Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

Machine Learning Comparative Analysis of SVR Method with RBF Kernel and Random Forest for Bitcoin Price Prediction

Miko Septa Pratama[1,*], Heni Sulistiani[2]

[1] Fakultas Teknik dan Ilmu Komputer, Program Studi Informatika, Universitas Teknokrat Indonesia, Lampung, Indonesia
[2] Fakultas Teknik dan Ilmu Komputer, Magister Ilmu Komputer, Universitas Teknokrat Indonesia, Lampung, Indonesia
Email: [1,*]miko_septa_pratama@teknokrat.ac.id, [2]henisulistiani@teknokrat.ac.id,
Correspondence Author Email: miko_septa_pratama@teknokrat.ac.id
Submitted: **04/02/2026**; Accepted: **05/03/2026**; Published**: 06/03/2026**

Abstract−This study aims to determine how accurate machine learning predictions are for predicting Bitcoin prices using the SVR With RBF Kernel and Random Forest methods. This study was conducted because Bitcoin's volatility is so high that it is difficult to predict. Therefore, this study uses two different methods to allow for a more objective evaluation of model characteristics on volatile data. The dataset was obtained through Kaggle with a Bitcoin price dataset from 2018 to October 2025, totaling 2,856 datasets in CSV format. After training both methods on the same dataset, price prediction results were obtained. Support Vector Regression (SVR) With RBF Kernel achieved a relatively high data evaluation result with an MAE of 10866.882878735294, MSE of 204836847.5591309, and RMSE of 14312.12239883138, while the Random Forest method achieved a low data evaluation result with an MAE of 19342.47, MSE of 659671833.13, and RMSE of 25684.08. The result of these two methods show a significant difference, with Random Forest more closely aligning with the acual data, with a lower evaluation value and producing values closer to the actual data. This research was conducted to determine the accuracy of the Support Vector Regression (SVR) with RBF Kernel and Random Forest algorithms. It is concluded that both methods make good predictions, only the Random Forest method is closer to the actual Bitcoin price.

Keywords: Machine Learning; Cryptocurrency Forecasting; Support Vector Regression; Random Forest; Time Series Analysis

1. INTRODUCTION

The cryptocurrency Bitcoin (BTC) was created to function as money and as a form of payment without being owned by a single person, group, or entity. This eliminates the need for a reliable third party (such as a bank or money printing company) to participate in financial transactions [1]. Machine learning is well suited to predicting Bitcoin prices because it can handle complex, non-linear, and highly volatile data patterns. Many factors, such as market demand, investor sentiment, global news, and trading activity, rapidly affect Bitcoin prices, and its movement patterns are difficult to predict using traditional machine learning methods [2]. Machine learning can study complex historical patterns and discover hidden relationships in data that conventional statistical approaches cannot capture. Furthermore, it can adapt to changing market characteristics over time [3].

Machine learning has become an effective, flexible, and more accurate method for predicting Bitcoin prices compared to traditional methods due to its ability to handle large-scale data and high volatility. Models such as neural networks, LSTMs, and modern regression algorithms have the ability to process large amounts of data, recognise long-term and short-term trends, and automatically adjust parameters to improve prediction accuracy [4]. Since each algorithm has different features, advantages, and disadvantages, no single method works universally well for all types of problems or data sets. The main reasons for this comparative analysis are to find the most accurate model, understand data characteristics, identify overfitting and underfitting, resource efficiency, decision-making based on characteristics, and decision-making based on information [5]. Based on specific evaluation metrics, such as MAE, MSE, RMSE, or prediction accuracy, research can determine which algorithm performs best by comparing various methods [4]. Furthermore, this comparative analysis enhances understanding of model stability, parameter sensitivity, generation capability, and computational time efficiency. Therefore, the comparison of machine learning methods provides a better basis for selecting the most ideal Bitcoin price prediction model more accurately and can be used as a reference for further research or investment decision-making [6].

Previous research has used various algorithms and has not comprehensively compared Support Vector Regression (SVR) with RBF Kernel and Random Forest. However, although both algorithms have different features in capturing non-linear Bitcoin price patterns, using SVR with RBF Kernel offers significant advantages because RBF Kernel can map the data to a larger space. The combination and comparison of these two methods provides new perspectives on which algorithm is superior in the context of current Bitcoin data [7].

In most cases, more robust algorithms, such as LSTM, GRU, or Ensemble Models, are better at learning long-term patterns and non-linear relationships in time series data. Meanwhile, simpler methods, such as linear regression or decision trees, can only find basic patterns and will not be robust enough to handle rapid changes in the crypto market. Furthermore, certain methods may be superior in specific situations due to differences in noise handling, generalisation, and sensitivity to parameters [8].

Although deep learning-based models such as LSTM and GRU have proven superior in learning long-tail dependencies in time series data, their implementation requires large datasets, complex hyperparameter tuning, and higher computational resources. In the context of this study, the Support Vector Regression (SVR) with RBF Kernel and Random Forest methods were chosen because both are capable of handling non-linear data patterns and high

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

volatility with more efficient computational requirements. Support Vector Regression (SVR) is effective in mapping data to a high-dimensional space so it can capture non-linear patterns without requiring too much data, while Random Forest excels in reducing overfitting through ensemble learning mechanisms. In addition, these two methods are easier to implement, more stable on medium-sized datasets, and provide better interpretability compared to deep learning models. Therefore, the use of Support Vector Regression (SVR) and Random Forest in this study is a rational choice to provide a robust performance evaluation, efficiency, and relevance before considering more complex deep learning models.

Although extensive research on Bitcoin price prediction has been conducted, several questions remain unanswered. Most studies focus on deep learning models, while comprehensive comparisons between SVR with RBF Kernel and Random Forest are still rare, even though both algorithms have potential for non-linear and volatile data. In addition, many previous studies used datasets with shorter time spans and without uniform preprocessing standards, resulting in inconsistent algorithm comparison results. This study contributes by conducting a comprehensive comparison between Support Vector Regression (SVR) with RBF Kernel and Random Forest using Bitcoin datasets from 2018 to 2025 that include high volatility phases, sharp uptrends and downtrends, and complex non-linear patterns. Furthermore, this study shows that although deep learning models require large datasets and high computational requirements, SVR and Random Forest methods are still able to provide competitive predictions with better computational efficiency and a higher level of interpretability.

2. RESEARCH METHODOLOGY

2.1 Research Stages

This method has research stages that show how the Bitcoin price prediction process using the SVR method with RBF Kernel and Random Forest works. Figure 1 below illustrates this:

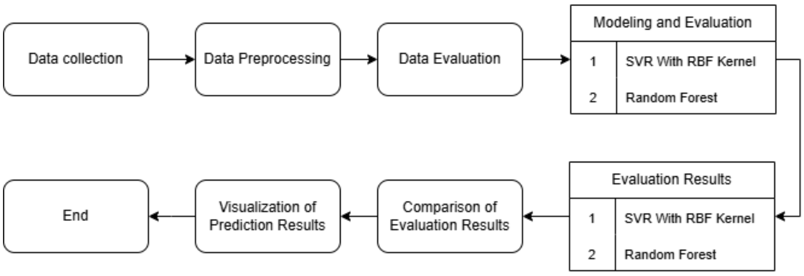

Figure 1. Research Stages

Figure 1 presents the research flowchart, outlining the steps in predicting Bitcoin prices using SVR with the RBF kernel and Random Forest. The process begins with collecting Bitcoin price data from Kaggle (2018–2025), followed by data preprocessing, including cleaning invalid entries, standardizing formats, converting dates into time series, and generating features such as lag. After preprocessing, the dataset is evaluated to identify patterns, trends, and volatility. The modeling stage then trains both SVR and Random Forest using the same dataset, and their performances are assessed using MAE, MSE, and RMSE. The results are compared to determine which model best approximates the actual data. Finally, the prediction outputs are visualized through graphs of 30-day forecasts and comparisons between actual and predicted values, concluding with an analysis of model performance and the recommendation of the most suitable model for Bitcoin price prediction.

2.2 Data Collection

The Bitcoin dataset from 2018–2025 includes attributes such as open time, open, high, low, close, volume, quote asset volume, number of transactions, and taker buy base asset volume. All attributes were used because each contributes to describing crypto market dynamics, such as daily volatility (high–low), market sentiment through trading activity (volume and number of transactions), and buying pressure through taker buy volume. To prevent bias due to relationships between variables, researchers conducted a correlation analysis on all features. Although some features showed high correlation, this was not a significant problem because the SVR and Random Forest models are relatively tolerant of multicollinearity. Furthermore, all features were normalized to the same scale, allowing the models to process the data more stably. Thus, using all attributes provides more comprehensive information without compromising the quality of predictions. This dataset was taken from Kaggle, where the dataset is updated daily. Here is the link to the dataset from Kaggle that was taken from: https://www.kaggle.com/datasets/novandraanugrah/bitcoin-historical-datasets-2018-2024

2.3 Data preprocessing

Preprocessing data is an important process that aims to clean, reformat, and prepare data so that it is easier and more accurate to analyse [9]. The initial data consisted of 2,856 datasets with attributes such as opening time, closing time,

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

high, low, closing time, volume, closing time, asset price, number of trades, and number of assets purchased by buyers. This data will be sorted to improve data quality for more accurate analysis and decision making. During pre-processing, data with missing values and duplicate values were removed from the dataset. This process involved checking the number of rows, columns, data types, and finding the location of missing values and duplicate values in the initial data before performing missing value imputation, which amounted to 2,856, and the result after performing missing value imputation amounted to 2,856 datasets. In this study, missing and duplicate values were removed from the dataset to ensure consistent data quality before entering the model training stage. In time series research, imputation methods such as Linear Interpolation are often recommended because they maintain the continuity of data patterns through a straight-line approximation between two observation points. However, this method was not applied in this study because after the cleaning process, there were no missing values in the dataset, making imputation unnecessary. In this study, MAE, MSE, and MAPE serve as evaluation metrics to measure model performance.

Preprocessing is performed to clean and prepare the data, including primary type extraction, label encoding, feature selection, and feature standardization [10]. Since data quality greatly affects the performance and accuracy of machine learning models, data must undergo pre-processing before being used in Bitcoin price prediction models. Bitcoin price data typically contains unstructured data patterns, noise, missing values, scale differences, and outliers [8]. Models can produce biased, unstable, or inaccurate predictions if these conditions are not corrected first. Models can better understand patterns through preprocessing processes such as normalization, data cleaning, outlier removal, missing value interpolation, and time series data transformation. Preprocessing is also necessary to improve computational efficiency, simplify the training process, and help the algorithm better capture relationships between variables [2].

In this study, the data were split using the Temporal Split method to prevent data leakage. Unlike random splits, this approach ensures the model learns only from past data. Bitcoin price data from January 2018 to September 2025 served as training data, while the last 30 days of October 2025 were used for testing. This setup aligns with the goal of forecasting the next 30 days based on long-term historical patterns. It also ensures that the testing data represents unseen future conditions, allowing for a more reliable model evaluation and realistic forecasting performance.

2.4 Data Evaluation

Data evaluation was conducted to measure the performance of the model in predicting Bitcoin prices using the SVR with RBF Kernel and Random Forest methods. At this stage, the model is tested using three main evaluation metrics, namely Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). These three metrics are widely used in machine learning research because they are able to comprehensively describe the level of prediction error [11].

The MAE metric is used to average the absolute value of the difference between the actual value (y_i) and the predicted value ($\hat{y}_i$). This metric provides a direct indication of the magnitude of the error without considering the direction of the error.The MAE formula is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

Furthermore, MSE is used to measure the mean square error between actual values (y_i) and predicted values ($\hat{y}_i$). MSE imposes a greater penalty on errors with large differences, making it effective in detecting extreme prediction deviations [11]. The MSE formula is written as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i))^2 \tag{2}$$

The third metric, RMSE, is the square root of the MSE. RMSE provides evaluation results in the same units as the original data, making it easier to interpret in the context of price values. By taking the square root of the MSE value, RMSE provides easier interpretation than MSE, as its value returns to the same units as the target data. This metric describes the overall error of the model. The RMSE formula is:

$$RMSE = \sqrt{MSE} \tag{3}$$

Overall, the evaluation using MAE, MSE, and RMSE provides a clear picture of the strengths and weaknesses of each model, and serves as the main basis for the performance comparison process in this study.

3. RESULT AND DISCUSSION

3.1 Method SVR With RBF Kernel

SVR with RBF Kernel is an SVM component used for classification. This component is highly sought after for various applications due to its ability to handle non-linear data and control overfitting [11]. However, with fluctuating and non-stationary Bitcoin price data, SVR often struggles to produce accurate predictions. One of the main reasons is the characteristics of the RBF kernel, which is heavily influenced by the C, epsilon, and gamma parameters. If these

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

parameters are not optimal, the model can experience overfitting or underfitting. In addition, the highly volatile nature of Bitcoin data and its lack of consistent patterns make it difficult for SVR to establish a stable margin. SVR is more suitable for data with relatively stable patterns that do not experience extreme changes in a short period of time. In this case, the non-stationary nature of Bitcoin makes it difficult for the model to find the optimal prediction hyperplane [12].

The SVR method has been widely used in previous studies. The Grid Search algorithm using a linear kernel showed an accuracy of 92.47% for training data and 83.39% for testing data. Furthermore, Crude Palm Oil (CPO) forecasting produced an accuracy of 98.71738% for training and 83.45659% for testing [13]. To ensure the accuracy of predictions, the quality of the dataset is thoroughly checked at the beginning of the analysis process. Key features (open, high, low, close, volume) and time attributes are checked to detect missing values. Based on the verification results, no missing data was found, so all rows were retained. With the fulfillment of the time series modeling completeness standard, the amount of data after the pre-processing stage remained at 2,856 rows. This was done to clean up the dataset so that the resulting data would be more optimal and accurate [9].

Furthermore, in the context of time series podcasting, the temporal order of the data is also examined to ensure continuity. All features are then normalized using the Min–Max Scaling method to prevent differences in value distances (such as price and volume) from biasing the kernel calculation process. This normalization is crucial for improving the stability and performance of the SVR on non-linear Bitcoin data.

The gamma parameter in RBF determines how far the influence of a single data point extends. An excessively high gamma causes the model to be highly sensitive to minor changes in the data, leading to overfitting, especially in volatile datasets such as Bitcoin. Conversely, an excessively low gamma prevents the model from capturing the short-term patterns that characterize Bitcoin price movements [14]. The C and epsilon parameters also play a significant role. Too large a C value forces the model to be small but sacrifices generalizability, while an inappropriate epsilon value hinders the model's ability to capture the details of price changes [14].

The next step in the modeling process is to perform hyperparameter tuning using GridSearchCV to obtain the best parameter configuration. This process is performed using GridSearchCV, testing various combinations of values for the C, epsilon, and gamma parameters. These parameters play a crucial role, as described in Table 1 below:

Table 1. The Role of Each Parameter

C	Controlling model complexity
Epsilon	Determine the error tolerance limit
Gamma	Determining the influence of data on the shape of the model curve

After going through the parameter exploration process in Table 1, the best combination was obtained as shown in Table 2 below:

Table 2. Best Parameter Results Obtained

C : 10	Epsilon : 0.2	Gamma : 1

Table 2 selects this combination because it yields the lowest error value compared to other configurations. These optimal parameters indicate that the model requires a moderate level of complexity (C=10), a small error tolerance (epsilon=0.2), and moderate sensitivity to data points (gamma=1). This configuration aligns with the volatile nature of Bitcoin prices. After running the parameters, the resulting graph is shown in Figure 2 below:

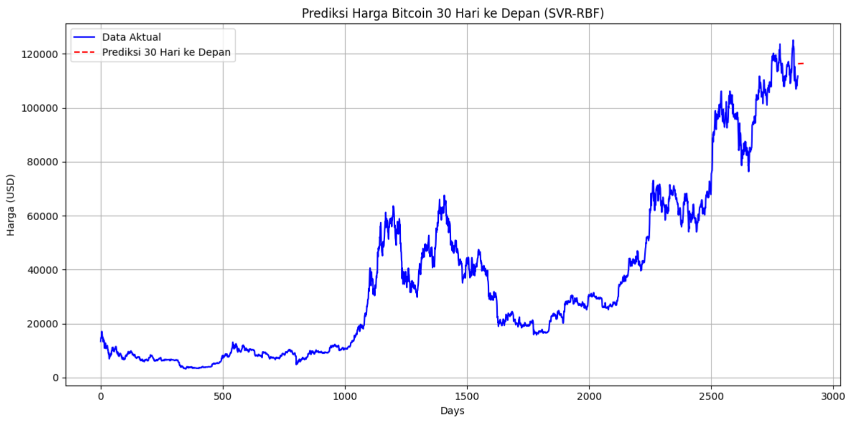

Figure 2. Prediction Results For The Next 30 Days

Figure 2 shows that the 30-day predicted trend graph shows that Bitcoin prices tend to fluctuate with an unstable up-and-down pattern. The SVR with RBF Kernel model appears to produce a smoother prediction curve and does not fully follow actual movements. This suggests that the SVR with RBF Kernel tends to maintain a stable separator function, making it less responsive to sharp price changes. At some points, especially when there are drastic price

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

increases or decreases, SVR appears to fail to keep up with the trend. The graph shows several segments where the SVR with RBF Kernel prediction line lags significantly behind the actual data. This condition can be attributed to the highly non-stationary nature of Bitcoin data, where linear kernel-based models such as SVR with RBF Kernel are not always able to learn the dynamics of rapid changes [15].

Nevertheless, the model is still able to capture the general trend direction of price movements, although its sensitivity to sudden spikes remains limited due to the smoothing effect of the RBF Kernel. The actual data and future predictions have been combined into a single graph to clearly demonstrate how the SVR with the RBF Kernel predicts a Bitcoin price of $116,399.41 for the next 30 days. This results in the evaluation values shown in Table 3 below:

Table 3. SVR Evaluation Value with RBF Kernel

MAE	10866.882878735294
MSE	204836847.5591309
RMSE	14312.12239883138

The results in Table 3 show that while the SVR with RBF Kernel is capable of capturing Bitcoin price movement patterns, its error rate is still relatively high. Pattern analysis shows that the SVR with RBF Kernel predictions tend to follow the price trend, but there are significant deviations from the actual values, reducing the model's accuracy.

3.2 Method Random Forest

In computational statistics, Random Forest is a classification method used to analyze various functions that map each selected data point to one of the predefined class categories. The Random Forest method is chosen because it produces lower errors, provides high classification accuracy, can handle very large amounts of data, and is effective in handling irregular data [16]. Random Forest generally provides more stable performance on complex, non-linear, and noisy datasets, such as Bitcoin prices. This is due to its ensemble nature, which combines many decision trees. Each tree learns a different part of the dataset, so that overall the model is able to capture more pattern variations [17].

Another advantage of Random Forest is its ability to follow short-term patterns through feature engineering, such as lag, moving average, or return. The collection of trees in Random Forest is able to learn the interactions between these features without requiring specific mathematical assumptions, unlike SVR, which depends on kernel selection [18]. In addition, Random Forest has proven to be robust against noise, where extreme fluctuations or price spikes do not significantly affect the overall prediction results due to the bootstrap aggregation (bagging) process. This makes Random Forest very suitable for volatile data such as Bitcoin, where price anomalies often occur [4]. The non-parametric nature of Random Forest makes it more flexible in modeling complex non-linear relationships without requiring sensitive parameter tuning as in SVR. This causes Random Forest prediction results to be closer to actual values and have a lower error rate.

In the Random Forest method, the steps for predicting prices are the same, but there are significant differences in the results. Previous research comparing the Random Forest and Decision Tree algorithms for predicting the success of immunotherapy yielded an accuracy of 85.5% [19], Then, there is also a study on Hybrid Machine Learning Models to predict heart disease using the Logistic Regression and Random Forest methods with an accuracy rate of 83.16% [20].

In the initial stage, data preprocessing was performed, which included checking all columns for missing values. The results were consistent: the dataset contained no missing values or duplicate data, so all data, totaling 2,856 rows, could be used without modification. In the Random Forest method, there is a crucial step called Feature Engineering, which aims to enrich the information available to the model. This step includes:

a. The date column becomes a year, month, day, and day of the week feature.
b. Lag features such as H-1, H-2, and H-3 prices
c. Moving average value

In this study, feature engineering, including lag and moving average features, was applied specifically to the Random Forest model. This approach was chosen because decision tree-based algorithms lack the inherent ability to understand time dependencies, requiring additional features to learn historical patterns more effectively. Meanwhile, the SVR model did not implement lag and moving average features. This difference is part of the research design, where each algorithm is optimized according to its own characteristics. Random Forest requires explicit feature engineering to capture time series patterns, whereas SVR with an RBF kernel is able to learn non-linear relationships between variables without adding lag features due to the kernel's ability to map data to a high-dimensional space. Therefore, the comparison of the two models remains scientifically valid, as each method was processed using the most appropriate approach to achieve optimal performance. The next step in the modeling process is hyperparameter tuning using GridSearch to find the best parameter configuration.This process is performed using GridSearch, testing various combinations of values for the parameters n_estimators, max_depth, max_features, and min_samples_split. These parameters play a crucial role, as explained in Table 4 below:

Table 4. Random Forest parameter explanation

n_estimators	This parameter determines the number of decision trees to be built in the forest.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

max_depth	This parameter limits the maximum depth of each tree in the forest.
max_features	This parameter determines the number of features (input features) that each tree randomly considers when splitting a node.
min_samples_split	This parameter determines the minimum number of samples required for an internal node (tree branch) to be split/split again.

After the parameter exploration process in Table 4, the best parameter configuration was tested with the parameter values in Table 5 below:

Table 5. Random Forest Optimal Value Parameters

n_estimators	200
max_depth	15
max_features	1
min_samples_split	2

Based on the optimal parameters shown in Table 5, the Random Forest model achieved its best performance using 200 estimators, a maximum depth of 15, one feature considered at each split, and a minimum split size of 2. These parameters indicate that the model benefits from a relatively deep tree structure and a sufficiently large ensemble size, thus capturing complex patterns in Bitcoin price movements while maintaining stability. The low min_samples_split and max_features values indicate that the model performs well when given greater flexibility during the splitting process, thus being able to learn detailed variations in the data without overfitting. After running the parameters, the prediction graph shown in Figure 3 below appears:

Figure 3. 30-Day Price Prediction Chart

Figure 3 shows the graph generated by the Random Forest prediction model for the next 30 days. Random Forest is able to follow the pattern of spikes and declines more accurately. At extreme points, the model's response is more accurate. At extreme points, the model's response is closer to the actual value, demonstrating resilience to noise and the ability to capture complex non-linear patterns. This is consistent with the Random Forest's nature of combining multiple decision trees, enabling it to learn from highly variable data. Chart analysis also revealed outliers, or nutrient movements that deviated from the general pattern. SVR with RBF Kernel tended to flatten the area, resulting in significant prediction errors, while Random Forest was more flexible in adapting to these changes. This finding further confirms that ensemble models are more suitable for volatile datasets like Bitcoin. This model predicted a price of USD 71,620.67 on day 30, which is visually and numerically closer to the actual price than SVR with RBF Kernel. The results of the evaluation are shown in Table 6 below:

Table 6. Random Forest evaluation score

MAE	19342.47
MSE	659671833.13
RMSE	25684.08

The results in Table 6 show that although the MSE and RMSE values for Random Forest are higher than those for SVR with RBF Kernel, the MAE value is lower in the final prediction context, making the prediction results more relevant and closer to the actual value. This is in line with the characteristics of Random Forest, which is stable, robust to noise, and able to capture interactions between features well.

3.3 Model Comparison Analysis

Model comparison analysis is a method for evaluating the performance of different models by comparing their features, performance, or results to determine which model is superior [21]. The performance comparison between SVR with RBF Kernel and Random Forest in this study can be explained through the theory of the basic characteristics

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

of both models and the general theory of time series modeling. Theoretically, SVR with RBF Kernel works optimally on data with structured patterns and relatively low noise levels. SVR with RBF Kernel uses the concepts of margin and kernel trick to find the best mapping function that minimizes prediction error [22].

On datasets with high stability, SVR with RBF Kernel can produce precise models because the decision margin is relatively consistent. However, this theory doesn't fully apply to crypto data like Bitcoin, which has extreme volatility and high noise. Kernel-based models tend to struggle to predict non-stationary and fluctuating data. [23], thus explaining why SVR with RBR Kernel in this study was less accurate. On the other hand, Random Forest is theoretically a more stable model and resistant to noise because it uses an ensemble learning approach through bagging and the formation of multiple decision trees. The combination of hundreds of trees produces a decision distribution that is more robust to extreme changes in data values. Research results show that Random Forest is able to follow the fast and irregular trends of Bitcoin. Ensemble models tend to be more accurate for non-linear and rapidly fluctuating financial data because they are able to learn various short-term patterns in parallel [23].

Based on time series theory, time series data often contains autocorrelation, which is the relationship between current values and previous values. Therefore, the use of features such as lag and moving average is very important to help the model understand the dynamics of price movements [24]. Lag helps models recognize short-term patterns, such as price reactions over the previous 1-7 days, while moving averages smooth out fluctuations so that trend patterns are clearer. Random Forest is naturally capable of handling such correlations because decision trees can break down lag features non-linearly, whereas SVR with RBF Kernel can only utilize them if the kernel parameters are truly optimal. Thus, theoretically and empirically, Random Forest is more suitable for Bitcoin datasets that are non-stationary, highly volatile, and non-linear, while SVR with RBF Kernel is more suitable for datasets with stable patterns and minimal noise. The results of both methods produce data evaluations as shown in Table 7 below:

Table 7. Comparison of Method Performance

Metode	MSE	MAE	RMSE
SVR With RBF Kernel	204836847.5591309	10866.882878735294	14312.12239883138
Random Forest	659671833.13	19342.47	25684.08

The comparison results shown in table 7 show the performance of the two Support Vector Regression (SVR) with RBF Kernel and Random Forest show that the Support Vector Regression (SVR) with RBF Kernel method is the most superior method by producing an MSE of 204836847.5591309, MAE of 10866.882878735294, and RMSE of 14312.12239883138. Support Vector Regression (SVR) with RBF Kernel produces more accurate and stable data modeling values, as indicated by lower prediction error values. The method Random Forest produces data evaluation on MAE of 19342.47, MSE of 659671833.13, and RMSE of 25684.08, which shows less than optimal performance in predicting data because it produces a relatively high error value compared to Support Vector Regression (SVR) with RBF Kernel so that this model comparison shows that the Random Forest method is less than optimal for several datasets tested.

4. CONCLUSION

The results of the analysis of Bitcoin price predictions for the next 30 days with a dataset from January 1, 2018 to October 25, 2025, based on the research results, it can be concluded that the SVR model with RBF Kernel shows better performance than Random Forest. This is shown by lower prediction error values in all evaluation metrics, namely MAE, MSE, and RMSE. The SVR model obtained an MAE value of 10866.882878735294, MSE of 204836847.5591309, and RMSE of 14312.12239883138, which indicates that SVR is able to produce predictions that are more stable and closer to the actual data. In contrast, Random Forest produced an MAE of 19,342.47, an MSE of 6,5967,1833.13, and an RMSE of 25,684.08, indicating that this model produces higher errors, primarily due to large errors at some extreme points. Although Random Forest theoretically excels at modeling nonlinear data and is robust to noise, the results of this study indicate that the highly volatile and non-stationary characteristics of Bitcoin data are more suited to the kernel regression approach of SVR. Thus, SVR with RBF Kernel can be declared a more optimal model for predicting Bitcoin prices on the dataset used in this study. These findings also demonstrate that the theoretical superiority of an algorithm does not necessarily guarantee better actual performance.

REFERENCES

[1] B. P. Guide, “Why Bitcoin ’ s Price Is So Volatile : Key Factors Explained Understanding Bitcoin ’ s Supply and Demand Dynamics The Influence of Investor Actions on Bitcoin Prices,” *investopedia*, pp. 1–12, [Online]. Available: https://www.investopedia.com/ask/answers/100314/why-do-bitcoins-have-value.asp

[2] M. K. Anam and D. A. Jakaria, “Sistem Prediksi Harga Kripto Dengan Metode Regresi,” *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 2, 2023, [Online]. Available: jurnal.mdp.ac.id

[3] D. F. Sulianta, “Dasar dan Konsep Machine Learning,” 2025

[4] S. A. Khoiri and A. Wahid, “Jurnal Sistem dan Teknologi Informasi Indonesia Analisis Kinerja Algoritma Machine Learning dalam Prediksi Harga Cryptocurrency Performance Analysis of Machine Learning Algorithms in Cryptocurrency Price Prediction,” *J. Sist. dan Teknol. Inf. Indones.*, vol. 9, no. 2, pp. 133–141, 2024, doi: 10.32528/justindo.v9i2.1965 Analisis.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2451−2458
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9348

[5] A. Wantoro *et al.*, “Evaluasi Kinerja Algoritma Machine Learning (ML),” *J. Inform. Polinema*, vol. 11, no. 3, pp. 311–316, 2025, [Online]. Available: https://www.researchgate.net/profile/Syazili-Mustofa/publication/392315470.
[6] V. W. Purwanto Simamora, Sutrisno Arianto Pasaribu, “Tampilan Peningkatan dan Optimalisasi Prediksi Harga Emas Menggunakan Metode Combine Machine Learning Random Forest dan Gradient Boosting,” *J. Mahkota Inform.*, vol. 1, no. 1, [Online]. Available: https://mojs.mtu.ac.id/index.php/mi
[7] I. P. Putri and N. Arminarahmah, “Comparative Analysis of Machine Learning Algorithms for Predicting Child Stunting Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak,” *Inst. Ris. dan Publ. Indones.*, vol. 4, no. 1, pp. 257–265, doi: 10.57152/malcom.v4il.1078.
[8] A. Bouteska, M. Z. Abedin, P. Hajek, and K. Yuan, “International Review of Financial Analysis Cryptocurrency price forecasting – A comparative analysis of ensemble learning and deep learning methods,” *Int. Rev. Financ. Anal.*, vol. 92, no. 1, doi: 10.1016/j.irfa.2023.103055.
[9] A. Agung, A. Daniswara, and I. K. D. Nuryana, “Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru,” *J. Informatics Comput. Sci.*, vol. 5, no. 1, pp. 97–100, 2023, doi: 10.26740/jinacs.v5n01.p97-100.
[10] N. Hidayanah, Y. Simarmata, I. W. Permadani, and N. K. Destianto, “Klasifikasi Tipe Pokemon Berdasarkan Statistik Tempur Menggunakan Algoritma Random Forest,” *J. Comput. Informatics Res.*, vol. 4, no. 3, pp. 362–367, doi: 10.47065/comforch.v4i3.2083.
[11] I. F. Riziq, A. R. Dzikrillah, and T. Informatika, “Impelementasi Algoritma LSTM Dan SVR Untuk Prediksi Harga Bitcoin Menggunakan Data Yahoo Finance,” *METIK J.*, vol. 9, no. 2, pp. 284–291, 2025, doi: 10.47002/metik.v9i2.1077.
[12] J. Khatib and S. Dalam, “Indonesian Journal of Computer Science,” *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 407–419, 2024, doi: 10.33022/ijcs.v13i1.3722.
[13] R. A. Maulana, I. Permana, and F. N. Salisah, “Comparison of the Support Vector Regression Kernel Algorithm on the Performance of Palm Production Prediction Perbandingan Kernel Algoritma Support Vector Regression Terhadap Performa Prediksi Produksi Kelapa Sawit,” *Inst. Ris. dan Publ. Indones.*, vol. 5, no. 1, pp. 405–413, doi: 10.57152/malcom.v5i1.1410.
[14] Y. Zhao, W. Zhang, and X. Liu, “Grid search with a weighted error function : Hyper-parameter optimization for financial time series forecasting,” *Appl. Soft Comput.*, vol. 154, 2024, doi: 10.1016/j.asoc.2024.111362.
[15] M. Er, K. Bayaz, S. Ümit, and O. Fırat, “SVR-Based Cryptocurrency Price Prediction Using a Hybrid FISA-Rao and Firefly Algorithm for Feature and Hyperparameter Selection,” *MDPI*, vol. 15, no. 24, 2025, doi: 10.3390/app152413177.
[16] S. Amaliah and M. Nusrang, “Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi Di Kedai Kopi Konijiwa Bantaeng,” *J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 2, pp. 121–127, 2022, doi: 10.35580/variansiunm31.
[17] L. Htet, “Random Forest Analysis for Key Factors in Bitcoin Price Prediction,” *J. Sist. Inf. dan Komput. Terap. Indones.*, vol. 6, no. 2, pp. 430–439, 2023
[18] V. Saha, “Predicting Future Cryptocurrency Prices Using Machine Learning Algorithms,” *J. Data Anal. Inf. Process.*, vol. 11, pp. 400–419, 2023, doi: 10.4236/jdaip.2023.114021.
[19] V. P. R. Fandi Yulian Pamuji, “View of Komparasi Algoritma Random Forest dan Decision Tree untuk Memprediksi Keberhasilan Immunotheraphy,” *J. Teknol. dan Manaj. Inform.*, vol. 7, no. 1, 2021
[20] N. S. S. Silmi Ath Thahirah Al Azhima, Dwicky Darmawan, Nurul Fahmi Arief Hakim, Iwan Kustiawan, Mariya al Qibtiya, “Tampilan Hybrid Machine Learning Model untuk memprediksi Penyakit Jantung dengan Metode Logistic Regression dan Random Forest,” *J. Teknol. Terpadu*, vol. 8, no. 1, 2022
[21] D. R. A. P. Muchamad Taufiq Anwar, “View of Perbandingan Performa Model Data Mining untuk Prediksi Dropout Mahasiwa,” *J. Teknol. dan Manaj.*, vol. 19, no. 2, 2021, doi: 10.52330/jtm.v19i2.34.
[22] S. Reddy and G. Nellore, “Improving Cryptocurrency Price Prediction Accuracy with Multi- Kernel Support Vector Regression Approach,” *Asian Res. Assoc.*, vol. 6, no. 4, pp. 20–31, 2024, doi: 10.54392/irjmt2443.
[23] G. Dudek, P. Fiszeder, and W. Orzeszko, “Forecasting Cryptocurrencies Volatility Using Statistical And Machine Learning Methods : A Comparative Study,” *ELSEVIER*, vol. 151, 2024, doi: 10.1016/j.asoc.2023.111132.
[24] X. Yang, “Machine Learning in Financial Time Series Forecasting : A Systematic Review,” Journal of Word Economy, vol. 4, no. 4, 2025, doi: 10.56397/JWE.2025.08.14.