IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 19, 2021, accepted February 26, 2021, date of publication March 2, 2021, date of current version March 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3063243

A Bayesian Regularized Neural Network for Analyzing Bitcoin Trends

R. SUJATHA[1], V. MAREESWARI[1], JYOTIR MOY CHATTERJEE[2,3], ABD ALLAH A. MOUSA[4], AND ABOUL ELLA HASSANIEN[3,5]

[1]SITE, Vellore Institute of Technology, Vellore 632014, India
[2]Department of IT, LBEF, Kathmandu 44600, Nepal
[3]Scientific Research Group in Egypt (SRGE), Giza 12613, Egypt
[4]Department of Mathematics and Statistics, College of Science, Taif University, Taif 21944, Saudi Arabia
[5]Faculty of Computers and Information, Cairo University, Giza 12613, Egypt

Corresponding author: Abd Allah A. Mousa (a.mousa@tu.edu.sa)

This work was supported by the Taif University Researchers Supporting Project Number (TURSP-2020/48), Taif University, Taif.

ABSTRACT Bitcoin is a decentralized digital currency without a central bank or single administrator sent from user to user on the peer-to-peer bitcoin blockchain network without intermediaries' need. In this Bitcoin trend analysis work, initial attributes are considered from five sectors based on financial, social, token, network, and that count to thirteen attributes. The thirteen attributes considered are price, volume, market cap, a mean dollar invested age, social volume, social dominance, development activity, transaction volume, token age consumed, token velocity, token circulation, market value to realized value, and realized cap. We apply the attribute selection and trend analysis mapped with potential seven attributes: Price, Volume, Market Cap, Social Dominance, Development Activity, Market Value to Realized Value & Realized Cap. We have conducted Nonlinear Autoregressive with External Input analysis considering seven attributes. The work employed three training algorithms to train a neural network as Levenberg-Marquard, Bayesian Regularization, and Scaled Conjugate Gradient algorithm. The Error histogram and regression plots results indicate that the Bayesian Regularized Neural Network is showing good performance and thus provides a better forecast.

INDEX TERMS Bitcoin, market cap, neural network, realized cap, nonlinear autoregressive with external input (narx), neural network (NN), Levenberg-Marquard (LM), Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG), Bayesian Regularized Neural Network (BRNN).

I. INTRODUCTION

In recent years' passion for blockchain technology research is increasing. The idea of Blockchain comes from transferring digital things securely from one to another without trusting the central authority. Blockchain is the collection of blocks of verified digital information that stores the transaction details. Every block has a unique hash code; unalterable once added on the Blockchain. This record-keeping cryptographic technology is a secure, decentralized, distributed, transparent public ledger. It has been used in many applications such as banking sectors, crypto-currency, healthcare record management, supply chain, enhancing education and learning, entertainment, social networking, and voting system. Cryptocurrency refers to the digital cash value depending on the supply and demand of any product, which stores the Blockchain's transaction details. In a few decades, the buyers will do all financial transactions without ready cash. The stock market, Climate forecast, Gold market, and so on are the various vital sectors that rely on the data.

Bitcoin is one of the crypto-currencies powered by blockchain developed by Satoshi Nakamoto in 2008 [1]. The bitcoin protocol is the predominant real-world application that elevates the blockchain technology as in height. This new electronic cash system built over on hash-based proof of work. Hence this secure currency exchange booms the crypto-currency arena and is turned to the investment assets. The growth of crypto-currencies is increased on a large scale. In general, other than the well-known bitcoin is called altcoins such as Ethereum, Tether, and Litecoin. Hence, the new

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan Bu.

#	Name	Market Cap	Price	Volume (24h)	Circulating Supply	Change (24h)	Price Graph (7d)
1	Bitcoin	$170,119,760,175	$9,353.64	$33,403,180,697	18,187,550 BTC	3.62%	
2	Ethereum	$19,436,862,433	$177.55	$11,765,943,635	109,474,151 ETH	2.95%	
3	XRP	$10,490,625,843	$0.240139	$2,012,339,233	43,685,558,183 XRP *	3.21%	
4	Bitcoin Cash	$7,101,645,520	$389.16	$4,159,410,408	18,248,613 BCH	6.80%	
5	Bitcoin SV	$5,459,641,734	$299.61	$3,004,058,111	18,222,577 BSV	-0.01%	
6	Tether	$4,650,517,806	$1.00	$44,099,301,064	4,642,367,414 USDT *	0.13%	
7	Litecoin	$3,933,722,849	$61.50	$4,373,828,318	63,960,974 LTC	2.73%	
8	EOS	$3,856,951,336	$4.06	$3,717,099,272	950,699,438 EOS *	1.12%	
9	Binance Coin	$2,834,978,253	$18.23	$248,758,229	155,536,713 BNB *	2.94%	
10	Cardano	$1,455,183,145	$0.056126	$249,546,261	25,927,070,538 ADA	15.27%	

FIGURE 1. **Top-10 Crypto-Currencies by Market Capitalization [28].**

coin is generated for a verified transaction asset and added to the chain as on the witness of proof. The global blockchain market size is an upswing of around $65 billion by 2025. IBM is one of the biggest companies to drive this distributed ledger to reach a peak point. IBM, Microsoft, and Accenture are the top market-leading companies for trading their financial clients to blockchain services.

On 22nd May 2010, the first bitcoin was used for the documented transaction of 2 Pizza purchase and delivery at the rate of $41. For which 10,000 bitcoins were used for $41. But, as of now, the price of 1 bitcoin is $9,353. For example, if an investor invests $1000 in the market for 24.39 bitcoin on that date of pizza purchasing, they have earned $228,119.67 (24.39 $*$ $9,353) within ten years. The investment in crypto-currency returns a high impact over the minimum time interval. But it varies over on different coins [2]. Hence, a lot of research is going on crypto-currency based in different sectors with various problems.

As of 29th January 2020, there are 5,075 crypto-currencies registered in the popular [28]. Figure 1 states that the top 10 cryptocurrencies by market capitalization. As of date, the total market cap is $257,091,894,877.

The market cap is calculated by the product of the circulating supply and price of a coin. For instance, the Bitcoin coin has around 18,187 coins in circulation, and the price of each coin is $9,573, so the market cap of bitcoin is around $170 billion. Compared with the Ethereum coin, even the circulation of bitcoin is lower, but it is higher in price, and hence the market cap is high. In the reverse case, even the Tether coin's price is lower than Litecoin, but it gets a higher market cap depending on higher circulation. So, market capitalization is an important attribute than the price for an indication of a specific asset. Likewise, there is a need for determining the importance of features and selection.

Most of the researches has been done by employing the strength of machine learning (ML) algorithms. Moreover, such predictive models of the stock market are suitable for analyzing crypto-currency coins' features [3]. Meanwhile, the ML algorithms have been applied partially to analysis on crypto-currency coins [4], NN-based methods as Bayesian neural network (BNN) [5], long-short term memory (LSTM), and recurrent NN (RNN) [6], and other algorithms. Among these and other concerns, research studies suggest that the NN algorithm provides a better result in predicting Bitcoin price and analysis on the crypto-currency

Market [7]. Hence this proposal practiced the time series analysis on Bitcoin price prediction using NN. Bayesian models demonstrated their impact on the monetary subordinate protection examination [39]. The arrangement of Blockchain, the main innovation of Bitcoin, recognizes Bitcoin from other fiat monetary forms and is straightforwardly identified with Bitcoin's market interest. To the most amazing aspect, our insight, notwithstanding macroeconomic factors, direct utilization of Blockchain data, for example, hash rate and block creation rate, has not been explored to depict the cycle of Bitcoin cost.

The primary contribution of this analysis is:

1) Bitcoin Trends plugs out significant features of Bitcoin and how it will affect the Bitcoin market.
2) Based on our condensed work, the past historical values are playing a vital role in this kind of non-linear time series analysis and found the NN stands as the optimal algorithm.

3) We contribute the comparative examination of the LM, BR, and SCG algorithm and conclude the BR has functioned as satisfactory with a low error rate.
4) Lastly, this proposed work will encourage predicting factors that make the trend in the cryptocurrency market.

II. RELATED WORK

There is a lot of interest in analyzing crypto-currencies growth and contemporary issues in present interdisciplinary research studies. Most investors invest in gold to increase their assets and are considered safe. Now it's turned into bitcoin aside since it is also a surrogate currency to buy and sell products in a secure digital system. This decentralized system has isolated the banking system and monetary strategies; however, the gain and loss are subject to contenders' decisions in the market. Onside, the participants can view the number of bitcoins, daily transactions, and predictability of price and such related to making the right decision. Researchers in [8] analyzed gold and bitcoin's co-movement using DCC-GARCH and Wavelet coherence methods since these two investments' uncertainty. They determined that in previous years gold returns headed than bitcoin later 2014-2015, this movement is to reverse in the short running of time.

Reference [22] introduced a point-by-point conversation of the key trust issues in the whole digital money environment & proposed various prompt, present moment & long-term arrangements. Reference [23] presented a novel methodology for de-anonymizing Bitcoin by utilizing supervised ML to anticipate the kind of yet-unidentified elements. Reference [24] suggested a two-level pipeline hardware engineering for the SHA256. Reference [25] investigated the data stream among Bitcoin & its split markets. Reference [26] attempted to make scientists & professionals aware of the present status of consensus protocols examination & meant to investigate the exploration presenting new consensus protocols to empower a more brought together treatment. Reference [27] investigated data streams between bitcoin costs & other monetary resources in 27 nations, demonstrating that bitcoin collaborates with monetary resources.

Meanwhile, many researchers analyzed many crypto-currencies and determined that it is a credible financial asset and its price has fluctuated in a short interval [9]. Authors in [10] discussed whether the bitcoin is either currency or property concerned with a taxable income and follows the IRS guidelines of the secondary sources. Like, the nominated proof-of-assets verification for exchange in a secure network is to be measured on cyber criminality basis [11]. Hence, plenty of research is on crypto-currency such as sustainability, regulation, taxation, cyber criminality, diversification benefits, digital currency efficiency, market share, market volatility, price volatility, price discovery, and product attention in social media. For which, the information is gathered from many data sources such as [29]–[37].

Many kinds of research are especially about crypto-currency price prediction models and associated experimental analysis on price volatility. The flexible mechanisms have been used for the investigation with several factors. The various ML techniques, specific deep learning (DL) techniques, and other traditional methods have been projected on crypto-currency-based research. Reference [12] initiated to analyze the relationship between bitcoin terms with Google Trends and Wikipedia. The query-based on 'Bitcoin' or 'bitcoin' is searched roar at both servers around 2013, still, it is increasing. These searches could be made by researchers, fund managers, developers, market analyzers, or regulators.

The price of Bitcoin is varied in a minute; hence this time series data has been investigated by BNN to forecast the log price and volatility. It provides better accuracy than Support Vector Regression [5]. This research combines the statistical model GARCH and the ML technique SVR and analyzes the daily and hourly frequencies of crypto-currency prices such as Bitcoin, Ethereum, and Dash. Also, they compared the frequency of prices in the US dollar of fiat currency such as Euro, Pound, and Yen. They proved that SVR-GARCH provides better accuracy than other GARCH methods by obtaining less error value of the Diebold-Mariano predictive test.

Researchers in [37], [38] gave a best-in-class review over Bitcoin-related advances & summarize different difficulties.

Researchers in [13] developed the prediction model for the Ethereum coin, in which the Linear Regression (LR) model does not provide better accuracy in prediction. Then, the Support Vector Machine (SVM) employed a radial basis kernel to produce better accuracy. They suggest that the best month for investment is founded on experimental results. According to the following attributes, authors in [14] proposed a Bitcoin prediction model such as open, close, low, and high exchange data daily. Approximately, out of 86 input attributes, the best 20 attributes are picked out. These attributes are selected based on five different correlation analyses, relief technique, Information Gain method, Principal Component Analysis (PCA), and correlation-based feature subset selection. Further, they achieved high accuracy prediction by SVM rather than RNN with Tree classifier.

Authors in [15] employed the DL approaches such as CNN (Convolutional Neural Network), hybrid CNN-LSTM (Long-Short Term Memory), MLP (Multi-Layer Perceptron), and radial basis NN to predict the high trend of six crypto-currencies including Bitcoin, Dash, Ethereum, Litecoin, Ripple and Monero. They suggest that LSTM provides better results than others. Moreover, the trend of Ether and Bitcoin is easily predictable for the rest of them.

Much of the research related to employing the prediction of Bitcoin price and compared different models. In parallel, such a proposal was raised to analysis on Bitcoin versus altcoins with specific methods. Sometimes a unique model doesn't produce better results in the prediction of crypto-currencies problems. Hence, a hybrid model has been raised to achieve

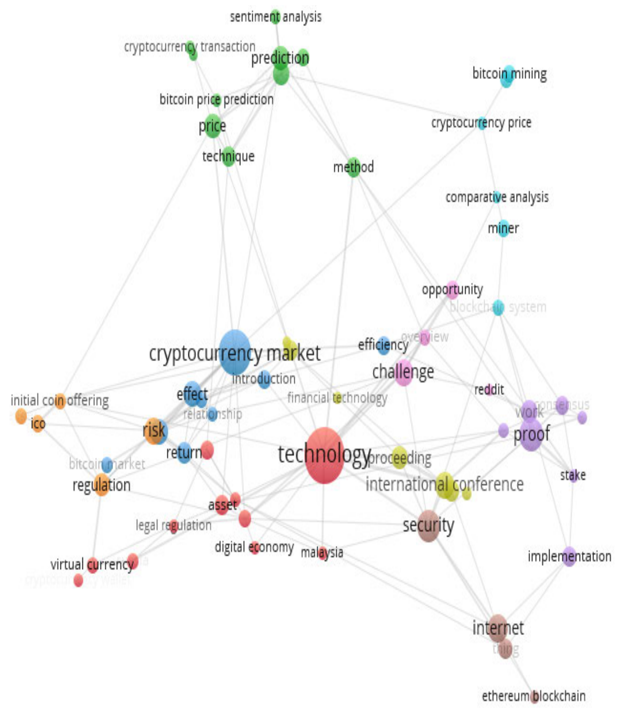

FIGURE 2. Bibliometric map for cryptocurrency journal.

the best solution like LSTM, EWT (Empirical Wavelet Transform) decomposition techniques, and cuckoo search algorithm optimization [16].

These researchers in [17] proposed a prediction model for finding the closing price of Bitcoin using NARX. In this model, selected attributes except closing price are considered as an external input. Hence, the number of input nodes depends on the number of input attributes; one output node designates the closing price on the next day. The number of hidden nodes selected depends on trial and error but limited between the count of input and output nodes. They practiced both one-step-ahead prediction (next day) and long-term prediction (31 days). Here, 31 output nodes were used for forecasting the closing price of Bitcoin for the next 31 days. Meanwhile, the ARIMA (Auto-Regressive Integrated Moving Average) model is practiced for predicting the closing price of the next day. The performance was measured by MSE (Mean Squared Error) and MAE (Mean Absolute Error) to compare these three models.

Moreover, they identified the best model for Bitcoin prediction as NARX proved by analyzing autocorrelation, cross-correlation, and residual histogram plot. Finally, they suggested that the NARX model is the best model for long-step forecasting. A similar ARIMA model is best suited for one-step prediction since it requires fewer data and cost-efficient [18].

Recent research studies concentrate on the bitcoin price; they still, need accurate prediction models. Figure 2 shows the bibliometric map drawn with Vosviewer of cryptocurrency-related journals and their related fields as potential work to be done.

III. PROPOSED WORK

Here we are presenting our approach with all the details.

A. DATASET DESCRIPTION

The data set is generated from an application sanbase in (https://santiment.net/). Santiment is the platform for both beginners and experts of investors as well as researchers. It provides the market data of crypto-currency, Blockchain, and different analyses. From where the time series data of the year 2019 is collected for this proposed work. Here, financial-based attribute price, volume and market cap, coin-based development activity, social-based social dominance, and social activity, token-based attributes such as transaction volume, token age consumed, token velocity and token circulation, and network value-based market value to realized value and realized cap is considered to generate the dataset for this proposed work. Market cap is an important metric to measure the value of the security. It can be defined as the product of the current price and existing shares on hand. A similar, realized gap is defined as the total amount expended to buy the coin by entire investors. Then profit or loss for an investor can be calculated by the attribute market value to realized value. For instance, this value is 2 means an investor or coin purchaser will double their opening purchase amount. Nowadays, most of the discussion of everything has happened on social media such as Twitter, Facebook, Reddit, Bitcointalk.org, etc. It is the best platform to promote business value for investors. Some online tools such as volume and crypto-metrics have reduced the difficulty of collecting social volume metrics from different websites. In crypto-currency, a token represents the encoded record of the transaction on a blockchain block. Token velocity represents the average number of times that a token changes wallet each day. If the token velocity is high, a particular token is used in transactions frequently at certain intervals.

Meanwhile, the distinctive tokens are counted to measure the token circulation of each day. Table 1 describes the attributes that are collected from sentiment for this proposed work.

Figure 3 narrates the workflow of the proposed method over the cryptocurrency data of Table 1.

B. ATTRIBUTE SELECTION

Attribute selection is the predominant step in data set analysis. Generally, it employs two steps which are attribute evaluator and search mechanism. Among the set of attributes, certain attributes are significant concerns to the target output or categorized labels. Likewise, the search mechanism analyses and directs the relevant group of attributes to work together and seeded as significant. Meanwhile, certain attribute evaluators are employed with certain search mechanisms. For instance, Correlation Attribute Evaluator (CAE) is compatible with the Ranker search method. CAE method estimates the Pearson correlation between the specific attribute and class and measures an attribute's wealth.

TABLE 1. Attribute Description.

Sector	Attribute	Description
Financial	Price	Daily price in US dollars
	Volume	The total amount of trading of coin on a particular exchange
	Market cap	The current price of a coin * circulating supply (total number of publicly accessible coins in a market)
	A mean dollar invested age	The average age of buyers of the cryptocurrency
Social	Social volume	Coins mentioned several times in more than 1000 social media
	Social dominance	The total share of coins related terms used in social media
Development	Development activity	Novel trading strategy that talks about business proportion.
Token Flow / Movement / Activity	Transaction volume	Total number of tokens for all transaction in a day
	Token age consumed	The number of tokens modifying their addresses on a particular timestamp * number of blocks created on the Blockchain later on the last movement.
	Token velocity	Total transaction volume / Average network value
	Token circulation	Number of distinctive tokens used in each day
Network value	Realized cap	Total acquisition cost i.e. circulating supply* price of a coin since it last moved.
	Market value to realized value	Measures the amount of profit or loss for an investor.

When low correlation or close to 0 comes, then drop that certain attribute. Hence remaining positive (nearest to +1) or negative (nearest to -1) correlated with specific attributes will be considered. The Ranker Search Method (RSM) works to sort the attributes according to the result of CAE. While practicing these combinations in a dataset of this present work, some attributes are correlated such as date, market cap, price, social dominance, development activity, market value to realize the value, and other attributes left out from this selection. This sensitive result is achieved fast because it removes the irrelevant attributes and retains the rest of them.

Moreover, the RSM works with such evaluation methods as Relief Attribute Evaluator (RAE), Gain Ratio, and Entropy. This present work practices the following combination as RSM and RAE to identify the significant attributes. By default, nothing has been discarded. The Ranker method rate the attribute high to low depends on the evaluation method's evaluation score. The RAE method detects the attribute dependencies by deriving the indirect interactions among all the nearest neighbor concept attributes. The attribute score is decreased when the value of the difference between attributes and neighbor instances in the same class.

Similarly, the evaluation score is increased in the neighbor instance, indifferent class. This filtering algorithm works without pairwise comparison; moreover, it can detect relevant attributes in large features of the dataset [19]. RAE can drive in both discrete and continuous data. It requires low polynomial time and also doesn't rest on a heuristic. While applying the combination of RSM and RAE in the present work evaluation, it recognized the relevant attributes as date, mean dollar invested age, development activity, market value to realized value, volume, and market cap.

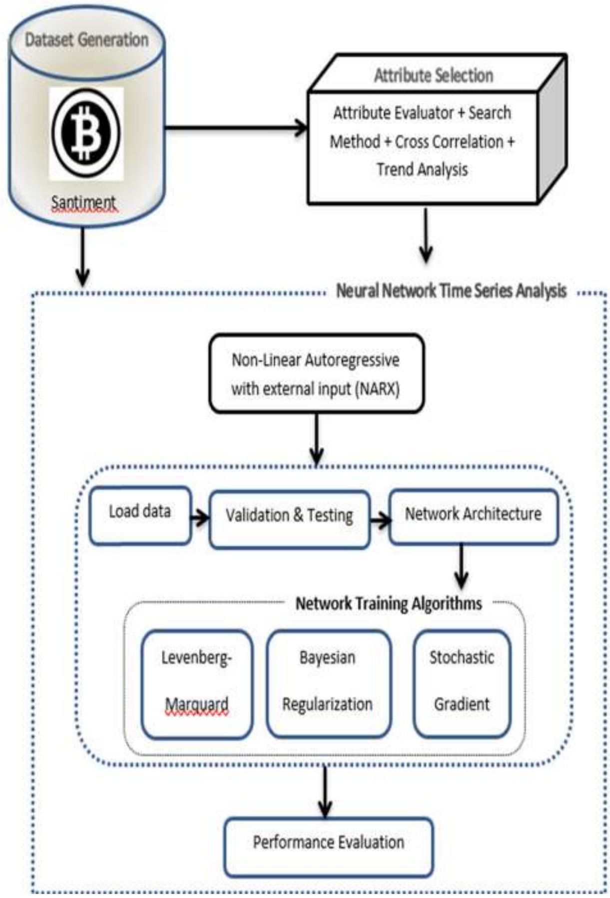

FIGURE 3. Proposed workflow.

Correlation-based Feature Subset Evaluator (Cfs Subset Eval) detects the worth of a subset of an attribute by deliberating the distinct predictive skill of an attribute along with the grade of redundancy between them. It prefers the low inter-correlation at the subset of attributes that are highly correlated with the target class. It ranks the attributes by evaluating the heuristic function based on correlation [20]. Low correlated attributes with class labels are discarded and relevant attributes are recognized with high correlation. When practicing this evaluation method with the Best First search method which lists out the relevant attributes such as date, social dominance, development activity, transaction volume, token velocity, and market value to realized value. According to the above evaluation mechanism, some attributes are selected as more significant: date, price, volume, market cap, social dominance, development activity, market value to realized value, and realized cap, by referring to the table. Especially, the realized cap attribute has been selected based on Kendall trend analysis and

TABLE 2. Attribute selection approach.

BestFirst+ CfsSubsetEval	Ranker+ CorrelationAttributeEval	Ranker+ ReliefAttributeEval
Date	Date	Date
Social Dominance	Market Cap	Mean Dollar Invested Age
Development Activity	Price	Development Activity
Transaction Volume	Social Dominance	Market Value To Realized Value
Token Velocity	Development Activity	Volume
Market Value to Realized Value	Market Value to Realized Value	Market Cap

cross-correlation. Other attributes are discarded in further processing. Table 2 illustrates the selection of attributes by various approaches. Potential attributes that contribute to the perfect model with the right combination are mandated.

C. TREND ANALYSIS

Generally, the Mann-Kendall Trend (MKT) test analyses the statistically significant trend in time series data, especially in meteorology [21]. It is a non-parametric test; hence the outlier data will not affect the results.

This rank-based MKT is applied to analyze the trend of specific crypto-currency attributes for a series of observations over time in this proposed work. The null hypothesis (H_0) denotes no statistical trend on particulars in the time series data. An alternative hypothesis (H_a) defines the monotonic increasing or decreasing trend specific over time. Consider $d_1, d_2, \ldots, d_t$ are t number of time series of data in particulars. For instance, d_j refers to the price value at time j. Next, the Mann-Kendall statistics (S) is calculated by the given equation 1 and 2.

$$S = \sum_{i=1}^{t-1} \sum_{j=i+1}^{t} sign(d_j - d_i) \tag{1}$$

$$sign(d_j - d_i) = \begin{pmatrix} 1, & (d_j - d_i) > 0 \\ -1, & (d_j - d_i) < 0 \\ 0, & (d_j - d_i) = 0 \end{pmatrix} \tag{2}$$

Here, we calculate the difference between every data at time t-1 and preceding data at time t. if the difference is a positive value, it will increase the trend, else if it is negative, it will constantly decrease the trend. It will be calculated for $t(t-1)/2$ pairs of data. Hence, the MKT statistics identify the significance of the trend. Next, calculate the variance of statistics is calculated by the following equation 3.

$$Variance(S) = \frac{1}{18}[t(t-1)(2t+5) - \sum_{k=1}^{m} g_k(g_k-1)(2g_k+5)] \tag{3}$$

In equation (3), t refers the total number of time series data, m refers the number of a tied group which contains a set of samples having identical value, g_k represents the total amount of data in the k^{th} group. Followed by, the normalized test statistic is calculated by applying Z-transformation like an equation 4.

$$Z = \begin{pmatrix} \frac{S-1}{Varience\,(S)^{\frac{1}{2}}}, & if\ S > 0 \\ 0, & if\ S = 0 \\ \frac{S+1}{Varience\,(S)^{\frac{1}{2}}}, & if\ S < 0 \end{pmatrix} \tag{4}$$

This MKT test is applied in the generated dataset. Table 3 lists the result of MKT for Bitcoin. Each row consists of Bitcoin and corresponding statistics (S), normalized statistics (Z), probability, and significance of trend respectively. As per the result, the value of the price attribute is increasing continuously over time. At the other end, the social volume is a decreasing trend, and no trend for token age consumed attribute. But the result might be varied in other crypto-currency coins.

D. CROSS-CORRELATION TEST

A cross-correlation test is applied to test the dependency among each attribute of Bitcoin. The table shows the correlation test result, highlighting the positive correlation (> 0.5) between two attributes. For instance, a timestamp is the dependent attribute for many attributes such as price, market cap, social dominance, development activity, and realized cap. An interesting thing is market cap and price has a strong relationship (=1) between them. Moreover, the market cap positively correlates with social dominance, market value to realized value, and realized cap. Table 4 provides insights into the relationship between the attributes based on the cross-correlation approach, and highlighted values show that more correlation persists for that attribute. The selected features are based on attribute evaluation + search method + Kendall trend analysis + cross-correlation.

The features considered based on the analysis are Price, Volume, Market Cap, Social Dominance, Development Activity, Market Value to Realized Value & Realized Cap.

IV. RESULTS AND DISCUSSIONS

Here we are presenting the experimental results achieved with all the details.

A. CORRELOGRAM

By visualizing the auto-correlation for the taken 7 variables by plotting 95% significance interval is too evident and supports the vital impact of this for analysis. As per the study, coefficients outside intervals are considered significant.

B. NEURAL NETWORK TIME SERIES ANALYSIS

A NN is an excellent choice to analyze this time series dataset. Especially, this kind of non-linear problem means variation of output is not proportionate to variation of inputs. NN is the collection of interconnected layers with enough neurons to model the dynamic systems with random accuracy. This work uses NARX, the recurrent dynamic and feed-forward

TABLE 3. Mann Kendall Test on Bitcoin.

Attribute of Bitcoin	S	Z	Variance	Trend prediction
Price	5649	7.8166	5.43E-15	statistically significant increasing trend
Volume	5113	7.0748	1.50E-12	statistically significant increasing trend
Market Cap	5751	7.9578	1.75E-15	statistically significant increasing trend
Mean Dollar Invested Age	2021	2.7956	0.00518	statistically significant increasing trend
Social Volume	-2380	3.2924	0.000993	statistically significant decreasing trend
Social Dominance	6427	8.8933	5.93E-19	statistically significant increasing trend
Development Activity	4776	6.6086	3.88E-11	statistically significant increasing trend
Transaction Volume	-3759	5.2009	1.98E-07	statistically significant decreasing trend
Token Age Consumed	229	0.3155	0.75235	no statistically significant trend
Token Velocity	-3163	4.3761	1.21E-05	statistically significant decreasing trend
Token Circulation	-1733	2.397	0.016529	statistically significant decreasing trend
Market Value to Realized Value	4085	5.6521	1.59E-08	statistically significant increasing trend
Realized Cap	11366	15.729	9.61E-56	statistically significant increasing trend

TABLE 4. The relationship among the attribute.

	Timestamp	*Price*	*Volume*	*Market Cap*	*MDIA*	*SV*	*SD*	*DA*	*TV*	*TAC*	*TVE*	*TC*	*MVRV*	*RC*
Timestamp	1													
Price	0.65	1												
Volume	0.47	0.64	1											
Market Cap	0.67	1	0.64	1										
Mean Dollar Invested Age (MDIA)	0.14	-0.3	0.23	-0.3	1									
Social Volume (SV)	-0.2	0.23	0.38	0.22	0.13	1								
Social Dominance (SD)	0.66	0.71	0.67	0.72	-0.1	0.2	1							
Development Activity (DA)	0.53	0.26	0.16	0.27	-0.1	-0	0.3	1						
Transaction Volume (*TV*)	-0.3	0	0.18	0	0.17	0.6	0	-0	1					
Token Age Consumed (TAC)	0.06	0.17	0.2	0.17	0	0.1	0.1	0	0	1				
Token Velocity (TVE)	-0.2	-0.1	0	-0.1	0.06	0.2	0	-0	0.4	-0.1	1			
Token Circulation (TC)	-0.2	0	0.36	0	0.15	0.3	0	-0	0.3	0.53	-0.1	1		
Market Value to Realized Value (MV-RV)	0.48	0.97	0.67	0.96	-0.3	0.4	0.7	0.1	0.1	0.18	0.02	0.1	1	
Realized Cap (RC)	0.95	0.69	0.36	0.7	-0.2	-0	0.6	0.6	-0	0.06	-0.2	-0	0.49	1

network. It predicts the next value of given input and is used for the non-linear filtering of a dynamic system. Commonly, this prediction model is employed by historical values of known input and output data series. This model's training phase uses the output as one of the inputs to train a model and repeats as a static backpropagation system to acquire an

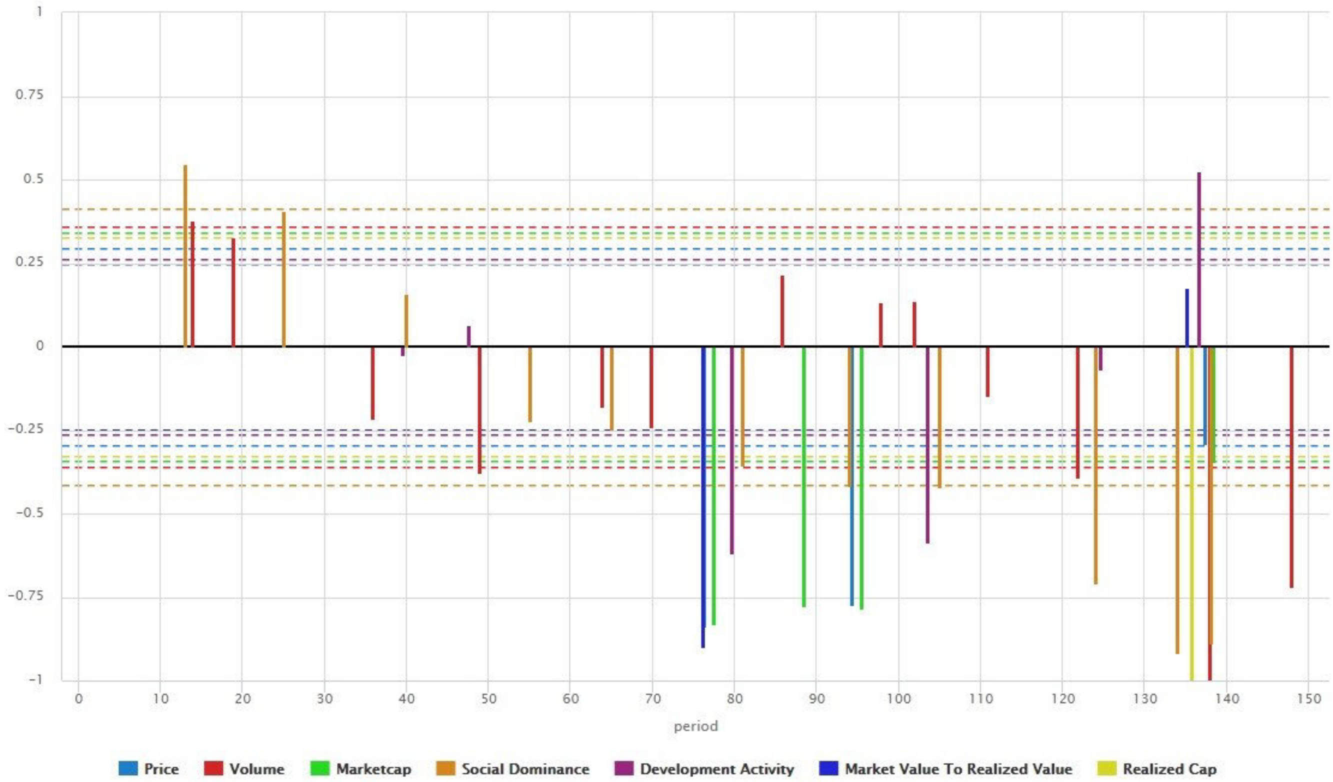

FIGURE 4. Correlogram for 7-attributes.

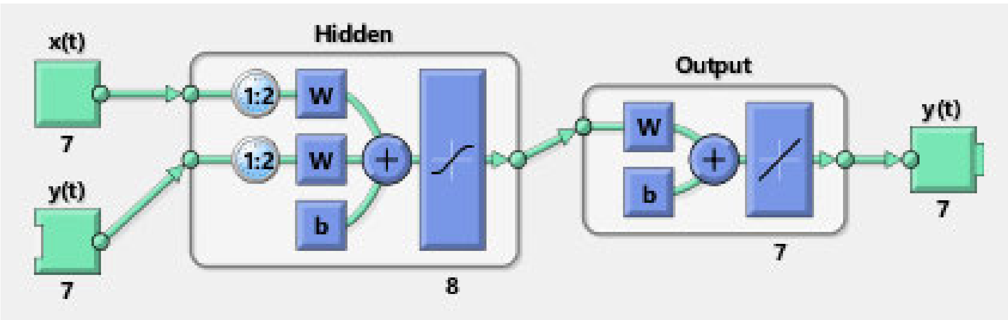

FIGURE 5. NARX on NN.

accurate result. The figure depicts the general NARX model recurrent connections of inputs and feedbacks. Here, n number of input time series in the Input layer as passive nodes just given single input into multiple outputs to the hidden layer. Hidden layers and output layers consist of active nodes activated by activation functions such as linear, sigmoid (0-1), and hyperbolic tangent. In general, a hyperbolic tangent is applied in a hidden layer in the range between -1 and 1, and linear in the output layer. This NN is trained at each stage and produces different outputs depending on weight, bias, and other sets of data series employed for training, validation, and testing. Especially, weight is the deciding factor to decide the class label of input, and bias is the adjusting factor. NARX predicted the target given by historical values of output time series and input time series and depicted. Figure 5 illustrates the structure of our proposed work in the NARX environment.

Figure 4 depicts the visualization of the auto-correlation.

This research study employed three training algorithms to train a NN as LM, BR, and SCG algorithm. Generally, LM is used to solve the non-linear least square curve-fitting problem. It is the combination of two minimization methods which are Gauss-Newton and gradient descent methods. It is one of the fastest algorithms with high memory utilization. But SCG is popular because of its simplicity but slow execution in nature and suitable for the simple objective function. BR algorithm is used to regularize the NN using Bayesian techniques and determines the optimal parameters. These

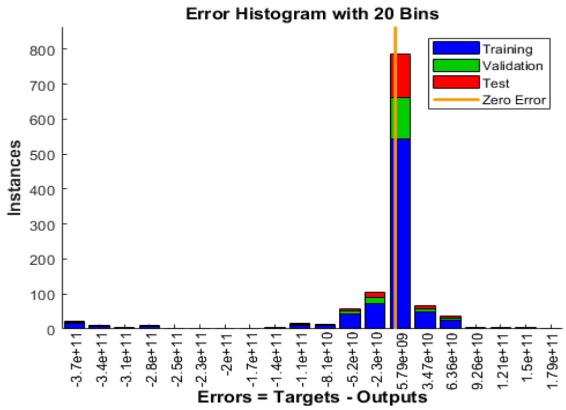

FIGURE 6. Error Histogram - LM.

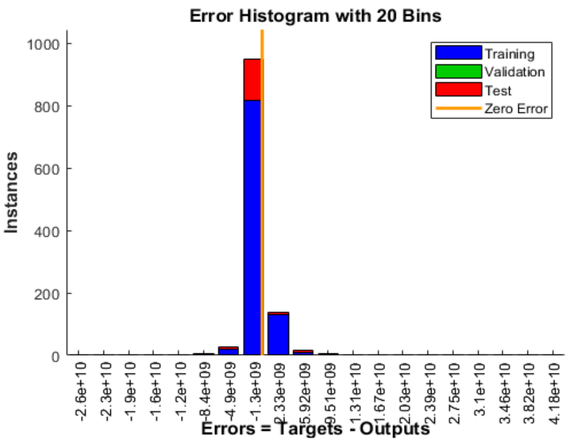

FIGURE 7. Error Histogram – BR.

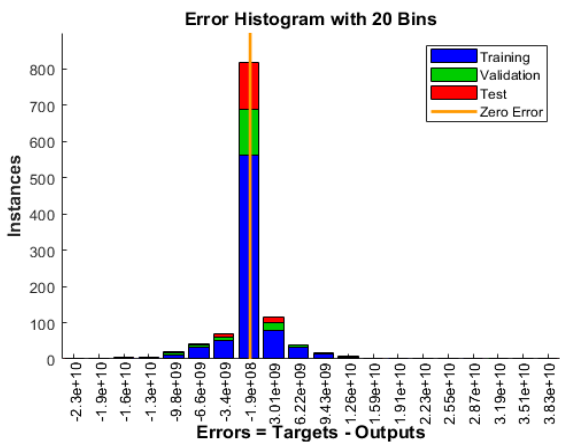

FIGURE 8. Error Histogram – SCG.

three training algorithms are practiced for this research study, and further results are discussed in the next section. The training process will stop automatically with no improvement in generalization, which is directed by increasing the validation trials' mean square error. When doing the initial configuration and data samples changes, the NN is trained and produced different results.

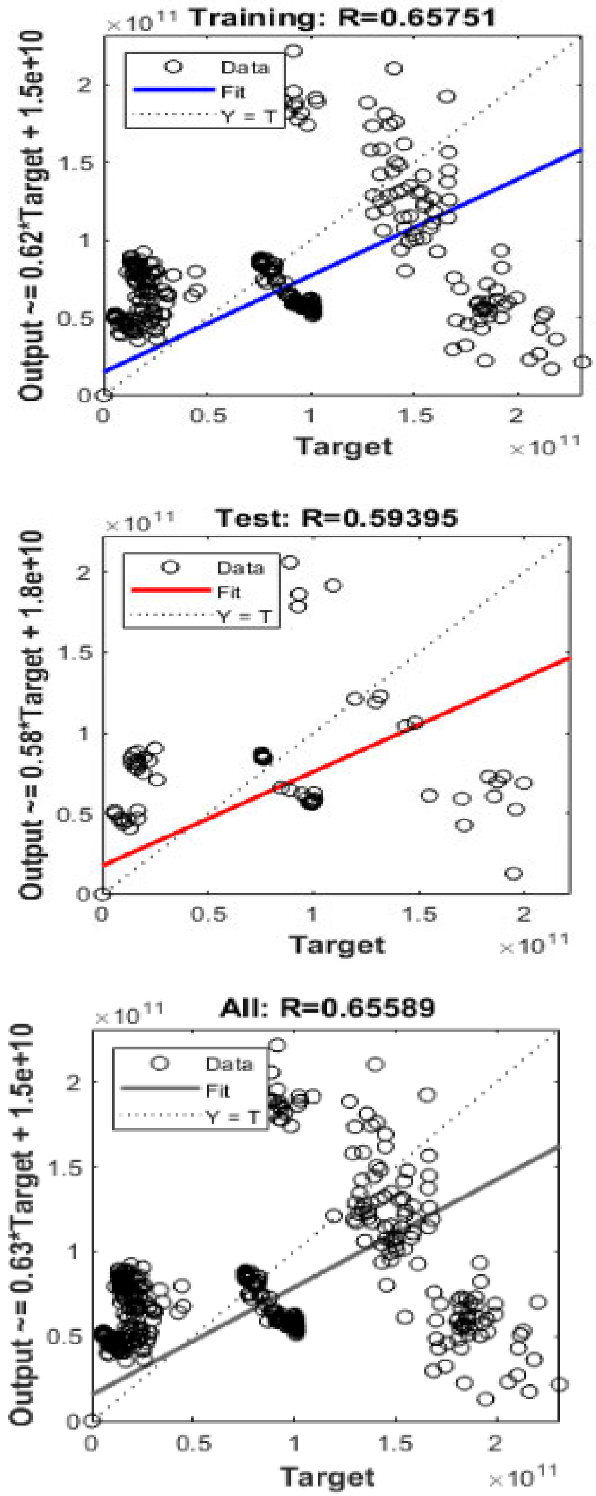

FIGURE 9. Regression – LM.

C. PERFORMANCE EVALUATION

For experimentation purposes, we have considered MATLAB R2020a. The original dataset and the reduced dataset are taken into consideration to make the analysis. For validation and testing purposes, randomly select the data as 70%

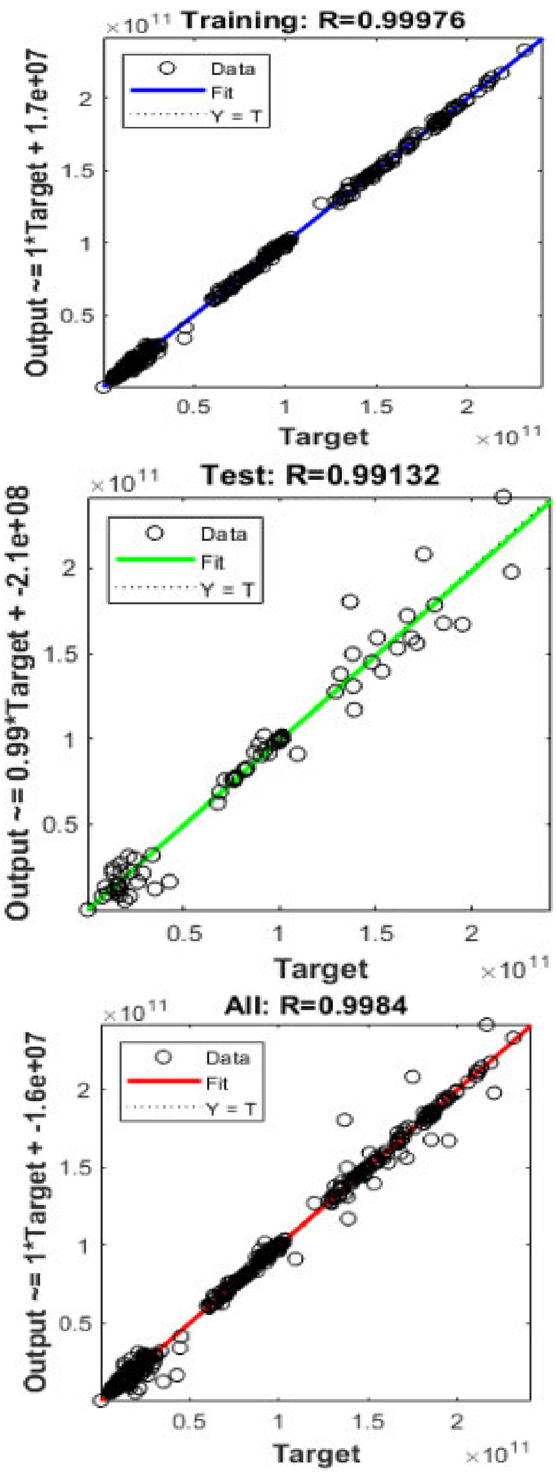

FIGURE 10. Regression – BR.

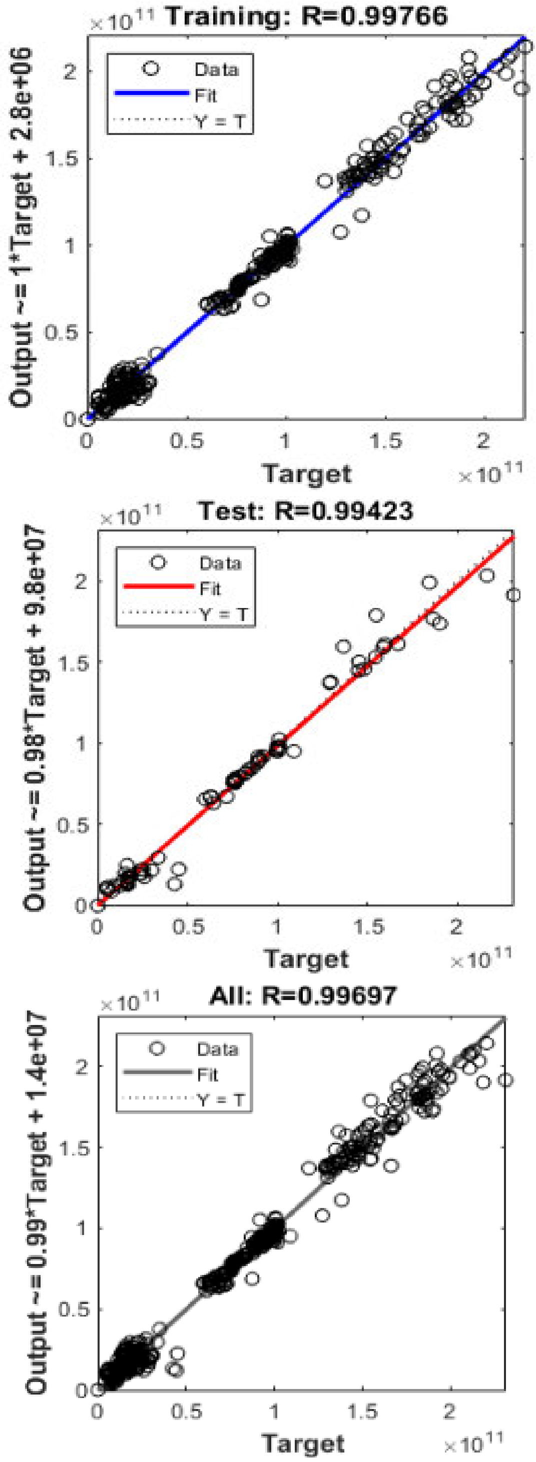

FIGURE 11. Regression – SCG.

for training and 15% for validation and testing individually. In our proposed approach, eight hidden neurons and two delays were utilized.

The training data, testing data, and validation data are indicated as blue, red, and green colors, respectively. As shown in Figure 6, for LM, all validation testing errors lie between −5.2e+10 and 6.36e+10. Similarly, in Figure 7, for BR −4.9e+09 and 5.9e+09 and in case of Figure 8. for SCG −9.8e+09 and 9.43e+09. It's very evident from the plots that BR has minimal errors, and this NARX-BR model is suitable for this system.

The error histogram for LM is presented in Figure 6.

The error histogram for BR is presented in Figure 7.

The error histogram for SCG is presented in Figure 8. For analysis perspective training, testing, and all plots are shown. The performance of BR and SCG is overwhelming the LM by the way of best fit around 0.99 in all the cases. On overall consideration, the NARX-BR model is the best suit for the system. The regression plot for LM, BR, and SCG are illustrated in Figure 9 to Figure 11.

V. CONCLUSION

Data-driven analysis is the success of many industries working with the help of historical datasets. Many of the markets use the hefty segregate data for deciding the near future. The bitcoin protocol is the key real-world application that raises the height of blockchain technology. Various attributes are playing a role, and in the present work, applying the attribute selection methods and trend analysis narrowed down to seven attributes for analysis. The non-linear autoregressive network with exogenous inputs is used, and it is a recurrent dynamic network. For BR, LM, and SCG all validation, testing errors lie between −4.9e+09 and 5.9e+09, for LM, −5.2e+10 and 6.36e+10, −9.8e+09 and 9.43e+09 respectively. BR possesses minimal errors. Even the regression value on the overall perspective, BR is showing a higher value of 0.9984. Error histogram and regression plots indicate that the BR NN is showing good performance.

REFERENCES

[1] S. Nakamoto. (2019). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Manubot. [Online]. Available: https://bitcoin.org/en/bitcoin-paper

[2] K. Prewett, R. W. Dorsey, and G. Kumar, "A primer on taxation of investment in cryptocurrencies," *J. Taxation Investments*, vol. 36, no. 4, pp. 3–16, 2019.

[3] A. F. Sheta, S. E. M. Ahmed, and H. Faris, "A comparison between regression, artificial neural networks and support vector machines for predicting stock market index," *Soft Comput.*, vol. 7, no. 8, p. 2, 2015, doi: 10.14569/IJARAI.2015.040710.

[4] L. Alessandretti, A. ElBahrawy, L. M. Aiello, and A. Baronchelli, "Anticipating cryptocurrency prices using machine learning," *Complexity*, vol. 2018, p. 16, Nov. 2018.

[5] 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, 2017.

[6] D. H. Kwon, J. B. Kim, J. S. Heo, C. M. Kim, and Y. H. Han, "Time series classification of cryptocurrency price trend based on a recurrent LSTM neural network," *J. Inf. Process. Syst.*, vol. 15, no. 3, pp. 694–706, 2019.

[7] K. Zbikowski, "Application of machine learning algorithms for Bitcoin automated trading," in *Machine Intelligence and Big Data in Industry*. Cham, Switzerland: Springer, 2016, pp. 161–168.

[8] S. H. Kang, R. P. McIver, and J. A. Hernandez, "Co-movements between Bitcoin and Gold: A wavelet coherence analysis," *Phys. A, Stat. Mech. Appl.*, vol. 536, Dec. 2019, Art. no. 120888.

[9] S. Corbet, B. Lucey, A. Urquhart, and L. Yarovaya, "Cryptocurrencies as a financial asset: A systematic analysis," *Int. Rev. Financial Anal.*, vol. 62, pp. 182–199, Mar. 2019.

[10] A. Gross, J. Hemker, J. Hoelscher, and B. Reed, "The role of secondary sources on the taxation of digital currency (Bitcoin) before IRS guidance was issued," *J. Accounting Educ.*, vol. 39, pp. 48–54, Jun. 2017.

[11] H. Wang, D. He, and Y. Ji, "Designated-verifier proof of assets for Bitcoin exchange using elliptic curve cryptography," *Future Gener. Comput. Syst.*, vol. 107, pp. 854–862, Jun. 2020.

[12] L. Kristoufek, "BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era," *Sci. Rep.*, vol. 3, no. 1, pp. 1–7, 2013.

[13] M. Poongodi, A. Sharma, V. Vijayakumar, V. Bhardwaj, A. P. Sharma, R. Iqbal, and R. Kumar, "Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system," *Comput. Electr. Eng.*, vol. 81, Jan. 2020, Art. no. 106527.

[14] D. C. A. Mallqui and R. A. S. Fernandes, "Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques," *Appl. Soft Comput.*, vol. 75, pp. 596–606, Feb. 2019.

[15] 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 Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113250.

[16] A. Altan, S. Karasu, and S. Bekiros, "Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques," *Chaos, Solitons Fractals*, vol. 126, pp. 325–336, Sep. 2019.

[17] N. I. Indera, I. M. Yassin, A. Zabidi, and Z. I. Rizman, "Non-linear autoregressive with exogeneous input (NARX) Bitcoin price prediction model using PSO-optimized parameters and moving average technical indicators," *J. Fundam. Appl. Sci.*, vol. 9, no. 3, pp. 791–808, 2017.

[18] N. A. Bakar and S. Rosbi, "Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of Bitcoin transaction," *Int. J. Adv. Eng. Res. Sci.*, vol. 4, no. 11, pp. 130–137, 2017.

[19] R. J. Urbanowicz, M. Meeker, W. L. Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," *J. Biomed. Informat.*, vol. 85, pp. 189–203, Sep. 2018.

[20] A. Wosiak and D. Zakrzewska, "Integrating correlation-based feature selection and clustering for improved cardiovascular disease diagnosis," *Complexity*, vol. 2018, pp. 1–11, Oct. 2018.

[21] I. Ahmad, D. Tang, T. Wang, M. Wang, and B. Wagan, "Precipitation trends over time using Mann-Kendall and Spearman's Rho tests in Swat river basin, Pakistan," *Adv. Meteorol.*, vol. 2015, pp. 1–15, Jan. 2015.

[22] M. H. U. Rehman, K. Salah, E. Damiani, and D. Svetinovic, "Trust in blockchain cryptocurrency ecosystem," *IEEE Trans. Eng. Manag.*, vol. 67, no. 4, pp. 1196–1212, Nov. 2020, doi: 10.1109/TEM.2019.2948861.

[23] H. H. Sun Yin, K. Langenheldt, M. Harlev, R. R. Mukkamala, and R. Vatrapu, "Regulating cryptocurrencies: A supervised machine learning approach to de-anonymizing the Bitcoin blockchain," *J. Manage. Inf. Syst.*, vol. 36, no. 1, pp. 37–73, Jan. 2019.

[24] L. V. T. Duong, N. T. T. Thuy, and L. D. Khai, "A fast approach for Bitcoin blockchain cryptocurrency mining system," *Integration*, vol. 74, pp. 107–114, Sep. 2020.

[25] E. Yi, Y. Cho, S. Sohn, and K. Ahn, "After the splits: Information flow between Bitcoin and Bitcoin family," *Chaos, Solitons Fractals*, vol. 142, Jan. 2021, Art. no. 110464.

[26] S. Bouraga, "A taxonomy of blockchain consensus protocols: A survey and classification framework," *Expert Syst. Appl.*, vol. 168, Apr. 2021, Art. no. 114384.

[27] S. M. Jang, E. Yi, W. C. Kim, and K. Ahn, "Information flow between Bitcoin and other investment assets," *Entropy*, vol. 21, no. 11, p. 1116, Nov. 2019.

[28] CoinMarketCap. *Cryptocurrency Prices, Charts and Market Capitalizations*. Accessed: Jan. 10, 2021. [Online]. Available: https://coinmarketcap.com/

[29] BitCoinCharts. *Bit Coin Charts*. Accessed: Jan. 10, 2021. [Online]. Available: https://bitcoincharts.com/

[30] Brave New Coin. *Brave New Coin*. Accessed: Jan. 10, 2021. [Online]. Available: https://bravenewcoin.com/

[31] CoinDesk. *CoinDesk—Leader in Crypto and Blockchain News and Information*. Accessed: Jan. 10, 2021. [Online]. Available: https://www.coindesk.com/

[32] Quandl. *Quandl*. Accessed: Jan. 11, 2021. [Online]. Available: https://www.quandl.com/

[33] Blockchain *Explorer | BTC | ETH | BCH. Blockchain Explorer*. Accessed: Jan. 11, 2021. [Online]. Available: https://www.blockchain.com/explorer

[34] CryptoCompare. *Cryptocurrency Prices, Portfolio, Forum, Rankings*. Accessed: Jan. 11, 2021. [Online]. Available: https://www.cryptocompare.com/

[35] *Winklevoss Index*. Accessed: Jan. 11, 2021. [Online]. Available: https://winkdex.com/

[36] BitcoinAverage. *BitcoinAverage*. Accessed: Jan. 11, 2021. [Online]. Available: https://bitcoinaverage.com/

[37] J. M. Chatterjee, L. H. Son, S. Ghatak, R. Kumar, and M. Khari, "BitCoin exclusively informational money: A valuable review from 2010 to 2017," *Qual. Quantity*, vol. 52, no. 5, pp. 2037–2054, Sep. 2018.

[38] R. Agrawal, J. M. Chatterjee, A. Kumar, and P. S. Rathore, Eds., *Blockchain Technology and the Internet of Things: Challenges and Applications in Bitcoin and Security*. Boca Raton, FL, USA: CRC Press, 2020.

[39] R. Gencay and M. Qi, "Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging," *IEEE Trans. Neural Netw.*, vol. 12, no. 4, pp. 726–734, Jul. 2001.

R. SUJATHA received the B.E. degree in computer science from Madras University, in 2001, the M.E. degree in computer science from Anna University, in 2009, with university ninth rank, the master's degree in financial management from Pondicherry University, in 2005, and the Ph.D. degree in the area of data mining from the Vellore Institute of Technology (VIT), Vellore, in 2017. She has 15 years of teaching experience and has been serving as an Associate Professor with the School of Information Technology and Engineering, VIT. She has organized and attended several workshops and faculty development programs. She actively involves herself in the growth of the institute by contributing to various committees at both academic and administrative levels. She gives technical talks in colleges for the symposium and various sessions. She acts as an advisory, editorial member, and technical committee member in conferences conducted in other educational institutions and in-house too. She has published a book *Software Project Management* for college students. Also, she has published research articles and papers in reputed journals. She used to guide projects for undergraduate and postgraduate students and currently guides doctoral students. She is interested in learning upcoming things and gets herself acquainted with the student's level. Her areas of research interests include data mining, machine learning, software engineering, soft computing, big data, deep learning, and blockchain.

V. MAREESWARI is currently working as Assistant Professor (Senior) with the School of Information Technology and Engineering (SITE), Vellore Institute of Technology (VIT), Vellore, India. She has published several national and international articles in reputed journals and conferences. Her areas of interests include programming in web technologies, web services, cloud computing, image processing, and data analytics.

JYOTIR MOY CHATTERJEE received the B.Tech. degree in computer science and engineering from Dr. MGR Educational and Research Institute, Maduravoyal, Chennai, in 2013, and the M.Tech. degree in computer science and engineering from the Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar, Odisha, in 2016. He worked as an Assistant Professor with the Department of Computer Science Engineering, G. D. Rungta College of Engineering and Technology (Chhattisgarh Swami Vivekananda Technical University), Bhilai, India. He is currently working as an Assistant Professor with the Department of Information Technology, Lord Buddha Education Foundation (Asia Pacific University of Innovation & Technology), Kathmandu, Nepal. He is the Young Ambassador of the Scientific Research Group of Egypt (SRGE) (2020-2021). He has been selected as Top 1% Peer Reviewer in Computer Science on Publons global reviewer database 2019 powered by the Web of Science Group and is also serving as the Publons Academy Mentor. His research interests include machine learning and the Internet of Things.

ABD ALLAH A. MOUSA received the M.Sc. and Ph.D. degrees in engineering mathematics from Menofia University, Egypt, in 2003 and 2006, respectively. He is currently a Full Professor of Engineering Mathematics with the Department of Basic Engineering Sciences, Faculty of Engineering, Menofia University, and at the Department of Mathematics and Statistics, Faculty of Science, Taif University, Saudi Arabia. He is the author of more than 90 scientific articles and textbooks in refereed journals and international conferences. He has served as an editorial board member and a referee for many reputed international journals. His research interests include theory of optimization and its application, numerical analysis, mathematical programming, metaheuristics, and computational intelligence.

ABOUL ELLA HASSANIEN is the Founder and Head of the Egyptian Scientific Research Group (SRGE) and a Professor of Information Technology at the Faculty of Computer and Artificial Intelligence, Cairo University. His research areas include computational intelligence, medical image analysis, security, animal identification, and multimedia data mining.