The Indonesian Journal of Computer Science
www.ijcs.net
Volume 13, Issue 4, August 2024
https://doi.org/10.33022/ijcs.v13i4.4161

Cryptocurrencies Price Estimation Using Deep Learning Hybride Model of LSTM-GRU

Ulul Azmiati Auliyah[1], Ema Utami[2], Dhani Ariatmanto[3]
ululazmiauliyah@gmail.com[1], ema.u@amikom.ac.id[2], dhaniari@amikom.ac.id[3]
[1,2,3]Master of Informatics, Universitas Amikom Yogyakarta, Indonesia

Article Information

Received : 20 Jun 2024
Revised : 2 Jul 2024
Accepted : 25 Jul 2024

Keywords

Cryptocurrency, Price, Estimation, LSTM, GRU

Abstract

One of the financial assets in currency exchange is now cryptocurrency. The public is drawn to cryptocurrency trading because it is considered a lucrative form of investing. For cryptocurrency investors to maximize their earnings, accurate price forecasting is crucial. As price forecasting involves time series analysis, a hybrid deep learning model is suggested to project cryptocurrency prices in the future. Long Short-Term Memory and Gated Recurrent Unit (LSTM-GRU) networks are integrated into the hybrid model. Three cryptocurrency datasets are evaluated using the suggested hybrid model: Ethereum, Ripple, and Bitcoin. According to experimental results, the suggested LSTM-GRU model may provide the lowest MSE and RMSE values on the Bitcoin dataset (0.0611 and 0.2472), the Ethereum dataset (0.0369 and 0.19222), and the Ripple dataset (0.0006 and 0.0247).

A. Introduction

Cryptocurrency is an encrypted and decentralized digital exchange tool with a blockchain system [1]. Crypto is a digital currency that can transferred instantly around the world. Many people consider crypto a digital financial tool capable of providing material benefits and freedom of transactions without a third party: security and decentralization. However, crypto commodities still have high risks due to their high volatility. Encourages researchers to make price forecasts to minimize the risks of this commodity.

Crypto coin price forecasts can be done using several types of algorithms. The author [2] compared four methods consisting of three machine learning and one deep learning. These methods are ARIMA, SVM, ANFIS, and LSTM for forecasting Bitcoin prices. The research results show that the LSTM model outperforms the conventional method used for comparison. Deep learning algorithms can provide better solutions than traditional machine learning [3].

Researchers [4] compared the 1DCNN-GRU hybrid model with non-hybrid models such as ARIMA, RNN, LSTM, GRU, Bi-LSTM, and XGBoost. The lowest RMSE value of 43.933 on the Bitcoin dataset, 3.511 on the Ethereum dataset, and 0.00128 on the Ripple dataset indicates that this method is used to predict time series prices for Bitcoin, Ethereum, and Ripple coins. The results prove that the hybrid model provides better prediction results with lower errors than other methods. In a different study by [5], a hybrid approach was used to predict the prices of Monero and Litecoin by merging LSTM and GRU. The results obtained hybrid performance better than the single LSTM model with the lowest MSE value for Litecoin 4.12 in a prediction window size of 3 days and Monero 10.7 in a prediction window size of 1 day. Several researchers compared his hybrid approach with the LSTM method alone, and the results showed that the hybrid model had higher accuracy in predicting the price of the selected cryptocurrency.

Hybrid LSTM-GRU research was also carried out by [6] on Litecoin and Zcash coins. Testing also carried out similar techniques: trials at window lengths of 1 day, three days, seven days, and 30 days. The research results obtained even lower MSE values, namely Litecoin with an error value of 0.02 and Zcash with 0.004 at a window size setting of 1 day. Differences in the MSE value results for Litecoin coins between the two studies above can occur due to using different hyperparameters, especially in batch size and epoch.

Due to the excellent performance of the hybrid method in predicting crypto prices, the author was encouraged to research crypto coin price predictions using the hybrid method. The hybrid techniques used are LSTM and GRU. as stated by researchers [7] in their research Systematic Literature Review, where from the papers reviewed between 2018-2022, the LSTM and GRU methods were popular in cryptocurrency price prediction. The performance of the LSTM and GRU methods has been proven to be good, with minimal error results [8]–[12], provides the idea of combining both methods or called a hybrid to improve overall model performance.

Based on the background and previous research above, the researcher conducted research to predict prices from three datasets of historical cryptocurrency prices, namely Bitcoin, Ethereum, and Ripple. The dataset was then delivered to the LSTM-GRU model for representation learning and price prediction after undergoing basic pre-processing, including normalization. Next, the root mean square error is computed to compare the anticipated price with the actual price. Setting the activation function and different batch size values will also be the testing focus of this research. It will optimize the proposed LSTM-GRU method in estimating the prices of the studied cryptocurrencies. This paper's main contribution is to highlight the potential of deep learning for cryptocurrency market predictions. In doing so, we found that proper hyperparameter settings play a critical role in creating a combination of recurrent neural networks that are effective in classifying cryptocurrencies' daily relative performance.

B. Research Method

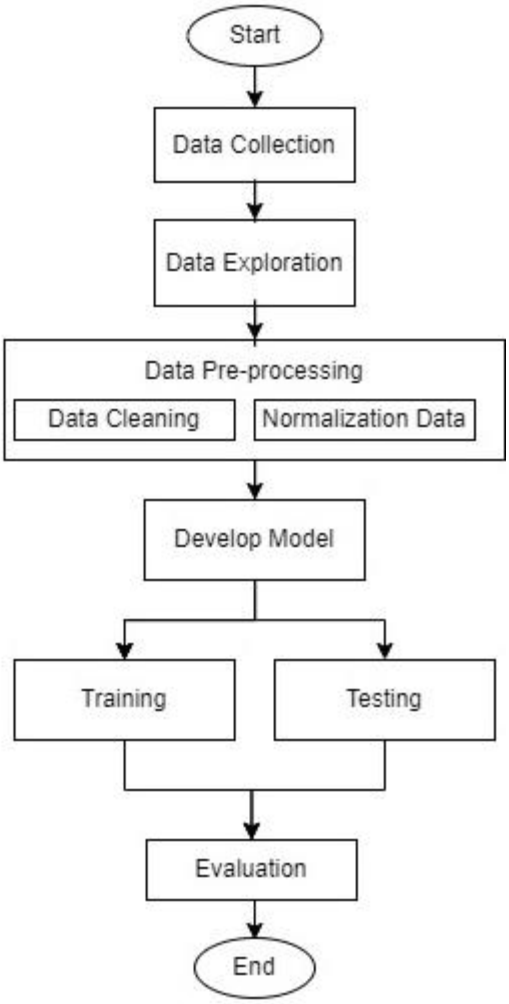

Figure 1. Research Stages

Figure 1 depicts the steps of the research process flow diagram, including data collecting, data exploration, data pre-processing, creating the suggested algorithm model, and assessing the model of research outcomes to make inferences. The dataset was obtained from the Yahoo Finance website in the first stage. The second stage is data exploration, where the data is checked for completeness and the amount of data. The third stage is pre-processing, which includes cleaning unused data columns and normalizing the data with a min-max scalar. In the fourth step, known as the LSTM-GRU deep learning scenario, a model is built, and its number of layers, neurons, and hyperparameters are adjusted to achieve optimal performance. Evaluation and analysis, the final step, entails analyzing the findings and formulating conclusions in light of the experiments.

1. **Data Collection**

For the purpose of forecasting cryptocurrency prices, three datasets are used: Ethereum, Ripple, and Bitcoin. Data collection begins by accessing Yahoo Finance at https://finance.yahoo.com/ in the crypto menu section. Then, select the target coin and select the historical data menu to see the cryptocurrency price. In the period menu, select a period with a daily timeframe ranging from 1 December 2017 to 31 March 2024. The collected crypto price data is in .csv file format, containing several variable columns: date, open, close, high, low, adj close, and volume. The Bitcoin closing price is shown in Figure 2, the Ethereum closing price is shown in Figure 3, and the Ripple closing price is shown in Figure 4.

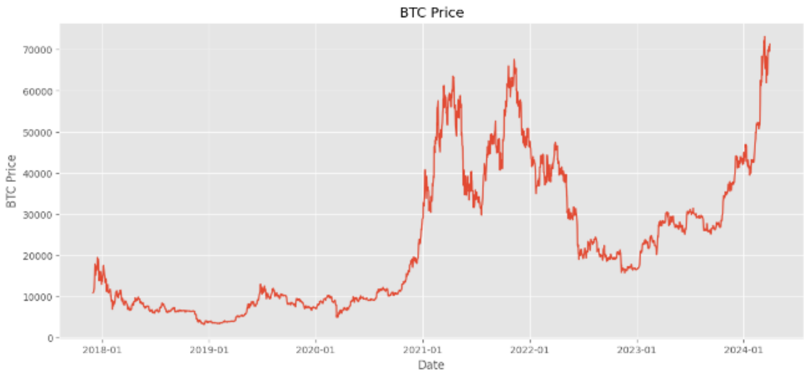

Figure 2. The historical price of Bitcoin (2017–2024)

Figure 3. The historical price of Ethereum (2017–2024)

Figure 4. The historical price of Ripple (2017–2024)

2. **Data Exploration**

At this stage, we check all data, data dimensions, and data distribution and check whether there are columns that do not contain/NULL. However, there is no data imbalance in the dataset, so there is no need to handle data imbalances uniquely, such as oversampling or undersampling.

3. **Data Pre-Processing**

The data that has been collected will be processed by eliminating variables that will not be used and only using date and close as variables in this research. The dataset is saved in CSV format and then processed using the proposed method to obtain the predicted value of crypto coin prices. Next, the data is normalized with a min-max scalar to scale it into the number range 0-1 to make it easier for the program to read the data to be predicted. This research carried out the data normalization process on all daily datasets for the LSTM-GRU method. The data normalization used is min-max normalization. The following is the min-max normalization (1) equation.

$$x' = \frac{x - X\ minimum}{X\ maximum - X\ minimum} \tag{1}$$

x : value in the dataset
X minimum: the minimum value in the dataset
X maximum: the maximum value in the dataset
x' : value after normalization into the range 0 to 1

4. **The Proposed Model**

Figure 5 shows the suggested methodology for predicting the prices of cryptocurrencies such as Bitcoin, Ethereum, and Ripple. LSTM and GRU can carry out sequence-related tasks like time series prediction because they possess sequential memory. These two deep learning algorithms are RNN variations created to address the disappearing gradients that RNNs encounter. In time series prediction, LSTM and GRU are superior in several earlier research [8]–[12]. The study suggests a way to take advantage of both models concurrently.

LSTM consists of an input layer, one or more hidden layers, and an output layer. In the hidden layer, there are special blocks called state cells, which effectively overcome long-term dependency problems. The state cell contains the memory cell because it represents the entire history of information stored in the memory cell, which effectively acts as the "memory" of the layer. One essential part is the memory cell, which has the dual functions of controlling the information flow via the gate and storing long-term data. Meanwhile, GRU has an architecture like LSTM with a chain of repeating modules and non-separated state cells but with a more straightforward design. Even though it simplifies the computational process in LSTM, GRU still has equivalent performance and is compelling enough to avoid the vanishing gradient problem [13].

Figure 5 shows the architecture of the suggested LSTM-GRU model. The suggested LSTM-GRU has one LSTM layer with 150 neurons and a 20% dropout rate to prevent overfitting. A dropout of 20% comes after the output from the dropout, which is sent to one GRU layer with 130 neurons. The dense layer of one neuron receives the output from this layer. After that, followed by an activation function. This activation function will be tested with several types, such as Sigmoid, Relu, Tanh, and Linear.

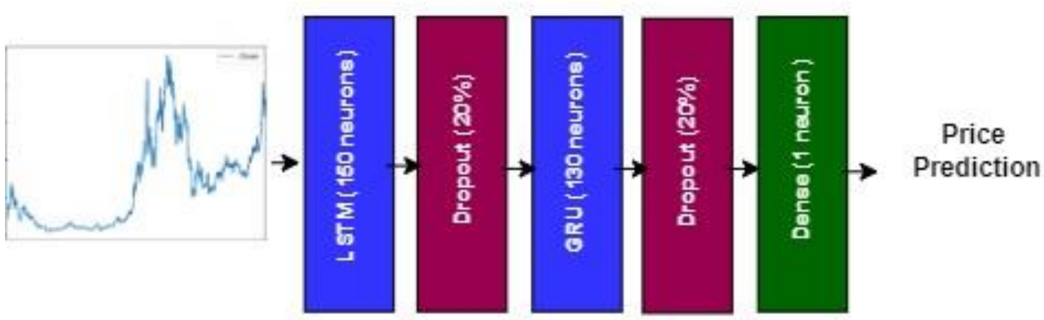

Figure 5. The architecture of the proposed LSTM-GRU model.

The model from this research was implemented in the Python platform using the Keras library. Apart from that, as a supporting element, the built LSTM-GRU model will use the Adam optimizer, epoch setting 30, and window size 1. Meanwhile, the activation function and batch size will be tested using several scenarios.

Table 1. Performance Parameters

Parameters	Values
Programming Language	Phyton
Platform	Google Colab
Framework	Tensorflow
Total data Points	@ 2313
Train data points	1850
Validation data points	231
Test data points	232
Window lengths	1
Learning rate	0.001
Epochs	30
Optimizer	Adam
Activation Function	Sigmoid, Relu, Tanh, Linear
Batch size	16, 32, 64, 128
Metrics	MSE, RMSE, MAPE

Table 1 includes values for the parameters: optimizer, number of splitting data points, programming language, and other parameters.

5. Evaluation Metric

The study uses MSE and RMSE to evaluate the proposed model, which provides a quadratic loss function and measures forecasting uncertainty. Mean Squared Error (MSE) is the average of the squared differences between the actual values in the dataset and the predicted values produced by the model (equation 2). Meanwhile, Root Mean Squared Error (RMSE) calculates the average squared difference between the actual value in the dataset and the predicted value produced by the model and then squares it (equation 3). The lower the MSE and RMSE values, the more minor the forecast error level [14].

The difference between the actual and predicted values is calculated divided by the actual data, and the result is an absolute value known as the Mean Percentage Absolute Error (MAPE), which is used as an evaluation metric. The prediction model built has high accuracy in forecasting new values, as evidenced by the reduced MAPE findings achieved [14]. Formula four below shows the results of the MAPE computation.

$$MSE = \frac{\sum(actual\ values - predicted\ value)^2}{total\ number\ of\ data} \quad (2)$$

$$RMSE = \sqrt{\sum \frac{(actual\ values - predicted\ value)^2}{total\ number\ of\ data}} \quad (3)$$

$$MAPE = \frac{1}{n}\sum\left(\frac{actual\ values - predicted\ values}{actual\ values}\right)x100 \quad (4)$$

Table 2 provides information that can be used to calculate the final value of the MAPE computation. The resulting forecast is better if the % MAPE value is lower.

Table 2. MAPE Calculation Range

Range MAPE	Information
<10%	Highly accurate forecasting
10%-20%	Good forecasting
20%-50%	Reasonable forecasting
>50%	Inaccurate forecasting

C. Result and Discussion

The hyperparameter settings that will be tested in this research are the activation function and batch size. Activation functions in the LSTM and GRU layers transform input, allowing the model to learn and perform more complex tasks. Activation functions are used in neural networks to calculate the sum of input weights and biases and decide whether a neuron can be activated [15]. The main challenge is to choose an activation function that suits the problem definition and considers model performance and loss function convergence. The type of activation function selected in the hidden layer will significantly impact the performance of the neural network model. In the output layer of the neural network, an activation function is also needed [16]. Four activation functions are explored: sigmoid, relu, tanh, and linear. Meanwhile, batch size determines the number of samples used for error gradient calculations in each model weight update. The four batch sizes explored were 16, 32, 64, and 128.

1. Result for Bitcoin (BTC)

Table 2 shows the experimental results of different hyperparameter values on the Bitcoin dataset. The Tanh activation function and batch size 128 are the best hyperparameter settings in the proposed model for Bitcoin price forecasting. The MAPE value for Bitcoin price forecasting is 37.79%, which means the proposed model's performance is quite good in predicting the price of the Bitcoin dataset. The lowest MSE value obtained was 0.0611, and the lowest RMSE was 0.247.

Figure 6 illustrates the performance of the LSTM-GRU model in forecasting Bitcoin's actual price. Bitcoin has a very high actual price of 30,000-70,000 USD, which is tremendous compared to the prices of Ethereum and Ripple cryptocurrencies. This affects the error value obtained. Here, Bitcoin gets an error value greater than the error value from the Ethereum and Ripple price dataset.

Table 2. Bitcoin dataset evaluation results

Activation	Batch Size	MSE	RMSE	MAPE
Sigmoid	16	0.0638	0.2526	40.05 %
	32	0.0653	0.2555	39.65%
	64	0.0655	0.2559	39.99%
	128	0.07	0.2646	41.31%
Relu	16	0.0671	0.259	38.69%
	32	0.0695	0.2636	39.93%
	64	0.0705	0.2655	40.40%
	128	0.0708	0.2661	40.09%
Tanh	16	0.0652	0.2553	39.27%
	32	0.0647	0.2544	38.99%
	64	0.0632	0.2515	38.79%
	128	0.0611	0.2472	37.97%
Linear	16	0.0675	0.2597	38.61%
	32	0.0681	0.261	39.20%
	64	0.0723	0.269	41.12%
	128	0.0707	0.2658	40.09%

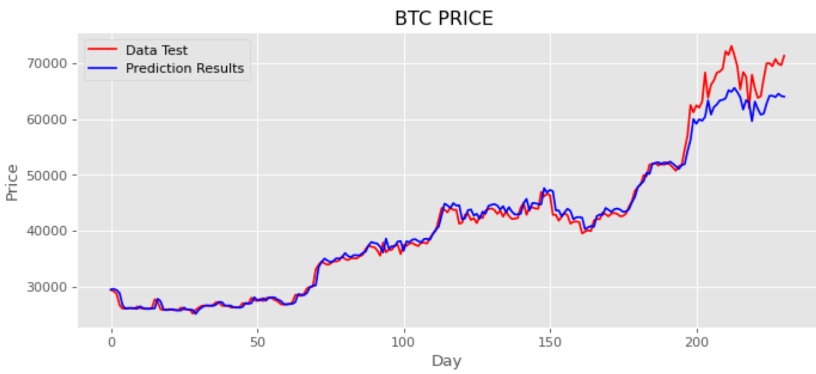

Figure 6. Actual and predicted result by LSTM-GRU model on the Bitcoin dataset.

2. Result for Ethereum (ETH)

Table 3 shows the experimental results of different hyperparameter values on the Ethereum dataset. The Tanh activation function and batch size 128 are the best hyperparameter settings in the proposed model for Ethereum price forecasting. The MAPE value for Bitcoin price forecasting is 33%, which means the proposed model's performance is quite good in predicting the price of the Ethereum dataset. The lowest MSE value obtained was 0.0369, and the lowest RMSE was 0.1922.

Table 3. Ethereum dataset evaluation results

Activation	Batch Size	MSE	RMSE	MAPE
Sigmoid	16	0.0423	0.2057	35.55%
	32	0.0417	0.2042	35.36%
	64	0.0421	0.2051	35.42%
	128	0.0449	0.212	36.29%
Relu	16	0.0392	0.1979	33.37%
	32	0.0393	0.1983	33.07%
	64	0.0386	0.1965	32.73%
	128	0.0394	0.1985	32.77%
Tanh	16	0. 0408	0.2019	33.59%
	32	0.0402	0.2005	34.08%
	64	0.0394	0.1984	33.56%
	128	0.0369	0.1922	33.00%
Linier	16	0.0395	0.1987	33.26%
	32	0.0391	0.1977	32.86%
	64	0.0397	0.1992	33.06%
	128	0.039	0.1975	32.64%

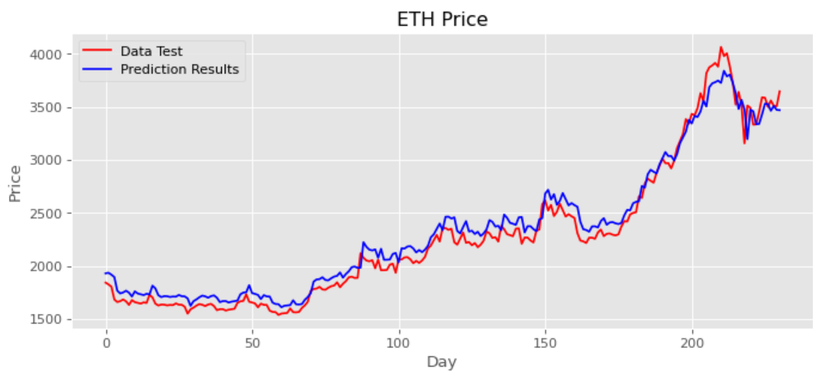

Figure 7. Actual and predicted result by LSTM-GRU model on the Ethereum dataset

Figure 7 illustrates the LSTM-GRU model's performance in forecasting Ethereum's price compared to the actual cost. Ethereum has a relatively high actual price in the 2,000-4,000 USD range and tends to experience an upward price trend in each period.

3. Result for Ripple (XRP)

Table 4 shows the experimental results of different hyperparameter values on the Ripple dataset. The Tanh activation function and batch size 64 are the best hyperparameter settings in the proposed model for Ripple price forecasting. The

MAPE value for Bitcoin price forecasting is 14.49%, which means the proposed model performs well in predicting the price of the Ripple dataset. The lowest MSE value obtained was 0.0006, and the lowest RMSE was 0.0247.

Table 4. Ripple dataset evaluation results

Activation	Batch Size	MSE	RMSE	MAPE
Sigmoid	16	0.0007	0.0269	15.84 %
	32	0.0007	0.0266	15.65 %
	64	0.0007	0.0271	15.93 %
	128	0.0008	0.0276	16.08 %
Relu	16	0.0007	0.0261	16.75 %
	32	0.0006	0.0252	15.23 %
	64	0.0006	0.0249	15.59 %
	128	0.0006	0.025	15.29 %
Tanh	16	0.0006	0.0251	15.64 %
	32	0.0006	0.025	15.46 %
	64	0.0006	0.0247	14.98 %
	128	0.0007	0.0255	15.54 %
Linier	16	0.0007	0.0268	15.98 %
	32	0.0006	0.0252	15.19 %
	64	0.0006	0.0249	15.11 %
	128	0.0006	0.0248	15.09 %

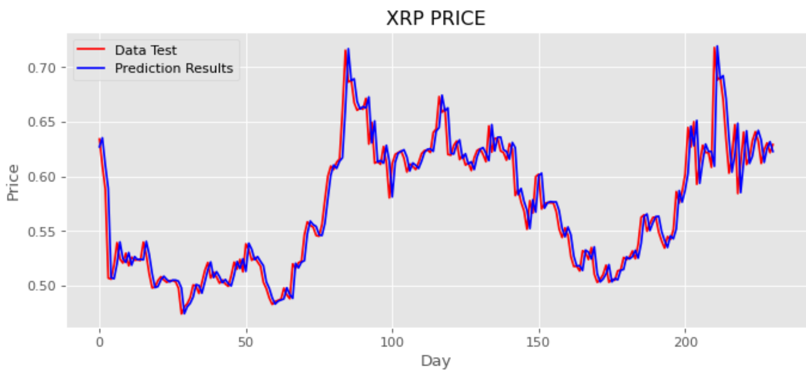

Figure 8. Actual and predicted result by LSTM-GRU model on the Ripple dataset

Figure 8 illustrates the performance of the LSTM-GRU model in predicting the price of Ripple compared to the actual price. Ripple's exact price is still low, less than 1 USD. Apart from that, the price of Ripple has a very dynamic trend and has not shown a very volatile price increase.

D. DISCUSSION

The model proposed in this study is a reliable and accepted model for forecasting cryptocurrency prices. Testing has been carried out, and optimal results have been provided on each dataset by setting the Tanh activation function and batch size to 128 on the Bitcoin and Ethereum coin datasets and 64 on the Ripple coin dataset. The experimental settings for each dataset can be seen in Table 5.

Table 5. The experimental settings of the proposed LSTM-GRU

Dataset/Hyperparameter	Configurations
Bitcoin	optimizer = Adam, activation function = Tanh, batch size = 128
Ethereum	optimizer = Adam, activation function = Tanh, batch size = 128
Ripple	optimizer = Adam, activation function = Tanh, batch size = 64
Learning rate	0.001
Epoch	30

In general, the MSE and RMSE of all methods on the Bitcoin dataset are the highest, followed by the Ethereum and Ripple datasets, which produce the lowest error values. This is due to price differences where higher prices produce higher MSE and RMSE. Experimental results show that the proposed LSTM-GRU performs quite well, as demonstrated by the low error values. Both LSTM and GRU have different strengths, and both work well in a variety of applications that require storing historical information through a gateway mechanism [9].

These models are considered reliable and suitable based on the evaluation method and the results obtained. However, remember that this model has limitations that may affect its accuracy in predicting cryptocurrency prices. First, cryptocurrency prices are highly dependent on many variables, and neither LSTM nor GRU may cover all of these dependencies, resulting in suboptimal forecasts.

E. Conclusion

This research presents a hybrid deep-learning model that leverages the power of LSTM and GRU for cryptocurrency price prediction. The historical prices of three cryptocurrencies are obtained: Bitcoin, Ethereum, and Ripple. The collected data is cleaned and normalized first to make the model's work easier. Next, the previously processed data is passed to the hybrid LSTM-GRU model. The LSTM and GRU models can capture patterns in both the long and short term. Experimental results show that the proposed LSTM-GRU can produce the lowest MSE error value of 0.1147 on the Bitcoin dataset, 0.0803 on the Ethereum dataset, and 0.0213 on the Ripple dataset. This research only utilizes historical data on crypto daily closing prices. In the future, it can be upgraded to a more complex model by using other factors, such as sentiment data or seasonal trends, as input for price prediction models.

F. References

[1] Kate Ashford and Benjamin Curry, “What Is Cryptocurrency?,” Forbes Advisor, Feb. 2023. https://www.forbes.com/advisor/investing/cryptocurrency/what-is-cryptocurrency/

[2] W. Chen, H. Xu, L. Jia, and Y. Gao, “Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants,” Int. J. Forecast., vol. 37, no. 1, pp. 28–43, 2021, doi: 10.1016/j.ijforecast.2020.02.008.

[3] A. Bilal Wardak and J. Rasheed, “Bitcoin Cryptocurrency Price Prediction Using Long Short-Term Memory Recurrent Neural Network,” Eur. J. Sci. Technol., Jun. 2022, doi: 10.31590/ejosat.1079622.

[4] C. Y. Kang, C. P. Lee, and K. M. Lim, “Cryptocurrency Price Prediction with Convolutional Neural Network and Stacked Gated Recurrent Unit,” Data, vol. 7, no. 11, p. 149, Oct. 2022, doi: 10.3390/data7110149.

[5] M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, "A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions," J. Inf. Secur. Appl., vol. 55, no. May, p. 102583, 2020, doi: 10.1016/j.jisa.2020.102583.
[6] S. Tanwar, N. P. Patel, S. N. Patel, J. R. Patel, G. Sharma, and I. E. Davidson, "Deep Learning-Based Cryptocurrency Price Prediction Scheme with Inter-Dependent Relations," IEEE Access, vol. 9, pp. 138633–138646, 2021, doi: 10.1109/ACCESS.2021.3117848.
[7] D. Ariatmanto, R. Sari, and Y. R. Mahdika, "SLR Metode Paling Populer Prediksi Harga Cryptocurrency dengan Machine Learning dan Deep Learning," Semin. Nas. Corisindo, pp. 543–548, 2022, [Online]. Available: https://corisindo.stikom-bali.ac.id/penelitian/index.php/semnas/article/view/114%0Ahttps://corisindo.stikom-bali.ac.id/penelitian/index.php/semnas/article/download/114/85
[8] N. T. Cao, D. Q. Nguyen, and A. H. Ton-That, "A Combination of Technical Indicators and Deep Learning to Predict Price Trends for Short-Term Cryptocurrency Investment," Proc. IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2022, pp. 1–5, 2022, doi: 10.1109/CSDE56538.2022.10089300.
[9] M. J. Hamayel and A. Y. Owda, "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms," Ai, vol. 2, no. 4, pp. 477–496, 2021, doi: 10.3390/ai2040030.
[10] X. Jiang, "Bitcoin Price Prediction Based on Deep Learning Methods," J. Math. Financ., vol. 10, no. 01, pp. 132–139, 2020, doi: 10.4236/jmf.2020.101009.
[11] D. H. Kwon, J. B. Kim, J. S. Heo, C. M. Kim, and Y. H. Han, "Time series classification of cryptocurrency price trend based on a recurrent LSTM neural network," J. Inf. Process. Syst., vol. 15, no. 3, pp. 694–706, 2019, doi: 10.3745/JIPS.03.0120.
[12] S. Marne, D. Correia, S. Churi, and J. Gomes, "Predicting Price of Cryptocurrency-A Deep Learning Approach," Int. J. Eng. Res. Technol., vol. 9, no. 3, pp. 387–393, 2021, [Online]. Available: www.ijert.org
[13] Suyanto, K. N. Ramadhani, and S. Mandala, Deep Learning: Modernisasi Machine Learning untuk Big Data, Pertama. Bandung: Informatika Bandung, 2019.
[14] I. Saputra and D. A. LKristiyanti, Machine Learning Untuk Pemula, 1st ed. Bandung: Informatika Bandung, 2022.
[15] G. Minotti, "Cryptocurrencies Price Prediction using LSTM Neural Network model," 2022.
[16] C. E. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation Functions : Comparison of Trends in Practice and Research for Deep Learning," pp. 1–20, 2018.Z. Zhang, F. Wu, H.J.W. Zandvliet, B. Poelsema, H. Metiu, & M.G. Lagally, "Paper Title," *Physical Review Letter,* vol. 74, pp. 3644-3677. 1995.