

Research Article

Two-Stage Hybrid Machine Learning Model for High-Frequency Intraday Bitcoin Price Prediction Based on Technical Indicators, Variational Mode Decomposition, and Support Vector Regression

Samuel Asante Gyamerah ^{1,2}

¹Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

²Laboratory for Interdisciplinary Statistical Analysis, Kwame Nkrumah University of Science and Technology (KNUST-LISA), Kumasi, Ghana

Correspondence should be addressed to Samuel Asante Gyamerah; asante.gyamerah@knust.edu.gh

Received 26 July 2021; Revised 13 October 2021; Accepted 20 November 2021; Published 7 December 2021

Academic Editor: Ning Cai

Copyright © 2021 Samuel Asante Gyamerah. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the inherent chaotic and fractal dynamics in the price series of Bitcoin, this paper proposes a two-stage Bitcoin price prediction model by combining the advantage of variational mode decomposition (VMD) and technical analysis. VMD eliminates the noise signals and stochastic volatility in the price data by decomposing the data into variational mode functions, while technical analysis uses statistical trends obtained from past trading activity and price changes to construct technical indicators. The support vector regression (SVR) accepts input from a hybrid of technical indicators (TI) and reconstructed variational mode functions (rVMF). The model is trained, validated, and tested in a period characterized by unprecedented economic turmoil due to the COVID-19 pandemic, allowing the evaluation of the model in the presence of the pandemic. The constructed hybrid model outperforms the single SVR model that uses only TI and rVMF as features. The ability to predict a minute intraday Bitcoin price has a huge propensity to reduce investors' exposure to risk and provides better assurances of annualized returns.

1. Introduction

Bitcoin, which is considered the largest cryptocurrency with a market capitalization of about \$125 billion [1], has experienced its largest-ever Bitcoin inflows and also seen significant plunges in value during the COVID-19 pandemic period. This has caused an unstable intraday price leading to price uncertainties, threatening its potential to be used as currency, and thus regarded as a highly volatile digital currency [2]. Different factors contribute to the volatility in the Bitcoin price, which includes the small market size of Bitcoin trading in contrast to conventional financial assets such as stocks, fiat currencies, and bonds; unmonitored mining activity; news events; its availability to trade 24/7; low liquidity which increases price fluctuations; shifting sentiment; decentralized and high speculations. This high price volatility makes it difficult to efficiently predict its price. Even more, there are structural changes in the price of

Bitcoin as a result of the effect of COVID-19. The article "What is going on with the Bitcoin Market" published on the website of Chainalysis [3] indicates the response of Bitcoin to the COVID-19 pandemic. From March 9, 2020, Bitcoin exchange markets have cumulatively received 1.1 million Bitcoin over an eight-day period, which peaked at 319,000 Bitcoin on March 13, 2020. This significantly differed from the average of 52,000 Bitcoins per day before March 9, 2020. Also, the daily average amount of Bitcoin that was sent to different Bitcoin exchange markets to be sold within March 12, 2020, to March 13, 2020, also increased by nine times. This selling pressure caused the price of Bitcoin to reduce to approximately 37%.

Bitcoin market is an inefficient market; hence, the market does not incorporate all available information to determine a fair price for Bitcoin [4–6]. Reference [4] concluded that Bitcoin returns do not satisfy the efficient market hypothesis. Using data at different frequencies,

overlapping and nonoverlapping window analysis, [5] examined the dynamics of informational efficiency of Bitcoin and concluded that the Bitcoin market is an inefficient market. Reference [6] supported this claim by investigating the efficiency of the top 31 cryptocurrencies by market capitalization. This suggests that it is possible to uncover price predictability based on historical information.

Three types of time-series prediction models have been proposed in the literature: statistical models, artificial intelligence models, and hybrid models. In the past decades, researchers have used conventional statistical models such as autoregressive (AR) models [7], autoregressive moving average (ARMA) [8], autoregressive integrated moving average (ARIMA) [9], and multivariate linear regression [10] for forecasting the price of Bitcoin. Statistical models are not appropriate for chaotic systems (such as the cryptocurrency market) with many uncertainties because they require the time-series data to be subject to specific a priori assumptions, such as stationarity [11]. However, Bitcoin price series are nonstationary and nonlinear [12]. As stated in [13], conventional forecasting models such as regression models can hardly capture nonlinear dynamics in most time-series data. These imply that there are inherent properties of nonstationarity and nonlinearity in some of the technical indicators that are used as features for prediction. As such, predicting the price of Bitcoin using statistical models is bound to large errors. Time-series data with these stylized facts can be effectively explored using nonlinear models such as machine learning models and decomposition techniques. Machine learning models, which are a subset of artificial intelligence, have received a lot of attention as a result of the advancement of computational intelligence. They are data-driven methods that do not rely on any assumptions made a priori and so have diverse applications. Support vector regression (SVR), a type of machine learning model, has become a popular algorithm for different forecasting problems because of its strong nonlinear learning capability [14, 15]. But, the performance of machine learning models is highly dependent on the nature and characteristics of data [16] which makes it difficult for machine learning models to deeply mine the inner characteristics of the Bitcoin series. Also, due to the nonlinear dynamics of the Bitcoin series, which include its inherent fractality and chaoticity, single-stage prediction models are not sufficient to forecast the price of Bitcoin with very high accuracy.

Reference [14] proposed an SVR model based on empirical mode decomposition (EMD) and AR for forecasting electric load. Using unbalanced data, the proposed hybrid model outperformed the original SVR model. By employing moving average technical indicators as input for a multilayer perceptron-based nonlinear autoregressive with exogenous inputs, [17] predicted the price of Bitcoin. Reference [12] predicted intraday Bitcoin price series using the ensemble model of VMD and generalized additive model (GAM). The VMD-GAM model was compared to an ensemble model of EMD and GAM. The author concluded that VMD-GAM performed better than EMD-GAM. As noted by [18], VMD performs better than other signal decomposition tools because it eliminates modal aliasing and noise robustness.

Reference [19] constructed a hybrid forecasting model for stock price indices using VMD and Gated Recurrent Units (GRU) network. The constructed model outperformed the single models using VMD or GRU. Reference [20] concluded in their studies that the proposed hybrid model (EMD-SVM) outperformed the individual forecasting model. Clearly, in chaotic time-series prediction, it has been proven that hybrid models outperform their single counterparts.

Further, [21] indicated that the price of Bitcoin is heavily fueled by market sentiments and momentum instead of underlying economic fundamentals. Hence, we employ technical indicators as features for predicting the one-minute intraday Bitcoin price. The benefit of technical indicators as features is because they require only historical data of Bitcoin price and do not depend on any economic fundamentals. Reference [22] constructed a classification tree-based model for Bitcoin return prediction using 124 technical indicators. They concluded that big data and technical analysis are efficient in predicting Bitcoin returns. More recently, [23] explored the suitability of neural networks with a convolutional component as a classification model for six popular cryptocurrencies (Bitcoin, Dash, Ether, Litecoin, Monero, and Ripple) based on technical indicators. The results indicate the suitability of technical indicators as inputs for predicting cryptocurrencies.

A new framework based on a decomposition algorithm (variational mode decomposition) that is able to capture the chaotic nature of the Bitcoin series has been proposed in the literature for time-series prediction. In variational mode decomposition (VMD), selecting an optimal model is contingent on the model's potential to sample the fundamental dynamics (variational modes) from the original series and the intensity of noise it carries [24]. Hence, Bitcoin price is decomposed into stationary variational mode function's (VMFs) components. These VMFs are reconstructed into new series, which give a better forecasting performance than the original VMFs. The reconstructed series are combined with the high-dimensional technical indicators as features for SVR predicting model. This hybrid model will help increase forecasting accuracy.

As one of the most volatile assets, is Bitcoin price predictable out-of-sample in the midst of the COVID-19 pandemic? Will a hybrid of technical indicators and VMF series provide an optimal feature for high-frequency intraday Bitcoin price prediction during this COVID-19 pandemic? Given different features as such a distinct training dataset, how do we measure the generalization performance of a support vector model trained to the data? These are the questions we seek to answer in this paper. The contribution of this paper is fourfold: (1) defining a new performance metric to evaluate the effectiveness of the reconstructed VMF in selecting an optimal mode (K) value called signal average absolute difference (SAAD), (2) evaluating the predictability of intraday price of Bitcoin out-of-sample in the midst of COVID-19 pandemic by using a hybrid of technical indicators (TI) and variational mode functions (rVMF) as features for SVR prediction model, (3) evaluating and comparing the predictive performance of two

features (TI and rVMF) to the hybrid model in the midst of COVID-19, and (4) adding to scarce empirical evidence of hybrid model using SVR, TI, and rVMF in predicting one-minute intraday Bitcoin price.

The rest of the paper is organized as follows: Section 2 provides the materials and methods for constructing the predictive models data, experimental results, and discussion of the study are presented in Section 3; and the conclusion is outlined in Section 4.

2. Materials and Methods

In this section, the theoretical concepts for the implementation of the SVR-TI-rVMF prediction model are described in detail. The methodology used for the proposed prediction model and the evaluation metrics are also described in this section.

2.1. Technical Indicators and Feature Selection Using Boruta Algorithm. Technical analysis of an asset is based on the premise that all the important information about that asset is contained in its price and/or other market data like the price low, price high, open price, and the volume traded. In this paper, 30 technical indicators (TI) are used as the initial feature space that characterizes Bitcoin price. Table 1 presents the list of initial technical indicators before feature selection. A comprehensive review of these technical indicators is given in [25].

Feature selection is an important step that finds a subset of features that minimize redundant technical indicator features and maximize relevance to the prediction. It helps to avoid the curse of dimensionality and thus improves the accuracy of prediction. In this paper, the Boruta algorithm is used for feature selection. Boruta algorithm (BA) is a feature selection algorithm that uses a wrapper approach built on Random Forest (RF) algorithm to select the most important features for a prediction model. BA can effectively handle the interactions between variables and consider all features, which are relevant to the outcome variable. TBA is implemented as follows:

- Step 1.* Generate duplicates of technical indicators
- Step 2.* Randomly shuffle the original and duplicate technical indicators to take out their correlations with the outcome variable
- Step 3.* Using the RF algorithm, search for the key technical indicators based on higher mean values
- Step 4.* Using the mean/standard deviation, compute the Z score
- Step 5.* From the duplicates and technical indicator feature, find the maximum Z score
- Step 6.* For Z less than the technical indicator feature, remove that technical indicator feature
- Step 7.* Repeat Steps 1–6 until iteration completes

2.2. Variational Mode Decomposition (VMD). VMD algorithm is used to decompose an actual valued input signal into sets of modes (μ_l), also called variational mode functions (VMFs), where each VMF has a unique property. Each VMF has a unique frequency range derived from the input signal. Each VMF is also assumed to be mainly compact across a center pulsation (τ_l), and the modes are extracted concurrently from a convex optimization perspective. The VMD algorithm can be stated as a constrained variation formulation,

$$\min_{\{\mu_l\}, \{\tau_l\}} \left\{ \sum_{l=1}^L \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times \mu_l(t) \right] e^{-j\tau_l t} \right\|_2^2 \right\}, \quad (1)$$

subject to $\sum_{l=1}^L \mu_l = x,$

where $\{\mu_l\} = \{\mu_1, \dots, \mu_l\}$ and $\{\tau_l\} = \{\tau_1, \dots, \tau_l\}$ are the center frequencies of all nodes.

Expressing equation (1) as an augmented Lagrangian results in

$$L(\{\mu_l\}, \{\tau_l\}, \beta) := \gamma \sum_{l=1}^L \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times \mu_l(t) \right] e^{-j\tau_l t} \right\|_2^2 + \left\| x(t) - \sum_{l=1}^L \mu_l(t) \right\|_2^2 + \left\langle \beta(t), x(t) - \sum_{l=1}^L \mu_l(t) \right\rangle, \quad (2)$$

where γ is the balancing parameter and β is the Lagrange multiplier. For the detailed algorithm, see [27].

2.2.1. Signal Average Absolute Average Difference for Variational Mode Determination and Reconstruction of VMF. The number of variational modes (K) is set from 4 to a maximum possible number (in this study, the maximum value set is 14) depending on the signal conditions. For each K value, VMD is used to decompose the original signal into k VMFs. For each K , all the VMFs are aggregated into a single signal (reconstructed VMF). In Figure 1, we present the proposed reconstruction of the variational mode function.

In this paper, a new performance metric that can be used to evaluate the effectiveness of the reconstructed VMF to help in selecting the best K value is introduced. The signal average absolute difference (SAAD) is a metric that computes the average absolute difference between the aggregate VMFs obtained from VMD and the original signal. A very small SAAD indicates that the signals are very similar, and a large SAAD is evidence of information loss of the original signal. From [28], the signal average absolute difference is given as

$$\text{SAAD} = \frac{1}{N} \sum_{n=1}^N |AVMF_n - y_n|, \quad (3)$$

where N is the total number of sampling points in the signals, y_n is the original input signal, and $AVMF_n$ is the aggregate VMFs.

TABLE 1: Technical indicators.

Number	Technical indicator	Number	Technical indicator
1	Stochastics oscillator	17	Fast percent- K stochastics
2	Exponential moving average	18	Fast percent- D stochastics
3	Bias indicator	19	Slow percent- D stochastics
4	Medium price indicator	20	Averaged percent- D stochastics
5	Typical price indicator	21	William's percent R
6	Oscillator indicator	22	Relative strength index
7	Rate of change	23	Williams accumulation/distribution
8	Weighted close indicator	24	Simple moving average
9	Zig zag	25	Exponentially weighted moving average
10	Momentum indicator	26	Vertical horizontal filter
11	MACD indicator	27	The ultimate oscillator
12	MACD signal line	28	Bollinger bands
13	MACD oscillator	29	Welles wilder's directional movement index
14	High/low volatility	30	Relative strength index
15	Volatility ratio		
16	William's % R		

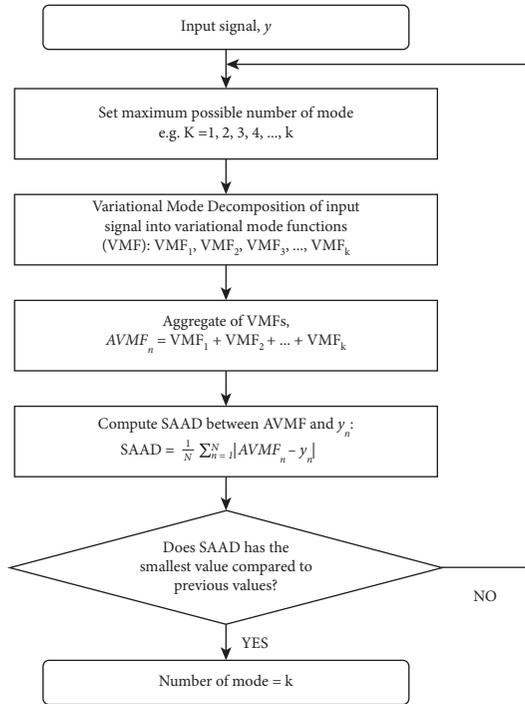


FIGURE 1: Proposed reconstruction of variational mode function.

2.3. *Support Vector Regression (SVR)*. Initially constructed by Vapnik as a classifier [29], SVR is designed with the ability to capture nonlinear relationships in the feature space. It is a machine learning technique that is highly regarded as an

effective technique in regression analysis (i.e., functional approximation) [30].

For a set of training patterns $H = \{(\mathbf{x}_i, y_i)\}_{i=1}^m$ obtained from an unknown function $y = f(\mathbf{x})$ with noise, a function $y = h(\mathbf{x})$ has to be established which completely depends on H and can reduce the difference between h and the unknown function f . Suppose f is a linear relationship between \mathbf{x} and y for linear regression; then

$$y = h(\mathbf{x}, \mathbf{w}, b) = \mathbf{w} \cdot \mathbf{x} + b = \sum_{k=1}^n w_k x_k, \quad (4)$$

where \mathbf{x} is the feature vector and lives in a space \mathbf{X} called the feature space, y is described as the label for each (\mathbf{x}, y) , n is the dimension of \mathbf{x} and \mathbf{X} , and y is described as the label for each (\mathbf{x}, y) . However, the assumption of linear regression is very simple in describing the dynamics of most time-series data. Consequently, it is important to take into consideration a nonlinear f . The basis of SVR for nonlinear regression is to construct a mapping $\mathbf{x} \rightarrow \phi(\mathbf{x})$ from the original n dimension of \mathbf{X} to a new \mathbf{X}' . The dimension of \mathbf{X}' relies on the mapping scheme, and it is not necessarily finite. The nonlinear form is given as follows:

$$y = h(\mathbf{x}, \mathbf{a}, b) = \sum_k \alpha_k y_k K(\mathbf{x}_k, \mathbf{x}) + b, \quad (5)$$

where \mathbf{x}_k 's are support vectors in the given training patterns H , y_k 's are the corresponding labels, and $K(\mathbf{x}_k, \mathbf{x}) = \phi(\mathbf{x}_k) \cdot \phi(\mathbf{x})$ and is defined as the inner product in \mathbf{X}' .

Some of the commonly used kernels are

$$\text{Radial basis function (RBF) kernel: } K(\mathbf{x}, \mathbf{y}) = \exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2), \quad \gamma > 0, \quad (6)$$

$$\text{Polynomial: } K(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x} \cdot \mathbf{y} + r)^d, \quad \gamma > 0,$$

where $K(\mathbf{x}, \mathbf{y})$ is the kernel density, γ is the gamma term in the kernel function, r is the bias term of the polynomial and sigmoid kernel, and d is the polynomial degree term

for the polynomial kernel. Generally, the performance of SVR depends on the settings of the global parameters γ , d , and c .

SVR always gives the same results when the same dataset is processed at any given time. We train Technical Indicator-SVR (TI-SVR), reconstructed variational mode functions-SVR model (rVMF-SVR), and TI-rVMF-SVR model using the parameter settings shown in Table 2. This is to help in deciding on optimal parameter values.

2.4. Data Preprocessing and Evaluation Metrics. To make the data more relevant for the SVR prediction model, the intraday Bitcoin price data are preprocessed and normalized. To make learning easier for the support vectors, heterogeneous time-series data (time series that have different scales)

should always be converted to homogeneous data (time series with similar scales). Hence, the Bitcoin time-series data should take small values (normally be between 0 and 1) and must be homogeneous (all features should possibly take values in the same range). In this study, the mean normalization method (see equation (7)) is used as the data normalization technique. For the mean normalization technique, all features are guaranteed to have the exact same scale. This technique allows data to have values between the range of 0-1. The normalized data values are changed to the magnitude of the original data values via the anti-normalization technique as given in equation (8).

$$y_{\text{normalization}} = \frac{y - \text{mean}(Y)}{\text{maximum}(Y) - \text{minimum}(Y)}, \quad (7)$$

$$y = y_{\text{normalization}} (\text{maximum}(Y) - \text{minimum}(Y)) + \text{mean}(Y), \quad (8)$$

where y and $y_{\text{normalization}}$ are the value of the inputs and the normalized input value, respectively. R statistical software was used in implementing the data normalization.

Table 3 shows the evaluation/performance metrics used in evaluating the prediction model.

2.5. Proposed Two-Stage Hybrid Model for Predicting the Price of Bitcoin. Let A_t represent the one-minute closing price of Bitcoin and P_t represent the final price prediction of A_t . A two-stage hybrid model for high-frequency Bitcoin price is proposed (see Figure 2). The two-stage approach is presented as follows:

Stage One. Selection of technical indicators and variational mode decomposition.

Step 1. Technical indicators are filtered using a correlation matrix filter (technical indicators with more than 0.7 correlations are removed).

Step 2. Boruta algorithm is used to select the most important technical indicator for predicting the Bitcoin series. These technical indicators are used as a feature set 1.

Step 3. Intraday Bitcoin price is decomposed into K ($k = 4, 6, 8, 10, 12, 14$) relatively stable variational mode functions via variational mode decomposition.

Step 4. Using evaluation metrics (SAAD and NRMSE), the best K value is selected and used as a feature set 2.

Stage Two. Aggregating feature set 1 and feature set 2 as input for SVR.

Step 1. Preprocessing the data (data normalization, data partitioning into validation, training, and testing data set).

Step 2. Hyperparameter optimization. A set of optimal hyperparameters for the SVR algorithm is selected using grid search. The SVR model is trained using

feature set 1 and 2. Using the testing data, SVR predicts the Bitcoin series.

Step 3. Single-stage prediction models (Figure 3) are constructed and used as competitor models. Evaluation metrics (MAE, RMSE, NMRSE, and MAPE) are used to verify the performance of the proposed two-stage model and the single-stage model.

3. Results and Discussion

In this section, we present the data and experimental results using technical indicators, reconstructed variational mode functions, a hybrid of technical indicators, and reconstructed variational mode functions as inputs for the support vector radial kernel regression model. The results obtained are also compared in the Discussion section.

3.1. Data. The dataset used for this study is downloaded from <https://www.cryptodatadownload.com/data/bitstamp/>, a publicly available source of data. The data are a high-frequency intraday data sampled at a one-minute time interval from 29/03/2020 to 22/11/2020, making a total of 143464 data points. Each data point contains the minute open, high, low, close, and the trading volume for Bitcoin in United State Dollars (USD). However, after data cleaning, final samples of 136465 data points were retained. The intraday Bitcoin closing price (close) is used as a measure of the price of Bitcoin in this paper. Figure 4 shows the Bitcoin price dynamics over the selected period under study. From the figure, the price of Bitcoin can be seen as highly volatile. Table 4 presents the descriptive statistics of the data.

The dataset was divided into three: training, validation, and testing datasets. The ratio between the training, validation, and testing is approximately 3:1:1. The training dataset was used to tune the parameters of the SVR model. The validation dataset was used in validating

TABLE 2: Parameter settings for SVR.

Parameter	Level (polynomial)	Level (RBF-Gaussian)
Degree of kernel function (d)	1, 2, 3, 4	—
Gamma coefficient in kernel function (γ)	0, 0.1, 0.2, ..., 5.0	0, 0.1, 0.2, ..., 5.0
Regularization parameter (c)	1, 10, 100	1, 10, 100

TABLE 3: Evaluation metrics.

Abbreviation	Metric name	Metric formula
MAE	Mean absolute error	$(1/N) \sum_{t=1}^N e_t $
RMSE	Sum of squared error	$\sqrt{(\sum_{t=1}^N e_t^2)/N}$
NMRSE	Normalized root mean squared error (normalized to the standard deviation of the original data)	$\sqrt{(\sum_{t=1}^N (A_t - \bar{A})^2)/(N-1)}$
MAPE	Mean absolute percentage error	$(100/N) \sum_{t=1}^N (e_t / A_t)$

$n, A_t, \bar{A}, P_t, e_t = A_t - P_t$ are the size of the data set, original data values, mean of the original data values, predicted values, and the error, respectively.

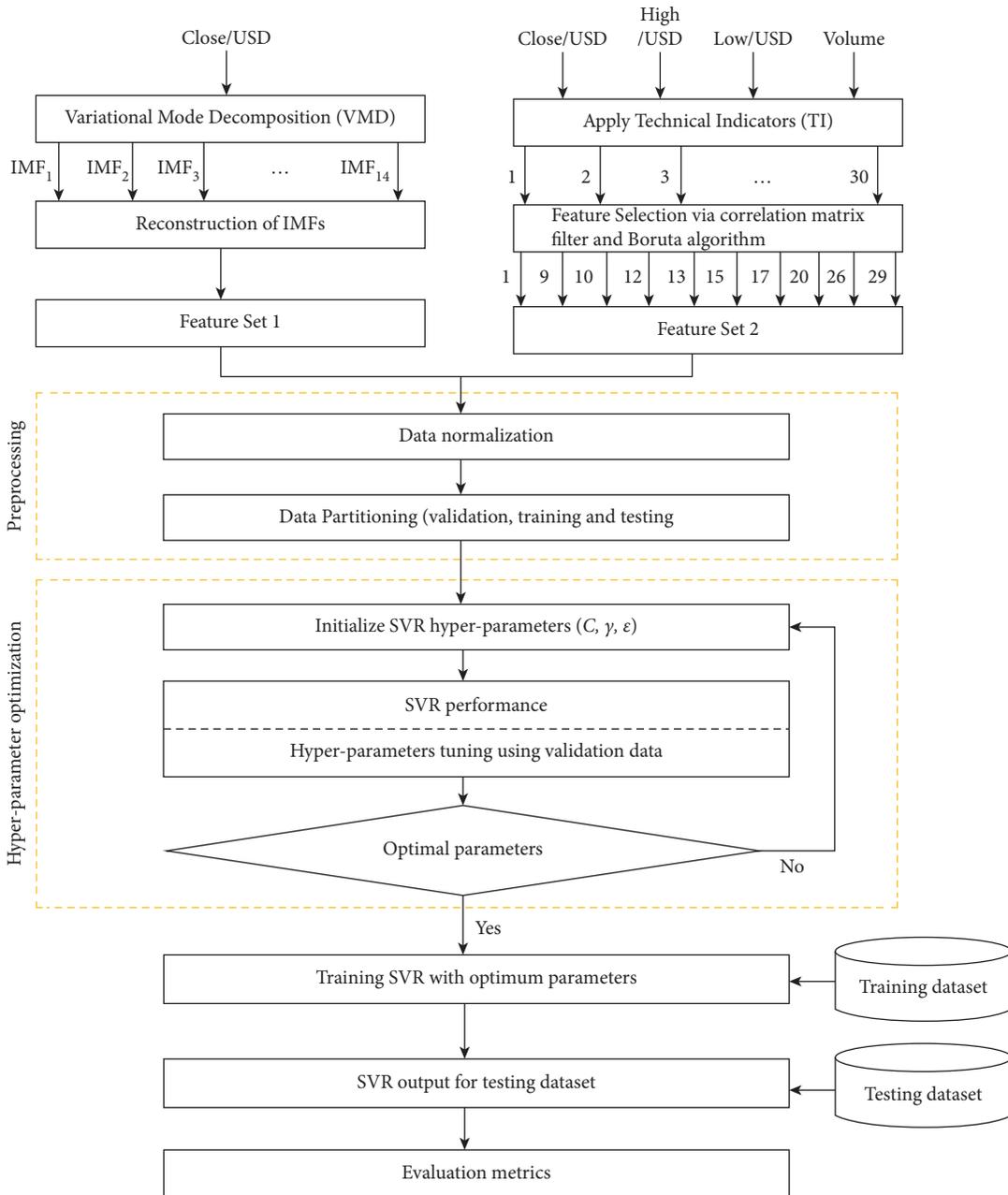


FIGURE 2: The structure of the proposed SVR-TI-rVMF model.

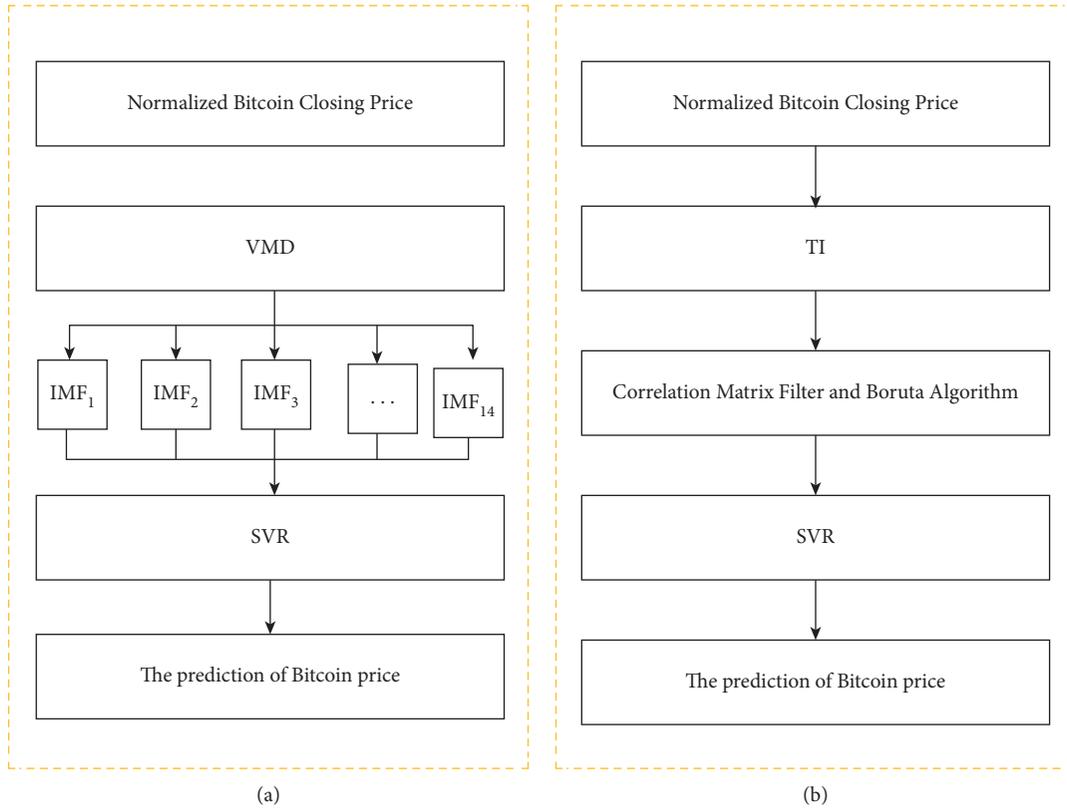


FIGURE 3: Details of SVR-rVMF and SVR-TI. (a) SVR-rVMF. (b) SVR-TI.

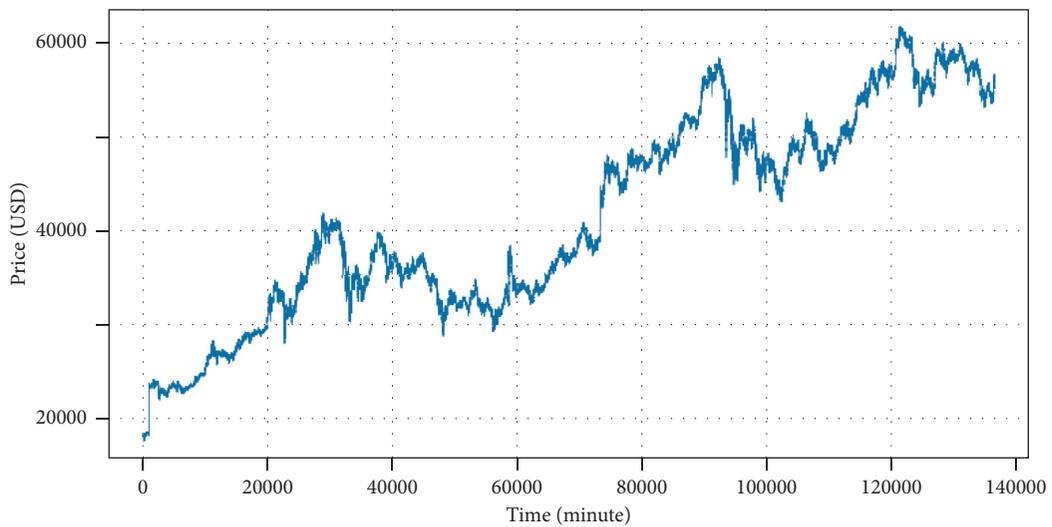


FIGURE 4: Original Bitcoin price series from 29/03/202006: 17 to 22/11/202007: 57.

TABLE 4: Descriptive statistics of closing price (price) of Bitcoin.

Timestamp	Mean	Std.	Skewness	Kurtosis
29/03/202006: 17–22/11/202007: 57	42201.22	10940.86	−0.0823	−1.2012

the optimal parameter values selected for the SVR model, and the testing data were used to test the constructed SVR model.

3.1.1. Data Preprocessing. The historical dataset downloaded is transformed into an acceptable format as inputs for the machine learning technique. The following preprocessing steps were used:

Data Transformation. The original historical time-series data is transformed into a set of technical indicators. In this study, 30 technical indicators are computed for each data point. Using the correlation matrix filter, the technical indicators are filtered and the Boruta algorithm is then used to select important features (10 technical indicators) as inputs (see Figures 5 and 6 and Table 5). The closing price of the data is also decomposed into 8 different IMFs.

Data Cleaning. The original historical data is complete, and, as such, there is no need for data cleaning. However, data cleaning is carried out because of missing data points in the technical indicators derived from the historical data. Overall, 6999 data points that represent 4.8785% of all the data points were missing due to data transformation into technical indicators.

Data Normalization. The data sets are normalized after the data was cleaned so that each technical indicator has zero mean and unit variance.

3.2. Selection of Technical Indicators. Before applying the Boruta algorithm to identify the most relevant technical indicators for the prediction model, we filter the technical indicators using a correlation matrix filter. That is, we remove technical indicators with more than 0.7 correlation. This helps to discard all irrelevant and redundant information. Figure 5 presents the correlation matrix for the 32 technical indicators selected as initial features for the prediction model. Positive and negative correlations are shown in blue and red colours, respectively. The correlation matrix is reordered according to the correlation coefficient. This is important to identify the hidden structure and pattern in the matrix. Colour intensity is proportional to the correlation coefficients. From the correlation matrix (Figure 5), some of the technical indicators are highly correlated. A higher correlation between two technical indicators indicates the redundancy of one feature with regard to the other feature. The resulting correlation matrix is as shown in Figure 6. The filtered technical indicators using the Boruta algorithm are given in Table 5,

Boruta algorithm is used to select the most important features for the prediction model. The algorithm performed 10 iterations in 3.126292 hours. All the 10 attributes were confirmed important; that is, there were no attributes deemed unimportant.

The graphical summary of the Boruta algorithm run is shown in Figure 7 for the features. The boxplots show the distribution of features' importance over the Boruta run, using colours to mark final decisions. Green boxplots represent Z scores of confirmed technical indicator attributes, and blue boxplots represent the minimum, mean, and maximum Z score of a shadow technical indicator attribute. Feature statistics (see Table 6) presents the value for mean importance (meanImp), medianImp (median importance), minImp (minimum importance), maxImp (maximum importance), and the decision for each feature for the complete iterations. The mean importance value of the Zig Zag

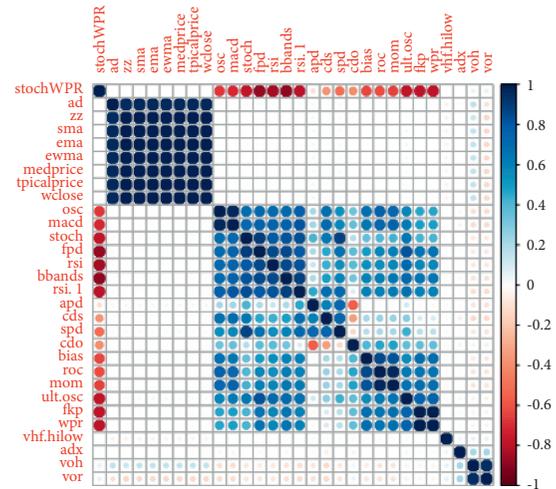


FIGURE 5: Highly correlated matrix for 32 features.

indicator is the maximum among the selected features. This indicates the importance of the Zig Zag indicator as a feature.

3.3. Decomposition of Bitcoin Price Series via VMD and Reconstruction of Variational Mode Function. Intraday Bitcoin price is decomposed into K ($k = 4, 6, 8, 10, 12, 14$) relatively stationary variational mode functions using VMD as depicted in Figure 1. With a SAAD value of 0.0006 and NMRSE value of $7.1537e-04$, $k = 14$ is selected as the best K -value (see Table 7). The VMD results of the optimal variational mode ($K = 14$) are shown in Figure 8. Compared to other VMFs, VMF 2 and VMF 11 have the largest and lowest errors, respectively. The reconstructed signal of the effective mode ($K = 14$) is given in Figure 9, and from the figure, VMD was able to avoid information losses during the decomposition process.

3.4. Discussion. Single-stage models (SVR-TI and SVR-rVMF) are constructed and compared to the SVR-TI-rVMF model to demonstrate the reliability and efficiency of the constructed SVR-TI-rVMF model in improving the performance of Bitcoin price prediction. SVR-TI and SVR-rVMF models are constructed using technical indicators and reconstructed variational mode functions as features (inputs) for SVR, respectively. Figures 10 and 11 illustrate the visual performance using the selected technical indicators and reconstructed variational mode functions as features for the SVR model. The visual performance of the proposed SVR-TI-rVMF is presented in Figure 12. Table 8 presents the performance measure of the constructed models using the testing data. The lower the performance metrics (MAE, RMSE, NMRSE, and MAPE), the more accurate the model. It is evident from the table that the constructed SVR-TI-rVMF hybrid model has the lowest MAE (748.4339 USD), RMSE (993.6821 USD), and NMRSE (0.0919 USD). However, SVR-TI performed well when MAPE was used as the evaluation metrics. The proposed model shows the

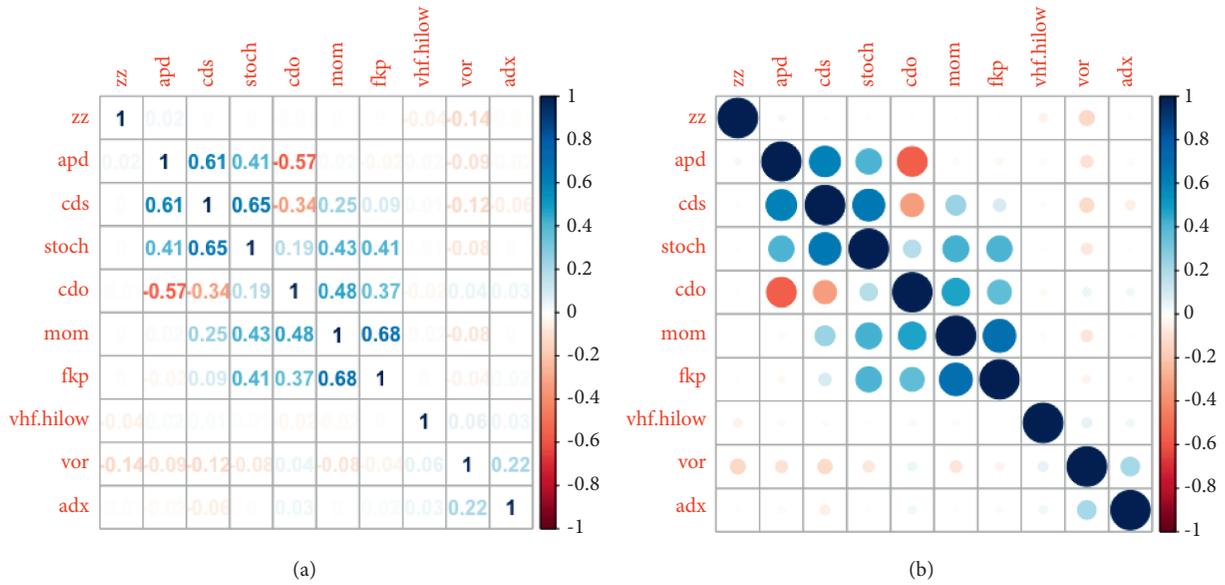


FIGURE 6: Correlation matrix after filter. (a) Visualization in numbers. (b) Visualization in circles.

TABLE 5: Filtered features for Boruta algorithm feature selection.

Number	Feature	Number	Feature
1	Stochastics oscillator (stoch)	15	Volatility ratio oscillator (vor)
9	Zig zag (zz)	17	Fast percent- K stochastics (fkp)
10	Momentum indicator (mom)	20	Averaged percent- D stochastics
12	MACD signal line (cds)	26	Vertical horizontal filter (vhf.hilow)
13	MACD oscillator (cdo)	29	Welles wilder's directional movement index (adx)

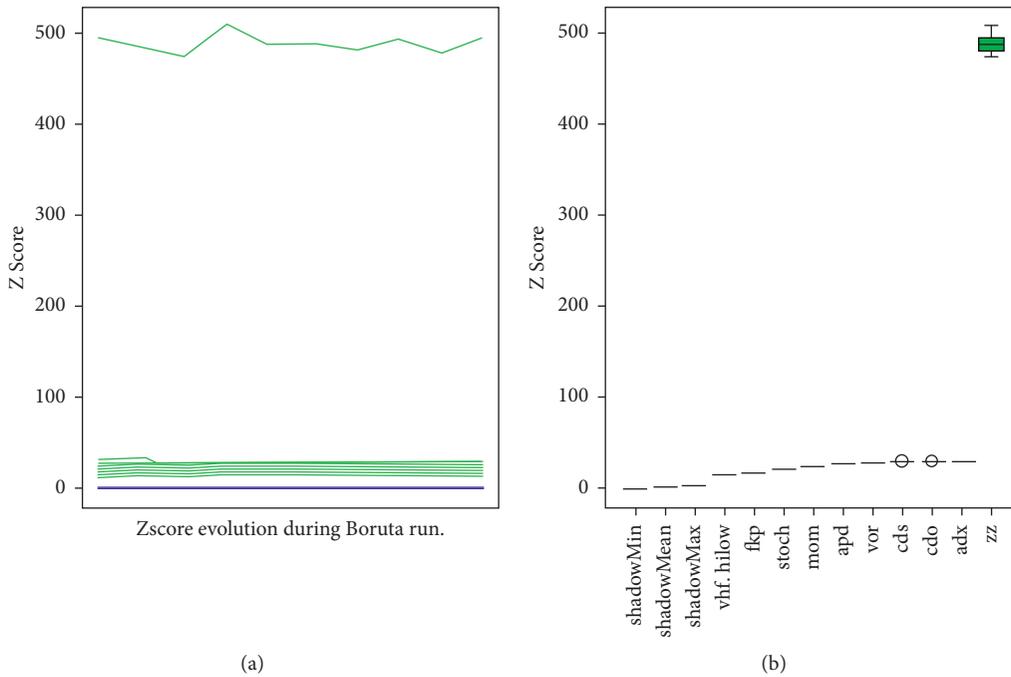


FIGURE 7: Boruta result plot (a) and plot of important history (b) on the Boruta on technical indicators object.

TABLE 6: Feature statistics.

Feature	meanImp	medianImp	minImp	maxImp	Decision
mom	22.7694	22.7558	21.7428	23.5120	Confirmed
cds	27.7633	27.6324	26.8433	29.7796	Confirmed
cdo	27.6667	27.7196	25.9650	28.8714	Confirmed
vor	26.5702	26.7220	25.4257	27.6595	Confirmed
stoch	19.8252	20.0266	18.2012	21.1991	Confirmed
fkp	15.8405	15.6331	13.8536	17.7204	Confirmed
apd	25.3169	25.6653	23.3705	26.3669	Confirmed
zz	488.3230	487.7775	474.3944	508.1439	Confirmed
vhf.hilow	13.3824	13.6999	11.1022	14.9703	Confirmed
adx	28.4670	28.3218	25.8814	31.7259	Confirmed

TABLE 7: Evaluation metrics for different variational mode functions.

	$K = 4$	$K = 6$	$K = 8$	$K = 10$	$K = 12$	$K = 14$
SAAD	27.04299	20.0539	15.25486	10.72741	7.466652	5.313273
NRMSE	$3.727546e-03$	$2.7367e-03$	$2.0681e-03$	$1.4455e-03$	$1.0028e-3$	$7.1537e-04$

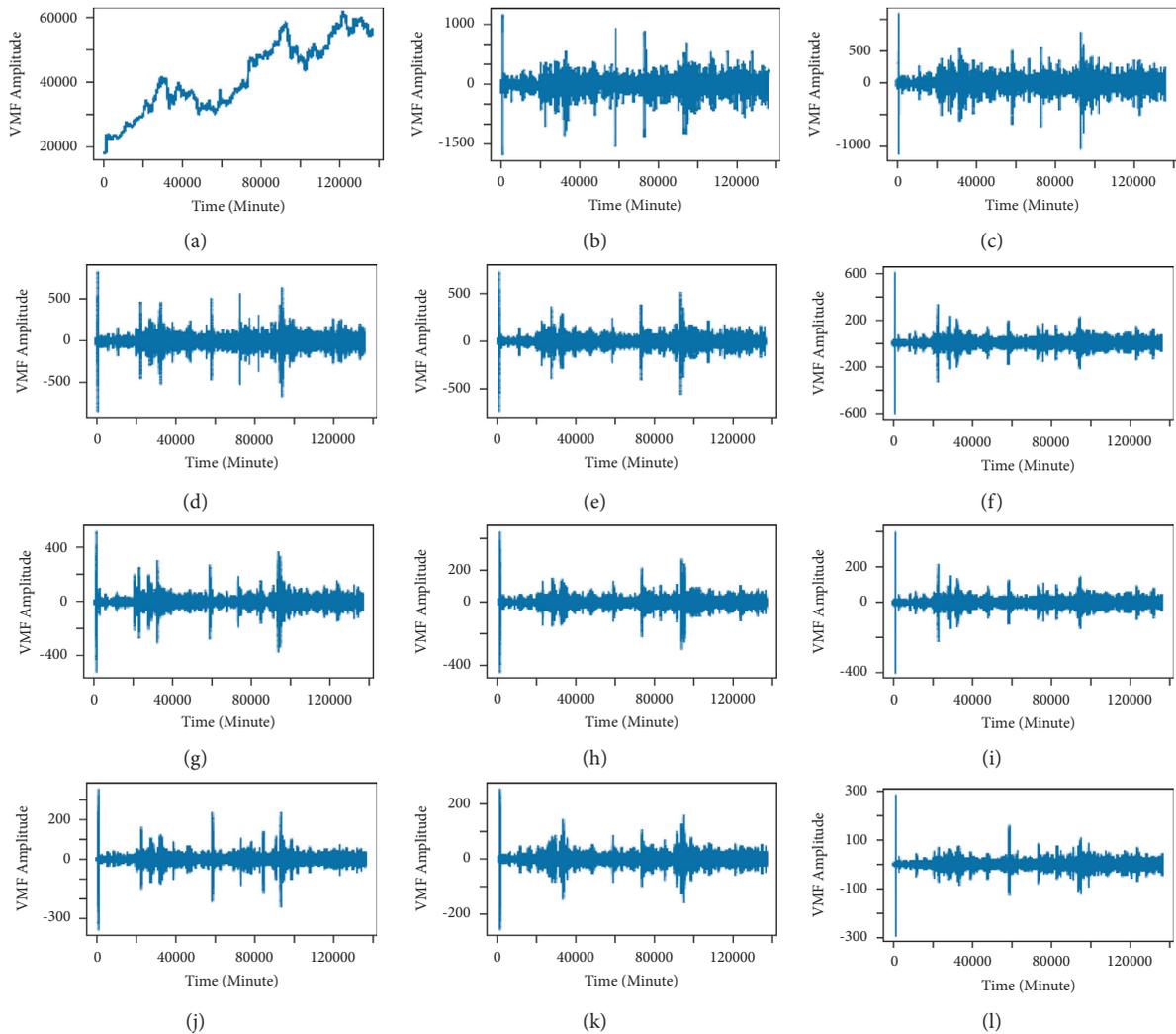


FIGURE 8: Continued.

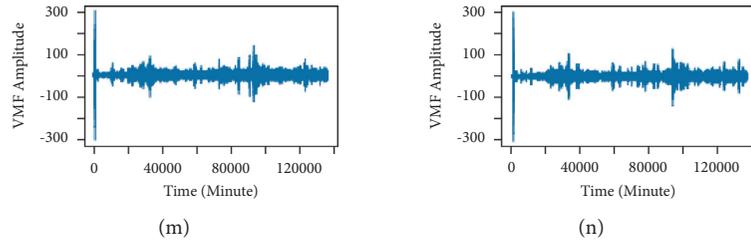


FIGURE 8: Variational mode decomposition of intraday Bitcoin price with 14 variational mode. (a) VMF #1. (b) VMF #2. (c) VMF #3. (d) VMF #4. (e) VMF #5. (f) VMF #6. (g) VMF #7. (h) VMF #8. (i) VMF #9. (j) VMF #10. (k) VMF #11. (l) VMF #12. (m) VMF #13. (n) VMF #14.

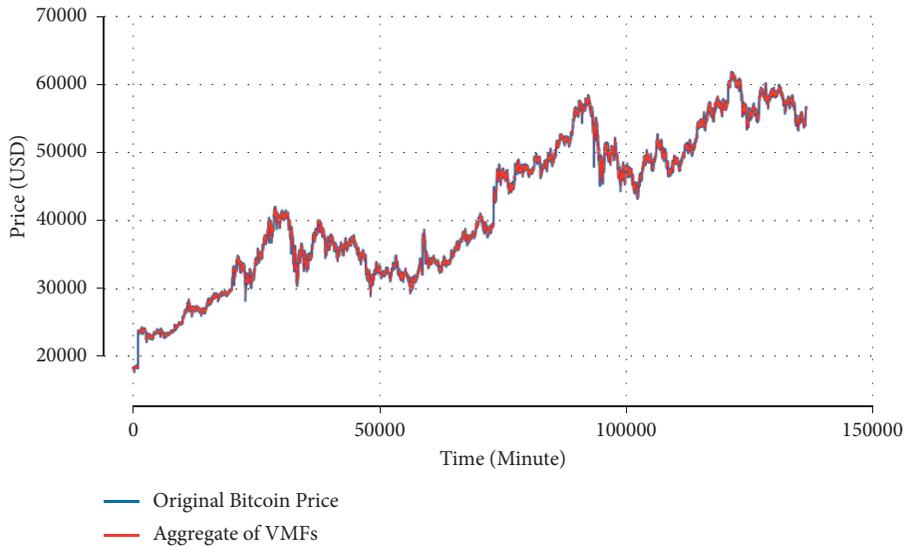


FIGURE 9: Reconstructed signal of VMF_{14} .

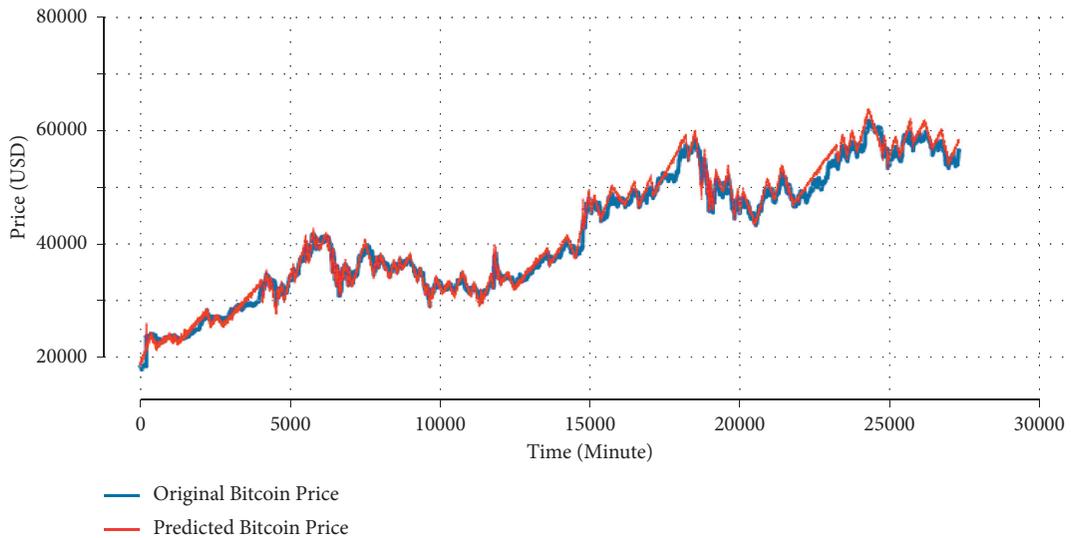


FIGURE 10: Prediction results via SVR-TI model.

highest forecasting accuracy in relation to three different statistical metrics. Figure 12 further shows the satisfactory visual performance of the constructed model. By observing Figure 12, it is clear that most of the

predicted intraday prices fall in line with the original intraday price.

In Figure 13, we present the comparison results with respect to the evaluation metrics of all three

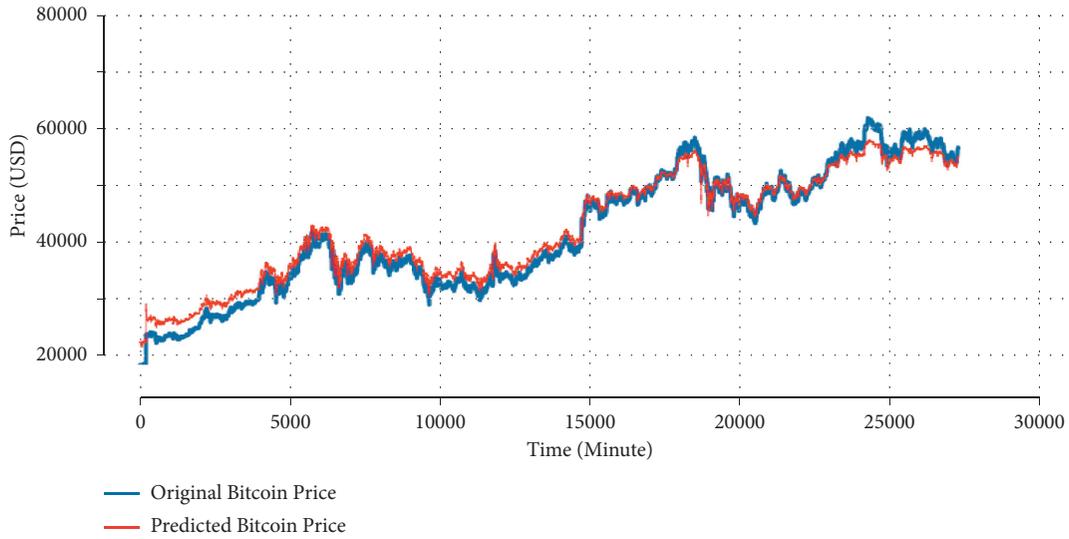


FIGURE 11: Prediction results via SVR-rVMF model.

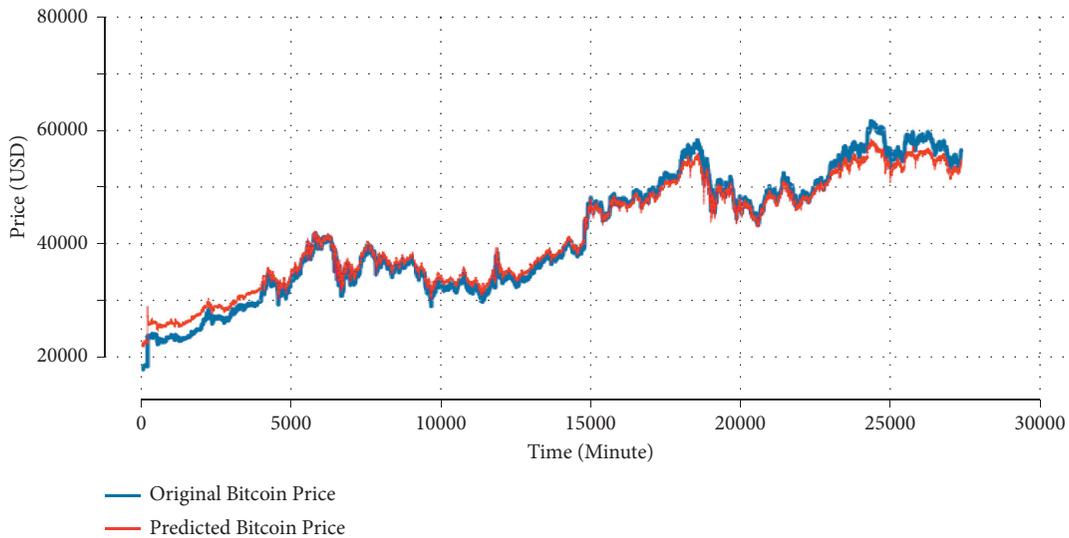


FIGURE 12: Prediction results via SVR-TI-rVMF model.

TABLE 8: Evaluation metrics of SVR-TI, SVR-rVMF, and SVR-TI-rVMF.

Model	MAE	RMSE	NMRSE	MAPE
SVR-TI	1005.8640	1330.1570	0.1230	0.0243
SVR-rVMF	1502.8970	1687.0840	0.1560	4.2626
SVR-TI-rVMF	748.4339	993.6821	0.0919	1.7828

models. In view of the model effectiveness and efficiency, on the whole, we can conclude that the proposed model is quite competitive against two single-stage models, the SVR-TI and SVR-rVMF models. In other words, the

hybrid model leads to better accuracy. Furthermore, the results from Table 8 show that hybrid models outperform their single counterparts in chaotic time-series prediction.

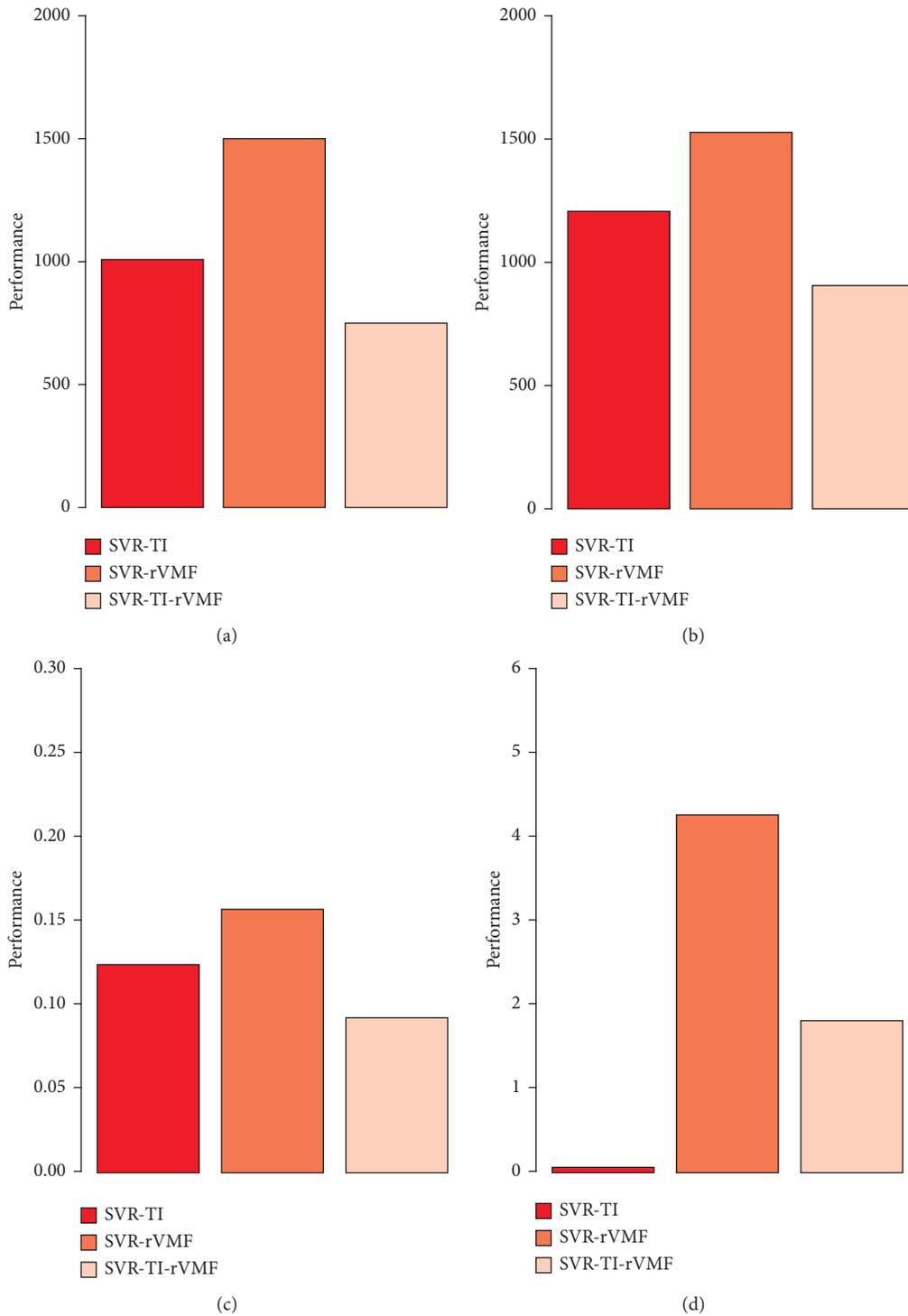


FIGURE 13: Bar chart of MAE, RMSE, NMRSE, and MAPE of actual and predicted intraday Bitcoin price by SVR-TI, SVR-rVMF, and SVR-TI-rVMF using testing data. (a) MAE. (b) RMSE. (c) NMRSE. (d) MAPE.

From Figure 14, the maximum and minimum values of the original price testing data values fall in the probability density curve constructed using the

predicted values of the constructed model. This proves the superiority of the constructed SVR-TI-rVMF model.

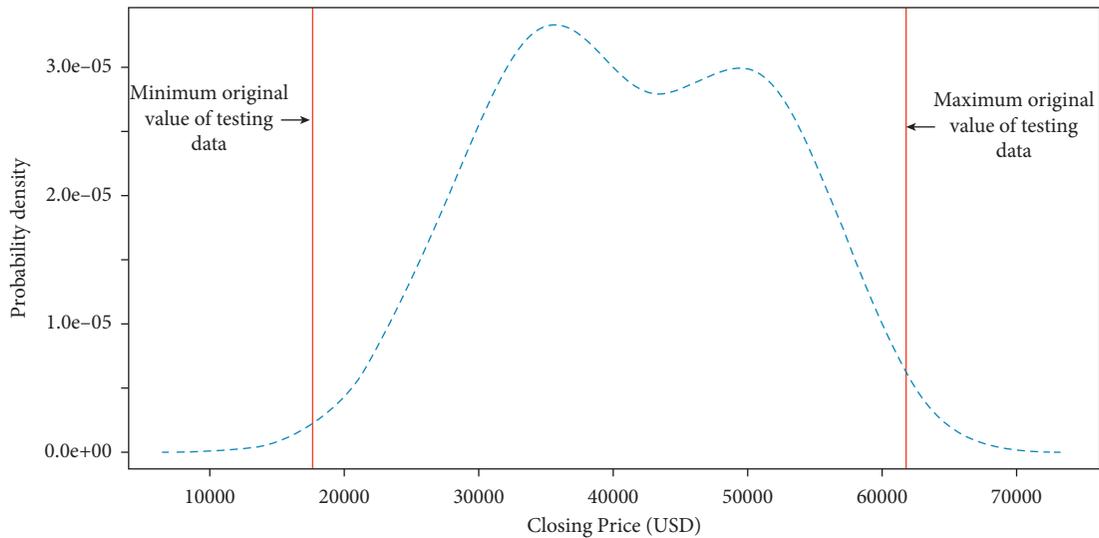


FIGURE 14: Probability density curve based on SVR-TI-rVMF model.

4. Conclusion

In this paper, we combine the advantage of technical indicators (TI) and reconstructed variational mode functions (rVMF) obtained from variational mode decomposition (VMD) to construct a hybrid support vector prediction model for one-minute intraday Bitcoin price. The model (SVR-TI-rVMF) reveals the fact that decomposition methods (VMD) and TI can be used together, yet separately, to construct a hybrid model to predict the Bitcoin price in the cryptocurrency market. Our contribution in this paper is as follows: (1) defining a new performance metric to evaluate the effectiveness of the reconstructed VMF in selecting an optimal mode (K) value called signal average absolute difference (SAAD), (2) predicting the out-of-sample intraday price of Bitcoin in the midst of COVID-19 pandemic via a hybrid of TI and rVMF as features for SVR prediction model, (3) evaluating and comparing the predictive performance of two features (TI and rVMF) to the hybrid model in the midst of COVID-19, and (4) adding to scarce empirical evidence of hybrid model using SVR, TI, and rVMF in predicting one-minute intraday Bitcoin price. Based on these studies, investors can decide whether to buy Bitcoin or not. The findings are important for practitioners, such as traders and investors, as well as policymakers, who want to learn more about the cryptocurrency market.

Data Availability

Data used for this work is available from the corresponding author upon request. Other supporting data can also be downloaded from <https://www.cryptodatadownload.com/data/bitstamp/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] N. Reif, *Top 5 Cryptocurrencies by Market Cap*, 2019, <https://www.investopedia.com/news/top-5-cryptocurrencies-market-cap/>.
- [2] S. A. Gyamerah, "Modelling the volatility of bitcoin returns using garch models," *Quantitative Finance and Economics*, vol. 3, no. 4, pp. 739–753, 2019.
- [3] P. Gradwell, *Wtf Is Going on in the Bitcoin Market?*, 2020, <https://blog.chainalysis.com/reports/bitcoin-market-march-2020>.
- [4] A. Urquhart, "The inefficiency of bitcoin," *Economics Letters*, vol. 148, pp. 80–82, 2016.
- [5] F. N. Zargar and D. Kumar, "Informational inefficiency of bitcoin: a study based on high-frequency data," *Research in International Business and Finance*, vol. 47, pp. 344–353, 2019.
- [6] Y. Hu, H. G. A. Valera, and L. Oxley, "Market efficiency of the top market-cap cryptocurrencies: further evidence from a panel framework," *Finance Research Letters*, vol. 31, pp. 138–145, 2019.
- [7] H. Zhang and N. Rudholm, *Modeling and Forecasting Regional Gap in Sweden Using Autoregressive Models*, Dalarna University, Högskolegatan, Sweden, 2013.
- [8] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas, and P. Liatsis, "Electricity demand loads modeling using autoregressive moving average (arma) models," *Energy*, vol. 33, no. 9, pp. 1353–1360, 2008.
- [9] S. Garg and Anupriya, "Autoregressive integrated moving average model based prediction of bitcoin close price," in *Proceedings of the 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 473–478, IEEE, Tirunelveli, India, December 2018.
- [10] Z. Ismail, A. Yahya, and A. Shabri, "Forecasting gold prices using multiple linear regression method," *American Journal of Applied Sciences*, vol. 6, no. 8, pp. 1509–1514, 2009.
- [11] X. Xu and W. Ren, "A hybrid model based on a two-layer decomposition approach and an optimized neural network for chaotic time series prediction," *Symmetry*, vol. 11, no. 5, p. 610, 2019.
- [12] S. A. Gyamerah, "On forecasting the intraday bitcoin price using ensemble of variational mode decomposition and generalized additive model," *Journal of King Saud University-Computer and Information Sciences*, 2020, In press.

- [13] W.-M. Hung and W.-C. Hong, "Application of svr with improved ant colony optimization algorithms in exchange rate forecasting," *Control and Cybernetics*, vol. 38, 2009.
- [14] G.-F. Fan, S. Qing, H. Wang, W.-C. Hong, and H.-J. Li, "Support vector regression model based on empirical mode decomposition and auto regression for electric load forecasting," *Energies*, vol. 6, no. 4, pp. 1887–1901, 2013.
- [15] Z. Rustam and P. Kintandani, "Application of support vector regression in Indonesian stock price prediction with feature selection using particle swarm optimisation," *Modelling and Simulation in Engineering*, vol. 2019, Article ID 8962717, 5 pages, 2019.
- [16] I. H. Sarker, "Machine learning: algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, pp. 1–21, 2021.
- [17] N. Indera, I. Yassin, A. Zabidi, and Z. Rizman, "Non-linear autoregressive with exogeneous input (narx) bitcoin price prediction model using pso-optimized parameters and moving average technical indicators," *Journal of Fundamental and Applied Sciences*, vol. 9, pp. 791–808, 2017.
- [18] S. Wu, F. Feng, J. Zhu, C. Wu, and G. Zhang, "A method for determining intrinsic mode function number in variational mode decomposition and its application to bearing vibration signal processing," *Shock and Vibration*, vol. 2020, Article ID 8304903, 16 pages, 2020.
- [19] H. Niu, K. Xu, and K. Xu, "A hybrid model combining variational mode decomposition and an attention-gru network for stock price index forecasting," *Mathematical Biosciences and Engineering*, vol. 17, no. 6, pp. 7151–7166, 2020.
- [20] L. J. Xian, S. Ismail, A. Mustapha, M. H. Abd Wahab, and S. Z. S. Idrus, "Crude oil price forecasting using hybrid support vector machine," in *Proceedings of the International Conference on Technology, Engineering and Sciences (ICTES)*, vol. 917, October 2020, Article ID 012045.
- [21] E.-T. Cheah and J. Fry, "Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin," *Economics Letters*, vol. 130, pp. 32–36, 2015.
- [22] J.-Z. Huang, W. Huang, and J. Ni, "Predicting bitcoin returns using high-dimensional technical indicators," *The Journal of Finance and Data Science*, vol. 5, no. 3, pp. 140–155, 2019.
- [23] S. Alonso-Monsalve, A. L. Suárez-Cetrulo, A. Cervantes, and D. Quintana, "Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators," *Expert Systems with Applications*, vol. 149, Article ID 113250, 2020.
- [24] E. Bradley and H. Kantz, "Nonlinear time-series analysis revisited," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 9, Article ID 097610, 2015.
- [25] S. B. Achelis, *Technical Analysis from A to Z*, 2001.
- [26] M. B. Kursat and W. R. Rudnicki, "Feature selection with the boruta package," *Journal of Statistical Software*, vol. 36, pp. 1–13, 2010.
- [27] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, pp. 531–544, 2013.
- [28] M. F. Isham, M. S. Leong, M. H. Lim, and Z. A. Ahmad, "Variational mode decomposition: mode determination method for rotating machinery diagnosis," *Journal of Vibroengineering*, vol. 20, no. 7, pp. 2604–2621, 2018.
- [29] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer New York Google Scholar, New York, 1995.
- [30] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," *Advances in Neural Information Processing Systems*, vol. 9, pp. 155–161, 1997.