Time Series Forecasting of Top 3 Ranking Cryptocurrencies

Ridwan Setiawan[1], **Indri Tri Julianto***[2], **Fikri Fahru Roji**[3]
[1,2] *Department of Computer Science, Institut Teknologi Garut, Garut*
[3] *Department of Digital Business, Universitas Garut, Garut*
E-mail : ridwan.setiawan@itg.ac.id [1]*, indritrijulianto@itg.ac.id**[2]*, fikri@uniga.ac.id* [3]
*Corresponding Author

Abstract – Cryptocurrency has become a phenomenon worldwide. Although not all countries have legalized it, it is considered a promising investment asset. Currently, there are three top-ranking cryptocurrencies: Bitcoin, Ethereum, and Tether. This research aims to compare the performance of five forecasting algorithms, namely Autoregressive Integrated Moving Average (ARIMA), Neural Network, Support Vector Machine, Linear Regression, and Generalized Linear Model, using the dataset of Bitcoin, Ethereum, and Tether cryptocurrencies. The research methodology employed is Knowledge Discovery In Databases (KDD). The technique involves assessing the performance based on the Root Mean Square Error (RMSE) and comparing the results to find the most optimal model performance. The research findings indicate that for Bitcoin cryptocurrency, the Neural Network algorithm produced the most optimal results with an RMSE of 9180.534. For Ethereum cryptocurrency, the Neural Network algorithm demonstrated the best performance with an RMSE value of 537.528. Furthermore, for Tether cryptocurrency, the ARIMA algorithm yielded the best performance with an RMSE value of 0.003.

Keywords – bitcoin, cryptocurrency, ethereum, forecasting, tether

1. INTRODUCTION

Cryptocurrency is a technology that utilizes cryptography with the purpose of ensuring security and regulating authority through decentralized systems. The primary objective of this system is to manage the creation of new units, record transactions, and provide security that cannot be replicated or forged [1]. Cryptocurrency serves as a medium of transaction for purchasing goods and services, and it can also be utilized as a long-term investment with the goal of obtaining future profits [2]. One of the advantages of the cryptocurrency market is its availability, operating 24 hours a day. This allows users to monitor the movements of the crypto market on a daily basis. As a result of these advantages, cryptocurrency is now gaining recognition in several countries worldwide [3]. The movement of cryptocurrency can be monitored through a website called Yahoo.Finance.com. The interface of this website is presented in Figure 1.
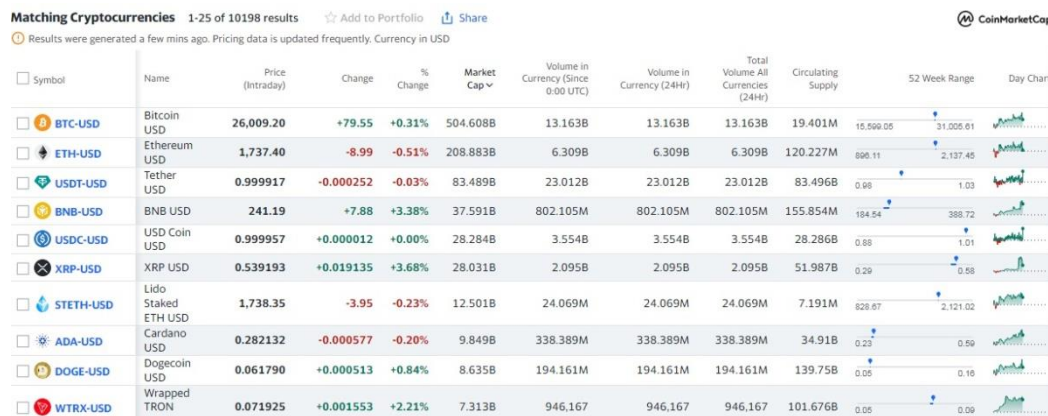
Figure 1. Cryptocurrency Market Cap [4]

Figure 1 displays the ranking of cryptocurrencies from highest to lowest. The top 3 ranked cryptocurrencies based on market capitalization are Bitcoin, Ethereum, and Tether. Bitcoin, ranked first, was introduced by Sakoshi Nakamoto as the first digital currency. Its unique characteristic lies in the control maintained through an open-source software system, granting anyone the ability to influence or modify the system [5]. Ethereum, ranked second, is a digital currency built on the innovations of Bitcoin but with significant differences. Ethereum serves not only as a means of payment but also as a marketplace for financial services, games, and applications that safeguard user data from security threats [6]. Tether, ranked third, is a type of cryptocurrency categorized as a stablecoin, with its value pegged to the U.S. dollar. Tether is frequently used as an intermediary when traders switch between different cryptocurrencies [6].

There have been several previous studies focusing on Cryptocurrency Forecasting. The first study examined the forecasting of Bitcoin prices using the Random Forest Algorithm [7]. The research findings revealed an RMSE value of 0.010 and a Mean Absolute Error (MAE) of 0.008 when employing the Random Forest Algorithm for Bitcoin forecasting. The second study explored the forecasting of Ethereum prices using the Backpropagation Neural Network Algorithm [8]. The results showed that utilizing this method with a learning rate of 0.001 and 1000 epochs resulted in Mean Absolute Percentage Error (MAPE) values of 1.4694, 1.4839, and 1.4727. The third study focused on forecasting Cryptocurrency prices using the Long-Short Term Memory (LSTM) method [9]. The dataset used in this study was DOGE, and the application of the LSTM method yielded an RMSE value of 0.0630 for DOGE cryptocurrency. The fourth study involved comparing the use of Linear Regression, Neural Network, Deep Learning, and K-Nearest Neighbor algorithms for forecasting Bitcoin prices [10]. The research findings demonstrated that the best models for predicting Bitcoin prices were the Linear Regression and Neural Network algorithms, resulting in RMSE values of 296.227 +/- 60.125 (micro average: 301.655 +/- 0.000) and 338.988 +/- 47.837 (micro average: 342.000 +/- 0.000), respectively. Lastly, the fifth study focused on forecasting Bitcoin prices using the Random Forest Algorithm [5]. The study showed that implementing this method yielded an MAPE value of 1.50% with an accuracy of 98.50%. A summarized comparison of the previous studies is presented in Table 1.

Table 1. Research Gap

Research	Algorithms	Cryptocurrency	Outcome
[7]	Random Forest	Bitcoin	Forecasting Price

Research	Algorithms	Cryptocurrency	Outcome
[8]	Backpropagation Neural Network	Ethereum	Forecasting Price
[9]	Long-Short Term Memory	DOGE	Forecasting Price
[10]	Linear Regression, Neural Network, Deep Learning, and K-Nearest Neighbor	Bitcoin	Forecasting Price
[5]	Random FOrest	Bitcoin	Forecasting Price
Present	ARIMA, Neural Network, Support Vector Machine, Linear Regression, Generalized Linear Model	Bitcoin, Ethereum, and Tether	Forecasting Price

This study fills a gap in previous research by conducting forecasting for three cryptocurrencies simultaneously, namely Bitcoin, Ethereum, and Tether, making it an essential contribution. Additionally, the five algorithms are used as benchmarks to obtain the most optimal model. The evaluation stage will utilize the Root Mean Square Error to measure the performance generated by each model.

2. RESEARCH METHOD

The research methodology employed in this study is Knowledge Discovery In Databases (KDD), a technique used to uncover and analyze patterns in data, as well as interpret and predict future events [11]. The proposed research framework is in chart form, as shown in Figure 2.
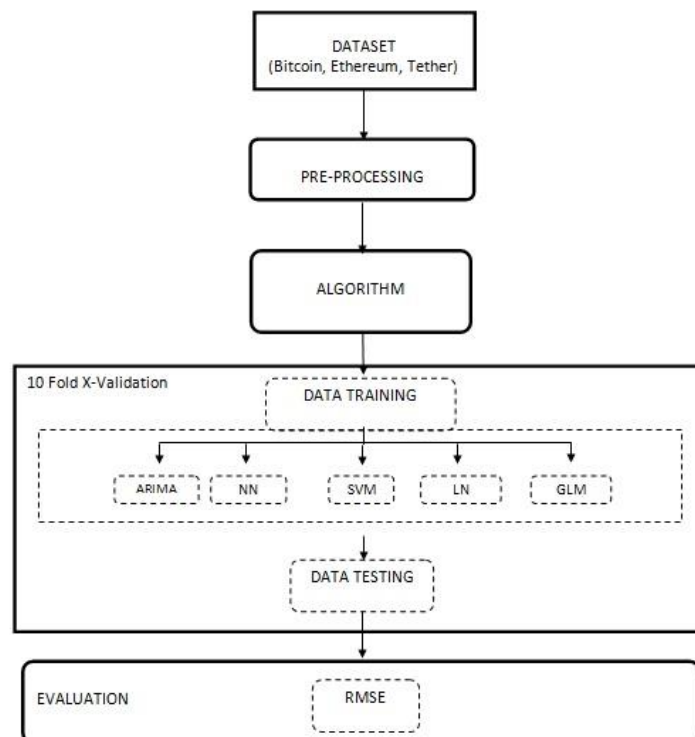

Figure 2. Research Framework

The first step involves discussing the Dataset. The Dataset used consists of the top three Cryptocurrencies based on market cap, namely Bitcoin, Ethereum, and Tether. The Dataset is gathered from the website Yahoo.Finance.com [4]. The trading charts for the three datasets are presented in the form of images, as shown in Figure 3 for Bitcoin, Figure 4 for Ethereum, and Figure 5 for Tether.

Figure 3. Chart Bitcoin [4]

Figure 4. Chart Ethereum [4]

Figure 5. Chart Tether [4]

The dataset represents the population data starting from the beginning of trading until the date of this research, which is 13/06/2023. The initial dataset results gathered from the website are presented in the form of tables, as shown in Table 2 for Bitcoin, Table 3 for Ethereum, and Table 4 for Tether.

Table 2. Initial Bitcoin Dataset

Date	Open	High	Low	Close	Adj Close	Volume
10/1/2014	387.427	411.698	289.296	338.321	338.321	902994450
11/1/2014	338.65	457.093	320.626	378.047	378.047	659733360
12/1/2014	378.249	384.038	304.232	320.193	320.193	553102310
1/1/2015	320.435	320.435	171.51	217.464	217.464	1098811912
....	....	....	....	....	....	....
6/13/2023	25901.96	26132.44	25853.21	26055.02	26055.02	11681516544

Table 3. Initial Ethereum Dataset

Date	Open	High	Low	Close	Adj Close	Volume
12/1/2017	445.209	881.944	414.411	756.733	756.733	79862276544
2/1/2018	1119.37	1161.35	574.419	855.199	855.199	83393441152
3/1/2018	856.012	880.302	368.633	396.457	396.457	53716179584
4/1/2018	397.254	708.875	363.805	669.924	669.924	60669918656
....	....	....	....	....	....	....
6/13/2023	1742.624	1753.694	1731.621	1749.001	1749.001	5931706880

Table 4. Initial Tether Dataset

Date	Open	High	Low	Close	Adj Close	Volume
12/1/2017	1.0072	1.088	0.963089	1.01226	1.01226	62323380736
1/1/2018	1.01255	1.06989	0.971403	0.990274	0.990274	94125940224
2/1/2018	0.991558	1.04818	0.925144	0.998903	0.998903	82122949376
3/1/2018	0.999847	1.02412	0.979633	1.00077	1.00077	69470580224
....	....	....	....	....	....	....
6/13/2023	0.999852	1.000152	0.999555	0.999827	0.999827	20658845696

All the gathered datasets consist of several attributes, including:
1. Date: an attribute that represents a time series [12];
2. Open: the opening price of Crypto [13];
3. High : the highest price of Crypto for one day [14];
4. Low: the lowest price of Crypto for one day [15];
5. Closing: the closing price of Crypto for one day [14];
6. Volume: trading volume of Crypto in USD [16];
7. Adj Close : Closing price adjusted for corporate actions such as rights issue, stock split or stock reverse [13].

The seven attributes in the raw dataset will not all be used. Only the Date and Close attributes will be utilized in this research, while the rest will be eliminated from the dataset. The second stage is Pre-Processing, which is a crucial step before entering the modeling process and before simulating using the Algorithms [17], [18]. Proper pre-processing techniques can enhance recognition accuracy and expedite the subsequent processes [19]. The techniques employed in this stage include :
1. Data Cleansing: the process of cleaning data from empty values, inconsistent, empty attributes such as missing values and noisy data [20];
2. Data Integration: merging data into one archive [21];
3. Data Reduction: eliminating unnecessary attributes [22].

The third stage is Validation. In the Validation stage, the accuracy performance of the developed model is assessed. The algorithms employed for validation include ARIMA, Neural

Network, Support Vector Machine, Linear Regression, and Generalized Linear Model, using the K-Fold Cross-Validation (KCV) technique. Cross-validation is a method used to evaluate the generalizability of statistical analysis outcomes to new and unseen datasets [23]. It involves assessing how well the model or analysis results can be applied to novel data. KVC divides the dataset into k parts and performs k iterations. In each iteration, one part of the dataset is used as testing data, while the remaining k - 1 parts are used for training. This process is repeated k times, and the average deviation (error) value is calculated based on the different test results obtained in each iteration [12]. The formulas for each algorithm are presented as follows:

1. ARIMA Algorithm [24]

   $$y_t = c + \varphi_1 y_{(t-1)} + ... + \varphi_p y_{(t-p)} + \theta_1 \varepsilon_{(t-1)} + ... + \theta_q \varepsilon_{(t-q)} + \varepsilon_t \qquad (1)$$

   In this formula, φ determines the coefficient of the AR process, and θ shows the coefficient for the MA process. As in the literature, the ARIMA model is determined by ARIMA (p, d, q). The "p" refers to the order of the AR process, which means how many previous lags of the variable are included in the model. The "d" represents the level of differencing is needed to transform the series into stationary. Finally, the "q" shows the level of the regression residuals' previous residuals that are included in the model.

2. Neural Network Algorithm

   Neural networks, also known as ANNs (artificial neural networks), are a specific branch within the realm of machine learning. Drawing inspiration from the intricate workings of the human brain, they imitate the exchange of signals between biological neurons [10]. These networks are constructed by interconnecting various layers, such as the input layer, hidden layers, and output layer. A visual representation of the structural model of a neural network can be observed in Figure 6.
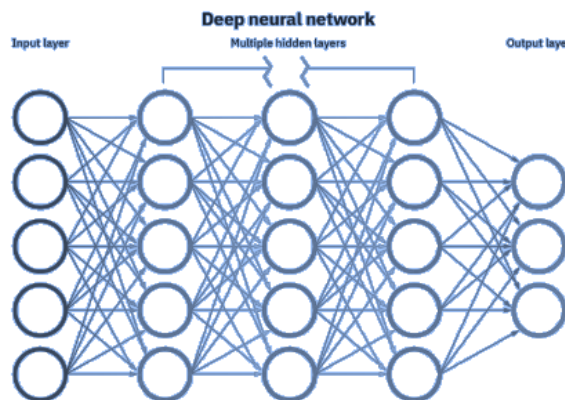

Figure 6. Neural Network Structure

3. Support Vector Machine Algorithm

   $$f(x) = \text{sign}\,(w.x + b) \qquad (2)$$

   If the value according to w.x + b > 0, then it is classified into class +1. And if w.x + b < 0, the data is classified into class -1 [25].

4. Linear Regression Algorithm

   Regression modeling is an analytical method employed to estimate the value of the dependent variable 'y' based on a range of independent variable 'x'. Multivariate linear regression, on the other hand, is a statistical approach utilized to forecast the outcome of a response variable by considering multiple explanatory variables [10].

$$y = \beta_0 + \beta_{1X_1} + \cdots + \beta_{mXm} + \varepsilon \qquad (3)$$

5. Generalized Linear Model

The generalized linear model (GLM) is an expansion of the Linear Regression model. It assumes that the predictors have a linear impact on the response variable but does not make assumptions about a particular distribution for the response variable. GLM is particularly useful when the response variable belongs to the exponential family. Common distributions within the exponential family include normal, Poisson, binomial, gamma, and inverse Gaussian distributions [26].

The final stage is Evaluation. The evaluation phase encompasses the assessment of the results obtained from applying the model in order to determine if the research objectives have been met. Based on this evaluation, a decision is made regarding the utilization of the modeling outcomes [27]. The Parameter is use Root Mean Square Error. RMSE (Root Mean Square Error) represents the magnitude of the error rate in a prediction. A smaller RMSE value indicates a higher level of accuracy in the resulting prediction [28], [29]. The formula for RMSE is as follows:

$$RMSE = \sqrt{\frac{\sum(y\prime - y)^2}{n}} \qquad (4)$$

Note:
$y\prime$ = Predicted Value
y = True Value
n = Total Data

3. RESULTS AND DISCUSSION

The results of the Dataset and Pre-Processing stages are presented in Table 5.

Table 5. Dataset

Bitcoin		Ethereum		Tether	
Date	Close	Date	Close	Date	Close
10/1/2014	338.321	12/1/2017	756.733	12/1/2017	1.01226
11/1/2014	378.047	2/1/2018	855.199	1/1/2018	0.990274
12/1/2014	320.193	3/1/2018	396.457	2/1/2018	0.998903
1/1/2015	217.464	4/1/2018	669.924	3/1/2018	1.00077
….	….	….	….	….	….
6/13/2023	26055.02	6/13/2023	1749.001	6/13/2023	0.999827

Table 5 displays the results of the pre-processing stage, where the selected attributes are "Attribute Date" and "Attribute Close". The "Attribute Close" is designated as the target or class variable. Each dataset is now prepared to proceed to the next stage. The subsequent stages are Validation and Evaluation. These stages involve using Rapidminer to build the process model. RapidMiner offers multiple benefits. In addition to its cross-platform compatibility due to being programmed in Java, it stands out with its agile error correction capabilities. It demonstrates excellence in data transformation, modeling, and visualization techniques [30]. Consequently, RapidMiner is an ideal choice for this research as it provides comprehensive graph-based data visualization capabilities. Validation is performed using the Validation operator in Rapidminer, incorporating five algorithms: ARIMA, Neural Network, Support Vector Machine, Linear Regression, and Generalized Linear Model. The output of the validation process is the Root Mean Square Error (RMSE) value. In the first process model, the

Validation operator is connected to the ARIMA algorithm. The process model is presented in Figure 7.
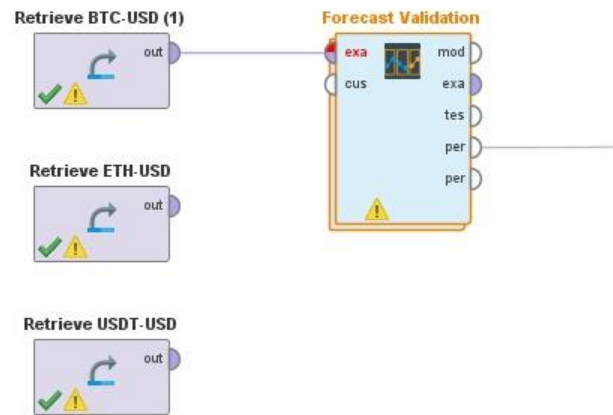

Figure 7. ARIMA Model Process

Figure 7 illustrates the sequential connection of the three cryptocurrency datasets to obtain their respective performance values. The Forecast Validation operator serves as a sub-process operator, which means it contains additional operators such as the ARIMA algorithm operator and the Performance operator. Furthermore, process models are constructed for the other four algorithms, namely Neural Network, Support Vector Machine, Linear Regression, and Generalized Linear Model, as shown in Figure 8.
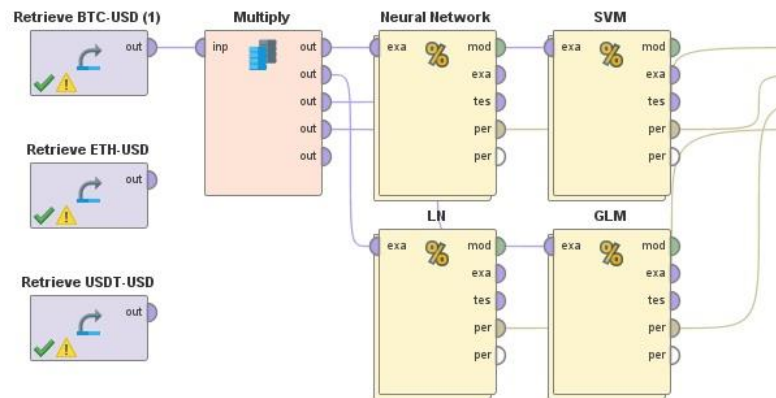

Figure 8. Model Process (NN,SVM,LR,GLM)

The results of the process model validation shown in Figures 7 and 8 are presented in tabular form, as depicted in Table 6.

Table 6. Evaluation Result

Cryptocurrency	Algorithm	Root Mean Square Error
Bitcoin		14521.738
Ethereum	ARIMA	1243.786
Tether		0.003
Bitcoin		9180.534
Ethereum	Neural Network	537.528
Tether		0.004

Cryptocurrency	Algorithm	Root Mean Square Error
Bitcoin		15506.083
Ethereum	Support Vector Machine	1130.260
Tether		0.004
Bitcoin		15536.218
Ethereum	Linear Regression	1121.165
Tether		0.004
Bitcoin		14836.417
Ethereum	Generalized Linear Model	910.516
Tether		0.004

Based on the evaluation results, it can be analyzed that for Bitcoin cryptocurrency, the most optimal RMSE value is shown by the Neural Network algorithm, with an RMSE value of 918.534. Furthermore, for Ethereum cryptocurrency, the most optimal RMSE value is indicated by the Neural Network algorithm with an RMSE value of 537.528. Lastly, for Tether cryptocurrency, the most optimal RMSE value is demonstrated by the ARIMA algorithm with an RMSE value of 0.003. To assess the conformity between the conducted forecasting and the actual data, it is presented in chart form, as shown in Figure 9 for Bitcoin, Figure 10 for Ethereum, and Figure 11 for Tether.
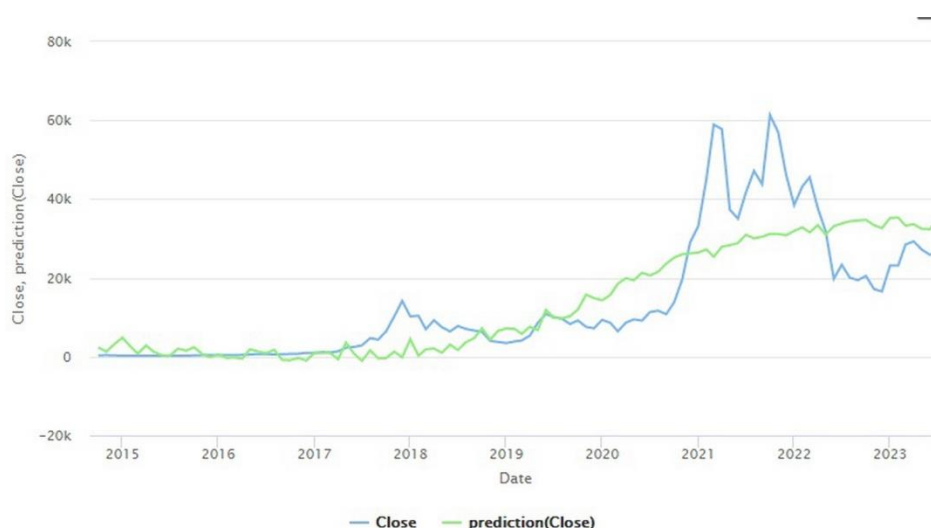

Figure 9. Chart Bitcoin using Neural Network

Based on the chart, the accuracy level between the forecasting values and the actual data can be observed. The highest price movement in reality occurred on October 1, 2021, with a value of 61,318 USD. However, the highest forecasted value occurred on January 1, 2023, reaching 35,193 USD.
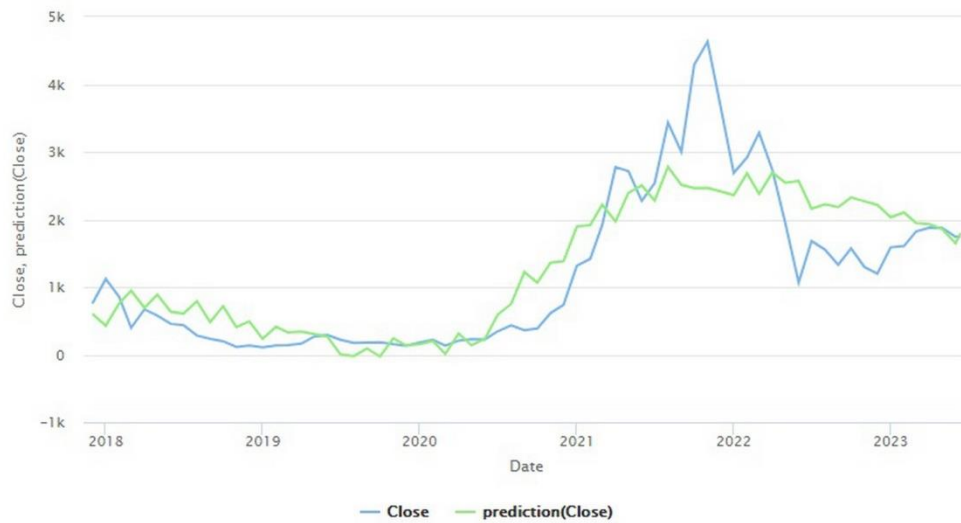
Figure 10. Chart Ethereum using Neural Network

Based on the chart, the accuracy level between the forecasting values and the actual data can be observed. The highest price movement in reality occurred on November 1, 2021, with a value of 4,631 USD. However, the highest forecasted value occurred on August 1, 2021, reaching 2,779 USD.
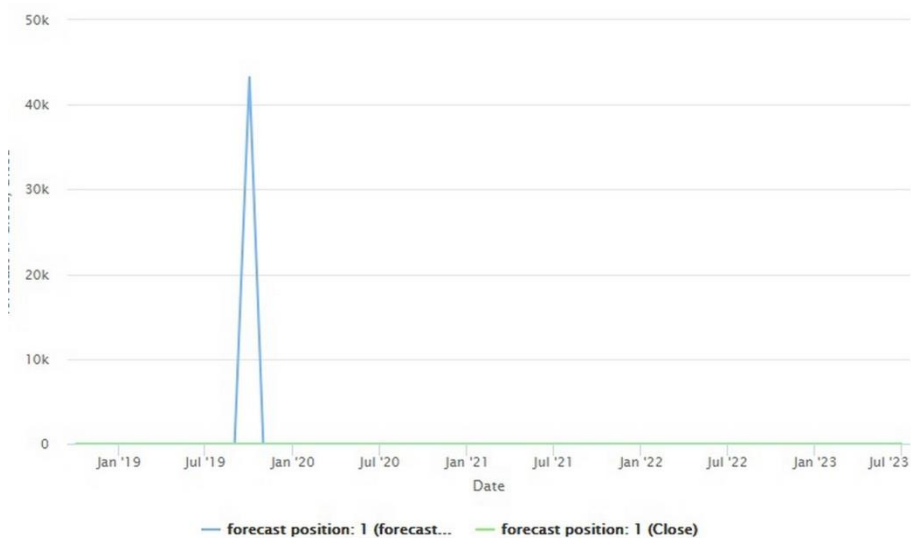

Figure 11. Chart Tether using ARIMA

Based on the chart, the accuracy level between the forecasting values and the actual data can be observed. The highest price movement in reality occurred on October 2, 2019, with a value of 43,327 USD. However, for the forecasting values, they seem to stabilize around 1,000 USD.

4. CONCLUSION

The research results indicate that each cryptocurrency has distinct characteristics in terms of the dataset. For Bitcoin, the most optimal result was achieved using the Neural

Network algorithm with an RMSE value of 918.534, while the lowest result was obtained with the Linear Regression algorithm, yielding an RMSE of 15536.218. Similarly, for Ethereum, the most optimal result was obtained using the Neural Network algorithm with an RMSE of 537.528, whereas the lowest result was demonstrated by the ARIMA algorithm with an RMSE of 1243.786. Regarding Tether, the most optimal result was exhibited by the ARIMA algorithm with an RMSE value of 0.003, while the other algorithms all obtained an RMSE of 0.004.

However, this study has some limitations as it focuses on finding suitable algorithms for cryptocurrency forecasting processes. Future research can overcome these limitations by incorporating algorithm improvement models such as feature selection to obtain more optimal RMSE values for cryptocurrency datasets.

REFERENCES

[1] M. N. Adhar, Y. Prayudi, and E. Ramadhani, "Analisis Artefak Digital Aplikasi Dompet Cryptocurrency Tokocrypto pada Android," *JIIP (Jurnal Ilm. Ilmu Pendidikan)*, vol. 6, no. 5, pp. 3552–3559, 2023, [Online]. Available: http://jiip.stkipyapisdompu.ac.id/jiip/index.php/JIIP/article/view/2024/1793.

[2] M. R. Sampurna, F. Y. Dharta, and D. Kurniansyah, "Komunikasi Persuasif Cuantomonologi Dalam Meningkatkan Pengetahuan Masyarakat Terhadap Cryptocurrency Di Kabupaten Karawang," *J. Ilm. Wahana Pendidik.*, vol. 9, no. 9, pp. 639–648, 2023, doi: doi.org/10.5281/zenodo.7969711.

[3] Z. L. Hudaaka and I. Hanifuddin, "Kejelasan sil ' ah Objektivikasi Cryptocurrency pada Aplikasi Pintu," *J. Ilm. Ekon. Islam*, vol. 9, no. 1, pp. 935–943, 2023.

[4] yahoo finance, "Matching Cryptocurrencies," *yahoo.finance.com*, 2023. https://finance.yahoo.com/crypto/.

[5] S. Saadah and H. Salsabila, "Prediksi Harga Bitcoin Menggunakan Metode Random Forest (Studi Kasus: Data Acak Pada Awal Masa Pandemic Covid-19)," *J. Komput. Terap.*, vol. 7, no. Vol. 7 No. 1 (2021), pp. 24–32, 2021, doi: 10.35143/jkt.v7i1.4618.

[6] M. F. Chania, O. Sara, and I. Sadalia, "Analisis Risk dan Return Investasi pada Ethereum dan Saham LQ45," *Stud. Ilmu Manaj. dan Organ.*, vol. 2, no. 2, pp. 139–150, 2021, doi: 10.35912/simo.v2i2.669.

[7] Indriyanti, N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin," *J. RESPONSIF*, vol. 4, no. 2, pp. 118–125, 2022.

[8] Y. D. Wibowo and R. Somya, "Penerapan Data Mining Untuk Memprediksi Harga Cryptocurrency Ethereum Menggunakan Metode Backpropagation Neural Network," *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 1, pp. 273–286, 2023.

[9] M. F. Rizkilloh and S. Widyanesti, "Prediksi Harga Cryptocurrency Menggunakan Algoritma Long Short Term Memory (LSTM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 25–31, 2022, doi: 10.29207/resti.v6i1.3630.

[10] Mikhael, F. Andreas, and U. Enri, "Perbandingan Algoritma Linear Regression, Neural Network, Deep Learning, Dan K-Nearest Neighbor (K-Nn) Untuk Prediksi Harga Bitcoin," *JSI J. Sist. Inf.*, vol. 14, no. 1, pp. 2450–2464, 2022, [Online]. Available: http://ejournal.unsri.ac.id/index.php/jsi/index.

[11] B. Molina Coronado, U. Mori, A. Mendiburu, and J. Miguel-Alonso, "Survey of Network Intrusion Detection Methods from the Perspective of the Knowledge Discovery in Databases Process," *IEEE Trans. Netw. Serv. Manag.*, vol. 17, no. 4, pp. 2451–2479, 2020, doi: 10.1109/TNSM.2020.3016246.

[12] I. T. Julianto, D. Kurniadi, and F. M. Khoiriyyah, "Price Prediction of Non-Fungible Tokens (NFTs) using Data Mining Prediction Algorithm," in *2023 International Conference on*

Computer Science, Information Technology and Engineering (ICCoSITE), Feb. 2023, pp. 633–637, doi: 10.1109/ICCoSITE57641.2023.10127679.

[13] I. T. Julianto, D. Kurniadi, F. A. Fauziah, and R. Rohmanto, "Improvement of Data Mining Models using Forward Selection and Backward Elimination with Cryptocurrency Datasets," *J. Appl. Intell. Syst.*, vol. 8, no. 1, pp. 100–109, 2023.

[14] I. T. Julianto, D. Kurniadi, M. R. Nashrulloh, and A. Mulyani, "Data Mining Algorithm Testing For SAND Metaverse Forecasting," *J. Appl. Intell. Syst.*, vol. 7, no. 3, pp. 259–267, 2022.

[15] A. Y. Pramudya, A. Kurniawati, and D. Agusten, "Implementasi Metode Autoregressive Integrated Moveing Average (ARIMA) pada Aplikasi Peramalan Harga Saham Berbasis Website," *J. Ilm. Komputasi*, vol. 22, no. 1, pp. 105–112, 2023.

[16] I. T. Julianto, D. Kurniadi, M. R. Nashrulloh, and A. Mulyani, "Comparison Of Data Mining Algorithm For Forecasting Bitcoin Crypto Currency Trends," *JUTIF*, vol. 3, no. 2, pp. 245–248, 2022.

[17] A. D. Savitri, F. A. Bachtiar, and N. Y. Setiawan, "Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan (Studi Kasus : Belle Crown Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 9, pp. 2957–2966, 2018.

[18] V. Mandailina, Abdillah, and Syaharuddin, "Analisis Tingkat Akurasi Variasi Algoritma Min-Max Backpropagation sebagai Pre-Processing Data Time Series," *Techno.COM*, vol. 22, no. 2, pp. 299–300, 2023.

[19] S. Ayu, R. Srg, and M. Zarlis, "Klasifikasi Citra Daun dengan GLCM (Gray Level Co-Occurence) dan Leaf Image Clasification with GLCM (Gray Level Co-Occurence) and K-NN (K-Nearest Neighbor)," *Matrik J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 2, pp. 477–488, 2022, doi: 10.30812/matrik.v21i2.1572.

[20] W. W. Kristianto and C. Rudianto, "Penerapan Data Mining Pada Penjualan Produk Menggunakan Metode K- Means Clustering (Studi Kasus Toko Sepatu Kakikaki)," *J. Pendidik. Teknol. Inf.*, vol. 5, no. 2, pp. 90–98, 2022.

[21] Abdussalam Amrullah, Intam Purnamasari, Betha Nurina Sari, Garno, and Apriade Voutama, "Analisis Cluster Faktor Penunjang Pendidikan Menggunakan Algoritma K-Means (Studi Kasus: Kabupaten Karawang)," *J. Inform. dan Rekayasa Elektron.*, vol. 5, no. 2, pp. 244–252, 2022, doi: 10.36595/jire.v5i2.701.

[22] A. Sukmawati, H. M. . Basri, and M. Akhir, "Pembentukan Karakter Berbasis Keteladanan Guru Dan Pembiasaan Murid Sit Al Biruni Jipang Kota Makassar," *Educ. Hum. Dev. J.*, vol. 5, no. 1, pp. 91–99, 2020, doi: 10.33086/ehdj.v5i1.1453.

[23] G. Feng, M. Fan, and Y. Chen, "Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining," *IEEE Access*, vol. 10, pp. 19558–19571, 2022, doi: 10.1109/ACCESS.2022.3151652.

[24] M. Benzekrı and H. Özütler, "On the Predictability of Bitcoin Price Movements: A Short-term Price Prediction with ARIMA," *J. Econ. Policy Res. / İktisat Polit. Araştırmaları Derg.*, vol. 8, no. 2, pp. 293–309, 2021, doi: 10.26650/jepr.946081.

[25] A. Tasari, D. Dinata Tarigan, E. Nia, D. B. Purba, K. Saputra, and I. Artikel, "Perbandingan Algoritma Support Vector Machine dan KNN dalam Memprediksi Struktur Sekunder Protein," *J. Inform.*, vol. 9, no. 2, pp. 172–179, 2022, [Online]. Available: https://raw.githubusercontent.com/JiayingLi/Pr.

[26] F. H. Hamdanah and D. Fitrianah, "Analisis Performansi Algoritma Linear Regression dengan Generalized Linear Model untuk Prediksi Penjualan pada Usaha Mikra, Kecil, dan Menengah," *J. Nas. Pendidik. Tek. Inform.*, vol. 10, no. 1, p. 23, 2021, doi: 10.23887/janapati.v10i1.31035.

[27] Yunitasari, H. S. Hopipah, and R. Mayasari, "Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-nearest neighbor (k-NN) and Naive Bayes," *Technomedia J.*, vol. 6, no. 1, pp. 99–110, 2021, doi: 10.33050/tmj.v6i1.1531.

[28] M. M. Sidabutar and G. Firmansyah, "Comparison of Linear Regression , Neural Net , and ARIMA Methods For Sales Prediction of Instrumentation and Control Products in PT . Sarana Instrument," *JRSSEM*, vol. 2, no. 8, pp. 1694–1705, 2023.

[29] B. Pradito and D. S. Purnia, "Komparasi Algoritma Linear Regression Dan Neural Network Untuk Memprediksi Nilai Kurs Mata Uang," *Evolusi J. Sains dan Manaj.*, vol. 10, no. 2, pp. 64–71, 2022.

[30] S. C. Dewi, A. M. Siregar, and D. S. Kusumaningrum, "Pengelompokan Jumlah Sumber Daya Manusia Kesehatan Puskesmas untuk Menunjang Pemerataan pada Provinsi Jawa Tengah Menggunakan Algoritma K-Means," *Sci. Student J. Information, Technol. Sci.*, vol. 1, no. 2, pp. 86–94, 2020.