Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 23, No. 2, March 2023, pp. 259∼272
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: 10.30812/matrik.v23i2.3267

Comparing Long Short-Term Memory and Random Forest Accuracy for Bitcoin Price Forecasting

Munirul Ula[1], Sariah[1], Veri Ilhadi[1], Zailani Mohamed Sidek[2]
[1]Universitas Malikussaleh, Aceh, Indonesia
[2]Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

Article Info

Article history:

Received August 08, 2023
Revised November 13, 2023
Accepted January 05, 2024

Keywords:

Accuracy
Bitcoin
Forecasting
Long Short-Term Memory
Random Forest

ABSTRACT

Bitcoin's daily value fluctuations are very dynamic. Understanding its rapid and intricate price movements demands advanced techniques for processing complex data. This research aims to compare the accuracy of two machine learning methods, Random Forest (RF) and Long Short-Term Memory (LSTM), in predicting Bitcoin price. This research employs RF and LSTM algorithms to forecast Bitcoin prices using a two-year Yahoo Finance dataset. The evaluation metrics used were accuracy based on Mean Absolute Percentage Error (MAPE) and computational power (CPU-Z). As a result of this research, the LSTM model demonstrates higher accuracy compared to the RF model. MAPE reveals LSTM's precision of 99.8% and RF's accuracy of 90.1%. Regarding computational time and resources, RF shows slightly better performance than LSTM. The visual comparison further emphasizes LSTM's better performance in predicting Bitcoin prices, highlighting its potential for informed decision-making in cryptocurrency trading. This research contributes valuable insights into the effectiveness, strengths, and weaknesses of LSTM and RF models in predicting cryptocurrency trends.

Corresponding Author:

Munirul Ula, +62 82367818878,
Engineering Faculty, Dept of Information Technology,
Universitas Malikussaleh, Aceh, Indonesia,
Email: munirulula@unimal.ac.id

How to Cite:
M. Ula, V. Ilhadi, and Z. Sidek, "Comparing Long Short-Term Memory and Random Forest Accuracy for Bitcoin Price Forecasting", *MATRIK: Jurnal Managemen,Teknik Informatika, dan Rekayasa Komputer*, Vol. 23, No. 2, pp.259-272 Comparing Long Short-Term Memory and Random Forest Accuracy for Bitcoin Price Forecasting, Mar, 2024.

1. INTRODUCTION

Since the start of the 4.0 era, a wave of technological innovations has changed many economies and every part of life. In the middle of all these big changes, the rise of cryptocurrencies has been a big deal for the financial world. Cryptocurrencies differ from traditional currencies because they are digital assets backed by cryptography that can be traded almost anywhere on the internet [1]. Bitcoin is a well-known example of one of these digital currencies. Bitcoin was created in 2009 by a mysterious person named Satoshi Nakamoto [4]. It is a digital currency that works without a central exchange system. Its unique design makes it easy for people to trade directly with each other without the need for middlemen [5, 6]. Every Bitcoin transaction is carefully recorded in the blockchain, a distributed ledger that a central bank or government cannot control [7, 8]. Because a central authority does not run Bitcoin, its price can grow without stopping or slowing down [9, 10]. With few and clear controls on inflation, the currency's value is inherently unstable. This quality is emphasized by Bitcoin's value changing daily, which shows how dynamic it is [11, 12]. The Indodax.com website says that the exchange rate for Bitcoin on February 11, 2023, is Rp. 329,188,000. In this situation, figuring out how Bitcoin's price moves so quickly and in a complicated way requires advanced techniques to separate and combine data into meaningful insights [13, 14]. However, data alone is not enough to bring out its hidden potential. People who want to invest in Bitcoin must deal with the fact that its value changes quickly, which means they need a system that can predict where it will go [14, 15].

In literature, researchers use many ways to predict the price of Bitcoin, including ARIMA, RNN (Recurrent Neural Network), Super Vector Regression, GRU, and ELM (Extreme Learning Machine) [16, 17]. This study looks at the potential of two different strategies, Long Short-Term Memory (LSTM) and Random Forest, among the many options [6, 18, 19]. The LSTM is a type of RNN. It has special gates that control how long memories are kept, which makes it easier to learn a lot from data sequences. Studies have shown that it is good at predicting time series data, with LSTM making average predictions that are 85% better than ARIMA [20, 1, 21]. On the other hand, the Random Forest technique stands out as a strong competitor that can handle complex and non-linear datasets well. It is very good at predicting Bitcoin prices affected by complex, nonlinear factors because it comprises a group of decision trees. Siti Saadah and Haifa Salsabilla's research proves how good Random Forest is. It can accurately predict Bitcoin prices from 95% to 98% for random data and 19% to 37% for non-random data [18]. The difference between this research and the previous one lies in its focused comparison of two distinct methodologies, Long Short-Term Memory (LSTM) and Random Forest, within a landscape that encompasses various established approaches like ARIMA, RNN, SVR, GRU, and ELM utilized for Bitcoin price prediction. The novelty of this study resides in its direct comparative analysis between LSTM and Random Forest, providing a comprehensive insight into their respective performances tailored specifically to cryptocurrency price prediction.

The contribution of this research is that it uses both the LSTM and Random Forest models to predict the price of Bitcoin. Random Forest uses its collection of decision trees to understand complex relationships that don't follow a straight line [22–24]. This contrasts LSTM, which takes advantage of its natural ability to recognize complex sequential patterns. By comparing how well these two methods can predict the future, this study aims to give an in-depth look at how good they are at predicting Bitcoin prices. This research aligns with the financial world's current needs because it bridges the gap between advanced data analysis techniques and the volatile world of cryptocurrency trading. By using cutting-edge methods, it wants to give investors the tools they need to navigate the complicated cryptocurrency market more accurately and confidently. This study makes a new contribution by combining LSTM and Random Forest models in a new way to predict the price of Bitcoin. This could lead to insights beyond the limits of traditional financial analysis. This research helps us learn more about how technology and finance merge in the 4.0 era by discovering the complicated relationships behind Bitcoin's price changes. This can help us make better investment decisions and better understand the ever-changing world of cryptocurrencies. There are several objectives for this study. Its main goal is to use the LSTM and Random Forest techniques to make accurate models for predicting Bitcoin prices. The models that are made will be judged on how well they can predict the future [4, 7, 10, 25]. This will show how well they can capture the complicated dynamics of Bitcoin's price changes. Also, the study tries to determine each method's pros and cons, which will add to the ongoing discussion about the best ways to predict cryptocurrency prices.

The following section discusses the research method. This study used a time series dataset of daily Bitcoin prices, focusing on the "Date" and "Close" features. The research involved data preparation, LSTM and Random Forest model implementations, and MAPE evaluation. LSTM and Random Forest were used for price prediction, and their accuracy was assessed using MAPE.

2. RESEARCH METHOD

The data used in this study is in the form of time series, which is secondary data. Secondary data is data collected indirectly by the researcher through intermediaries. The data collected in this research is the daily bitcoin price data for the last two years. The dataset obtained consists of 7 features: Date, Open, High, Low, Close, Adj Close, and Volume. The date represents the date of Bitcoin transactions. Open represents the initial price of Bitcoin on the respective date. High represents the highest price of bitcoin on the respective date. Low represents the lowest price of bitcoin on the respective date. Close is the closing price or the final price of bitcoin on the respective date. Adj Close is the closing price minus dividends. Volume is the total trading volume of bitcoin on the respective date. However, only Date and Close are used as the features in this research. The author can access bitcoin price data through Yahoo Finance Bitcoin USD (BTC-USD) Price History & Historical Data - Yahoo Finance. The detailed stages of the research are shown in Figure 1.

1. "Start" or "begin" is an initial system initialization process.
2. "Check missing Value" is the stage of examining whether there are any empty data in the dataset to be used.
3. "Data Allocation" is the step of dividing data into training data and testing data using the Python library that implements train/test split with a parameter in the form of data proportions for testing, namely Scikit-Learn. The dataset is divided into 80% for training data and the remaining 20% for testing data. A larger amount of training data is used to train the learning machine to understand the model better. This way, when the machine produces a model, it will provide more optimal predictions for the testing data.

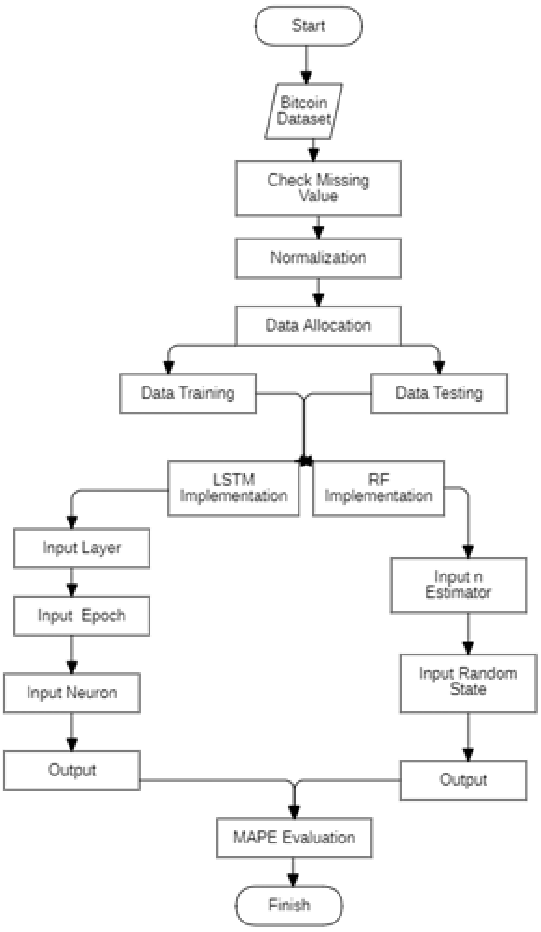

Figure 1. Research stages

4. LSTM Implementation is carried out as follows:
   a. Input the number of layers for LSTM construction, where Layers 1 and 2 consist of 50 Neurons each.
   b. Input the number of Epochs, where in this research, a total of 200 epochs are used. This means the training steps will be repeated 200 times.
   c. Neurons are input into each layer, where in this research, there are 100 neurons in the LSTM layer and 1 Neuron in the Dense layer.

5. Random Forest Implementation is carried out as follows:
   a. Input the number of Estimators, where n-Estimators will determine the number of decision trees used in the ensemble.
   b. Input the Random_state, which will be used to set the seed for initializing the random number generator, allowing reproducibility of results.
6. MAPE Evaluation, at this stage, is the step to evaluate the built models. In this stage, how well the models can predict the price of Bitcoin will be determined. The accuracy of both models will be obtained in this stage, and it can be determined which model is better at making predictions.
7. "End" is the conclusion of the System.

2.1. Normalization

A significant price range can lead to higher errors in the built model. To overcome this, normalization is needed to address the wide price range. The normalization process is done by converting the closing price values into values within the interval range of 0-1 based on the *xnorm* formula [26]. Equation 1 expresses the normalization formula.

$$xnorm = (X - x\_min)/(x\_max - x\_min) \tag{1}$$

Where x is the original value of the data, x_min is the minimum value in the dataset, and x_max is the maximum value in the dataset.

2.2. Long Short-Term Memory

The LSTM RNN model can learn long-term relationships and dependencies. This model can generate errors without vanishing gradients by introducing cell modes with constant errors. Unlike other RNNs, LSTM does not have just one neuron layer; instead, it possesses four interacting neuron layers structured in a specific manner. LSTM has four times the parameters and computational cost of a regular RNN. It comprises three gates and a network for computing memory inputs. The input or update gate functions to assess whether the information acquired at the present moment is worth storing in long-term memory, while the output gate is used to transfer necessary information selectively. The role of the forget gate is to inspect the age of memory within the memory cell and store unnecessary past information. Parameters to be determined using training data include the number of hidden layers and neurons, maximum epochs, and learning rate. Batch size can be determined algorithmically or randomly [27].

2.3. Random Forest

The Random Forest (RF) method can enhance results because the random generation of child nodes is performed for each node. This technique is used to construct a decision tree consisting of root, internal, and leaf nodes by randomly selecting attributes and data as needed [28]. The Random Forest algorithm has decision trees and logical trees that differentiate data. For instance, the grouped data can form a tree in a dataset containing two instances of the number 1 and five instances of the number 0, where each number is differentiated by color [29].

2.4. Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) is the average of absolute differences between predicted and actual observed values, expressed as a percentage of the actual observed values. The MAPE value can be calculated using the following equation [30]. Equation 2 express the MAPE value calculation formula:

$$MAPE = \Sigma \mid xt - yt \mid / \text{xt nt} = 1/n \times 100 \tag{2}$$

MAPE is calculated as the average absolute difference between the forecasted and actual values, expressed as a percentage of the actual values [31].

2.5. Calculation Time and Computational Power Measurement

The duration between the start and end of the model training process is carefully recorded to measure the total computation time required to train the model. The Python library' psutil' was used to evaluate the CPU load during the training and predicting phase. In particular, the 'cpu_percent' method in this library facilitates the extraction of real-time CPU load metrics, displaying the CPU load percentage during the training process. By utilizing 'psutil,' access to the system memory usage during the model training period is performed. Using this library's 'virtual_memory' function, the script retrieves comprehensive information about memory usage, presenting the memory used in megabytes (MB). Every metric, including calculation time, CPU load percentage, and memory usage, is carefully recorded and archived for later analysis and comparison. These recorded metrics become fundamental data points for evaluating the trained model's computational efficiency and resource utilization.

3. RESULT AND ANALYSIS

3.1. 3.1 Descriptive Analysis

On the coinmarketcap.com website on July 6, 2023, the overall cryptocurrency market capitalization was dominated by BTC at 49.8% and ETH at 19.3%. The Fully Diluted Market Cap amounted to $639,892,760,958, while the current market capitalization stood at $591,742,768,131. Thus, it can be concluded that Bitcoin remains highly sought after by traders/investors worldwide. Figure 2 shows the daily price analysis of Bitcoin from May 5, 2021, to May 5, 2023. Table 1 shows the description of each feature in the Bitcoin price dataset.

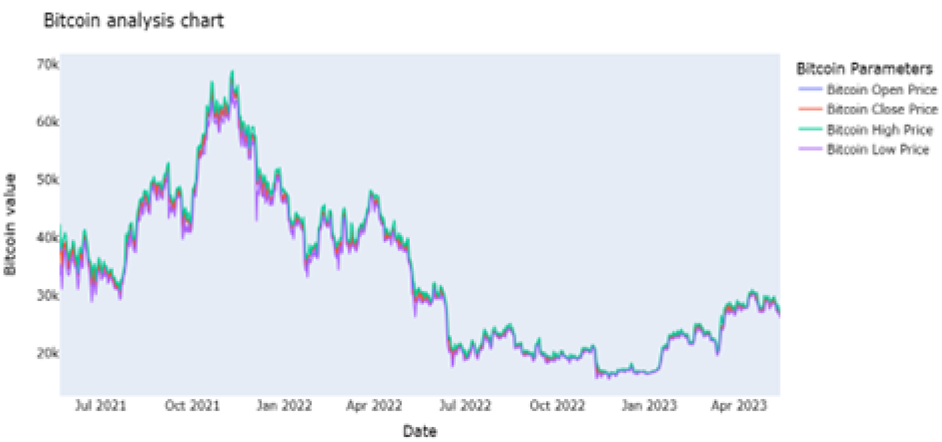

Figure 2. Bitcoin analysis chart

Table 1. Fitur Description

	Open	High	Low	Close	Adj Close	Volume
Count	731.000000	731.000000	731.000000	731.000000	731.000000	7.310000e+02
Mean	33491.305345	34243.519887	32622.839665	33444.692569	33444.692569	3.086349e+10
Std	12955.188166	13286.485661	12550.802398	12920.089178	12920.089178	1.312701e+10
Min	15782.300781	16253.047852	15599.046875	15787.284180	15787.284180	7.714767e+09
25%	21528.958985	21803.812500	20959.862305	21531.104492	21531.104492	2.274377e+10
50%	31151.480469	31935.945313	29944.802734	31022.906250	31022.906250	2.927904e+10
75%	42915.814454	43835.765625	41966.779297	42841.775390	42841.775390	3.633242e+10
Max	67549.734375	68789.625000	66382.062500	67566.828125	67566.828125	1.263581e+11

The data in Table 1 represent various statistical measures and summary statistics for a dataset related to Bitcoin price and trading volume in daily or periodic intervals. The description of the data based on the table as follows: the dataset contains 731 data points, the mean opening price is approximately 33,491.31, and the mean closing price is approximately 33,444.69, the standard deviation for the opening price is approximately 12,955.19, the lowest opening price recorded in the dataset is approximately 15,782.30, and the highest closing price recorded is approximately 67,566.83.

3.2. Forecasting Analysis by LSTM and Random Forrest

To determine the system requirements during the research, various analyses will be conducted on various factors that support the study. This research utilizes the Random Forest and Long Short Term Memory (LSTM) algorithms to predict the movement of Bitcoin prices. The dataset sample is extracted from Yahoo Finance and shown in Table 2.

Table 2. Sample Dataset

Date	Open	High	Low	Close	Adj Close	Volume
5/12/2021	56714.53125	57939.36328	49150.53516	49150.53516	49150.53516	75215403907
5/13/2021	49735.43359	51330.84375	46980.01953	49716.19141	49716.19141	96721152926
5/14/2021	49682.98047	51438.11719	48868.57813	49880.53516	49880.53516	55737497453
5/15/2021	49855.49609	50639.66406	46664.14063	46760.1875	46760.1875	59161047474
5/16/2021	46716.63672	49720.04297	43963.35156	46456.05859	46456.05859	64047871555
...	...	...	...	...	...	...
5/9/2023	27695.06836	27821.40039	27375.60156	27658.77539	27658.77539	14128593256
5/10/2023	27654.63672	28322.6875	26883.66992	27621.75586	27621.75586	20656025026
5/11/2023	27621.08594	27621.94141	26781.82617	27000.78906	27000.78906	16724343943
5/12/2023	26990.14844	27054.1582	26178.61719	26375.36719	26375.36719	18337546240
5/3/2023	28680.49414	29259.5332	28178.38867	29006.30859	29006.30859	19122972518
5/4/2023	29031.30469	29353.18555	28694.03906	28847.71094	28847.71094	15548678514

This dataset contains Bitcoin-related financial information from 12 May 2021 to 11 May 2023. It includes important metrics such as the opening, highest, lowest, and closing prices, as well as adjusted close values and daily trading volumes. This time-series dataset provides a valuable resource for analyzing Bitcoin's price trends, volatility, and trading patterns over this extensive time span, thereby facilitating diverse financial research and investment decision-making endeavors in the cryptocurrency market.

1. Checking Missing Values

The checking missing values process, as shown in Figure 3, is the checking of whether there are empty or missing data in the dataset used. Therefore, it is known that there are no missing data in the dataset used, allowing us to proceed with the next process.

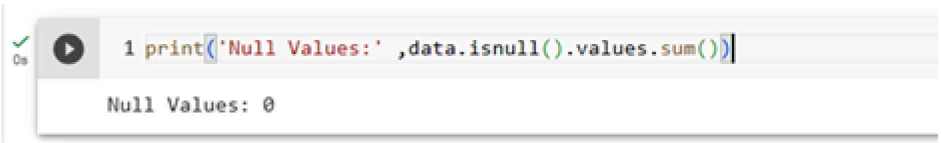

Figure 3. Checking missing values

2. Data Allocation

The research data is divided into training and testing data, with 80% allocated for training, 554 data, and 20% for testing, 117 data. The purpose of using larger training data is to help the machine learning or learning algorithms better understand the data patterns from the training data. The training data is used to train the LSTM and Random Forest methods. This training will result in a model, which will then be applied to the testing data to evaluate its performance. This procedure is continued until the model achieves the best accuracy.

3. LSTM Algorithm Implementation in Bitcoin Price Prediction

In this section, the LSTM algorithm is implemented for predicting Bitcoin prices. The parameters for constructing a Long Short-Term Memory (LSTM) neural network are summarized in Table 3. The table outlines the network's three layers: the "Tanh" activation function, the "Adam" optimization algorithm, 200 training epochs, and 32-batch size. These particulars offer crucial insights into the configuration of the LSTM network, thereby facilitating its replication and comprehension in machine learning applications.

Table 3. LSTM Construction

No	Type	Value
1	Layer	3
2	Activation	Tanh
3	Optimizer	Adam
4	Epoch	200
5	Batch Size	32

Table 4. Actual Prices and LSTM Predicted Prices

Actual	Predicted
21169.632813	21022.289062
21161.519531	21227.421875
20688.781250	21233.451172
21086.792969	20879.974609
22676.552734	20964.839844
22777.625000	22068.351562
22720.416016	22628.533203
22934.431641	22756.371094
22636.468750	22894.142578
23117.859375	22715.818359
...	...
29006.308594	28462.531250
28847.710938	28795.431641
29534.384766	28777.562500
28904.623047	29318.238281
28454.978516	28956.357422
27694.273438	28498.375000
27658.775391	27784.980469
27621.755859	27583.968750
27000.789063	27530.718750
26375.367188	27071.433594

The data in Table 4 compares actual prices to predicted prices, which a forecasting model like LSTM likely generated. The "Actual Prices" column contains the actual prices of financial assets recorded at various time intervals, while the "Predicted Prices" column contains the model's estimates of those prices. This dataset is used to evaluate the model's performance by measuring its price forecasts' accuracy. Analysts often calculate evaluation metrics to gauge the model's forecasting precision and make informed decisions or enhancements to improve predictive accuracy. For the comparison of those prices, refer to Figure 4.

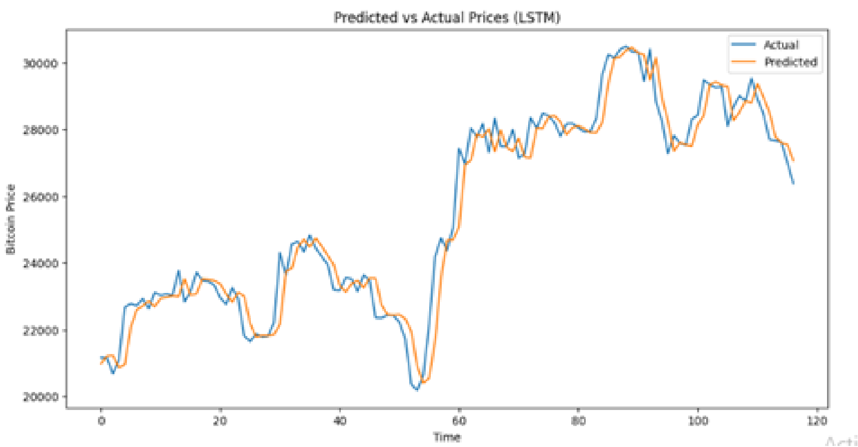

Figure 4. LSTM Price Comparison Graph

Figure 4, titled "LSTM Price Comparison Graph," visually represents the earlier-mentioned comparison of actual and predicted prices. This graph is likely a visual representation of how closely the predicted prices match the actual prices over the time period covered by the dataset. This type of visualization provides a simple way to evaluate the performance of a model, allowing analysts to observe trends, discrepancies, and the overall accuracy of the predictions at a glance.

4. Implementation of the Random Forest Algorithm

The prediction of Bitcoin price using the Random Forest method on the testing model will be built with 100 trees and a Random State of 42, ensuring that every time we run this model, we will obtain the same results if we use the same seed. This is essential to ensure reproducibility and compare results with consistency. Table 5 shows some sample prediction data.

Table 5. Actual Prices and Random Forest Predicted Prices

Actual	Predicted
21169.632813	20658.546914
21161.519531	20976.278457
20688.781250	20961.364395
21086.792969	19962.677852
22676.552734	20689.814727
22777.625000	22475.046133
...	...
29006.308594	27122.665820
28847.710938	28741.319414
29534.384766	28998.234356
28904.623047	29161.065352
28454.978516	28640.326504
27694.273438	29274.165410
27658.775391	25600.100313
27621.755859	24867.100078
27000.789063	24845.719336
26375.367188	23572.514688

Table 5 provides a comparison between observed actual prices and predicted prices generated by a Random Forest forecasting model. The "Actual" column contains prices recorded in the real world, representing the values of Bitcoin at various time points. The "Predicted" column displays price forecasts generated by the Random Forest model. This dataset is essential for evaluating the model's performance, enabling analysts to evaluate the accuracy of its price forecasts. Figure 5 shows the graph comparing the actual prices and predictions obtained from the built model.

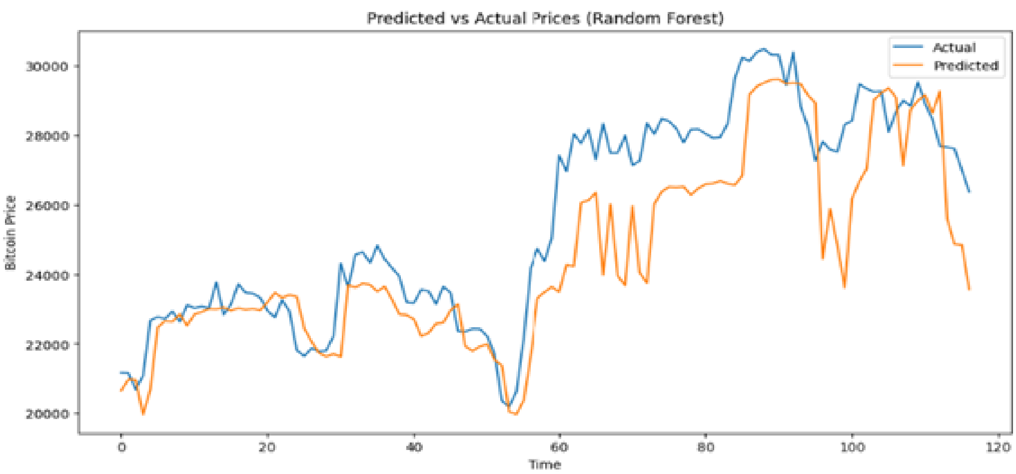

Figure 5. Random forest price comparison graph

The comparison between actual prices and predictions derived from a constructed model is depicted in Figure 5. Throughout a range of time periods, this graph likely depicts the degree to which predicted prices and actual prices correspond. Such visualizations are essential for evaluating the model's performance, as they allow viewers to quickly grasp the accuracy and alignment of the predictions with the actual price trends, thereby facilitating comprehension of the model's forecasting capabilities.

3.3. Comparing LSTM and Random Forest Performance Metrics

1. Accuracy

The accuracy results obtained using the MAPE Evaluation on the built model are LSTM MAPE: 0.020422925094496457, Random Forest MAPE: 0.9999935016178609. From the obtained values, the model's accuracy percentage can be calculated as follows: LSTM: (1 0.020422925094496457) * 100 = 99.8%, Random Forrest: (1 0.9999935016178609) * 100 = 90.1%.

Figure 6. Comparison Both of Model and Actual Price

Figure 6 is a visual comparison of the performance of the Long Short-Term Memory (LSTM) model and the Random Forest model in predicting Bitcoin prices. Upon closer inspection of the graph, it is evident that the LSTM model displays a significantly closer alignment with the actual price values than the Random Forest model. This observation leads to a clear and convincing conclusion: in the context of this study, the Long Short-Term Memory (LSTM) method is the superior choice for predicting Bitcoin prices. The ability of the LSTM model to closely track and approximate actual price trends demonstrates its efficacy and dependability in capturing the complex dynamics of the cryptocurrency market. These findings have significant implications for practical applications, as they suggest that investors and stakeholders may benefit from LSTM-based models when making informed decisions regarding digital asset trading. Moreover, this conclusion highlights the significance and promising potential of advanced deep learning techniques, such as LSTM, for enhancing financial forecasting and data-driven decision-making processes in the cryptocurrency industry.

2. Calculation Time and Computational Power

In evaluating model performance encompassing accuracy, calculation time, and computational power, this research examined both LSTM and Random Forest (RF) methodologies for computational modeling. Both LSTM and RF models were executed on a workstation equipped with high-end specs: 2 × AMD Ryzen^{TM} 9 5900X CPUs @ 3.7 GHz, 256 GB DDR4 memory, and NVIDIA® GeForce RTX^{TM} 3090 graphics card boasting 24 GB GPU memory. Table 6 presents detailed comparisons of calculation time and computing power consumption during training and prediction processes for both methods.

Table 6. Comparisons of Calculation Time and Computing Power Consumption During Training and Prediction

Method	Process	CPU Time (s)	CPU Load (%)	Memory Used (MB)
LSTM	Training	544.3	74.00	2676
Random Forest	Training	410.2	70.90	2465
LSTM	Prediction	27.55	4.30	248.5
Random Forest	Prediction	25.20	4.10	235.7

Long Short-Term Memory (LSTM) models in the training phase require significantly more computing resources than Random Forest (RF) models. LSTM requires 544.3 seconds of CPU time, exceeding the RF computing duration of about 134.1 seconds. This longer duration is accompanied by a higher CPU load of 74.00% compared to RF's 70.90%, indicating that LSTM may involve more intensive processing tasks. Additionally, LSTM consumes a larger amount of memory, consuming 2676 MB, compared to 2465 MB of RF during training. However, although LSTM continues to show slightly higher resource usage in the prediction phase, the difference between the two models is less significant. During prediction, LSTM requires 27.55 seconds of CPU time compared to RF's 25.20 seconds, with a slightly higher CPU load of 4.30% compared to RF's 4.10%. Additionally, LSTM uses slightly more memory at 248.5 MB compared to RF's 235.7 MB during prediction.

Overall, LSTM shows higher resource consumption, especially regarding CPU time, CPU load, and memory usage, during both training and prediction phases compared to Random Forest. Although the differences are more significant in the training phase, both models show efficient resource usage during prediction, with LSTM consistently showing slightly higher resource usage. This data suggests that Random Forest may have an advantage in some computational tasks due to its lower resource demands compared to LSTM in certain scenarios. However, the efficiency of each model can differ depending on the nature of the computational task and the dataset's characteristics.

3.4. Discussion

The findings of this research emphasize the effectiveness of LSTM and Random Forest algorithms in predicting Bitcoin prices. This study also indicates that LSTM demonstrates superior accuracy in forecasting Bitcoin prices compared to the Random Forest method. The LSTM model, with its ability to capture long-term dependencies in sequences and learn from complex data patterns, showcased an accuracy rate of approximately 99.8%. On the other hand, the Random Forest method achieved an accuracy rate of around 90.1%. This substantial discrepancy in accuracy underscores the LSTM's capability to provide more precise predictions for Bitcoin prices. The results of this research are in line with or supported by [14, 30], it has consistently highlighted the strength of LSTM in time series prediction, showcasing its outperformance compared to traditional methods like ARIMA and even other machine learning approaches. Furthermore, while Random Forest has proven adept at handling complex and nonlinear datasets, its performance in predicting Bitcoin prices, as indicated by studies (cite specific research, such as Siti Saadah and Haifa Salsabilla's work), falls behind the precision achieved by LSTM.

Theoretically, LSTM is more accurate than RF for predicting Bitcoin price movement. This can be explained as follows: LSTM is a model specifically designed for sequential data analysis. Therefore, it is more suited for time series datasets like Bitcoin prices, where the order of data points is crucial. LSTM can capture long-term dependencies and patterns in sequences, allowing it to discern intricate trends in the historical price movements of Bitcoin [16]. On the other hand, RF is more powerful in handling diverse datasets, but it may not capture sequential dependencies as effectively as LSTM. RF strength is in aggregating predictions from multiple decision trees, which are individually effective for certain data types but may struggle to capture the temporal aspects of time series data [17]. Another reason why LSTM has better accuracy is because LSTM is a type of recurrent neural network (RNN) that utilizes memory cells to store and retrieve information over long sequences. This architecture allows LSTM to learn complex patterns and adapt to varying trends in the data [18]. On the other hand, RF is more versatile and might not capture the complex, non-linear relationships in time series data as effectively as LSTM. RF works by constructing decision trees based on random subsets of features, which may not capture the patterns inherent in the sequential nature of Bitcoin price movements. LSTM's superiority in predicting Bitcoin price movements can be attributed to its ability to effectively handle sequential data, capture long-term dependencies, and adapt to complex patterns. The architecture and hyperparameter tuning of LSTM makes it well-suited for the dynamic and sequential nature of cryptocurrency price data.

In terms of computational speed and power, there are reasons why Random Forest (RF) may be faster at performing calculations to predict Bitcoin prices compared to Long Short-Term Memory (LSTM). The Random Forest algorithm is based on a collection of relatively simple decision trees [32], whereas LSTM involves a more complex neural network with multiple layers and more complicated connections [33]. Computations in RF usually involve making simple decisions on each tree, which may take less time than LSTMs and have to go through many layers of neurons and more complicated mathematical computations. In memory, the LSTM uses more memory, which can be a factor in computing speed. More memory usage can require extra time to access and manipulate data.

This study's complicated analysis and thorough evaluation led to a profound discovery. In the ever-changing and unpredictable world of cryptocurrency, the ability of advanced data analysis techniques to make predictions is of the utmost importance. The results of comparing the LSTM and Random Forest models not only show how well they can predict but also give a deep understanding of their strengths and weaknesses. The LSTM is a better model for predicting Bitcoin prices because it can figure out sequential patterns and pick up subtle trends. The model's tendency to be close to actual values makes it a useful tool for investors trying to figure out how to trade cryptocurrencies, which is a very complicated process.

The implications of the research results confirm that although LSTM has advantages in prediction accuracy, RF can be an alternative that is more efficient in using computing resources. However, these differences in efficiency depend on the nature of the computational task and the dataset's characteristics. Although LSTM requires more resources, its ability to analyze sequential data and capture long-term patterns in time series data, such as Bitcoin prices, remains a significant plus. The results show how data-driven insights have the power to change things and help investors make better decisions. Through the lens of cryptocurrency prediction, this study goes beyond the limits of traditional financial analysis. It opens the door to better strategies and a deeper understanding of how technology, finance, and the world of digital currencies work together and change over time.

4. CONCLUSION

The results demonstrated that the LSTM model consistently outperformed the Random Forest model regarding price prediction. LSTM demonstrated superior accuracy with a 99.8% accuracy rate, as reflected in its low Mean Absolute Percentage Error (MAPE) of 0.0204. In contrast, the Random Forest model achieved a lower accuracy of 90.1% with a MAPE of 0.9999. However, for calculation time and workload, the comparison between LSTM and RF models shows that the differences are insignificant. In the training stage, LSTM takes 544.3 seconds, 134.1 seconds longer compared to RF, with a CPU load of 74.00% and memory usage of 2676 MB, surpassing RF's load of 70.90% and memory usage of 2465 MB. In the prediction stage, LSTM takes 27.55 seconds, slightly above RF. Visual comparisons confirmed LSTM's ability to closely track actual price trends, making it the preferred choice for Bitcoin price prediction. The novelty of this study is that it highlights the potential of advanced deep learning techniques like LSTM for cryptocurrency market analysis and decision-making. Future research may investigate improvements to predictive models, additional factors influencing cryptocurrency prices, and the evolving relationship between technology and the financial industry.

5. ACKNOWLEDGEMENTS

Thank you to all colleagues who have taken the time to collect data on the research that researchers do, and thank you also to those who have helped in the research process.

6. DECLARATIONS

AUTHOR CONTIBUTION

All four authors contributed significantly to the conception, design, analysis, and interpretation of the research. Munirul Ula and Sariah were involved in data collection, while Veri Ilhadi and Zailani Mohamed Sidek contributed to the data analysis. All authors participated in drafting and revising the manuscript critically for important intellectual content. Each author has approved the final version of the manuscript and takes responsibility for the accuracy and integrity of the work.

FUNDING STATEMENT

This research was self-funded, and the authors did not receive any external financial support for the design, data collection, analysis, or interpretation of the study. All expenses related to this research were borne by the authors personally.

COMPETING INTEREST

The authors declare that there are no competing interests associated with the research presented in this manuscript.

REFERENCES

[1] R. Afrinanda, L. Efrizoni, W. Agustin, and R. Rahmiati, "Hybrid Model for Sentiment Analysis of Bitcoin Prices using Deep Learning Algorithm," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 2, pp. 309–324,

mar 2023.

[2] S. Velankar, S. Valecha, and S. Maji, "Bitcoin price prediction using machine learning," in *2018 20th International Conference on Advanced Communication Technology (ICACT)*. IEEE, feb 2018, pp. 144–147. [Online]. Available: https://ieeexplore.ieee.org/document/8323675/

[3] P. Ciaian, M. Rajcaniova, and d'Artis Kancs, "The economics of BitCoin price formation," *Applied Economics*, vol. 48, no. 19, pp. 1799–1815, apr 2016.

[4] R. Gupta and J. E. Nalavade, "Metaheuristic Assisted Hybrid Classifier for Bitcoin Price Prediction," *Cybernetics and Systems*, vol. 54, no. 7, pp. 1037–1061, oct 2023.

[5] Y. Li and W. Dai, "Bitcoin price forecasting method based on CNNLSTM hybrid neural network model," *The Journal of Engineering*, vol. 2020, no. 13, pp. 344–347, jul 2020.

[6] S. Tandon, S. Tripathi, P. Saraswat, and C. Dabas, "Bitcoin Price Forecasting using LSTM and 10-Fold Cross validation," in *2019 International Conference on Signal Processing and Communication (ICSC)*. IEEE, mar 2019, pp. 323–328.

[7] S. Ji, J. Kim, and H. Im, "A Comparative Study of Bitcoin Price Prediction Using Deep Learning," *Mathematics*, vol. 7, no. 10, p. 898, sep 2019.

[8] H. Kundra, S. Sharma, P. Nancy, and D. Kalyani, "A two level ensemble classification approach to forecast bitcoin prices," *Kybernetes*, vol. 52, no. 11, pp. 5041–5067, nov 2023.

[9] R. Sriwiji and A. H. Primandari, "An Empirical Study in Forecasting Bitcoin Price Using Bayesian Regularization Neural Network," in *Proceedings of the Proceedings of the 1st International Conference on Statistics and Analytics, ICSA 2019, 2-3 August 2019, Bogor, Indonesia*. EAI, 2020, pp. 1–12. [Online]. Available: http://eudl.eu/doi/10.4108/eai.2-8-2019.2290515

[10] Y. Zhu, J. Ma, F. Gu, J. Wang, Z. Li, Y. Zhang, J. Xu, Y. Li, Y. Wang, and X. Yang, "Price Prediction of Bitcoin Based on Adaptive Feature Selection and Model Optimization," *Mathematics*, vol. 11, no. 6, pp. 1–22, mar 2023. [Online]. Available: https://www.mdpi.com/2227-7390/11/6/1335

[11] N. Jagannath, T. Barbulescu, K. M. Sallam, I. Elgendi, A. A. Okon, B. Mcgrath, A. Jamalipour, and K. Munasinghe, "A Self-Adaptive Deep Learning-Based Algorithm for Predictive Analysis of Bitcoin Price," *IEEE Access*, vol. 9, no. February, pp. 34 054–34 066, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9359745/

[12] S. M. Raju and A. M. Tarif, "Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis," *arxiv*, pp. 1–14, jun 2020. [Online]. Available: http://arxiv.org/abs/2006.14473

[13] C. Dinshaw, R. Jain, and S. A. I. Hussain, "Statistical Scrutiny of the Prediction Capability of Different Time Series Machine Learning Models in Forecasting Bitcoin Prices," in *2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA)*. IEEE, oct 2022, pp. 329–336.

[14] H. Jang and J. Lee, "An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information," *IEEE Access*, vol. 6, pp. 5427–5437, 2018.

[15] S. A. Gyamerah, "Are Bitcoins price predictable? Evidence from machine learning techniques using technical indicators," sep 2019. [Online]. Available: http://arxiv.org/abs/1909.01268

[16] W. Riyadi and J. Jasmir, "Performance Prediction of Airport Traffic Using LSTM and CNN-LSTM Models," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 3, pp. 627–638, jul 2023.

[17] Y. Fang, T. Li, and H. Zhao, "Random Forest Model for the House Price Forecasting," in *2022 14th International Conference on Computer Research and Development (ICCRD)*. IEEE, jan 2022, pp. 140–143.

[18] 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, oct 2021.

[19] M. Shin, D. Mohaisen, and J. Kim, "Bitcoin Price Forecasting via Ensemble-based LSTM Deep Learning Networks," in *2021 International Conference on Information Networking (ICOIN)*. IEEE, jan 2021, pp. 603–608.

[20] C. Kaope and Y. Pristyanto, "The Effect of Class Imbalance Handling on Datasets Toward Classification Algorithm Performance," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 2, pp. 227–238, mar 2023.

[21] M. I. C. Rachmatullah, "The Application of Repeated SMOTE for Multi Class Classification on Imbalanced Data," *Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 1, pp. 13–24, 2022. [Online]. Available: https://creativecommons.org/licenses/by-nc-sa/4.0/

[22] S. Hartini, Z. Rustam, G. S. Saragih, and M. J. Segovia Vargas, "Estimating probability of banking crises using random forest," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 2, pp. 407–413, jun 2021. [Online]. Available: http://ijai.iaescore.com/index.php/IJAI/article/view/20880

[23] K. Gawthorpe, "Random Forest as a Model for Czech Forecasting," *Prague Economic Papers*, vol. 30, no. 3, pp. 336–357, jun 2021.

[24] P. Wang, K. Xu, Z. Ding, Y. Du, W. Liu, B. Sun, Z. Zhu, and H. Tang, "An Online Electricity Market Price Forecasting Method Via Random Forest," *IEEE Transactions on Industry Applications*, vol. 58, no. 6, pp. 7013–7021, nov 2022.

[25] Z. Ye, Y. Wu, H. Chen, Y. Pan, and Q. Jiang, "A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin," *Mathematics*, vol. 10, no. 8, pp. 1–21, apr 2022. [Online]. Available: https://www.mdpi.com/2227-7390/10/8/1307

[26] G. Budiprasetyo, M. Hani'ah, and D. Z. Aflah, "Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short-Term Memory (LSTM)," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 8, no. 3, pp. 164–172, jan 2023.

[27] C. Chen, Q. Zhang, M. H. Kashani, C. Jun, S. M. Bateni, S. S. Band, S. S. Dash, and K.-W. Chau, "Forecast of rainfall distribution based on fixed sliding window long short-term memory," *Engineering Applications of Computational Fluid Mechanics*, vol. 16, no. 1, pp. 248–261, dec 2022.

[28] V. W. Siburian and I. E. Mulyana, "Prediksi Harga Ponsel Menggunakan Metode Random Forest," in *Annual Research Seminar (ARS)*, 2019, pp. 144–147. [Online]. Available: https://api.semanticscholar.org/CorpusID:209969741

[29] R. Panggabean, Y. Dewi, and L. Widyasari, "A comparison between Super Vector Regression, Random Forest Regressor, LSTM, and GRU in Forecasting Bitcoin Price," in *Proceeding International Applied Business and Engineering Conference 2022*, 2022, pp. 17–19.

[30] I. Nabillah and I. Ranggadara, "Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut," *JOINS (Journal of Information System)*, vol. 5, no. 2, pp. 250–255, nov 2020.

[31] B. Putro, M. Tanzil Furqon, and S. H. Wijoyo, "Prediksi Jumlah Kebutuhan Pemakaian Air Menggunakan Metode Exponential Smoothing (Studi Kasus : PDAM Kota Malang)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 11, pp. 4679–4686, 2019. [Online]. Available: http://j-ptiik.ub.ac.id

[32] S. R. Polamuri*, D. K. Srinivasi, and D. A. K. Mohan, "Stock Market Prices Prediction using Random Forest and Extra Tree Regression," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 3, pp. 1224–1228, sep 2019.

[33] M. Rafi, Q. A. K. Mirza, M. I. Sohail, M. Aliasghar, A. Aziz, and S. Hameed, "Enhancing Cryptocurrency Price Forecasting Accuracy: A Feature Selection and Weighting Approach With Bi-Directional LSTM and Trend-Preserving Model Bias Correction," *IEEE Access*, vol. 11, pp. 65 700–65 710, 2023.

[This page intentionally left blank.]