International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

A Review on Forecasting Machine Learning Models for Predicting Crypto Prices

Vicky Soni[1], Prof. C.K.Tiwari[2], Mr. Devkinandan Nagar[3]
M.Tech Scholar, Patel College of Science and Technology, Indore[1]
Associate Professor, Patel College of Science and Technology, Indore[2]
Assistant Professor, Patel College of Science and Technology, Indore[3]

Abstract: Crypto price prediction is a category of time series prediction which extremely challenging due to the dependence of crypto prices on several financial, socio-economic and political parameters etc. Moreover, small inaccuracies in crypto price predictions may result in huge losses to firms which use crypto price prediction results for financial analysis and investments. Conventional statistical methods render substantially lesser accuracy compared to new age machine learning techniques. This machine learning based techniques are being used widely for crypto price prediction due to relatively higher accuracy compared to conventional statistical techniques. This paper presents a review on contemporary data driven approaches for crypto currency forecasting highlighting the salient attributes. Moreover, the identified non-trivial research gap in the existing approaches has been used as an underpinning for subsequent direction of research in the domain. The paper culminates with the performance metrics and concluding remarks.

Keywords: Crypto price Forecasting, Data Driven Models, Regression Analysis, Machine Learning, Performance Metrics.

1. Introduction

With increasing digitization and resource distribution, cryptocurrencies have gain significant importance. This has led to large scale investments in cryptocurrencies such as Bitcoin, Etherium etc.. However, crypto prices are extremely random, fluctuating and volatile in nature which makes investments risk prone. Moreover, previous crypto data often exhibits random fluctuations, volatility and deviation from a particular trend, which is often termed as noise [2]. This noisy behavior makes pattern recognition difficult leading to inaccuracies in forecasting results. Hence, it is necessary to filter out the baseline noise from the time series crypto data prior to applying the data to any machine learning or deep learning model for pattern recognition [3]. While crypto trend analysis is a time series regression problem, what makes it extremely complicated is the dependence on several non-numeric parameters such as socio economic conditions, political conditions of a country, political stability, financial crisis and trade wars, global slowdown and public sentiments pertaining to a company etc. This leads to variabilities in the stock trends often exhibiting non-coherent patterns with respect to historical data [4]. Figure 1.1 depicts the common crypto currencies.

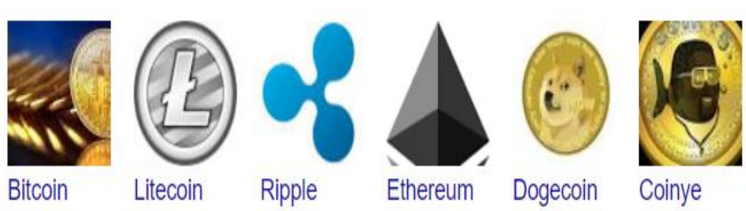

Fig.1 Common Crypto Currencies

Global events, such as natural disasters, geopolitical tensions, or pandemics, can also impact crypto trends. These events can create uncertainty in the markets and cause investors to become more risk-averse and hence it is essential to note that data trends are inherently variable and can be influenced by a wide range of factors It's also worth mentioning that past performance does not guarantee future results, so investors should exercise caution when making investment decisions based on historical trends. Cryptocurrency prediction is basically a time series prediction problem. Mathematically [5]:

$$\boldsymbol{P} = \boldsymbol{f}(\boldsymbol{t}, \boldsymbol{v}) \qquad \textbf{(1)}$$

Here,
P represents crypto price
f represents a function of
t is the time variable
v are other influencing global variables

The dependence of crypto prices over time makes it somewhat predictable under similar other conditions of global influencing variables. However, even the slightest of changes can derail the prediction completely. Hence crypto market prediction extremely challenging due to the dependence of crypto prices on several financial, socio-economic and political parameters etc. Moreover, small inaccuracies in crypto price predictions may result in huge losses to firms which use crypto price prediction results for financial analysis and investments. Hence it has remained a serous research problem for researchers since [6]:

1) Inaccuracies in crypto market prediction can lead to large financial losses for individuals and firms aiming to leverage crypto market data.
2) It is a complex time series problem
3) It is difficult to find trends in such diverse parameters which affect crypto prices.

Crypto price movement is often very random and dynamic in nature; the prices are varying in fraction of seconds. Additionally each time the changes in prices are unpredictable. These changes can be positive or can be negative according to the market outlooks.

II. Factors Affecting Crypto Price

A crypto price may depend on several factors operating in the current world and crypto market. We will try to take into account a combination of mainly two factors [7]:

- The impact of global influencing factors.
- The past performances and records of the target company.

Therefore in this proposed study the prediction of crypto price influencing feature is the key aim and objective. Additionally it is also required to utilize these features during the time prediction to improve the prediction accuracy of the systems. Therefore the proposed work involves the study of machine learning and data mining techniques (supervised and unsupervised) by which the prediction is feasible. In addition of that need to implement

the additional methodology that accurately analyze the crypto market influencing features and involve these factors to reduce the error in prediction data (i.e. error minimization techniques or optimization techniques [8]
Finally after implementation of the proposed technique, it is required that the error rate has to be squeeze out to justify the proposed work. Therefore the proposed work involves the comparative study of the proposed technique with the similar available techniques. In addition of a case study with a company crypto prices also involved with the proposed study work. This section provides the basic overview of the proposed crypto price prediction study the next section provides the core aim and objectives of the proposed work. Statistical techniques do not render high accuracy of prediction and hence are ineffective in prediction problems which need low errors in prediction. The emergence of machine learning (ML) has offered promising tools for understanding and predicting complex patterns in these markets. By analyzing vast datasets including historical prices, trading volumes, social media sentiment, and macroeconomic indicators, ML models can generate predictive insights. However, despite their potential, applying ML to crypto price prediction is fraught with numerous technical and conceptual challenges that limit model accuracy and reliability [9]

One of the primary challenges in crypto price prediction using ML is the market's inherent volatility. Prices of cryptocurrencies can fluctuate dramatically within minutes due to speculative trading, regulatory news, or even social media rumors. This high degree of unpredictability introduces substantial noise in the data, making it difficult for ML models to distinguish between meaningful patterns and random fluctuations. For cryptocurrency, relevant data may include price history, trading volume, blockchain metrics (on-chain data), and sentiment analysis from social platforms. However, collecting and preprocessing this data poses several challenges. Data sources may be inconsistent, unstructured, or contain missing values. These models often function as black boxes, making it difficult to understand the rationale behind a prediction. For financial forecasting, where explainability is crucial for risk assessment and decision-making, the lack of transparency can hinder user trust and model adoption. Moreover, identifying the right features that influence price movements is a non-trivial task. Poorly chosen features or irrelevant indicators can degrade model performance, while important but hidden variables may be ignored [10].

III. Related Work

This section presents the existing work in the domain:

Han et al. [11] proposed the multimodal fusion Bitcoin (MFB), an innovative generalized multimodal fusion approach that effectively integrates BiLSTM and BiGRU layers for complex feature extraction. The model employs the BorutaShap algorithm for feature selection and utilizes attention mechanisms and spatial dropout for optimization and generalization. MFB's training and validation use news and tweet data combined with Bitcoin technical indicators to explore the impact of time-lagged sentiment on price movements, leading to more accurate and timely market predictions. The MFB performs superior Bitcoin prediction performance, MAPE of 2.66%.

Patel et al. [12] proposed bidirectional LSTM with preprocessed data. We have performed data cleansing and sentiment analysis as part of the data pretreatment step. Recent 90-days data have been selected for the prediction which is then fed to a bidirectional LSTM model. It is a specialized version of recurrent neural network (RNN) architecture that is often used in sequence-to-sequence tasks, including time series prediction. It can be applied to Bitcoin price prediction by leveraging its ability to capture dependencies in both past and future data. MAPE of 2.7% obtained.

Rafi et al. [13] proposed a price forecasting model based on three vital characteristics (i) a feature selection and weighting approach based on Mean Decrease Impurity(MDI) features. (ii) Bi-directional LSTM and (iii) with a trend preserving model bias correction (CUSUM control charts for monitoring the model performance over time) to forecast Bitcoin and Ethereum values for long and short term spans. On a new test-set collected from January 01, 2020 to January 01, 2022 for the two cryptocurrencies we obtained an average RSME of 9.17, with model bias correction, Comparing with the prevalent forecasting models we report a new state of the art in cryptocurrency forecasting.

Kim et al. [14] proposed a novel framework that predicts the price of Bitcoin (BTC), a dominant cryptocurrency. For stable prediction performance in unseen price range, the change point detection technique is employed. In particular, it is used to segment time-series data so that normalization can be separately conducted based on segmentation. Furthermore, this work proposes self-attention-based multiple long short-term memory (SAM-LSTM), which consists of multiple LSTM modules for on-chain variable groups and the attention mechanism, for the prediction model. Experiments with real-world BTC price data and various method setups have proven the proposed framework's effectiveness in BTC price prediction. The results are promising, with the highest MAE, RMSE, MSE, and MAPE values of 0.3462, 0.5035, 0.2536, and 1.3251, respectively.

Shahbazi et al. [15] showed that during recent developments, cryptocurrency has become a famous key factor in financial and business opportunities. However, the cryptocurrency investment is not visible regarding the market's inconsistent aspect and volatility of high prices. Due to the real-time prediction of prices, the previous approaches in price prediction doesn't contain enough information and solution for forecasting the price changes. Based on the mentioned problems in cryptocurrency price prediction, authors proposed a machine learning-based approach to price prediction for a financial institution. The main focus of this system is on Litecoin and Monero cryptocurrencies. The results show the presented system accurate the performance of price prediction higher than another state-of-art algorithm.

Mudassir et al. [16] proposed a high-performance machine learning-based classification and regression models for predicting Bitcoin price movements and prices in short and medium terms. In previous works, machine learning-based classification has been studied for an only one-day time frame, while this work goes beyond that by using machine learning-based models for one, seven, thirty and ninety days. For daily price forecast, the error percentage is as low as 1.44%, while it varies from 2.88 to 4.10% for horizons of seven to ninety days. These results indicate that the presented models outperform the existing models in the literature.

A summary of existing literature in the domain with the salient points are presented.

Table 1. Summary of Literature Review

S.No.	Paper Title	Authors	Publication	Findings
1	MFB: A Generalized Multimodal Fusion Approach for Bitcoin Price Prediction Using Time-Lagged Sentiment and Indicator Features	Han et al. [11]	Elsevier 2025	Multimodal Fusion Approach combining layers of Bi-LSTM and GRU. It also included stochastic sentiment features. Best case MAPE of 2.66 obtained.
2	A deep learning framework for hourly bitcoin price prediction using bi-LSTM and sentiment analysis of Twitter data	Patel et al. [12]	Springer 2024.	Recurrent Neural Nets (LSTM and B-LSTM) models trained on historical bitcoin prices as well as Twitter Opinion Data. Base case MAPE of 2.7% obtained.
3	Enhancing Cryptocurrency Price Forecasting Accuracy: A Feature Selection and Weighting Approach With Bi-Directional LSTM and Trend-Preserving Model Bias Correction	Rafi et al. [13]	IEEE 2023	LSTM with model bias correction method for forecasting Bitcoin Prices. The Mean Absolute Percentage Error (MAPE) of 0.66 was achieved.
4	A Deep Learning-Based Cryptocurrency Price Prediction Model That Uses On-Chain Data	Kim et al. [14]	IEEE 2022	Self-Attention based Multiple LSTM (SAM-LSTM) model used to forecast Bitcoin Prices. MAPE of 1.3251% obtained.
5	Improving the Cryptocurrency Price Prediction Performance Based on Reinforcement Learning	Shahbazi et al. [15]	IEEE 2021	Forecasting crypto prices based on historical data and reinforcement learning (Q-Learning). Best case MAPE of 3.16% and 5.95% for a 3-day and 7-day forecast window
6.	Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach	Mudassir et al. [16]	Springer 2020	Random Forecast and Principal Component Analysis (PCA) for forecasting crypto prices. Best case MAPE of 1.44% for daily forecast.

IV. Baseline Models

The noteworthy contributions in the field of the proposed work can be summarized as:

1) Neural Networks: One of the most commonly adopted methodologies in the domain of crypto market forecasting is the use of neural networks which mimic the learning of the human brain. The most common neural network categories used off late have been:
a) Long Short Term Memory (LSTM) neural networks: These neural networks are capable to weight how much weightage is to be given to the recent data and the data of the far past. It is often shown to work effectively for time series prediction [17]-[18].
2) Incorporating sentiment analysis from platforms like Twitter, Reddit, and news articles adds another layer of complexity. Natural language processing (NLP) techniques must handle slang, sarcasm, multilingual posts, and noisy textual data [19]. Additionally, real-time processing of sentiment data with high temporal relevance is computationally demanding. Despite advancements in transformer-based models (e.g., BERT), accurately quantifying sentiment and linking it to price action remains a challenging task [20]

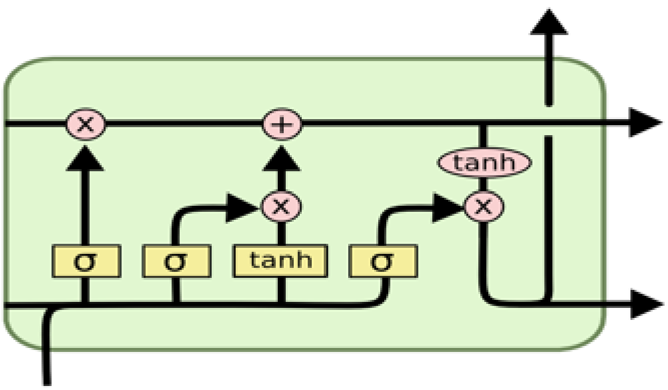

Fig.2 The Structure of the LSTM Neural Network

3) Back Propagation: This neural network structure has a feedback path for errors in prediction. It is often effective in time series prediction problems since the estimation of future samples is based on both previous data as well as the errors in the previously prediction samples [21].

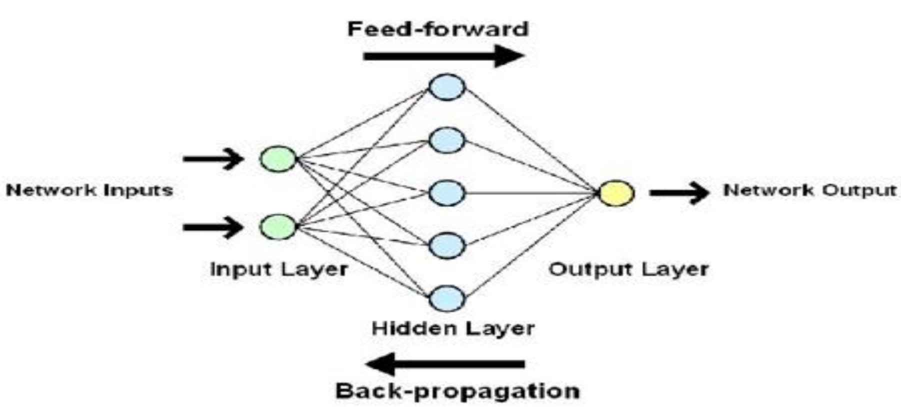

Fig.3 The Structure of the BP Neural Network

4) Regression Models: A variety of regression models have been used for time series prediction among which the most common ones are [22]:

a) Linear
b) Non-Linear
c) Polynomial
d) SVR (Support Vector Regression)

In this approach, the relationship between the independent and dependent variable is found using different mathematical or statistical models depending on the nature of the approach used [23]-[24]. The most common type of regression model can be thought of as the linear regression model which is mathematically expressed as [25]:

$$\boldsymbol{y = \theta_1 + \theta_2 x} \quad \textbf{(2)}$$

Here,
x represenst the state vector of inut variables
y rperesenst the state vector of output variable or variables.
Ɵ1 and Ɵ2 are the co-efficents which try to fit the regression learning models output vector to the input vector.
The aim of the approach is to attain the best fit regression line which is equivalent to saying that the co-efficient values θ1 and θ2 , should be adjusted such that to minimize the error between predicted y value (pred) and true y value (y). The cost function J is mathematically defined as [26]:

$$\boldsymbol{J = \frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2} \quad \textbf{(3)}$$

Here,
n is the number of samples
y is the target
pred is the actual output.

While machine learning models offer promising results in cryptocurrency price forecasting, several challenges remain [27].

1. The unpredictable nature of the crypto market, including sudden crashes and pumps, limits the accuracy of even the most advanced models [28].
2. Cryptocurrency markets are highly non-linear, meaning price movements often don't follow a simple, predictable pattern. Traditional regression models like linear regression assume a linear relationship between variables, which may not hold true in the complex and volatile crypto market. As a result, these models struggle to capture the non-linear dependencies between features like historical prices, market trends, or trading volume, leading to poor prediction accuracy [29].
3. Overfitting, where models become too tailored to historical data and fail to generalize to new data, is another significant issue.
4. Cryptocurrency data is often sparse and noisy, which presents a significant challenge for regression models. Noisy data can mislead regression models, causing inaccurate predictions [30].
5. Often sentiment analysis based exogenous inputs may be heavily biased, rigged or even the sources may be fake resulting in falsified training [31].

Conclusion:

It can be concluded that crypto price prediction is a category of time series prediction which extremely challenging due to the dependence of crypto prices on several financial, socio-economic and political parameters etc. Moreover, small inaccuracies in crypto price predictions may result in huge losses to firms which use crypto price prediction results for financial analysis and investments. Off late, soft computing techniques are being used widely for crypto market prediction due to relatively higher accuracy compared to conventional statistical techniques. This paper presents a comprehensive review on existing work in the domain of crypto price forecasting. Moreover, the existing challenges and research gaps have also been mentioned to aid future directions of research.

References

1. N. P. Patel et al., "Fusion in Cryptocurrency Price Prediction: A Decade Survey on Recent Advancements, Architecture, and Potential Future Directions," in IEEE Access, vol. 10, pp. 34511-34538, 2022.
2. S. Karasu, A. Altan, Z. Saraç and R. Hacioğlu, "Prediction of Bitcoin prices with machine learning methods using time series data," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2018, pp. 1-4.
3. Y. Indulkar, "Time Series Analysis of Cryptocurrencies Using Deep Learning & Fbprophet," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2021, pp. 306-311.
4. M. Saad, J. Choi, D. Nyang, J. Kim and A. Mohaisen, "Toward Characterizing Blockchain-Based Cryptocurrencies for Highly Accurate Predictions," in IEEE Systems Journal, 2020, vol. 14, no. 1, pp. 321-332.
5. P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar and M. Alazab, "Stochastic Neural Networks for Cryptocurrency Price Prediction," in IEEE Access, 2020, vol. 8, pp. 82804-82818.
6. P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar and M. Alazab, "Stochastic Neural Networks for Cryptocurrency Price Prediction," in IEEE Access, vol. 8, pp. 82804-82818, 2020.
7. S. Otabek and J. Choi, "From Prediction to Profit: A Comprehensive Review of Cryptocurrency Trading Strategies and Price Forecasting Techniques," in IEEE Access, vol. 12, pp. 87039-87064, 2024
8. H. Jang and J. Lee, "An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information," in IEEE Access, vol. 6, pp. 5427-5437, 2018.
9. B. H. A. Khattak et al., "A Systematic Survey of AI Models in Financial Market Forecasting for Profitability Analysis," in IEEE Access, vol. 11, pp. 125359-125380, 2023.
10. H. Shamshad, F. Ullah, A. Ullah, V. R. Kebande, S. Ullah and A. Al-Dhaqm, "Forecasting and Trading of the Stable Cryptocurrencies With Machine Learning and Deep Learning Algorithms for Market Conditions," in IEEE Access, vol. 11, pp. 122205-122220, 2023.
11. P Han, H Chen, A Rasool, Q Jiang, M Yang, “MFB: A Generalized Multimodal Fusion Approach for Bitcoin Price Prediction Using Time-Lagged Sentiment and Indicator Features”, Expert Systems with Applications, Elsevier 2025, vol.261, 125515.
12. R Patel, J Chauhan, NK Tiwari, V Upaddhyay, A Bajpai, “A deep learning framework for hourly bitcoin price prediction using bi-LSTM and sentiment analysis of Twitter data”, SN Computer Science, Springer 2024, vol.57, no.767.
13. M. Rafi, Q. A. K. Mirza, M. I. Sohail, M. Aliasghar, A. Aziz and S. Hameed, "Enhancing Cryptocurrency Price Forecasting Accuracy: A Feature Selection and Weighting Approach With Bi-Directional LSTM and Trend-Preserving Model Bias Correction," in IEEE Access, vol. 11, pp. 65700-65710, 2023
14. G. Kim, D. -H. Shin, J. G. Choi and S. Lim, "A Deep Learning-Based Cryptocurrency Price Prediction

Model That Uses On-Chain Data," in IEEE Access, vol. 10, pp. 56232-56248, 2022
15. Z. Shahbazi and Y. -C. Byun, "Improving the Cryptocurrency Price Prediction Performance Based on Reinforcement Learning," in IEEE Access, vol. 9, pp. 162651-162659, 2021.
16. M Mudassir, S Bennbaia, D Unal, M Hammoudeh, “Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach”, Neural computing and applications, Springer 2020.
17. J.-Z. Huang, W. Huang, and J. Ni, ‘‘Predicting bitcoin returns using highdimensional technical indicators,’’ J. Finance Data Sci., vol. 5, no. 3, pp. 140–155, Sep. 2019.
18. R. Adcock and N. Gradojevic, ‘‘Non-fundamental, non-parametric bitcoin forecasting,’’ Phys. A, Stat. Mech. Appl., vol. 531, Oct. 2019, Art. no. 121727.
19. D. Philippas, H. Rjiba, K. Guesmi, and S. Goutte, ‘‘Media attention and bitcoin prices,’’ Finance Res. Lett., vol. 30, pp. 37–43, Sep. 2019.
20. J. Abraham, D. Higdon, J. Nelson, and J. Ibarra, ‘‘Cryptocurrency price prediction using tweet volumes and sentiment analysis,’’ SMU Data Sci. Rev., vol. 1, no. 3, p. 1, 2018.
21. D. Shen, A. Urquhart, and P. Wang, ‘‘Does Twitter predict bitcoin?’’ Econ. Lett., vol. 174, pp. 118–122, Jan. 2019.
22. N. Aslanidis, A. F. Bariviera, and Ó. G. López, ‘‘The link between cryptocurrencies and Google trends attention,’’ Finance Res. Lett., vol. 47, Jun. 2022, Art. no. 102654, doi: 10.1016/j.frl.2021.102654.
23. Y. B. Kim, J. G. Kim, W. Kim, J. H. Im, T. H. Kim, S. J. Kang, and C. H. Kim, ‘‘Predicting fluctuations in cryptocurrency transactions based on user comments and replies,’’ PLoS ONE, vol. 11, no. 8, Aug. 2016, Art. no. e0161197.
24. D. Liang, F. Ma and W. Li, "New Gradient-Weighted Adaptive Gradient Methods With Dynamic Constraints," in IEEE Access, 2020, vol. 8, pp. 110929-110942.
25. X. Huang, J. Guan, B. Zhang, S. Qi, X. Wang and Q. Liao, "Differentially Private Convolutional Neural Networks with Adaptive Gradient Descent," 2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC), Hangzhou, China, 2019, pp. 642-648.
26. AC Wilson, R Roelofs, M Stern, “The marginal value of adaptive gradient methods in machine learning”, Proceedings in Advances in Neural Information Processing Systems 30 (NIPS 2017).
27. Y Xue, Y Wang, J Liang, “A self-adaptive gradient descent search algorithm for fully-connected neural networks”, Neurocomputing, Elsevier 2022. Vol.478, pp. 70-80.
28. I Khandelwal, R Adhikari, G Verma, “Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition”, Procedia in Computer Science, vol.48, pp. 173-179.
29. A Butorova, E Baglaeva, I Subbotina, “Application of the Wavelet Data Transformation for the Time Series Forecasting by the Artificial Neural Network”, Proceedings in New Trends in the Applications of Differential Equations
30. D. Vujicic, D. Jagodic, and S. Randic, ‘‘Blockchain technology, bitcoin, and ethereum: A brief overview,’’ in Proc. 17th Int. Symp. INFOTEHJAHORINA (INFOTEH), Mar. 2018, pp. 1–6.
31. L. Duan, F. Duan, F. Chapeau-Blondeau and D. Abbott, "Noise-Boosted Backpropagation Learning of Feedforward Threshold Neural Networks for Function Approximation," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-12, 2021, Art no. 1010612.