



Deep Learning Model Analysis and Web-Based Implementation of Cryptocurrency Prediction

Gege Ardiyansyah¹, Ferdiansyah², Usman Ependi^{3,*}

¹Information System Departement, Bina Darma University, Palembang, Indonesia

^{2,3}Informatics Departement, Bina Darma University, Palembang, Indonesia

Email: ¹gegeardiansyah@gmail.com, ²ferdi@binadarma.ac.id, ³u.ependi@binadarma.ac.id (*)

Abstract

Cryptocurrency is a digital asset designed by cryptography, such as Secure Hash Algorithm 2 (SHA-2) and Message Digest 5 (MD5). Cryptocurrency uses Blockchain technology to ensure security, transparency, ease of locating, and unchangeability. This makes cryptocurrency very popular in many sectors, especially in the financial industry. Although, the uncertainty and the dynamic change of cryptocurrency price make the risk for investment in this digital asset high. This is the reason why studies about cryptocurrency price prediction became popular globally. This study intended to predict cryptocurrency prices using hybrid GRU LSTM than setting up the epoch to get the most accurate prediction model. The researcher would make a web-based application that can be used by the public, especially those involved in cryptocurrency investment. The result was a web-based application that could predict the price of cryptocurrency for the next few days, which had been validated using data from the previous 7 days, 14 days, 30 days, 60 days, and 90 days.

Keywords: Prediction, Deep Learning, Cryptocurrency, Website, Regression

1. INTRODUCTION

Cryptocurrency is a digital asset that was first introduced on January 3, 2009. Unlike common currencies, Cryptocurrencies are designed based on cryptography, such as Secure Hash Algorithm 2 (SHA-2) and Message Digest 5 (MD5). Cryptocurrencies use Blockchain technology to guarantee security, transparency, traceability, and immutability. This is what makes cryptocurrencies very popular in various fields, especially in the field of finance. However, uncertainty and dynamic price changes make the level of risk in investing in this digital asset quite high.

Cryptocurrency prices with very high fluctuations can be seen on the Yahoo Finance website. The stock market is influenced by uncertain factors such as political and economic problems at the global level, such as on December 2, 2021, when the Fed accelerated the normalization of its monetary policy or tapering, which resulted in a significant decline in cryptocurrency prices. So, to describe



predictions accurately is a complicated thing, the problem that is also the basis of this research is price fluctuations that experience rapid changes every day. In addition, an application is needed to monitor and prevent transaction uncertainty. In Islam, it also aims to avoid Flowers (Usury), Gambling (Masyir), and Uncertainty (Ghahar)[1]. Therefore, Over the past few years, deep learning algorithms have been widely implemented in the financial field [2], replacing the previous regression algorithms [3] [4].

Research conducted by Islam and Hossain [5] Previous research using LSTM sean et all [6], The study proposed a method to predict the price of bitcoin using the Recurrent Neural Network (RNN) and combined with LSTM [7] proposed automatic cryptocurrency price predictions using machine learning techniques based on historical trends (daily trends). Di Bakar Raj et all [8], It was proposed that the prediction of bitcoin price based on the Artificial Neural Network Technique repeated by Twitter sentiment. The impressive results show the relationship between sentiment and LSTM results [9] and create a new bitcoin price prediction framework using LSTM. They proposed two different LSTM models (conventional LSTM and LSTM with Arima model with 208 dataset records, compared to Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

Based on the above background, researchers used deep learning models to predict cryptocurrency prices due to the success of previous research[10][11], which predicted cryptocurrency prices using deep learning with the LSTM layer. So that the results can be made recommendations and used by the public or crypto activists in reaching decisions, researchers also created a web-based application to find out the price of cryptocurrencies in real-time and can be used to predict them online.

This research contributes in several ways to the scope of cryptocurrency.

1. Researchers propose LSTM's hybrid GRU model in predicting cryptocurrency prices for the next 60 days. The decision is based on data validation that has been carried out using data from 7 days, 14 days, 30 days, 60 days, and 90 days in advance which are then juxtaposed to get the best accurate prediction results. Previous research has been carried out using data from the next two to 60 days [12].
2. Fixing different RMSE values in each cryptocurrency resulting from significant differences in cryptocurrency prices from each other and different levels of fluctuations in each cryptocurrency by normalizing their value
3. We creating a website-based application so that the proposed model can be used online by a general audience, especially cryptocurrency enthusiasts.

The website generated in this study can predict the price of cryptocurrencies with good accuracy. However, the high fluctuations of cryptocurrencies cannot be predicted only using historical data due to other factors that affect the changes. There is still a need for further development of the researched model and other influences to produce a cryptocurrency price prediction method with a higher level of accuracy.

2. METHODS

To get the best model, this research uses a literature study method by collecting data on books, research journals, articles, and others related to cryptocurrency price predictions with the deep learning model. The deep learning model used in this case is Recurrent Neural Networks (RNN). RNN uses layers such as LSTM and GRU, which can predict prices based on the previous price in order and are very suitable for time series-type data. The following are the layers that will be used in this study.

LSTM Networks [13] Are different types of Recurrent Neural Networks. LSTM can learn long-range dependencies and is very popular for concurrent data such as time series. LSTM cell uses an input gate i_t , a forget gate f_t , and an output gate o_t . This gate will define whether data can pass based on data priorities. Gates also enabled the network to learn what to keep, forget, remember, pay attention to, and produce. Cell state and hidden state are used to collect data that will be used for the next state.

The Gated Recurrent Unit (GRU) [13] is a Recurrent Neural Network, like LSTM. However, the GRU has a simpler structure when compared to LSTM. The GRU does not have an output gate but an update gate z_t and reset gate r_t . This gate is a vector that will determine whether information should be passed as output. The reset gate defines combining the new input with the previous memory. Limiting how much memory was last stored is done at the Update gate.

While making the application itself uses the Rapid Application Development (RAD) method, the helper features, namely Activity Diagrams, Data Flow Diagrams, and Entity Relationship Diagrams. In detail, the research consists of 6 stages. The stages consist of data preparation, modeling, evaluation, application design, creation, and testing. The stages can be seen in Figure 1.

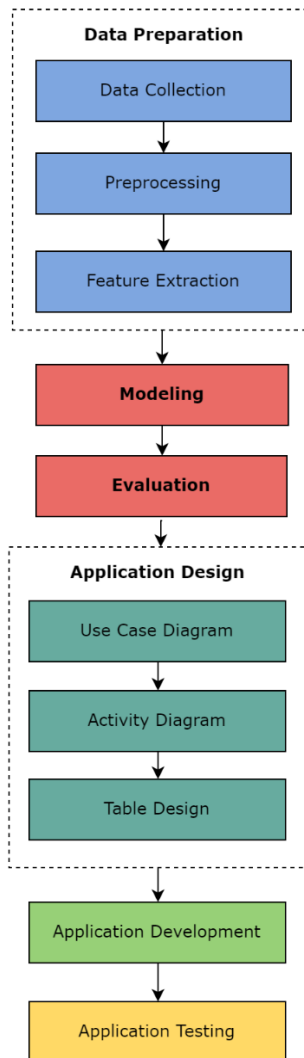


Figure 1. Research Framework

3. RESULTS AND DISCUSSION

3.1. Data Preparation

At this stage, data is collected through the yahoo finance website, which provides cryptocurrency data in the form of a time series. The historical data is five years apart, from November 9, 2017, to July 6, 2022. After the data is obtained, a preprocessing stage is carried out to clean the data from missing values or noise and determine the attributes to be used (feature extraction).

3.2. Modeling

The LSTM GRU Layer is used to create a hybrid model at the modeling stage, as shown in Figure 2.

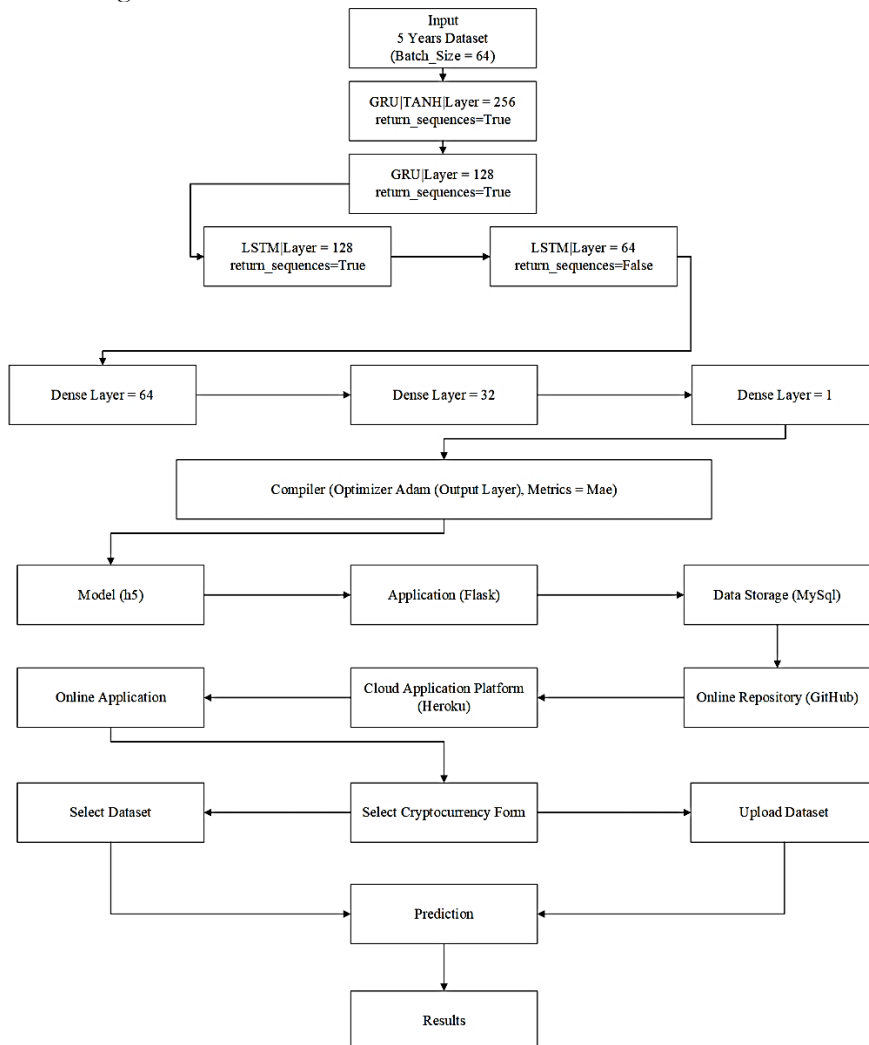


Figure 2. LSTM GRU Model Framework

The research framework in Figure 2. is a model created based on a combination of the GRU and LSTM algorithms that begins by entering a predetermined dataset. The data that has been processed then goes through 2 GRU layers with tanh activation, where the first layer uses units 256 [5] and the second layer uses units 128 based on experiments conducted using multiples of 64 to 512 units. In the LSTM algorithm, the number of layers used is the same, namely 2, with the

number of units of the first layer being 128 and the number of units of the second layer being 64. The return sequences used on each layer are true and false on the last layer.

Data that has gone through the GRU and LSTM layers is then processed through the dense layer, and the dense layer is a traditional neural network model that aims to classify data according to the class at the output. A dense layer has one input and an output whose number corresponds to the number of classes to be divided. The Dense layer in this model uses 3-layers with different numbers of units, namely 64, 32, and 1. Finally, all model layers are compiled using the model compiler with the Adam optimizer. The created model is encoded using the python programming language in the jupyter notebook application, as shown in Figure 3.

```
def create_hybrid():
    model = tf.keras.models.Sequential([
        tf.keras.layers.GRU(units=256, return_sequences=True,
            activation = "tanh",
            input_shape=(x_train.shape[1],1)),
        tf.keras.layers.GRU(units=128, return_sequences=True),
        tf.keras.layers.LSTM(128, return_sequences=True),
        tf.keras.layers.LSTM(64, return_sequences=False),
        tf.keras.layers.Dense(64),
        tf.keras.layers.Dense(32),
        tf.keras.layers.Dense(1)])

    lr_schedule = tf.keras.optimizers.schedules.ExponentialDecay(1e-4,
        decay_steps=100000,
        decay_rate=0.98,
        staircase=False)

    return model

model = create_hybrid()
```

Figure 3. Model coding using the python programming language

For each cryptocurrency tested, researchers used historical data over a span of 5 years. The results obtained for each cryptocurrency will certainly vary. Figure 4. Displays a line chart in the form of a comparison of the original price and a prediction of one of the predicted cryptocurrencies.

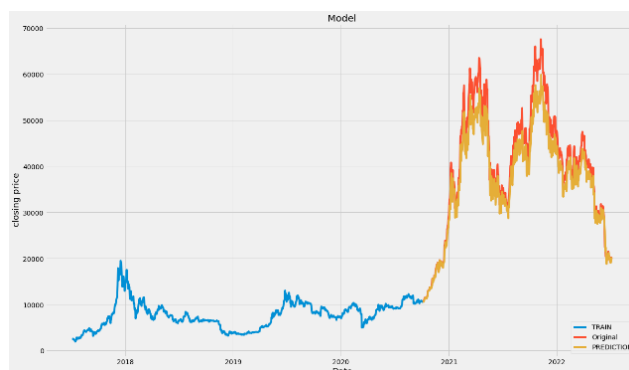


Figure 4. Line Chart Bitcoin price prediction results

3.3. Evaluation

At the evaluation stage, the quality and effectiveness of the model's prediction results on several cryptocurrencies can be known. Evaluation of the model is carried out using RMSE and MAPE calculations.

RMSE (Root Mean Square Error). RMSE will always be greater than or equal to MAE. The RMSE matrix evaluates how well the model can predict continuous values. RMSE units are the same as the data-dependent variables/targets (so if it's dollars, then the RMSE is also dollars), which is useful for understanding whether the error rate is high or low. The smaller the RMSE, the better the performance of the model [14][15][16]. The main benefit of using RMSE is that it can know the error rate and scale the score in units equal to the estimated value[17]. The RMSE formula can be seen in equation (1) where N is the number of observations, x_i is actual, and \hat{x}_i as the predicted value

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2} \quad (1)$$

The magnitude of the RMSE value depends on the value of the dataset used; basically, the greater the RMSE value, the better the accuracy of the model. However, the RMSE value's size also depends on how large the value of the dataset used [18]. The formula for normalizing RMSE can be seen in equation (2).

$$\text{Normalized RMSE} = RMSE \div (\text{Max Value} - \text{Min Value}) \quad (2)$$

Mean Absolute Percentage Error (MAPE) is used to measure relative error. MAPE is usually more effective than MAD because MAPE describes the error percentage in estimating or forecasting actual results over a certain period. In other words, MAPE is the average of absolute errors over a certain period which is then multiplied by 100% to obtain a percentage result [19][20]. The MAPE formula can be seen in equation (3), with A_t as the actual value of the data t , F_t as the forecasting value of the data t , and N as the sum of the data periods.

$$MAPE = \frac{100}{n} \sum \frac{|A_t - F_t|}{A_t} \quad (3)$$

Based on Lewis 1982 [21], MAPE values can be interpreted or interpreted into four categories, namely:

<10% = Highly accurate

10-20% = good

20-50% = Reasonable

> 50% = Inaccurate

The smaller the MAPE value, the smaller the error in the estimated result. On the other hand, the greater the MAPE value, the greater the error in the estimated result. The result of a prediction method has excellent forecasting ability if the MAPE value is less than 10% and has good predictability if the MAPE value is between 10% and 20%.

3.4. Evaluation Results

After training on the model using the dataset provided, the predicted results from the model can be evaluated. Using the sklearn library in Python or manually computing.

Table 1. Evaluation Results

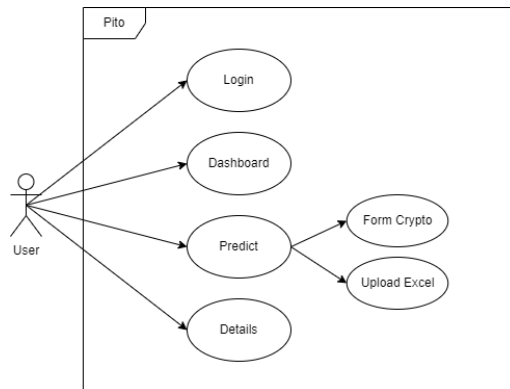
No	Cryptocurrency	Epoch	RMSE	Normalize RMSE	MAPE
1	Bitcoin	10	4457.3	0.078	10.33%
		50	2760.04	0.048	10.048%
		100	2943.55	0.0516	6.73%
2	Ethereum	10	109.8	0.025	8.47%
		50	348.6	0.080	12.79%
		100	224.75	0.0517	8.27%
3	Binance	10	33.64	0.051	11.66%
		50	20.71	0.031	8.02 %
		100	20.371	0.0314	6.87%
4	Ripple	10	0.081	0.050	11.83%
		50	0.0305	0.018	5.8%
		100	0.0151	0.009	5.62%
5	Tether	10	0.0015	0.097	0.16%
		50	0.00037	0.024	0.05 %
		100	0.0002	0.0129	0.04%

Table 1 shows the RMSE, Normalized RMSE, and MAPE values from the prediction results of the five selected cryptocurrencies. The RMSE value at epoch 100 has the most stable value among the three epochs tested, which makes the model with epoch 100 used in predicting the price of cryptocurrencies that will be displayed on the website later.

3.5. Application Design

After obtaining a suitable model and saved in h5 format, the next stage is the application design, where the model will be implemented.

A Use Case diagram is a depiction of how the interaction of the user and the system will be created and what everyone can do with the system. Figure 5. is a use-case diagram of the system that the researcher will create.



Gambar 5. Use Case Diagram

Figure 6. Below is an activity diagram that aims to describe the processes that occur in the system when used

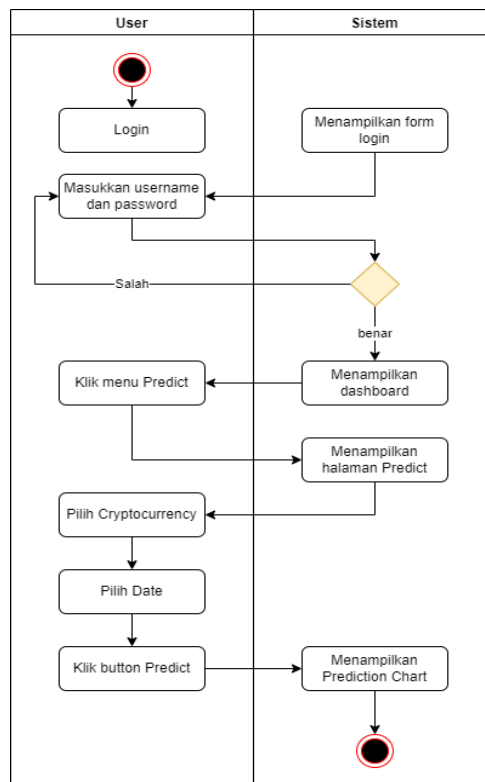


Figure 6. Activity Diagram

Researchers also design user tables and historical tables of predictions.

3.6. Application Development

In the application creation stage, researchers use the python programming language with the flask framework for front-end and back-end creation, then use Mysql to store the data.

3.7. Application Testing

Testing performed by the user on the internal scope of development. Alpha testing is performed with the Black Box Testing technique to determine if the application is working as the user expects [22]. Table 2. shows the test results for the website that has been developed.

Table 2. Alpha Test Results

No.	Test Scenarios	Expected results	Results
1	Pressing the registration button after entering the appropriate form data	Displaying the 'Account Created Successfully notification	Succeed
2	Pressing the login button after entering your username and password	View dashboard page	Succeed
3	Pressing Page Predict	Display the predictions page	Succeed
4	Pressing the Predict button on the prediction form	Display prediction results in the form of line charts and data tables	Succeed
5	Pressing the details page	View the details page	Succeed
6	Pressing the log out button on the user menu	Back to login page	Succeed

3.8. Application Interface

This section describes how to use and display the website interface of the previously created application design. Figure 7 displays the interface of the login page to the application that has been created.

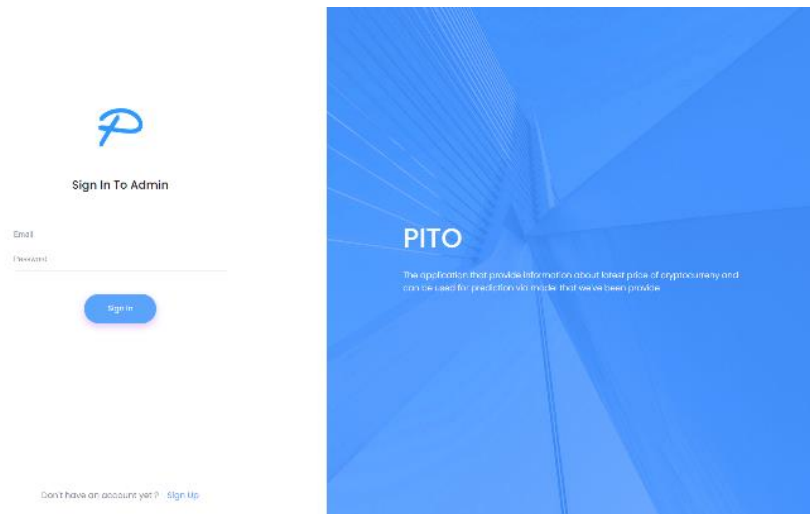


Figure 7. Login Page

Before being able to enter the application, users are required to register first on the registration page in Figure 8.

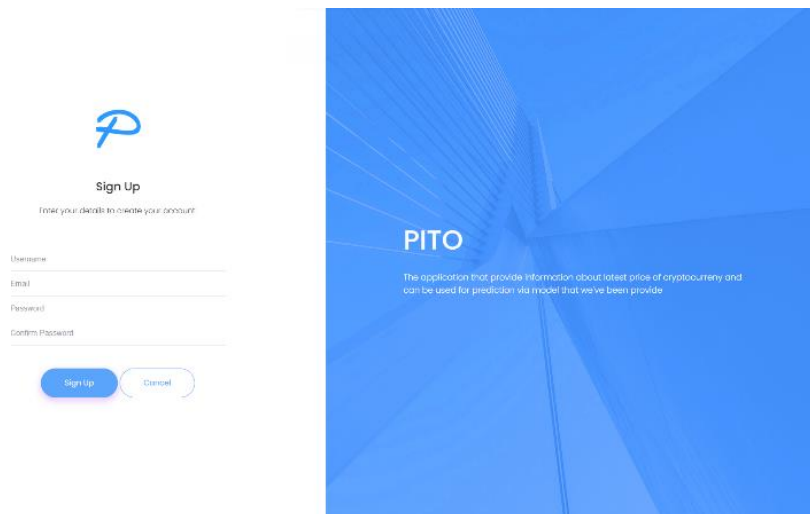


Figure 8. Registration Page

Suppose the user has registered and logged in through the login page. Users will be directed to a dashboard page that displays the latest prices of all cryptocurrencies recorded, as shown in Figure 9.

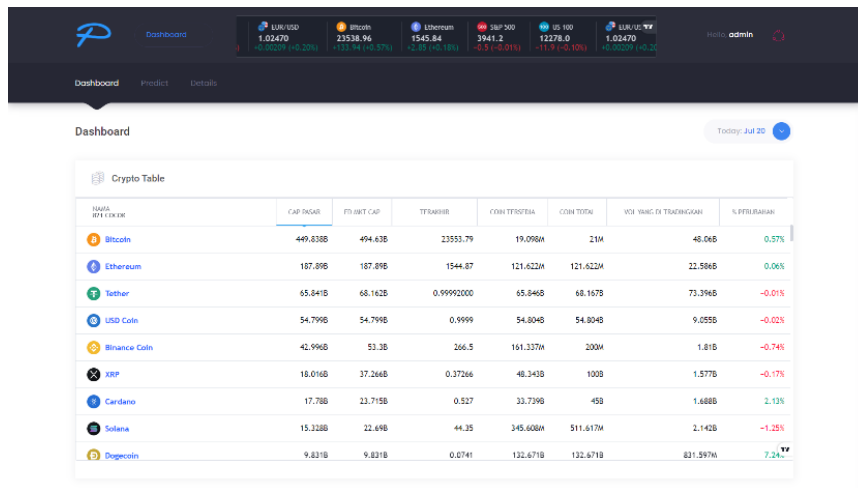


Figure 9. Dashboard Page

Furthermore, users can go to the Predict page provided to make predictions in Figure 10. The visible page provides two main menus for predictions. The form uses data provided by the system and a form that users can use by uploading their cryptocurrency data in CSV or XLSX file format.

The Prediction page contains two main sections:

- Prediction Chart:** Features a 'Menu' dropdown set to 'Cryptocurrency', a 'Group' field, and a 'Predict' button.
- Prediction Results:** Features a 'Menu' dropdown set to 'Cryptocurrency', a 'Group' field, and a 'Predict' button.

Figure 10. Prediction page with regular form and file form

After selecting the cryptocurrency to be predicted or entering the data, the user can predict by pressing the predict button provided. The system will display the prediction results after a while, as shown in Figure 11.

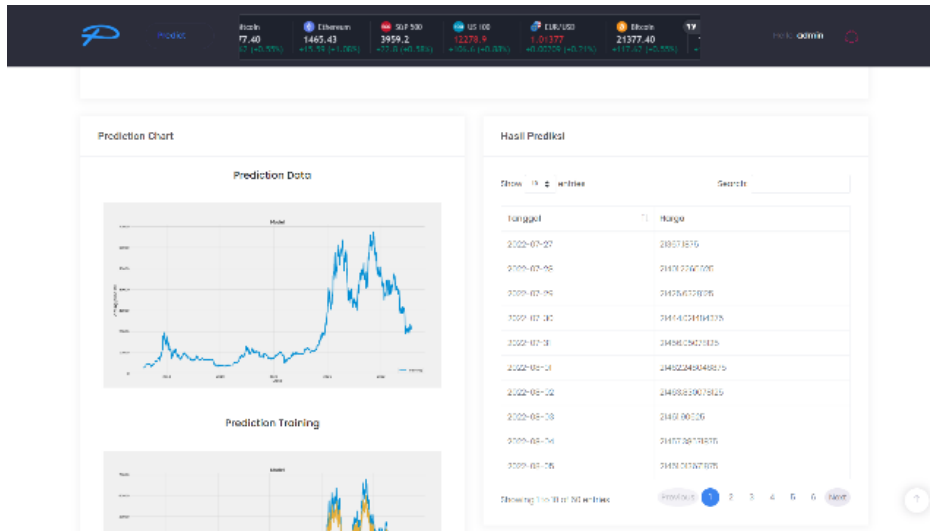


Figure 11. Cryptocurrency Prediction Results

The prediction results consist of a line chart that displays price movement patterns in the next 60 days and data tables that display the cryptocurrency's price for the next 60 days based on the last 60 days of data used. Line chart prediction training is the main line chart that describes the results of data predictions at the time of training and the results of predictions for the next 60 days of data, as shown in Figure 12.

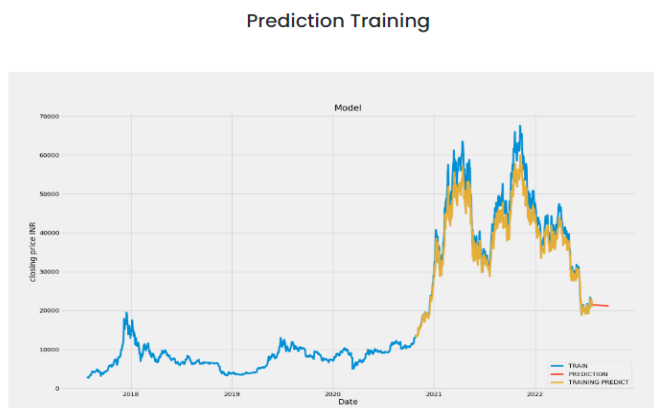


Figure 12. Prediction Training

Predicting the next 60 days of data is based on the validation of the previous 7 days, 14 days, 30 days, 60 days, and 90 days of data. The cryptocurrency's predicted price is juxtaposed with its predictions' results to determine the accuracy and stability of the LSTM hybrid GRU model that has been used. The smaller the distance between the original price and the predicted price, the better the model will predict the price within a predetermined period, as shown in Table 3.

Table 3. Prediction Validation Results

No	Cryptocurrency	Days	Price	Predictions
1	Bitcoin	7	19784.72	19787.20
		14	21085.87	19260.27
		30	31155.47	31782.96
		60	34059.26	33786.99
		90	42287.66	38931.26
2	Ethereum	7	1067.29	1126.69
		14	1143.38	1097.12
		30	1814.04	1769.31
		60	2517.45	2579.61
		90	3192.07	3115.24
3	Binance	7	219.30	230.05
		14	228.95	218.30
		30	290.34	294.73
		60	355.39	354.89
		90	421.67	400.28
4	Ripple	7	0.331448	0.322614
		14	0.335531	0.308053
		30	0.407702	0.396853
		60	0.566527	0.580934
		90	0.754056	0.760719
5	Tether	7	0.998819	0.999817
		14	0.99928	0.999568
		30	0.999414	0.999636
		60	0.999854	1.000525
		90	1.00028	0.999768

In the table, the difference in the predicted price and the original price of the cryptocurrency does not have a significant comparison. As in the bitcoin price, where the price difference is only around 10%, cryptocurrency price predictions using validation data from several other days give good results. However, with so many cryptocurrencies in force, using one of the best validation data from all five experiments is one of the most effective ways the model can be used on other cryptocurrencies.

The website also provides a details page that provides information about the latest predictable prices of cryptocurrencies, research evaluation results, and evaluation

results every time all users make predictions. The Details page can be seen in Figure 13.

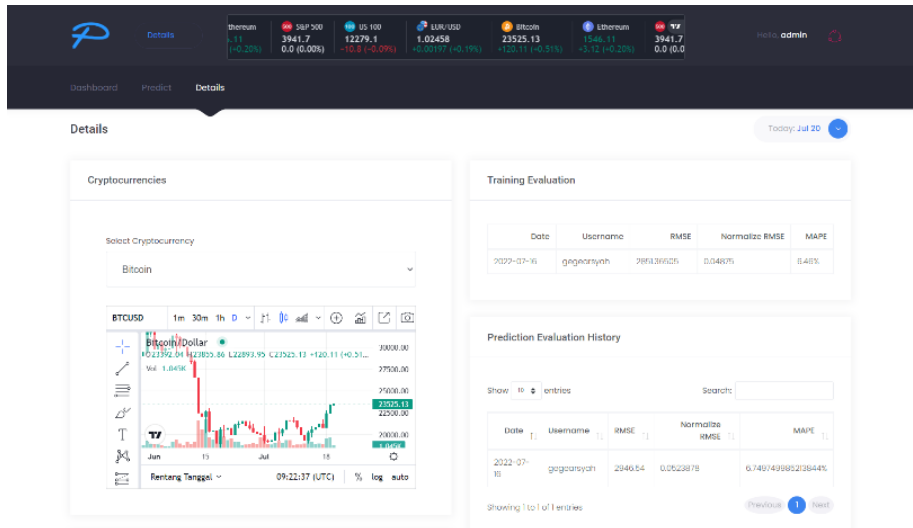


Figure 13. Details page

4. CONCLUSION

The LSTM GRU model can be used to improve and provide a new perspective for predicting cryptocurrencies compared to previous research that only used one type of deep learning layers, such as GRU or LSTM. Several approaches have been taken in this study to predict the price of cryptocurrencies; historical data was obtained from the yahoo finance website over a span of 5 years. Five different cryptocurrencies, namely Bitcoin, Ripple, Binance, Tether, and Ethereum, were used in the study. It is intended to determine whether the suggested model provides optimal results for each type of cryptocurrency. Each predicted data provides a MAPE and RMSE value, which is then juxtaposed to determine how accurate the model is. As a result, experiments on the LSTM GRU model by setting the epoch to 10, 50, and 100 gave good results, with epoch 100 providing the most maximum accuracy value for each data tested. For example, in bitcoin data, an epoch of 100 models can produce accuracy with RMSE values of 0.0516 and MAPE of 6.73%. The model was also tested using several validation data, namely 7, 14, 30, 60, and 90. Testing the model with 60 validation data provided the prediction value closest to the original price. In bitcoin data, the original price is 42287.66 while the predicted price is 38931.26; the price difference is 7.938%.

Implementing the model on a website-based application can also provide easy access for users to predict cryptocurrencies with the model that has been studied.

The website has several widgets that can provide information about the latest cryptocurrency prices, making it easier to get information about the desired cryptocurrency price. The website also has two main features in predicting the price of cryptocurrencies: using data provided online or using data in CSV or XLSX file formats from users. However, the Website is still unable to train the model using the latest data, so the model's accuracy will decrease over time. The challenge for further research is how to create a model capable of predicting each cryptocurrency with better accuracy. The high fluctuations of cryptocurrencies can not only be predicted using historical data due to other factors that affect the fluctuations. Further development is still needed regarding the model under study and other influences that were not applied to this study.

REFERENCES

- [1] F. Ferdiansyah, S. H. Othman, R. Z. M. Radzi, and D. Stiawan, 'A study of economic value estimation on cryptocurrency value back by gold, methods, techniques, and tools', *Journal of Information Systems and Informatics*, vol. 1, no. 2, pp. 178–192, 2019.
- [2] W. Hastomo and A. S. B. Karno, 'Kemampuan Long Short Term Memory Machine Learning Dalam Proyeksi Saham Bank BRI TBK', *Prosiding SeNTIK*, vol. 4, no. 1, pp. 229–236, 2020.
- [3] N. F. B. Pradana and S. Lestanti, 'Aplikasi Prediksi Jangka Pendek Harga Bitcoin Menggunakan Metode Arima', *Jurnal Ilmiah Informatika Komputer*, vol. 25, no. 3, pp. 160–174, 2021.
- [4] M. M. Hasan, 'Prediksi jangka pendek harga bitcoin menggunakan metode ARIMA', 2019.
- [5] M. S. Islam and E. Hossain, 'Foreign exchange currency rate prediction using a GRU-LSTM Hybrid Network', *Soft Computing Letters*, p. 100009, 2020.
- [6] S. McNally, J. Roche, and S. Caton, 'Predicting the price of bitcoin using machine learning', in *2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP)*, 2018, pp. 339–343.
- [7] R. Mittal, S. Arora, and M. P. S. Bhatia, 'Automated Cryptocurrencies Prices Prediction Using Machine Learning', *ICTACT Journal on Soft Computing*, vol. 08, no. 04, p. 4, 2018.
- [8] D. R. Pant, P. Neupane, A. Poudel, A. K. Pokhrel, and B. K. Lama, 'Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis', in *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, Kathmandu, Oct. 2018, pp. 128–132. doi: 10.1109/CCCS.2018.8586824.
- [9] C.-H. Wu, C.-C. Lu, Y.-F. Ma, and R.-S. Lu, 'A new forecasting framework for bitcoin price with LSTM', in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2018, pp. 168–175.

- [10] F. Ferdiansyah, S. H. Othman, R. Z. R. M. Radzi, D. Stiawan, Y. Sazaki, and U. Ependi, 'A lstm-method for bitcoin price prediction: A case study yahoo finance stock market', 2019, pp. 206–210.
- [11] A. D. Arisandi, L. Atika, E. S. Negara, and K. R. N. Wardani, 'Prediksi Mata Uang Bitcoin Menggunakan LSTM Dan Sentiment Analisis Pada Sosial Media', Jurnal Ilmiah KOMPUTASI, vol. 19, no. 4, pp. 559–566, 2020.
- [12] R. Albariqi and E. Winarko, 'Prediction of Bitcoin Price Change using Neural Networks', in 2020 International Conference on Smart Technology and Applications (ICoSTA), Feb. 2020, pp. 1–4. doi: 10.1109/ICoSTA48221.2020.1570610936.
- [13] M. De Caux, F. Bernardini, and J. Viterbo, 'Short-Term Forecasting in Bitcoin Time Series Using LSTM and GRU RNNs', in Anais do Symposium on Knowledge Discovery, Mining and Learning (KDMiLe 2020), Brasil, Oct. 2020, pp. 97–104. doi: 10.5753/kdmile.2020.11964.
- [14] Squark, 'ROOT MEAN SQUARE ERROR OR RMSE'. <https://squarkai.com/root-mean-square-error-or-rmse/#.XWVCLegzZPY> (accessed Aug. 28, 2019).
- [15] N. Vandeput, 'Forecast KPI: RMSE, MAE, MAPE & Bias', 2019. <https://medium.com/analytics-vidhya/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d> (accessed Apr. 10, 2020).
- [16] B. Santoso, A. I. S. Azis, and others, Machine Learning & Reasoning Fuzzy Logic Algoritma, Manual, Matlab, & Rapid Miner. Deepublish, 2020.
- [17] N. J. Salkind, Encyclopedia of research design, vol. 1. Sage, 2010.
- [18] Z. Bobbitt, 'What is Considered a Good RMSE Value?', 2021. Accessed: Jan. 16, 2022. [Online]. Available: <https://www.statology.org/what-is-a-good-rmse/>
- [19] V. Gaspersz, 'Production planning and inventory control', PT Gramedia Pustaka Umum, Jakarta, 2004.
- [20] U. Khair, H. Fahmi, S. Al Hakim, and R. Rahim, 'Forecasting error calculation with mean absolute deviation and mean absolute percentage error', in Journal of Physics: Conference Series, 2017, vol. 930, no. 1, p. 12002.
- [21] C. D. Lewis, Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. Butterworth-Heinemann, 1982.
- [22] S. U. Masruroh, S. Hanna, N. Azza, K. Kamarusdiana, and N. A. R. Vitalaya, 'Klasifikasi Mazhab Menggunakan Metode Naïve Bayes (Studi Kasus: Salat)', JEPIN (Jurnal Edukasi dan Penelitian Informatika), vol. 8, no. 1, pp. 74–79.