Comparative Bitcoin Price Prediction Using Multiple Machine Learning Techniques

Gowri Sree Lakshmi.N[1*],Ajaykumar.V[2], Ashish.B[3], Hemalatha.K[4], Manohar.T[5], Sivashankerreddy.V[6]

[1,2,3,4,5,6]Dept. of CSE, Godavari Institute of Engineering and Technology (A), Rajahmundry, A.P, India
gowrin.130@gmail.com[1*] , ajaykumar@giet.ac.in[2], ashishbaipalli369@gmail.com[3] , hemalatha.kariminji@gmail.com[4], talaganamanohar9@gmail.com[5], siva17062003@gmail.com[6]

Abstract. The cryptocurrency market is known for its inherent volatility, making accurate predictions a challenging endeavor. In this research study, investigate the efficacy of logistic regression, support vector machines (SVM), decision trees and random forests for the task of Bitcoin price prediction. To address this, conduct a thorough analysis and comparison of these machine learning models using historical Bitcoin price data. By rigorously assessing their performance and predictive capabilities, this study aims to provide valuable insights for both cryptocurrency traders and researchers operating in the dynamic digital asset landscape. These results illuminate the strengths and weaknesses of each model, shedding light on their respective abilities to forecast Bitcoin price movements. Through this research, contribute to the growing body of knowledge surrounding cryptocurrency market analysis and prediction techniques. This analysis can inform traders' decision-making processes and assist researchers in developing more robust models in the exciting and rapidly evolving realm of cryptocurrency investment and analysis.

Keywords: Bitcoin, Price Prediction, Machine Learning, Support Vector Machines, Logistic Regression, Random Forests, Decision Trees, Cryptocurrency Markets.

1 Introduction

Virtual currencies, a form of cryptocurrency, are a significant advancement in digital finance. While they offer benefits, they can't fully replace traditional currencies. This study explores topics like how money is calculated, currency governance, and the impact on financial systems. These days, virtual currencies are the most effective and extensively utilized for international business operations. Its success is a result of its unique traits, which include transparency, simplicity, and expanding global acceptance. The most successful and commonly utilized virtual money at the moment is bitcoin. Data on the virtual currency market is believed to be worth close to 90 billion dollars, albeit this amount is subject to change, based on information examined on April 19, 2019, on the website https://bitcoin.org. Peer-to-peer technology like Bitcoin allows for total decentralization of the transaction process, with no external authority controlling any aspect of it. It is not possible for third parties to intervene in customer disputes. It is a constantly fluctuating market price. Periodically, the market capitalization of bitcoin rises. Currently traded on public markets are over 71 billion dollars. Of all the virtual currencies in the world, its open-source nature makes it transparent, easy to use, clear, and time-saving. The most popular cryptocurrency worldwide, When Bitcoin was initially launched in 2008, a person going by the name of Satoshi Nakamoto exploited it as open-source software in 2009. Still, its popularity truly surged in 2017. With no involvement from outside parties, transactions are verified and recorded in a publicly accessible ledger known as the blockchain, making it possible for Bitcoin to function as a decentralized digital money exchange.

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_120

Made up of interconnected secure shell algorithms, transaction blocks function as non-editable data that is captured during the transaction. Following this, individuals started using virtual currencies, especially bitcoin, and the trend of the virtual currency industry kept growing. In a brief amount of time, bitcoin's popularity has grown. Bitcoin is used to connect various business enterprises and technological platforms. Various researchers note that since 2015, about 100,000 technology and business companies, including big names like Amazon and Microsoft, have entered the bitcoin market. While there have been studies on predicting bitcoin's value, deep learning models haven't been widely used for this purpose. This study aims to fill that gap by leveraging deep learning, known for its accuracy in predicting trends like bitcoin prices. It reviews existing literature on bitcoin price prediction and focuses on three key papers. Bitcoin's high price volatility is a challenge, requiring precise prediction methods. Understanding bitcoin's price involves considering factors like public relations, politics, and market policies. However, the lack of regulation in bitcoin exchanges makes forecasting difficult. This study aims to improve price estimation using deep learning, benefiting investors and policymakers by reducing risks.

2 Literature Survey

W. Yiying et al.., in 2019, Cryptocurrency is gaining popularity in reshaping finance, yet its unpredictability poses risks for investors. This study uses advanced AI models like ANN and LSTM to analyze Bitcoin, Ethereum, and Ripple price dynamics, showing LSTM's proficiency in short-term patterns and ANN's potential with enough historical data. Overall, the research highlights cryptocurrency market price predictability, influenced by the machine-learning model used [2].

G. Cheuque et al.., in 2019, The Bitcoin protocol and cryptocurrency have introduced new challenges in understanding digital currency pricing, particularly amid Bitcoin's price volatility. This study explores the influence of crypto-influencers on Bitcoin prices, using neural models to predict price movements. Preliminary findings suggest that Twitter data, incorporating sentiment analysis of influencers' posts, can enhance Bitcoin price prediction, although refining sentiment measurement remains an ongoing challenge [3].

M. Saad et al.., in 2020, In 2017, the blockchain-based cryptocurrency market experienced significant growth, particularly with Bitcoin reaching record highs. This paper investigates the factors contributing to Bitcoin's price surges by analyzing user and network activity data, identifying key features affecting demand and supply dynamics. Leveraging machine learning, the regression model achieves an impressive 99.4% price prediction accuracy with a 0.0113 root mean squared error (RMSE) [4].

P. Jay et al.., in 2020, Blockchain technology has driven a surge in cryptocurrency usage, but their volatile nature has deterred investment. To tackle this, this system simulates market volatility and identifies patterns of market behavior by proposing a stochastic neural network model based on random walk theory. This model outperforms deterministic approaches in Bitcoin, Ethereum, and Litecoin price prediction [5].

T. E. Koker et al.., in 2020 Active cryptocurrency trading research involves studies on trading strategies, risk-return dynamics, and machine learning applications in financial markets. These studies examine cryptocurrency price behavior, market capitalization, and challenges in price modeling. They also explore the direct reinforcement learning approach and performance metrics, highlighting the need for continued research in this evolving field [6].

E. Sin et al.., in 2017 The research uses a dataset with over 200 cryptocurrency variables spanning two years and a Selective Neural Network Ensemble based on Genetic Algorithms to predict the direction of Bitcoin prices the next day. The ensemble technique demonstrates its practicality and effectiveness in cryptocurrency trading by outperforming both a single MLP model strategy and the conventional "previous day trend following" approach, yielding an astonishing 85% return [7].

P. M. A. Sharma et al.., in 2020 This study explores Ether price prediction using machine learning, utilizing both Support Vector Machine (SVM) and Linear Regression (LR) with various window lengths on daily closing prices. SVM outperforms LR, and the model's performance can be further improved to near-perfection (95%) with the addition of extra features. The enduring appeal of cryptocurrencies, known for their untraceable and decentralized nature, continues to fuel research and investment in the field [8].

I. Madan et al.., 2015 also made use of Oncovin's datasets but chose to partition the data into 30, 60, and 120-minute intervals to anticipate Bitcoin prices, they used the Support Vector Machine (SVM), Random Forest and Binomial Logistic Regression techniques. Achieving a high 97% accuracy for short-term (10-minute) price predictions and 55% accuracy for 10-minute price forecasts. However, it's worth noting that their research did not include cross-validation, potentially leaving the models susceptible to overfitting [9].

P. N. Sureshbhai et al.., 2020 A public blockchain is used by the decentralized system KaRuNa to facilitate transparent transactions, reducing the risk associated with investing in cryptocurrencies. Sentiment analysis and a special hash address approach are integrated into its three-phase trust model, which feeds into an LSTM classifier that achieves 98.99% accuracy in risk assessment. This method improves scalability and drastically lowers fraud in bitcoin schemes by taking into account factors like price swings and sociological trends. By guaranteeing openness and reliability, it promotes trust among investors, crowdfunding platforms, and financial institutions. [11].

R. Gupta et al.., 2020 The system, called BATS, combines blockchain technology with artificial intelligence (AI) to improve security and dependability in telesurgery. It is enabled by 6G. It uses UAVs to carry medical supplies during operations efficiently and uses AI algorithms like XGBoost to categorize illnesses according to their criticality. BATS outperforms conventional surgical methods like as HaBiTs and AaYusH, demonstrating greater prediction accuracy, minimal packet loss, and decreased bandwidth consumption via IPFS [12].

3 Existing Model

The existing system relies on Linear Regression for Bitcoin price prediction. A statistical technique called linear regression is used to describe the link between the price of bitcoin, which is the dependent variable, and a number of independent variables, or predictor characteristics. It evaluates the correlation between changes in predictor variables and variations in the price of Bitcoin using a linear equation to depict this connection.

Disadvantages

Limited Complexity - By using linear regression, it is assumed that the predictor factors and the dependent variable (the price of bitcoin) have a linear relationship. The truth is that a multitude of complex factors influence the price of Bitcoin, and these relationships are typically nonlinear. As a result, predictive accuracy can decline.

Outliers Sensitivity - Outliers in the data could make linear regression susceptible. One extreme data point has the potential to significantly affect the model's coefficients and predictions, leading to unpredictable results.

Inability to Capture Complex Patterns - Numerous factors, including as macroeconomic variables, news events, and market state of mind, might impact the fluctuations in the price of bitcoin. It could be difficult for linear regression to accurately represent these complex nonlinear patterns.

4 Proposed Model

The main goal of the suggested system is to forecast Bitcoin values using past transaction data, particularly by exploiting the features of price and timestamps. For price prediction, this method uses four different machine learning techniques: Decision Tree, Logistic Regression, Support Vector Machine (SVM), and Random Forest.

Data Selection - Price at particular timestamps is one of the primary characteristics used in this study's selection of historical Bitcoin transaction data. System prediction models rely on this data as their basis.
Machine Learning Models - This application uses four distinct methods of machine learning.
Logistic Regression - This approach is renowned for being straightforward and easily understood. Although it is usually applied to classification problems, this has been modified to forecast Bitcoin values.
Support Vector Machine (SVM) - Classification and regression tasks benefit greatly from the use of SVM. For the purpose of predicting Bitcoin prices, this system uses SVM.
Random Forest - Random Forest is an ensemble learning method for prediction that makes use of several decision trees. It is quite adept at identifying complex patterns in data.
Decision Tree - Models for regression tasks that are straightforward but efficient are decision trees. Their interpretability is well-known.
Prediction Horizons - Assess the performance of these models for various prediction horizons. This includes both short-term and longer-term predictions. For example, this evaluates the models' abilities to predict Bitcoin prices for the next day, as well as for longer durations, such as the next 5-7 days.
Performance Evaluation - This application evaluates the models' performance using appropriate regression evaluation metrics, Mean Squared Error (MSE), This allows us to quantify how well each model performs for different prediction horizons.

Model-Specific Insights

Random Forest - Performs well for short-term predictions but may not excel for longer-term forecasts.
Decision Tree - Demonstrates consistent performance for up to 6 days.
Logistic Regression - Suitable only when a separable hyperplane exists, indicating limitations in its performance.

Advantages

Diverse Model Selection - By utilizing four distinct machine learning methods - Logistic Regression, Support Vector Machine (SVM), Random Forest, and Decision Tree - the proposed model capitalizes on the strengths of each approach. This diversity can improve the robustness of price predictions.
Enhanced Prediction Accuracy - The use of multiple machine learning methods allows for a comparative analysis of their performance in predicting Bitcoin prices. This can lead to improved accuracy by selecting the model or combination of models that provides the most reliable forecasts.
Flexibility - The model's ability to employ various machine learning techniques means it can adapt to different types of data and market conditions. This adaptability is crucial in the cryptocurrency market, known for its high volatility and evolving trends.
Interpretability - Certain machine learning methods, such as Decision Trees, offer high interpretability. This means that not only can the model make predictions, but it can also provide insight into the factors driving those predictions, which can be valuable for understanding market dynamics.
Real-World Application - Predicting crypto-currency prices has practical applications for both individual investors and institutions. The proposed model can serve as a valuable tool for financial decision-making in the rapidly evolving and high-stakes world of crypto-currencies.

5 Block Diagram

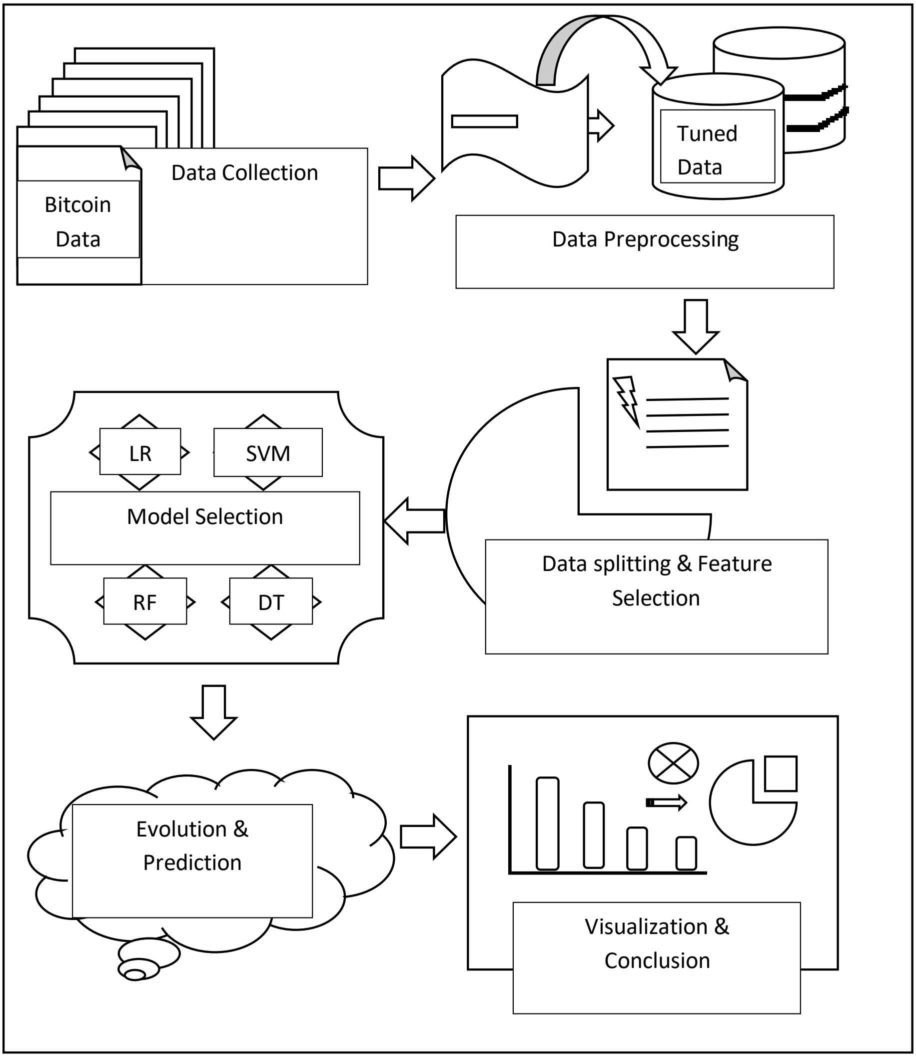

Fig 1: Block diagram of proposed model

6 Work Flow of Proposed System

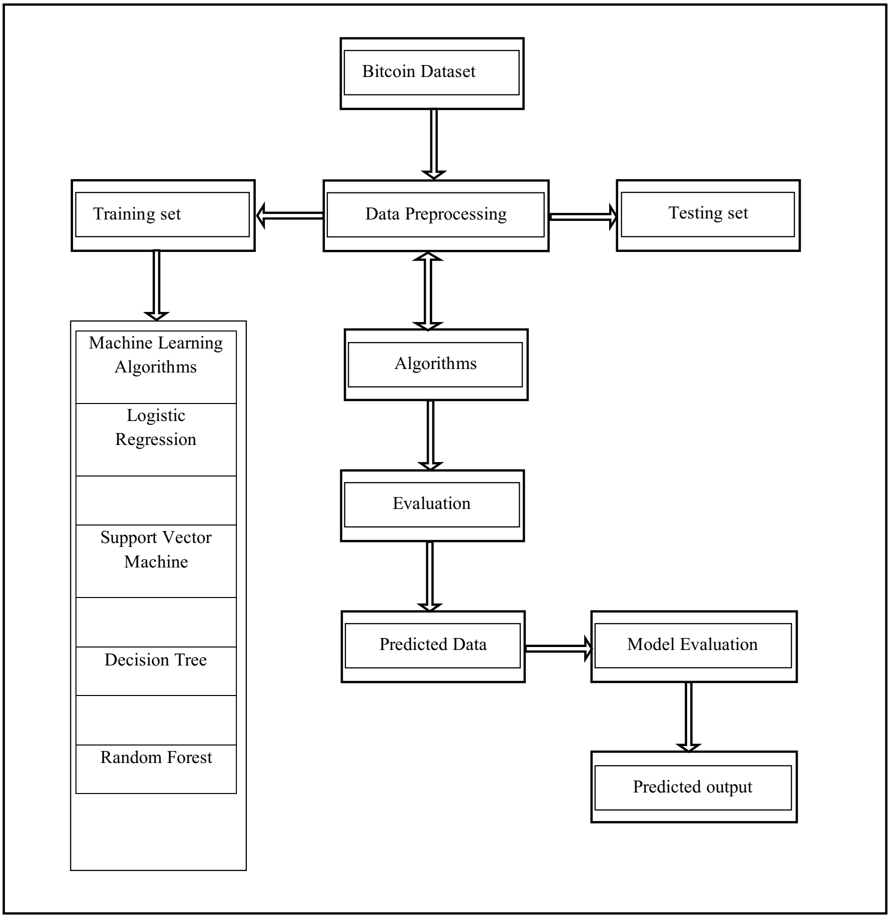

Fig 2: Architecture of Proposed system

7 Comparitive Study

Table-1: Difference In The Proposed System's Predicted Prices And Bitcoin Prices

Bitcoin Price (USD)	Logistic Regression Predicted Price (USD)	Support Vector Machines Predicted Price (USD)	Decision Tree Predicted price (USD)	Random Forest Predicted Price (USD)
671.622	671.229	672.278	671.104	671.092
681.150	698.689	692.551	689.652	686.541
685.809	708.762	705.896	703.311	702.048
713.239	711.672	710.032	711.922	708.559
702.076	742.917	736.291	730.106	726.585
704.955	727.145	728.352	733.373	729.774
727.063	732.721	730.322	728.022	724.511
726.052	756.212	752.015	747.953	744.237
692.705	728.023	732.672	739.585	735.932
704.181	719.905	720.058	720.418	716.989
700.377	753.725	752.170	754.145	750.336

8 Results

The Bitcoin price prediction project revealed that Support Vector Machine (SVM) and Random Forest models demonstrated strong performance for short-term price predictions, while Decision Tree offered consistent accuracy for up to 6 days. Logistic Regression, although interpretable, had limitations. Further analysis and enhancements are recommended to improve longer-term forecasting accuracy and explore additional data sources for enhanced predictions.

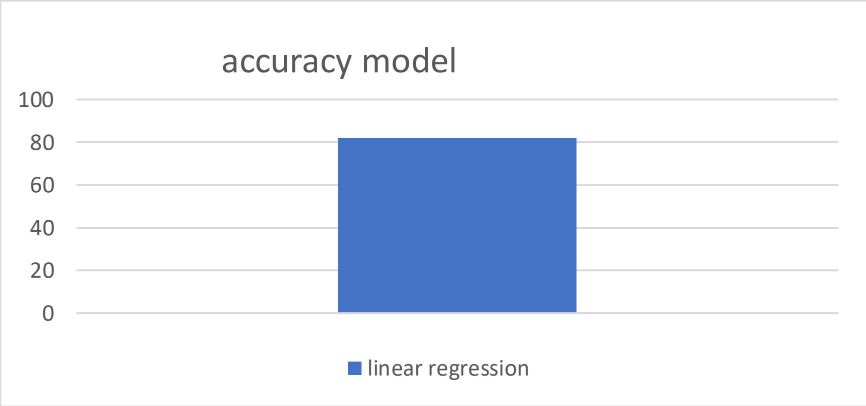

Fig 3: Existing Model Accuracy

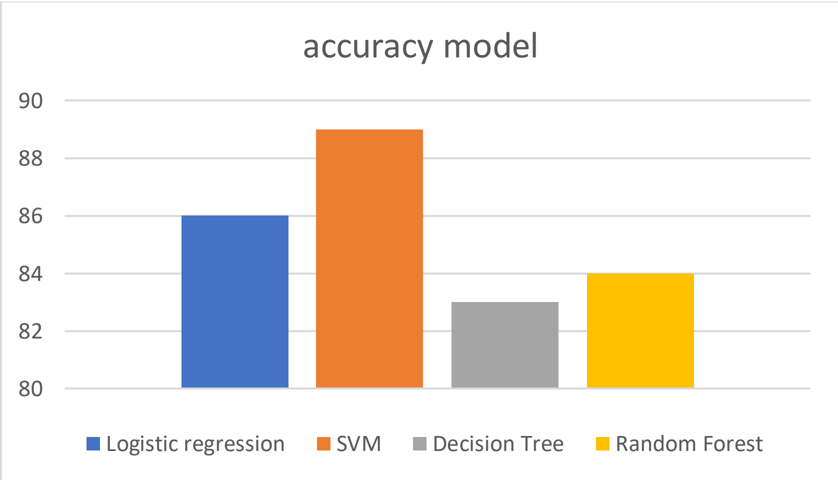

Fig 4: Proposed Model Accuracy

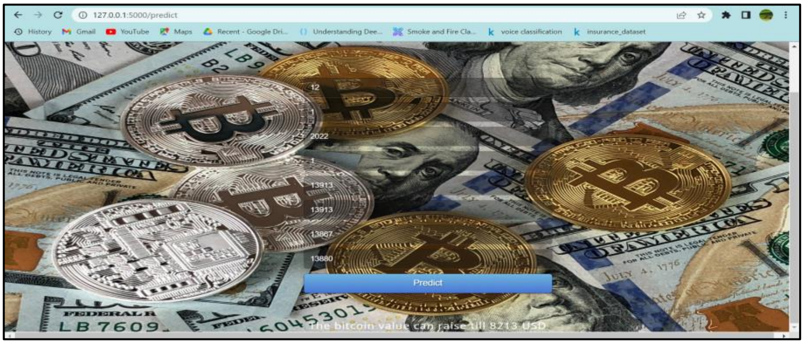

Fig 5. predicted prices

TABLE-2: Comparision of Accuracy

Existing System	**Technique Name**	**Accuracy**
	Linear Regression	83%
Proposed System	**Technique Name**	**Accuracy**
	Logistic Regression	92%
	Support vector machines	94.3%
	Decision Tree	93.2%
	Random Forest	90.7%

TABLE - 3: MSE of all implemented models

Model	**MSE**
Logistic regression	0.000375
Support vector machines	0.000373
Decision Tree	0.000485
Random Forest	0.00048

9 Conclusion

This analysis employed random forest regression, a machine-learning technique that demonstrated superior forecasting accuracy compared to traditional machine-learning algorithms commonly used in prior studies. Notably, random forest regression exhibited limitations in extrapolating beyond its training data; for instance, it could not predict prices surpassing historical record highs. During the initial price bubble phase (April 2015 to October 2018), significant predictors included the U.S. stock market indices (NASDAQ, DJI, and S&P500), Bitcoin mining difficulty metrics, oil prices, and Ethereum (ETH) prices. Conversely, during the subsequent price bubble (October 2018 to April 2022), key predictors were ETH prices and Japan's JP225 index. However, the model encountered challenges in forecasting Bitcoin prices exceeding USD 60,000 per coin by the close of 2021 due to its inherent limitations. Nonetheless, its accuracy remained commendable for price ranges below USD 60,000. Future refinements could potentially enhance the model's predictive capabilities as transaction history accumulates and market dynamics evolve.

References

1. A. Hashish, F. Forni, G. Andreotti, T. Facchinetti, and S. Darjani, ‘‘A hybrid model for bitcoin rices prediction using hidden Markov models and optimized LSTM networks,’’ in Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA), Sep. 2019, pp. 721–728.
2. W. Yiying and Z. Yeze, ‘‘Cryptocurrency price analysis with artificial intelligence,’’ in Proc. 5th Int. Conf. Inf. Manage. (ICIM), Mar. 2019, pp. 97–101.
3. G. Cheuque and J L. Reutter, ‘‘Bitcoin price prediction through opinion mining,’’ in Proc. Companion World Wide Web Conf., May 2019, pp. 755–762
4. M. Saad, J. Choi, D. Nyang, J. Kim, and A. Mohaisen, ‘‘Toward characterizing blockchain-based cryptocurrencies for highly accurate predictions,’’ IEEE Syst. J., vol. 14, no. 1, pp. 321–332, Mar. 2020.
5. P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar, and M. Alazab, ‘‘Stochastic neural networks for cryptocurrency price prediction,’’ IEEE Access, vol. 8, pp. 82804–82818, 2020
6. T. E. Koker and D. Koutmos, ‘‘Cryptocurrency trading using machine learning,’’ J. Risk Financial Manage., vol. 13, no. 8, p. 178, Aug. 2020
7. E. Sin and L. Wang, ‘‘Bitcoin price prediction using ensembles of neural networks,’’ in Proc. 13th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD), Jul. 2017, pp. 666–671.
8. P. M., A. Sharma, V. V., 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.
9. I. Madan, S. Saluja and A. Zhao. "Automated Bitcoin Trading via Machine Learning Algorithms," 2015.
10. M. K. B, M. S. Kumar, F. D. Shadrach, S. R. Polamuri, P. R and V. N. Pudi, "A binary Bird Swarm Optimization technique for cloud computing task scheduling and load balancing," 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2022, pp. 1-6, doi: 10.1109/ICSES55317.2022.9914085.
11. P. N. Sureshbhai, P. Bhattacharya, and S. Tanwar, ‘‘KaRuNa: A blockchainbased sentiment analysis framework for fraud cryptocurrency schemes,’’ in Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops), Jun. 2020, pp. 1–6.
12. R. Gupta, A. Shukla, and S. Tanwar, ‘‘BATS: A blockchain and AI-empowered drone-assisted telesurgery system towards 6G,’’ IEEE Trans. Netw. Sci. Eng., early access, Dec. 8, 2020, doi: 10.1109/ TNSE.2020.3043262.
13. Madhavi, K. Reddy, Padmavathi Kora, L. Venkateswara Reddy, Janagaraj Avanija, K. L. S. Soujanya, and Prabhakar Telagarapu. "Cardiac arrhythmia detection using dual-tree wavelet transform and convolutional neural network." Soft Computing 26, no. 7 (2022): 3561-3571.
14. S. Shanaev, S. Sharma, B. Ghimire, and A. Shuraeva, ‘‘Taming the blockchain beast? Regulatory implications for the cryptocurrency market,’’ Res. Int. Bus. Finance, vol. 51, Jan. 2020, Art. no. 101080.
15. V. Derbentsev, N. Datsenko, O. Stepanenko, and V. Bezkorovainyi, ‘‘Forecasting cryptocurrency prices time series using machine learning approach,’’ SHS Web Conf., vol. 65, Jan. 2019, Art. no. 02001.