonesia, saham menjadi salah satu investasi yang diminati oleh masyarakat [1]. Perkembangan teknologi menciptakan banyak platform digital dan perusahaan broker saham yang memberikan kemudahan sehingga menyebabkan berinvestasi di pasar modal semakin populer. Salah satunya adalah saham LQ45 yang merupakan indeks pasar saham untuk Bursa Efek Indonesia yang terdiri dari 45 perusahaan yang memenuhi kriteria tertentu. Peningkatan minat masyarakat berinvestasi dengan kesadaran bahwa lebih baik memiliki *passive income* daripada hanya menyimpan uang saja di bank. Tidak heran banyak masyarakat Indonesia yang mengubah gaya hidup dari *saving society* menjadi *investing society*. Meskipun demikian, berinvestasi saham perlu dilakukan dengan beberapa pemahaman agar mendapatkan banyak keuntungan. Analisa utama yang diperlukan oleh para investor adalah prediksi harga saham karena bersifat fluktuatif [2]. Informasi mengenai prediksi harga saham yang akurat dapat menguntungkan masyarakat yang menjadi investor atau yang hendak menjadi investor.

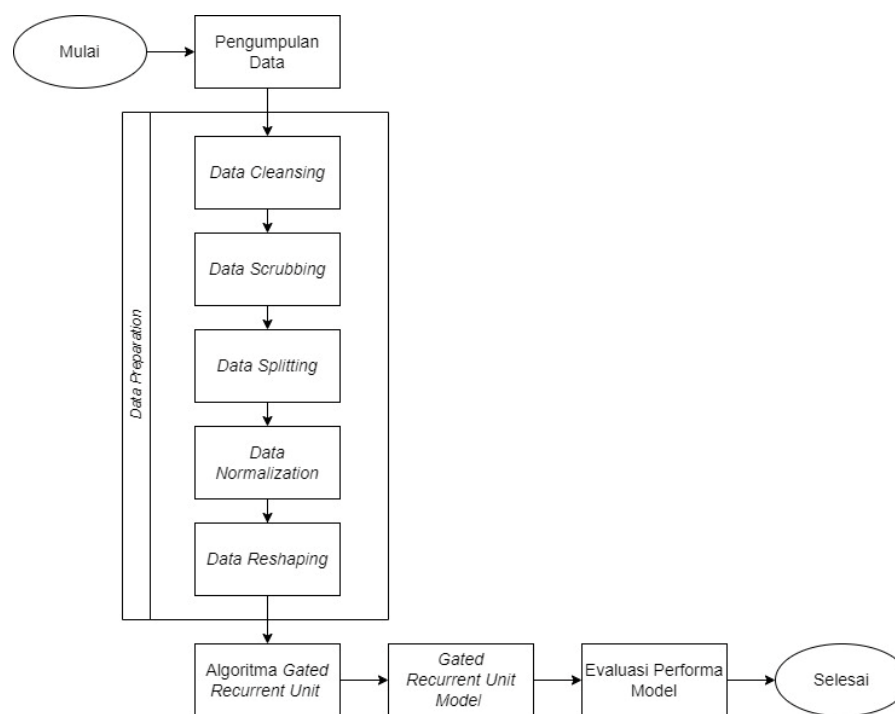
Beberapa penelitian prediksi harga saham dengan menggunakan algoritma *Gated Recurrent Unit* mendapatkan hasil yang akurat [3, 4, 5]. Algoritma ini merupakan varian dari algoritma *Long Short-Term Memory* atau jenis modul pemrosesan lain untuk algoritma *Recurrent Neural Network* [6, 7, 8]. Algoritma *Gated Recurrent Unit* memiliki keutamaan komputasi yang lebih sederhana karena memiliki parameter yang lebih sedikit, namun mempunyai akurasi yang setara dan efektif untuk menghindari permasalahan gradien yang menghilang [9, 10, 11, 12]. Sehubungan dengan itu, penggunaan jumlah himpunan data yang sedikit pada penelitian ini lebih tepat diimplementasikan dengan algoritma *Gated Recurrent Unit*. Pada prinsipnya, himpunan data yang sedikit, namun diterapkan dengan parameter yang banyak dapat menimbulkan *overfitting* [13]. Hal lain yang menarik, *Gated Recurrent Unit* merupakan algoritma yang tercipta kurang dari satu dekade.

METODE

Algoritma *Gated Recurrent Unit* merupakan pilihan yang tepat untuk memprediksi himpunan data *time series*. Himpunan data diperoleh dari *Yahoo Finance* yang berjumlah 6.506 data dari tanggal 3 Maret 1997 hingga 30 Desember 2022 dengan 7 fitur. Fitur-fitur tersebut berisi istilah yang digunakan untuk jual beli saham, yaitu *date*, *open*, *high*, *low*, *close*, *adj close* dan *volume*. Gambar 1 menunjukkan himpunan data awal sebelum dilakukan penelitian. Data tersebut dapat dihitung dan dibandingkan pada skala numerik karena bersifat kuantitatif.

	Date	Open	High	Low	Close	Adj Close	Volume
0	1997-03-03	148.250000	148.250000	146.529999	146.919998	146.919998	0.000000e+00
1	1997-03-04	146.869995	147.350006	145.149994	145.919998	145.919998	0.000000e+00
2	1997-03-05	146.190002	146.190002	142.300003	143.869995	143.869995	0.000000e+00
3	1997-03-06	143.929993	144.479996	142.759995	142.759995	142.759995	0.000000e+00
4	1997-03-07	142.869995	143.229996	142.240005	142.910004	142.910004	0.000000e+00
...
6501	2022-12-26	935.150024	935.150024	935.150024	935.150024	935.150024	0.000000e+00
6502	2022-12-27	935.150024	941.979980	935.119995	940.380005	940.380005	3.701200e+09
6503	2022-12-28	940.390015	942.559998	930.520020	933.690002	933.690002	2.936800e+09
6504	2022-12-29	933.690002	940.059998	926.799988	939.869995	939.869995	5.039200e+09
6505	2022-12-30	939.880005	944.830017	937.179993	937.179993	937.179993	3.950600e+09

Gambar 1. Himpunan Data Yahoo Finance



Gambar 2. Alur Penelitian

Prediksi harga indeks penutupan saham LQ45 melalui beberapa tahapan yang tertera pada Gambar 2. Tahapan yang dilakukan setelah pengumpulan data, yakni *Data Preparation*. Keakuratan dan konsistensi data saham LQ45 dipastikan agar dapat diprediksi dengan berbagai teknik yang diterapkan secara sistematis. Diawali dengan *data cleansing* dan diakhiri dengan *data reshaping*. Tahapan selanjutnya, membuat model untuk prediksi harga saham menggunakan *Gated Recurrent Unit* dan evaluasi model tersebut.

Data Cleansing dilakukan untuk memodifikasi atau menghilangkan data yang tidak diperlukan. Penelitian ini memprediksi harga penutupan atau fitur *close* sehingga seluruh data diabaikan kecuali fitur *date* dan fitur *close* yang ditunjukkan pada Gambar 3. Setelah itu, dilakukan pemeriksaan nilai yang kosong pada

himpunan data, seperti pada Gambar 4. Berdasarkan hasil pencarian, ditemukan 211 nilai kosong sehingga data tersebut dihapus. Penghapusan data dengan nilai kosong menjadi pendekatan yang sering dilakukan oleh para peneliti karena efektif dan efisien diterapkan ketika data tersebut berjumlah jauh lebih sedikit dibandingkan dengan jumlah data tanpa nilai yang kosong. Selain itu, teknik pengisian nilai kosong dengan nilai rata-rata tidak dipilih agar lebih sesuai dengan kenyataan.

	Date	Close
0	1997-03-03	146.919998
1	1997-03-04	145.919998
2	1997-03-05	143.869995
3	1997-03-06	142.759995
4	1997-03-07	142.910004
...
6501	2022-12-26	935.150024
6502	2022-12-27	940.380005
6503	2022-12-28	933.690002
6504	2022-12-29	939.869995
6505	2022-12-30	937.179993

Gambar 3. Himpunan Data Fitur *Date* dan *Close*

	Date	Close
19	1997-03-28	NaN
20	1997-03-31	NaN
27	1997-04-09	NaN
34	1997-04-18	NaN
48	1997-05-08	NaN
...
5622	2019-06-03	NaN
5623	2019-06-04	NaN
5624	2019-06-05	NaN
5625	2019-06-06	NaN
5626	2019-06-07	NaN

211 rows x 2 columns

Gambar 4. Data dengan Nilai Kosong

Data Scrubbing dilakukan agar tidak terkendala dalam pengoperasian numerik terhadap data-data. Tipe data fitur *date* perlu disesuaikan dari *object* menjadi *datetime64[ns]*. Gambar 5 dan Gambar 6 menunjukkan sebelum *data scrubbing* dan setelah *data scrubbing*.

```
Date      object
Close     float64
dtype: object
```

Gambar 5. Sebelum *Data Scrubbing*

```
Date      datetime64[ns]
Close     float64
dtype: object
```

Gambar 6. Setelah *Data Scrubbing*

Data Splitting diterapkan untuk membagi himpunan data menjadi dua bagian, yaitu data latih dan data uji. Data latih sebesar 70% dan data uji sebesar 30% dari himpunan data. Jadi, data latih berisi 4.406 data dan data uji berisi 1.889 data yang tertera pada Gambar 7.



Gambar 7. Hasil *Data Splitting*

Teknik *Min-Max* diterapkan untuk *Data Normalization*, yakni setiap nilai pada fitur *close* dikurangi dengan nilai minimum pada fitur *close*, kemudian dibagi dengan rentang nilai atau nilai maksimum dikurangi nilai minimum dari fitur *close*.

```
array([[0.10727683],
       [0.10617993],
       [0.10393128],
       [0.10271372],
       [0.10287827]])
```

Gambar 8. Sampel Data setelah *Data Normalization*

Data Reshaping merupakan proses tambahan yang dilakukan untuk mengubah dimensi array dari hasil *Data Normalization*. Array masukan yang direpresentasikan oleh *X* berupa array 3D sedangkan array keluaran atau target yang direpresentasikan oleh *Y* berupa array 2D. Himpunan data berbentuk *time series* maka *X* memiliki axis berupa *instance*, *timestamp*, *future* dan *Y* memiliki axis berupa *instance*, *future* yang ditampilkan pada Gambar 9.

```
Shape of X_train after reshape : (4346, 60, 1)
Shape of y_train after reshape : (4346, 1)
```

Gambar 9. *Data Reshaping*

Algoritma *Gated Recurrent Unit* yang dibangun menggunakan 4 *hidden layer*. Setiap *hidden layer* terdapat 64 *units*, menggunakan fungsi aktivasi *tanh* untuk setiap unit, menggunakan optimasi *Adam* untuk melakukan *update* bobot, dan melakukan 100 *epoch* atau iterasi *training*. Setelah proses ini, didapatkan *Gated Recurrent Unit Model*.

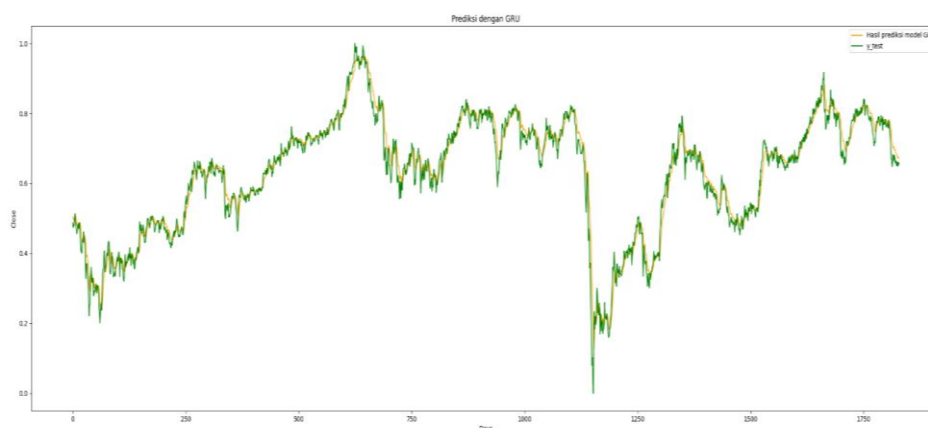
Pada umumnya, evaluasi performa model regresi menggunakan perhitungan dengan pendekatan numerik, yaitu *MSE* (*Mean Squared Error*), *RMSE* (*Root Mean Squared Error*) dan *MAE* (*Mean Absolute Error*). *MSE* dan *RMSE* digunakan untuk mengetahui besar kesalahan dari model yang dibandingkan dengan data aktual sedangkan *MAE* digunakan untuk mengetahui rata-rata kesalahan absolut antara hasil prediksi dengan nilai aktual.

HASIL DAN PEMBAHASAN

Berdasarkan penjelasan pada metode penelitian, model *Gated Recurrent Unit* yang dibangun menghasilkan perbandingan antara data *Date* dan *Close* dari hasil prediksi dan data aktual. Kemudian, evaluasi model *Gated Recurrent Unit* dilakukan untuk memastikan keandalan serta kualitas prediksi yang dihasilkan dan memberikan penilaian seberapa baik model dapat memprediksi data baru.

Hasil Prediksi

Nilai prediksi indeks penutupan saham LQ45 relatif dekat dengan nilai data aktual indeks penutupan saham LQ45 menunjukkan bahwa *Gated Recurrent Unit* memiliki tingkat keakuratan yang baik. Pada Gambar 10, hasil prediksi direpresentasikan oleh garis berwarna kuning dan data aktual direpresentasikan oleh garis berwarna hijau.



Gambar 10. Hasil Prediksi dan Data Aktual

Evaluasi Performa Model

Evaluasi model prediksi yang paling baik bernilai 0 untuk metrik regresi *MSE*, *RMSE* dan *MAE*. Ketiga metrik regresi tersebut digunakan. Hasil evaluasi performa *Gated Recurrent Unit Model* yang telah dibuat sangat baik karena mendekati 0 untuk setiap metrik regresi.

Tabel 1. Evaluasi Performa Model

Metrik Regresi	<i>Gated Recurrent Unit Model</i>
MSE	0,0009
RMSE	0,17325
MAE	0,0207

Salah satu keunggulan *Gated Recurrent Unit* dalam prediksi harga saham adalah kemampuannya untuk memahami dan memodelkan pola dalam data berurutan dengan baik. Dalam konteks harga saham, *Gated Recurrent Unit* dapat menangkap hubungan temporal yang kompleks, seperti fluktuasi harga saham dari waktu ke waktu. Selain itu, *Gated Recurrent Unit* mampu menangani masalah pemrosesan data yang berkelanjutan dan konsisten dengan sifat data harga saham yang terdiri dari serangkaian titik waktu berurutan. Mekanisme gerbang dalam *Gated Recurrent Unit* juga membantu dalam mengatasi masalah *vanishing gradient* yang umum terjadi dalam jaringan saraf rekuren tradisional, seperti *Long Short-Term Memory* sehingga memungkinkan *Gated Recurrent Unit* untuk mempertahankan informasi jangka panjang dengan lebih baik. Fleksibilitas *Gated Recurrent Unit* dalam menyesuaikan *hyperparameter* juga memungkinkan penyesuaian model untuk berbagai jenis data dan masalah prediksi sehingga meningkatkan kemampuannya dalam menerapkan analisis harga saham yang efisien dan responsif terhadap perubahan harga pasar. Dengan kombinasi keunggulan ini, *Gated Recurrent Unit* menjadi alat yang kuat dalam mengembangkan model prediksi harga saham yang akurat dan efektif.

PENUTUP

Fluktuasi harga saham sangat memungkinkan untuk diprediksi menggunakan algoritma *Gated Recurrent Unit* dengan baik. Nilai performa model pada penelitian ini mendekati 0 dengan perhitungan MSE, RMSE, dan MAE. Selain itu, penelitian-penelitian sebelumnya juga menghasilkan performa model yang baik, sehingga layak dicoba untuk memprediksi dengan himpunan data indeks saham yang lain. Oleh karena itu, dengan mempertimbangkan kemampuan *Gated Recurrent Unit* yang telah teruji serta hasil penelitian sebelumnya, pengembangan lebih lanjut terhadap model prediksi harga saham dengan dataset lain dapat menjadi langkah yang menjanjikan untuk meningkatkan pemahaman terhadap dinamika pasar.

Sebagai arah penelitian selanjutnya, perlu dilakukan penelitian yang lebih mendalam untuk mengeksplorasi integrasi *Gated Recurrent Unit* dengan teknik-teknik baru seperti *Attention Mechanism* atau penggabungan dengan data fundamental perusahaan untuk meningkatkan kinerja prediksi harga saham. Selain itu, memperluas cakupan analisis untuk menyertakan pasar global dan mengintegrasikan data non-keuangan seperti sentimen media sosial atau berita juga dapat menjadi bidang yang menarik untuk dieksplorasi. Dengan memperluas jangkauan dan meningkatkan keakuratan model prediksi harga saham, diharapkan dapat memberikan kontribusi yang berarti bagi praktisi pasar keuangan dalam mengambil keputusan investasi yang lebih baik di masa depan.

DAFTAR PUSTAKA

- [1] IDX Indonesia Stock Exchange, "Produk Saham," IDX - Indonesia Stock Exchange, Diakses pada 15 Maret 2024, Tautan: <https://www.idx.co.id/id/produk/saham>.
- [2] A. Nilsen, "Perbandingan Model RNN, Model LSTM, dan Model GRU dalam Memprediksi Harga Saham-Saham LQ45," *J. Stat. dan Apl.*, vol. 6, no. 1, pp. 137-147, 2022, doi: 10.21009/jsa.06113.
- [3] W. Hastomo, A. S. B. Karno, N. Kalbuana, E. Nisfiani, and L. ETP, "Optimasi Deep Learning untuk Prediksi Saham di Masa Pandemi Covid-19," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 2, p. 133, 2021, doi: 10.26418/jp.v7i2.47411.
- [4] A. I. Caniago, W. Kaswidjanti and Juwariah, "Recurrent Neural Network With Gate Recurrent Unit For Stock Price Prediction," vol. 18, no. 3, pp. 345-360, 2021, doi: 10.31515/telematika.v18i3.6650.
- [5] L. Azizah and Y.L. Sukestiyarno, "Metode Genetic Algorithm - Long Short-Term Memory pada Peramalan Harga Saham," vol. 11, no. 2, pp. 153-160, 2022.
- [6] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [7] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1724-1734, 2014, doi: 10.3115/v1/d14-1179.
- [8] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," pp. 1-9, 2014, [Online]. Available: <http://arxiv.org/abs/1412.3555>
- [9] 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.
- [10] B. W. Akramunnas, L. Hakim, D. M. Putri, Fahrizal, A. Rahmawati and Y. Purbolingga, "Comparison of MDKA Stock Price Prediction using Multi-Layer Perceptron , Long Short-Term Memory, and Gated Recurrent Unit," vol. 10, no. 1, pp. 738-743, 2023.
- [11] Idham, M. G. T. Akbar, S. Panggabean, and M. Noor, "Perbandingan Prediksi Harga Saham dengan Menggunakan LSTM GRU dengan Transformer," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 11, no. 1, pp. 44-47, 2022, doi: 10.30591/smartcomp.v11i1.3185.
- [12] A. Zhang, Z. C. Lipton, M. Li and A. J. Smola, *Dive into Deep Learning*. City: Cambridge, 2023.
- [13] V. Maini and S. Sabri, *Machine Learning for Humans*, 2017.