

PREDIKSI HARGA LUNA CLASSIC (LUNC) DENGAN METODE LONG SHORT TERM MEMORY (LSTM)

Abdul Hadi¹, Prihastuti Harsani², Mulyati^{3,*}

^{1,2,3}Department of Computer Science, Faculty of Mathematics and Natural Science, Pakuan University, Bogor, West Java, 16143, Indonesia

Abstrak

Penelitian ini melakukan analisis terhadap Prediksi Harga Luna Classic (LUNC) Dengan Metode Long Short Term Memory (LSTM). Metode penelitian yang digunakan adalah Knowledge Discovery in Database (KDD) memiliki serangkaian tahapan yaitu Data Cleaning & Integration, Data Selection & Transformation, Data Mining, dan Evaluation and Presentation. Tujuan dari penelitian ini adalah memprediksi Harga Luna Classic (LUNC) dengan Metode Long Short Term Memory (LSTM) dengan prediksi harga sesuai dataset yang dipakai untuk mencari pemodelan yang paling baik dan akurat. Data pada penelitian ini diperoleh dari website yahoo.finance.com yang merupakan data harian time series tanggal 3 Februari 2021 sampai dengan 11 Mei 2022. Proses pembuatan model prediksi dilakukan menggunakan software Jupyter Notebook. Untuk mencari model dengan tingkat kesalahan error terkecil dilakukan 4 eksperimentasi pada menggunakan rasio 90:10, 80:20, 70:30, dan 60:40, lalu dilakukan juga ekperimentasi pada perubahan hyperparameter yang berbeda pada parameter neuron 10, 20, 30, 40, 50 dan epoch 100, 500, 1000. Dari keseluruhan percobaan yang telah dilakukan didapatkan hasil yang paling optimal pada percobaan dengan rasio 80:20, parameter neuron 40 dan epoch 1000. RMSE yang didapatkan sebesar 0.638, perolehan nilai MAPE sebesar 1.66%. Dilakukan juga percobaan pada LSTM 1 variabel dengan rasio 80:20 parameter neuron 40 dan epoch 1000. RMSE yang didapatkan sebesar 8.579, dengan perolehan nilai MAPE sebesar 9.10%.

Kata Kunci: Luna Classic; Cryptocurrency; Data Mining; Prediksi; LSTM

Abstract

This research analyzes the Luna Classic (LUNC) Price Prediction using the Long Short Term Memory (LSTM) Method. The research method used is Knowledge Discovery in Database (KDD) which has a series of stages, namely Data Cleaning & Integration, Data Selection & Transformation, Data Mining, and Evaluation and Presentation. The aim of this research is to predict the price of Luna Classic (LUNC) using the Long Short Term Memory (LSTM) method with price predictions according to the dataset used to find the best and most accurate modeling. The data in this research was obtained from the website yahoo.finance.com which is daily time series data from February 3 2021 to May 11 2022. The process of creating a prediction model was carried out using Jupyter Notebook software. To find a model with the smallest error rate, 4 experiments were carried out using ratios of 90:10, 80:20, 70:30, and 60:40, then experiments were also carried out on different hyperparameter changes in neuron parameters 10, 20, 30, 40, 50 and epoch 100, 500, 1000. From all the experiments that have been carried out, the most optimal results were obtained in experiments with a ratio of 80:20, neuron parameters 40 and epoch 1000. The RMSE obtained was 0.638, the MAPE value obtained was 1.66%. Experiments were also carried out on LSTM 1 variable with a ratio of 80:20 for 40 neuron parameters and 1000 epochs. The RMSE obtained was 8,579, with a MAPE value of 9.10%.

Keywords: Luna Classic; Cryptocurrency; Data Mining; Prediction; LSTM

*Corresponding author. E-mail address: mulyati@unpak.ac.id

Received: xx xxxxx 20xx, Accepted: xx xxxxx 20xx and available online XX July 2022

DOI: <https://doi.org/10.33751/komputasi.v19i2.5260>

1. Introduction

Cryptocurrency merupakan sebuah mata uang virtual yang digunakan sebagai alat transaksi secara virtual atau melalui jaringan internet [1]. Sebagian besar Cryptocurrency bersifat desentralisasi yaitu menggunakan sistem yang tidak bergantung pada otoritas pusat seperti bank pusat atau institusi administratif lainnya, melainkan melalui sistem jaringan berbasis komputer dan berdasarkan pada teknologi peer-to-peer yang keseluruhan pengambilan keputusannya telah diberikan kepada para pengguna sistem tersebut [2]. Salah satu contoh dari Cryptocurrency yang sedang hangat diperbincangkan di dunia kripto adalah Luna Classic.

Luna Classic (LUNC) adalah native token dari jaringan atau rantai (chain) lama Luna setelah percabangan atau fork ke rantai Terra baru dan merupakan salah satu alternative token dalam Cryptocurrency [3]. Ekosistem Luna yang hancur pada Mei 2022 lalu tidak membuat token Luna Classic mati seketika, aset ini justru terpantau masih alami pergerakan. Koin LUNC dalam satu tahun terakhir mengalami peningkatan dan penurunan volume trading secara signifikan yang membuatnya menjadi salah satu koin yang diminati dalam dunia kripto. Banyak dari para trader maupun investor dalam dunia crypto telah berpendapat bahwa prediksi harga dari token LUNC dapat menyentuh diatas 1 dollar dan dengan prediksi dapat memudahkan trader dan investor crypto untuk menentukan kapan harus entry dan menginvestkan dana mereka kedalam crypto.

Prediksi (forecasting) adalah kegiatan yang memperkirakan suatu keadaan yang akan terjadi dimasa mendatang. Prediksi berguna sebagai acuan dalam pengambilan keputusan dengan memperkirakan keadaan dimasa mendatang dan sangat membantu untuk perencanaan kedepannya [4]. Di antara berbagai metode prediksi, algoritma ANN (Artificial Neural Network) biasa digunakan untuk memprediksi harga di masa mendatang, salah satu algoritma ANN adalah Long Short Term Memory (LSTM). LSTM menyaring informasi melalui struktur gerbang untuk memelihara dan memperbarui status sel memori [5]. Long short term memory network (LSTM) adalah salah satu modifikasi dari recurrent neural network atau RNN. Penelitian ini menggunakan algoritma LSTM untuk forecasting karena LSTM melengkapi kekurangan RNN yang tidak dapat memprediksi kata berdasarkan informasi lampau yang disimpan dalam jangka waktu lama.

2. Method

Metode yang digunakan dalam melakukan penelitian ini adalah Knowledge Discovery in Database (KDD). Metode ini memiliki serangkaian tahapan yaitu Data Cleaning & Integration, Data Selection & Transformation, Data Mining, dan Evaluation and Presentation [6] dapat dilihat Gambar 1.

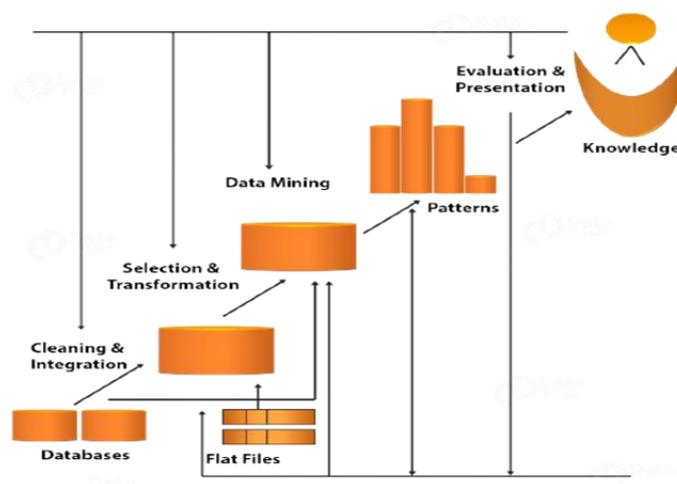


Figure 1. Tahapan Knowledge Discovery in Database (KDD).

2.1. Data Preparation and Data Integration

Tahapan ini merupakan proses operasional dasar dalam menghilangkan noise, menghilangkan data duplikat, data yang tidak konsisten dan data atau informasi yang tidak relevan [7].

2.2. Selection and Transformation Data

Data yang digunakan pada proses penambangan sebelumnya akan diseleksi terlebih dahulu, kemudian data terpilih tersebut akan ditransformasikan agar data tersebut dapat cocok untuk proses data mining [8]. Sebelum memasuki tahap data mining, dilakukan proses pemilihan dan transformasi data agar sesuai dengan kebutuhan data yang relevan untuk memprediksi harga koin Luna, tipe data diubah menjadi tipe data float32.

2.3. Data Mining

Process mining adalah proses pencarian informasi suatu model proses dengan menggunakan teknik atau metode tertentu. Proses penambangannya menggunakan algoritma Long Short Term Memory (LSTM) sebagai metode prediksi harga Luna Classic (LUNC). Data diambil dari website finance.yahoo.com yang kemudian akan melalui proses prediksi dan dilihat keakuratannya berdasarkan variabel pada data harga LUNC. Pertama data akan masuk ke tahap proses normalisasi untuk meminimalisir error, data aktual akan diubah ke dalam rentang interval [0,1] menggunakan penskalaan min-max dengan rumus sebagai berikut:

$$X' = \frac{(Xi - Xmin)}{(Xmax - Xmin)} \quad (1)$$

Dataset diolah menggunakan bahasa pemrograman Python. Proses pengolahan data ini menggunakan pendekatan Deep Learning dengan algoritma Long Short Term Memory (LSTM), pada tahap ini data yang sudah dinormalisasi akan diuji dan dihitung. Proses penambangan di LSTM melalui beberapa tahapan dengan beberapa rumus persamaan sebagai berikut:

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

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

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

$$C_t = f_t * c_{t-1} + i_t * \check{c}_t \quad (5)$$

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

$$h_t = o_t \tanh(c_t) \quad (7)$$

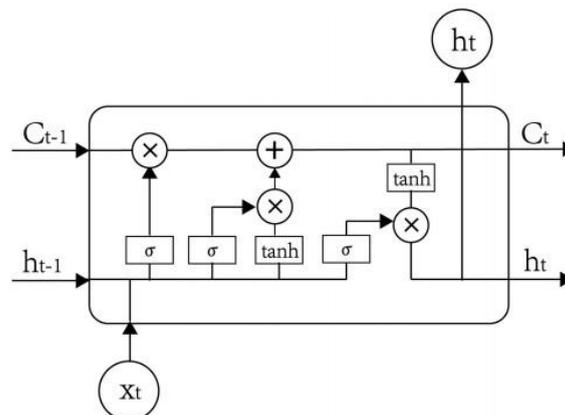


Figure 2. Arsitektur Long Short Term Memory (LSTM)

Dalam penelitian ini juga dilakukan skenario uji coba dalam membagi data latih dan data uji. Terdapat 4 skenario yang dilakukan untuk mendapatkan hasil rasio pembagian data terbaik.

Table 1. Tabel Skenario Uji Data

Skenario Percobaan	Data Training	Data Testing
1	416 data (90%)	46 data (10%)
2	370 data (80%)	92 data (20%)
3	323 data (70%)	139 data (30%)
4	277 data (60%)	185 data (40%)

Pada skenario uji coba data latih dan data uji, selain membagi data dan menentukan rasio terbaik, dilakukan perubahan parameter pada saat pembuatan model untuk mengukur tingkat kesalahan pada saat proses uji coba untuk memantau rasio kesalahan yang diperoleh. Parameter yang dimaksud adalah jumlah neuron dan jumlah epoch, hal ini dilakukan untuk mencari perbandingan dan mencari model dengan variasi jumlah epoch dan jumlah neuron terbaik sehingga menghasilkan bobot dan bias yang paling tepat untuk melakukan prediksi. Sesi pengujian menggunakan data pengujian dan hal ini terbukti dengan tingkat kesalahan yang rendah. Berikut ini adalah skenario perubahan parameter dan epoch neuron selama proses pelatihan.

1. Perubahan parameter neuron dengan nilai neuron: 10, 20, 30, 40, dan 50
2. Perubahan parameter epoch dengan nilai epoch: 100, 500, dan 1000

Karena dataset pada penelitian ini cukup besar maka diperlukan epoch yang lebih banyak agar model dapat melihat dan mempelajari sebagian besar variasi data, untuk itu digunakan epoch 100, 500, 1000 agar lebih efektif dalam pengujian..

2.4. Evaluation

Tahapan ini dilakukan untuk mengevaluasi nilai hasil prediksi yang telah dibuat agar lebih optimal. Untuk dapat mengetahui keakuratan hasil prediksi, dilakukan pengujian menggunakan Mean Square Error (MSE) dan Mean Absolute Percentage Error (MAPE) pada data yang dinormalisasi.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2}{n}} \quad (5)$$

$$MAPE = \text{absolute}((Y_i - \bar{Y}_i) / Y_i) * 100 \quad (6)$$

Root Mean Squared Error (RMSE) dan Mean Absolute Percentage Error (MAPE) adalah dua metrik evaluasi yang umum digunakan dalam analisis regresi dan prediksi untuk mengevaluasi kinerja model [9]. Nilai MAPE di bawah 10% memiliki kinerja yang sangat baik. Sedangkan nilai MAPE antara 10% hingga 20% mempunyai kinerja yang baik berdasarkan skala Lewis MAPE [10].

3. Result and Discussion

Telah dilakukan pengumpulan data dan integrasi data pada data harga LUNC untuk prediksi yang akan dilakukan, data yang diambil dari website yahoo.finance.com merupakan data time series harian mulai tanggal 3 Februari 2021 sampai dengan tanggal 11 Mei 2022. Dataset terdiri dari 6 variabel dengan jumlah 1 variabel berjumlah 462 data dan untuk seluruh variabel berjumlah 2778 data, terdiri dari variabel Open (harga pembukaan dalam periode tertentu), High (harga tertinggi yang dicapai dalam jangka waktu tertentu), Low (harga terendah yang dicapai dalam jangka waktu tertentu), Low (harga terendah yang dicapai dalam jangka waktu tertentu), dicapai dalam jangka waktu tertentu), Close (Harga penutupan dalam jangka waktu tertentu), Adj Close (Harga penutupan yang telah disesuaikan dengan mempertimbangkan faktor-faktor seperti dividen, pemecahan saham, atau perubahan lain yang mungkin mempengaruhi nilai aset kripto), dan Volume (Jumlah transaksi jual atau beli yang dilakukan selama bursa perdagangan dibuka).

Pada penelitian ini juga dilakukan pengujian terhadap prediksi harga LUNC dengan menggunakan metode Long Short Term Memory (LSTM) dengan 1 variabel untuk membandingkan dan menentukan model pengujian terbaik antara Multivariate LSTM dan 1 Variable LSTM. Pengujian dilakukan dengan rasio dan parameter model yang sama dengan Multivariate LSTM, yaitu menggunakan rasio 80:20 dengan data latih sebanyak 370 data yang diambil pada tanggal 3 Februari 2021 hingga 7 Februari 2022, kemudian untuk data uji sebanyak 92 data diambil pada tanggal 8 Februari 2021 hingga 11 Mei 2022, parameter yang digunakan adalah neuron 40 dan epoch 1000.

3.1. Modeling Process

Dalam penelitian ini sebelum dilakukan pengujian model, model akan dipilih berdasarkan skenario terbaik yang diperoleh, selanjutnya model akan dipilih dan diuji. Skenario uji coba pembagian data latih dan data uji memperoleh hasil terbaik yaitu rasio perbandingan data sebesar 80% (370 data) untuk data latih dan 20% (92 data) untuk data uji. Hasil eksperimen model terbaik dapat dilihat pada Tabel 2:

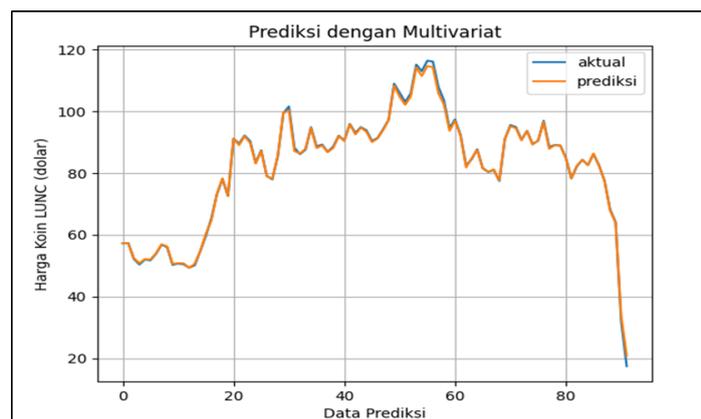
Table 2. Data Test Scenario Table

Skenario Uji Coba	Jumlah <i>Neuron</i>	Jumlah <i>Epoch</i>	RMSE	MAPE
2 (80% Data <i>Training</i> dan 20% Data <i>Testing</i>)	10	100	4.717	9.8047
		500	5.336	2.31
		1000	1.358	7.9405
	20	100	4.022	4.597
		500	1.529	3.269
		1000	1.538	4.044
	30	100	4.543	2.609
		500	2.052	2.84
		1000	0.644	3.296
	40	100	3.749	6.119
		500	0.958	1.759
		1000	0.638	1.688
	50	100	3.942	5.311
		500	0.741	2.66
		100	0.506	5.065

Hasil pengujian dari 4 skenario yang telah dilakukan menghasilkan model terbaik dengan perbandingan 80:20 (370 data latih dan 92 data uji) dengan jumlah 40 neuron dan 1000 epoch. Data latih (370 data) diambil pada tanggal 3 Februari 2021 hingga 7 Februari 2022, kemudian data uji (92 data) diambil pada tanggal 8 Februari hingga 11 Mei 2022.

3.2. Model Testing Accuracy Results

Hasil prediksi berdasarkan model yang telah diperoleh pada proses pelatihan dengan menggunakan 80% data pelatihan (370 data) dan 20% data pengujian (92 data) serta parameter terbaik yaitu 40 neuron dan 1000 epoch. Pada penelitian ini juga dilakukan percobaan untuk memprediksi harga LUNC dengan variabel tunggal LSTM, dengan menggunakan variabel 'Open' dengan rasio dan parameter yang sama yang digunakan yaitu rasio perbandingan data sebesar 80% dan 20% untuk data latih dan data uji. , parameter yang digunakan adalah neuron 40 dan epoch 1000 untuk membandingkan keakuratan hasil penelitian antara LSTM Multivariat dengan LSTM pada umumnya yang menggunakan variabel tunggal. Dalam penelitian ini untuk prediksi LSTM multivariat digunakan variabel terikat yaitu dari pasangan keluaran variabel terbuka yang dipengaruhi oleh variabel lain yaitu close, high, low, adj close dan volume, sedangkan untuk prediksi LSTM variabel tunggal independen. variabel yang digunakan tidak dipengaruhi oleh variabel tersebut. lainnya. Variabel yang digunakan sebagai pembanding keduanya bersifat terbuka, hasil yang diperoleh dengan memvisualisasikan grafik perbandingan data aktual dan data prediksi dapat dilihat pada Gambar 3 dan Gambar 4.

**Figure 3.** Grafik Data Aktual dan Prediksi Multivariate

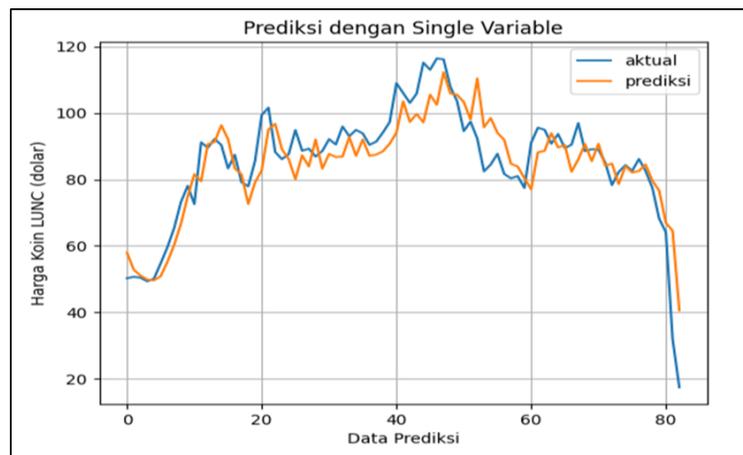


Figure 4. Grafik Data Aktual dan Prediksi Dengan 1 Variabel

Berdasarkan grafik pada Gambar 3 dan Gambar 4, model yang digunakan menghasilkan output yang sesuai, pola data aktual ditunjukkan dengan garis biru dan pola data prediksi ditunjukkan dengan garis oranye, data aktual adalah data harga koin LUNC sebelumnya sudah diprediksi dan data prediksi merupakan data aktual yang telah diprediksi, terlihat pola data prediksi mengikuti bentuk pola data sebenarnya. Pola garis data prediksi pada Gambar 3 terlihat seperti menyatu yang menunjukkan bahwa hasil prediksi antara data prediksi dan data aktual tidak jauh berbeda dan lebih akurat dibandingkan dengan pola data pada Gambar 4.

Hasil prediksi yang telah dibuat sebelumnya dievaluasi menggunakan Root Mean Square Error (RMSE) dan Mean Absolute Percentage Error (MAPE). Berikut pemaparan hasil nilai RMSE dan MAPE yang diperoleh.

1. Nilai RMSE yang diperoleh dari pengujian model prediksi harga LUNC dengan LSTM Single Variable yang dilakukan adalah sebesar 8,579, sedangkan tingkat akurasi yang diperoleh sebesar 90,9% dengan nilai MAPE sebesar 9,10% dengan jumlah 40 neuron dan 1000 epoch..

2. Nilai RMSE yang diperoleh dari pengujian model prediksi harga LUNC dengan Multivariate LSTM yang dilakukan adalah sebesar 0,638, sedangkan tingkat akurasi yang diperoleh sebesar 98,3% dengan nilai MAPE sebesar 1,66% dengan jumlah 40 neuron dan 1000 epoch termasuk dalam kategori sangat baik. rentang kategori berdasarkan tabel akurasi MAPE skala Lewis.

Dari hasil yang diperoleh terdapat perbedaan persentase kesalahan hasil akurasi prediksi antara LSTM Multivariat dan LSTM yang menggunakan 1 variabel, prediksi dengan 1 variabel LSTM menghasilkan tingkat kesalahan yang tinggi karena pengujian yang dilakukan hanya menggunakan variabel terbuka tanpa variasi. variabel lain, sedangkan prediksi dengan Multivariate menghasilkan tingkat error. lebih rendah karena memanfaatkan informasi lintas variabel, hubungan dan interaksi antar fitur antar variabel memberikan informasi tambahan untuk melakukan prediksi, menangkap pola dan hubungan yang lebih kompleks dalam data sehingga dapat meningkatkan akurasi prediksi [11].

3.3 Predictions for the Next 7 Days

Pada penelitian ini juga dilakukan percobaan prediksi selama 7 hari ke depan setelah tanggal 11 Mei 2022, pengujian dilakukan dengan model yang sama dengan pengujian prediksi sebelumnya yaitu diperoleh model terbaik dengan perbandingan 80:20 dan dengan nilai parameter yang sama. Tabel 2 menunjukkan hasil prediksi harga LUNC 7 hari ke depan.

Table 2. Tabel Skenario Data

Tanggal	Prediksi 7 Hari Berikutnya
5/12/2022	1.489555
5/13/2022	0.95662
5/14/2022	0.53840
5/15/2022	0.51743
5/16/2022	0.27309
5/17/2022	0.16679
5/18/2022	0.09342

Tabel 2 menampilkan prediksi harga LUNC 7 hari ke depan. Pengujian dilakukan dengan mengambil data pengujian 20 hari sebelumnya yaitu tanggal 8 April 2022 hingga 11 Mei 2022. Terlihat hasil prediksi terus mengalami penurunan dan pada tanggal 18 Mei 2022 harga LUNC menunjukkan 0,09342 dollar atau setara dengan Rp. 1455.59 per koin, karena pada saat itu pasar mata uang kripto sedang mengalami fase bearish atau menurun sehingga menyebabkan harga koin kripto turun, salah satunya koin alternatif seperti LUNC yang mengalami penurunan harga yang cukup drastis.

4. Conclusion

Penelitian mempunyai kesimpulan bahwa dari beberapa skenario uji coba rasio yang telah dilakukan dari 4 skenario diperoleh hasil terbaik yaitu pada data latih rasio 80% dan data uji 20% dengan 40 neuron dan 1000 epoch. Model terpilih diuji dengan LSTM Multivariat dan LSTM variabel tunggal sebagai pembanding. Berdasarkan hasil prediksi dari model terpilih yang telah dilakukan, algoritma Multivariate LSTM menghasilkan tingkat error dengan nilai error RMSE yang rendah sebesar 0,638 dan tingkat akurasi sebesar 98,3% berdasarkan nilai MAPE, lebih akurat dibandingkan LSTM yang menggunakan 1 variabel atau variabel tunggal dengan nilai error RMSE sebesar 8,579 dan tingkat akurasi sebesar 90,9% karena LSTM Multivariate memanfaatkan informasi lintas variabel yang membantu model dalam melatih data dalam proses prediksi agar menghasilkan tingkat error yang rendah dan menghasilkan tingkat akurasi yang lebih baik.

References

- [1] Kurniasih, Wida. "Apa Itu Cryptocurrency Jenis, Fungsi, Dan Cara Kerja", 2022, [Online]. Available: <https://www.gamedia.com/best-seller/apa-itu-cryptocurrency/>.
- [2] Indodax. "Yuk Kenali Apa Itu Desentralisasi pada Aset Kripto", 2022, [Online]. Available: <https://indodax.com/academy/apa-itu-desentralisasi/>.
- [3] Pamela, "Mengenal Luna Classic Sebagai Native Token dari Terra Classic", 2022, [Online]. Available: <https://kripto.ajaib.co.id/luna-classic/>.
- [4] Faizal, R., Setiawan, B. D., & Cholissodin, I, "Prediksi Nilai Cryptocurrency Bitcoin menggunakan Algoritme Extreme Learning Machine (ELM)", *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK)*, 3(9), 4226–4233, 2019.
- [5] Aldi, M. W. P., Jondri, & Aditsania, A, "Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin", *Jurnal Informatika*, 5, No(2), 3548, 2019.
- [6] DataFlair, "Data Mining and Knowledge Discovery Database(Kdd Process)", 2023, [Online]. Available: <https://data-flair.training/blogs/data-mining-and-knowledge-discovery>
- [7] Kusmawati., "Pengaruh Motivasi Terhadap Minat Berinvestasi Di Pasar Modal Dengan Pemahaman Investasi Dan Usia Sebagai Variabel Moderat". *Jurnal Ekonomi dan Informasi Akuntansi*. 1 (2), 2011.
- [8] Amri, MK, "Penerapan data mining untuk menentukan strategi penjualan pada toko buku Gamedia Palembang. Skripsi", Universitas Palembang, Palembang, 2013.

- [9] Morley, S. K., Brito, T. V., & Welling, D. T, “Measures of Model Performance Based On the Log Accuracy Ratio. *Space Weather*”, 16(1), 69–88. 2018, [Online]. Available <https://doi.org/10.1002/2017SW001669>.
- [10] Sudhan, S. P., Mathew, B. S., Rose, S., & Isaac, K. P, ‘Development of pavement deterioration prediction models for low volume roads using system dynamics’, *Journal of Stomatology*, 146(3), 2020
- [11] He, Xiaoyu., Suixiang Shi, Xiulin Geng, Jie Yu, & Lingyu Xu, “Multi-step forecasting of multivariate time series using multi-attention collaborative network”., *Expert Systems with Applications*, 211, 2023.