Computers and Electrical Engineering 81 (2020) 106527

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system[☆]

Poongodi M.[a,1], Ashutosh Sharma[b,*], Vijayakumar V.[c], Vaibhav Bhardwaj[c], Abhinav Parkash Sharma[c], Razi Iqbal[d], Rajiv Kumar[e]

[a] *Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa Univeristy, Doha, Qatar*
[b] *School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, Punjab, India*
[c] *Computer Science and Engineering, Vellore Institute of Technology, Chennai, India*
[d] *College of Computer Information Technology, American University in the Emirates, UAE*
[e] *Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan, Himachal Pradesh, India*

A R T I C L E I N F O

Article history:
Received 30 November 2018
Revised 22 October 2019
Accepted 25 November 2019
Available online 3 December 2019

Keywords:
Linear regression
SVM
Cryptocurrency
Ether
Industrial finance system

A B S T R A C T

Cryptocurrency has gained considerable popularity in the past decade. The untraceable and uncontrolled nature of cryptocurrency attracts millions of people around the world. Research in cryptocurrency is dedicated to finding the ether and predicting its price according to the cryptocurrency's past price inflations. In this study, price prediction is performed with two machine learning methods, namely linear regression (LR) and support vector machine (SVM), by using a time series consisting of daily ether cryptocurrency closing prices. Different window lengths are used in ether cryptocurrency price prediction by using filters with different weight coefficients. In the training phase, a cross-validation method is used to construct a high-performance model independent of the data set. The proposed model is implemented using two machine learning techniques. When using the proposed model, the SVM method has a higher accuracy (96.06%) than the LR method (85.46%). Furthermore, the accuracy score of the proposed model can be increased up to 99% by adding features to the SVM method.

1. Introduction

Over the past decades, the world has seen multiple technological advancements, such as the Internet of things (IoT), e-commerce, and digital payments [1]. These advancements not only affect how one interacts with the world but also change how one interacts and exchanges money for requested services. Nowadays, banking, transaction, and other e-commerce activities are becoming more advanced due to new developments in technology. In the past 20 years, due to the trend of moving from printed currency to virtual currency, a new type of currency called cryptocurrency has emerged [2]. Cryptocurrency is basically a digital currency that uses advanced encryption techniques for regulating its currency units. Cryptography

☆ This paper is for CAEE special section SI-bciot. Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. Shaohua Wan.

* Corresponding author.
E-mail addresses: sharmaashutosh1326@gmail.com (A. Sharma), vijayakumar.varadarajan@gmail.com (V. V.), abhinavparkash.2016@vitstudent.ac.in (A.P. Sharma), razi.iqbal@ieee.org (R. Iqbal).
[1] Indicates that author was with Computer Science and Engineering, Vellore Institute of Technology, Chennai, India at the time of submission.

https://doi.org/10.1016/j.compeleceng.2019.106527

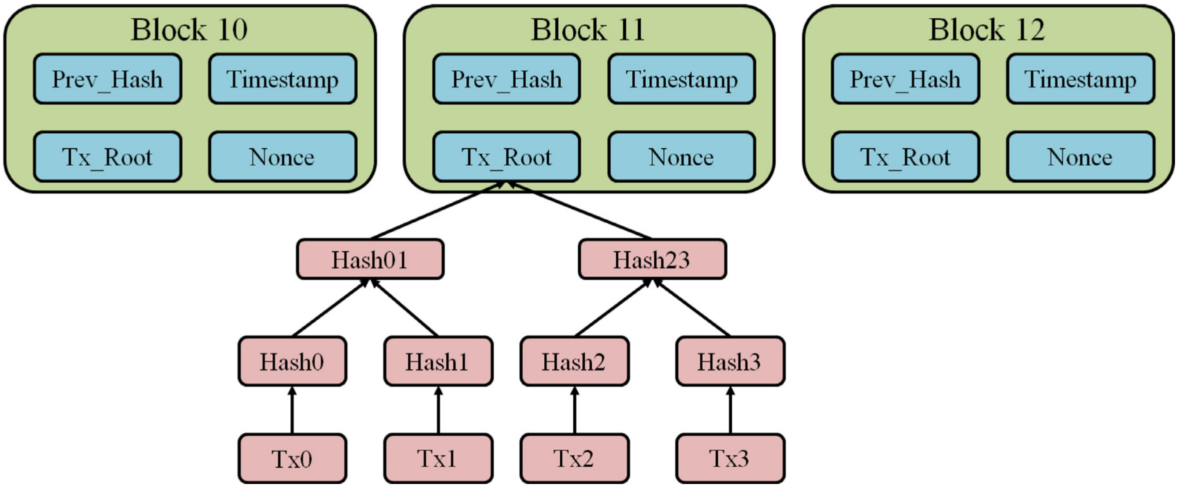

Fig. 1. Block data for a cryptocurrency [8].

has been applied to make payments more secure and to verify transactions. Verification of digital currency depends on the transaction process. Cryptocurrency is used to transfer funds electronically without any interference of a central entity, such as a bank [3]. The flow of such transactions is controlled with the help of cryptocurrency.

Initially, when industries adopted new technological developments, several efforts were made by researchers from companies such as Flooz, Beenz, and Digicash to create a digital currency [3]. Later, these emerging currency companies inevitably failed due to a lack of resources and adaptations in user societies and communication. The main causes behind the failure of these companies were fraud, financial crises, friction in competitive markets, and management scams. In addition, the flow of the trusted third-party approach also contributed to the failure of these systems. In the third-party approach, companies conducting financial transfers are carefully verified and facilitated to provide transaction services. Due to the failures of the aforementioned companies, the creation of a digital cash system was seen as a lost cause and considered hopeless for a very long time. Later, using advanced technology, a group of obscure developers attempted to introduce a currency popularly known as Bitcoin [4]. The abstract idea of Bitcoin was first proposed by Satoshi Nakamoto [5]. Bitcoin is a shared electronic framework in which cryptocurrency can be exchanged. The framework is similar to the distributed system used for record sharing. The Bitcoin framework is different from previous digital cryptocurrency frameworks in its use of decentralized innovation, that is, no server is used and there exists no central control over the currency. People with specialized software or suitable mathematical skills can distribute money according to their wishes. The major issue in Bitcoin was that any payment arrangement had to account for double spending, which refers to the situation where the same coin is assigned to two users. Double spending is a deceitful strategy of spending the same amount of money twice [6]. This issue has been solved using distributed time-stamping services [5]. In a decentralized system such as Bitcoin or Ether coin, each member must perform mining activities. These activities are performed on a blockchain, which is an associate open record of all exchanges that occur within the system for a specific purpose and is accessible to everybody. Hence, everybody in the system can see each record's equalization. This eliminates the necessity of a central authority and provides control to the authorities directly involved in the transaction [7].

Blockchain technology is a super-fast-growing technology in which an increasing number of records known as blocks are linked through cryptography. As displayed in Fig. 1, each block contains a cryptographic hash structure of the previous block, a timestamp, and transaction data [8]. Fig. 1 is used to explain the cryptocurrency used in a blockchain, which is insusceptible to information modification. Blockchain is essentially an exceptionally open, completely circulated record that is exchanged between two gatherings in an ensured, irrefutable, and perpetual manner [9].

A blockchain is commonly overseen by a distributed system, with fast communication between hubs and fast approval of new squares. In some random squares, the modifications in information are made retroactively without a change in every single subsequent square. Although blockchain records are not unalterable, a blockchain might be viewed as secure by plan and exemplify a distributed computing framework with high Byzantine tolerance to noncritical failure. In this manner, a decentralized consensus is guaranteed with a blockchain [10]. The cryptocurrency Bitcoin uses blockchain technology, which was invented by Satoshi Nakamoto. The application of blockchain technology in Bitcoin made it the main advanced money ledger that has taken care of the double-spending problem without any need for a trusted authority or a central server. The Bitcoin configuration has made other applications possible, and blockchains are widely used by digital forms of currency. Moreover, private blockchains have been proposed for business use, as depicted in Fig. 2.

In cryptocurrency, the exchange of every transaction is a ledger record that contains public keys of the sender or receiver and the number of transferred coins. These public keys are called wallet addresses. In each transaction, a personal key must be signed off by the sender. In due course, the transaction is broadcasted within the network after confirmation.

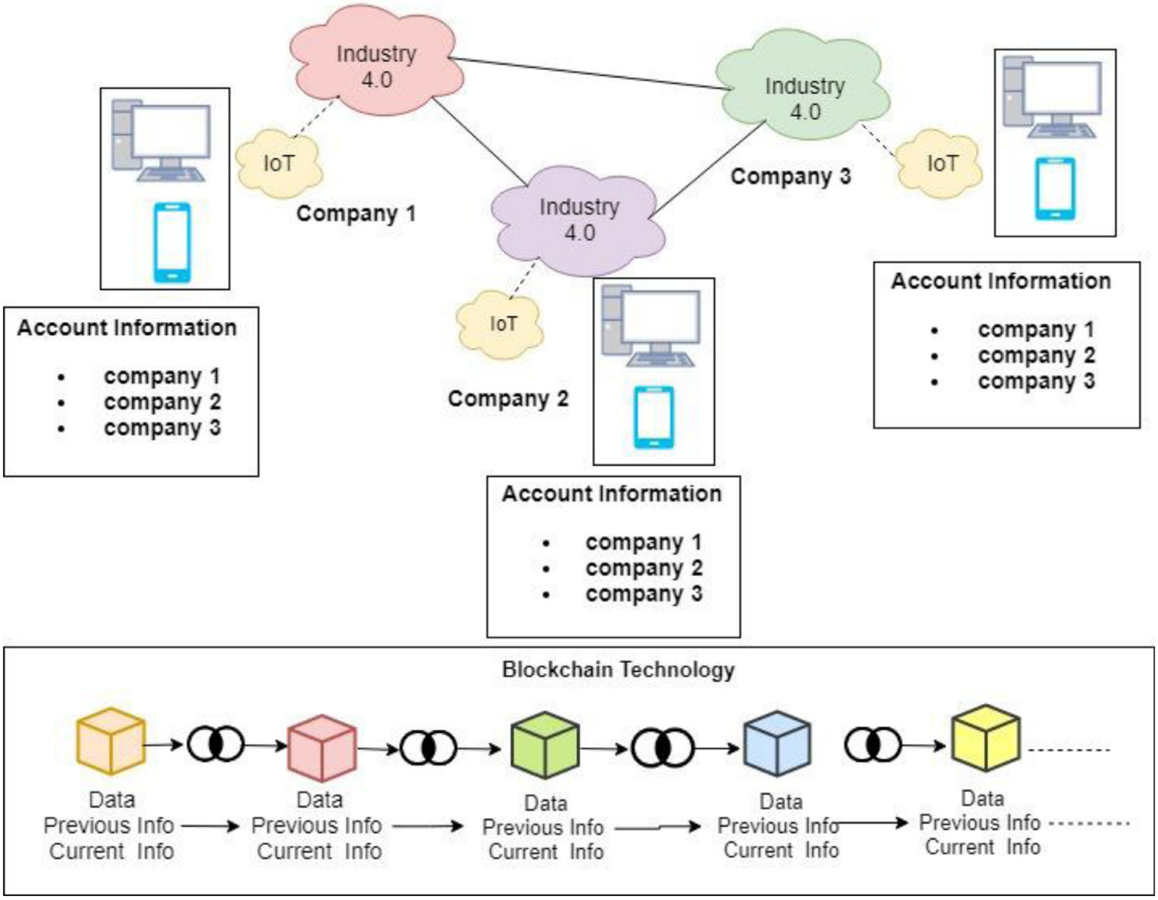

Fig. 2. Blockchain technology in industries.

In a crypto network, miners can confirm exchanges or transactions by unraveling a cryptographic riddle. After confirming the exchange or transaction, a miner must stamp it as genuine and distribute it over the system. Later, every hub adds the transactions that have occurred to their databases. When the exchange has been confirmed, the process of transaction stops without any problem. In this process, a miner or sender is rewarded in addition to the exchange of expense payments.

The development of a country depends mainly on the country's economy, and the role of currency in the economy cannot be ignored. In the literature, numerous approaches have been proposed for easy, efficient, and fast payment methods [11]. Initially, the concept of e-money and e-commerce was widely used in markets, and due to the growth of e-services, this concept has now gained importance. Various modes of digital currency and mobile wallets have attracted the attention of users. E-services provide easy, fast, and efficient methods for business growth and comfort of life. Ethereum is a type of cryptocurrency payment platform where transactions are completed by clients, with machines executing the requested operations. More broadly, ether is a crucial incentive that ensures that developers write and maintain quality throughout transactions.

1.1. Research contribution

This study investigates how ether cryptocurrency prices are predicted and what types of trends can be discovered over time in the prices of this currency.

- The main focus of this study is on a cryptocurrency called ether coin. Ether is a crucial and necessary element for the operation of the distributed application platform Ethereum. Moreover, developers pay attention to healthy network signals or the availability of specific executions.
- The proposed approaches are used for price prediction in an industrial finance system and exhibit suitable accuracy scores.

1.2. Organization

The remainder of this paper is structured as follows. The related work is presented in Section 2. A proposed system model is explained and formulated in Section 3. Section 4 presents the dataset and processing in the experiment. The experimental evolutions and results are presented in Section 5. Finally, Section 6 provides the conclusions of this study as well as future directions and recommendations.

2. Related literature

The world is now an extremely globalized place, and the business operations in the current market and the ways in which these operations are conducted are also globalized. The economy is crucial for any country to achieve success on the global stage. Therefore, currency cannot be ignored. Over the past few years, numerous approaches have been used to develop easy and fast payment methods (Fig. 3).

The concepts of e-money and e-commerce have a long history and have been used in the market for a long time. The present era of technology has been caused by the rising importance of mobile wallets, especially due to demonetization of India and the need for transferring money by using mobile applications. For these reasons, a type of currency that is transferable anywhere and can be easily used in business would be welcome. One such type of currency is the cryptocurrency ether. Cryptocurrencies such as Litecoin, Bitcoin, Z cash, Dash, and Ripple are other examples of newly introduced blockchain-based currencies. This survey provides an insight into the prediction of ether prices. Blockchain technology has recently attracted the attention of a broad range of industries, such as healthcare, real estate, government sectors, and finance [6,12,13]. Blockchain technology has progressed at an astonishing pace to provide transparency to its users. Moreover, this technology can ensure security or transparency to its users even when IoT objects or devices are hacked by intruders. Blockchain technology is capable of tracking, organizing, and supporting communications by storing data from a large number of devices and facilitating the formation of parties without any federal cloud. The blockchain method confines users' activity and traces product information from IoT objects close to yields as the consignment is transported from one place to another. The use of supply chains is a major component of the blockchain technology for resolving business issues caused by the requirement of traceability of products or information on users' legal or illegal activities [14,15].

In [16], Bitcoin prices were predicted using machine learning techniques, such as linear regression (LR) and support vector machine (SVM), by using a time series consisting of daily Bitcoin closing prices between 2012 and 2018. Different parameter combinations were used in the prediction model with least error, such as SVM with the linear function and polynomial kernel function. For different window lengths, different weight coefficients were used in filters. Bitcoin price prediction can be performed for different window lengths by using various filters with different combinations of weight coefficients. This method, which is known as the 10-fold cross-validation method and used mainly during the

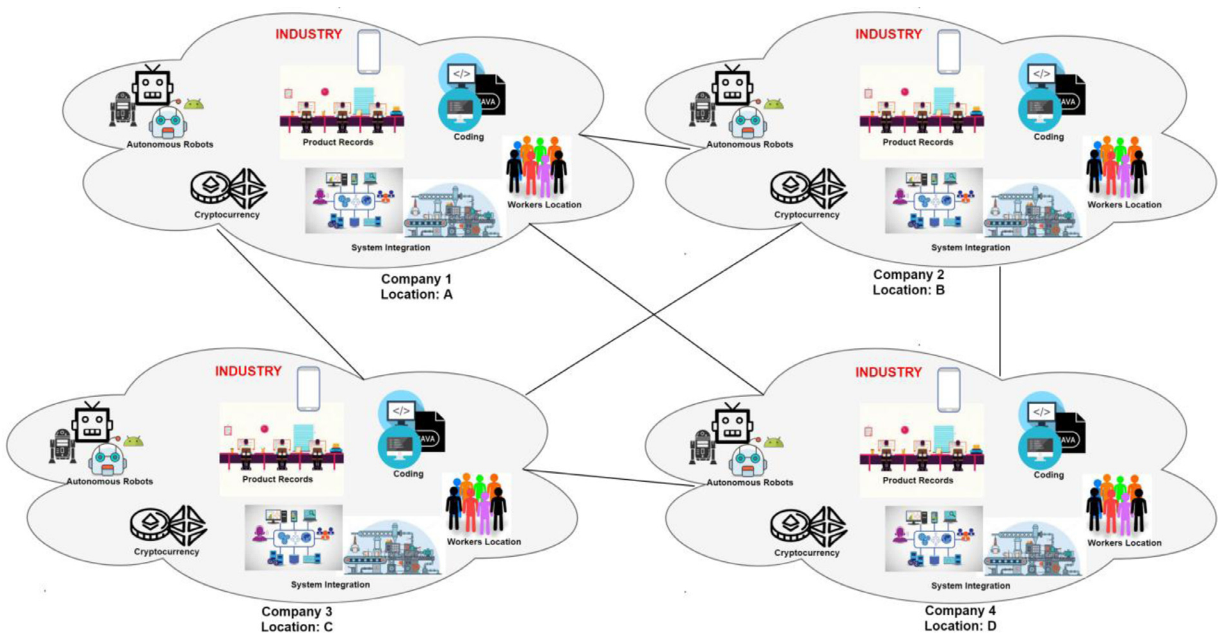

Fig. 3. Industrial architecture with IoT.

training phase, is purely used for the construction of a high-performance model, which is surprising given the independence of the dataset. A relationship between Bitcoin features and variation in daily price was found using the artificial neural network approach [17]. This approach is known as the genetic-algorithm-based selective neural network method. The selective neural network model is constructed for using a multilayer perceptron as a base model for each neural network.

Researchers have also proposed using a hidden Markov model for predicting cryptocurrency prices [18]. This model was previously used for detecting influenza epidemic outbreaks with the help of the behavior of novel online social media indicators. To validate the model, a trading strategy was adopted and tested on historical data. This strategy outperformed the buy-and-hold strategy. This validation indicated the broader effectiveness of epidemic-detecting hidden Markov models in the detection of bubble-like behavior in time series. Moreover, the proposed model is helpful for predicting valuable information pertaining to cryptocurrency price movements from social media data. Various authors have proposed other models to predict cryptocurrency prices. Plots have been drawn for the price change trend of five major cryptocurrencies over time [19,20]. The difference in the actual price value between each currency is high, and plotting all currencies in one graph is impossible. Therefore, min–max normalization is used to scale the data in the range [0,1], and the normalized price is plotted. The plot exhibits an exponential increase in the price of every cryptocurrency over the years. Authors have also proposed a method called the prospect theory for calculating the utility of different prospects through determining how many people behave differently by confronting them with the outcomes of the respective gains and losses. Therefore this phenomenon can be called the reflection effect [20,21]. Some people prefer a profitable loss over a certain lower loss. This phenomenon is known as the certainty effect.

Some major issues that online transactions and cryptocurrency transactions face are as follows:

Phishing and spoofing of payment information: This includes problems such as theft. When transferring money, a malware may replace the address in the clipboard of the user with another address even when they have copied their wallet address perfectly. It is very difficult to be vigilant and double-check the address of every user after copying an address, which is especially the case when the address is a combination of numerous characters. With ordinary e-money, users can be tricked into going to a phishing website where they upload their crypto wallet and enter a password.

Hacking of payment gateways: Hackers uses social engineering methods to take the host into confidence by pretending to be real domain owners. After gaining access, they can start intercepting cash flows.

User address error: A common risk occurring in cryptocurrencies is the error in the money transfer address, which causes the loss of considerable money. In the case of Ethereum, if a single digit is not copied, then the transferred money disappears.

Loss of a wallet file: In cryptocurrencies, the loss or theft of a wallet is mainly due to the storage of wallet files. These files are stored in hard disks and can be stolen using malware or when the hard disk crashes.

3. Proposed system model

Ether is the cryptographic currency of Ethereum, whose blockchain is produced by the Ethereum platform. Ether can be exchanged among records and used to repay member digging hubs for the calculations performed. Ethereum offers a decentralized Turing-finish virtual machine, namely the Ethereum virtual machine, which can execute contents by using a worldwide system of open hubs. This system is more or less enclosed, and a confidential exchange estimating system is used to moderate spam and designate assets on the system. Ethereum was proposed, presented, and created in late 2013 by Vitalik Buterin, who is an extraordinary digital currency scientist and developer [22]. Ether went live on 30 July 2015. The release was supported by the online community, which launched the currency with 72 million coins “pre-mined,” and these coins accounted for 70% of the total flow in 2018 [23]. Since its initial launch, Ethereum has undergone many changes and arranged convention redesigns. These modifications are critical changes that have influenced the fundamental usefulness and motivating force structures of Ethereum, utilizing records and equalizations through state changes. The state indicates current adjustments and additional information. A state is not placed away on the blockchain but is placed away in a different Merkle Patricia tree. A digital currency wallet stores the general population and private “keys” or “addresses,” which can be utilized to obtain or spend ether. These “keys” or “addresses” can be produced through BIP 39 style memory aides for a BIP 32 HD Wallet [24]. In Ethereum, this method is pointless because it does not work in an unspent transaction output (UTXO) plot. With a private key, it is possible to write in the blockchain, successfully making an ether exchange. To send ether to a record, you require the general population key of that account. Ether accounts are pseudonyms in the sense that they are not connected to a particular person but instead to at least one explicit location. Proprietors store these addresses in programs, on paper, and possibly in memory. Ether cryptocurrency is completely different from Bitcoin cryptocurrency. By contrast, Bitcoin’s UTXO framework more closely resembles the actions of spending money and accepting change. The aforementioned two frameworks have their advantages and disadvantages in terms of the storage room, multifaceted nature, and security or obscurity.

The proposed method for the price prediction of the ether cryptocurrency involves using the LR and SVM algorithms.

3.1. LR

LR is used to determine the relationship between two continuous variables. One of the variables is the predictor or independent variable, and the other variable is the response or dependent variable. The LR algorithm is used to determine a statistical relationship. In a machine learning algorithm, the main idea is to obtain a line that best fits the data. The line that best fits the data is the one for which the total prediction error (all data points) is as small as possible. The error is the distance between the point and the regression line.

LR is a simple, interesting, and attractive model due to the simple representation of its algorithm. This algorithm is a very basic machine learning algorithm but has numerous applications in many areas [25]. The algorithm is represented in the form of a linear equation with a polynomial degree of 1. The representation is a combination of a specific set of input values (x) and the solution to a particular problem. The prediction of the solution in the form of output values (y) is based on the set of input values (x). Therefore, both the input values (x), which represent various input features, and the output values, which are the price predictions in this project or the category predictions, are always given in a numeric form. With the help of this simple and effective LR method, a single input or dataset is given. This dataset can be used in statistical concepts to estimate the coefficients of the polynomial, which are required to predict the output with this algorithm. The statistical concepts used here involve properties of the data such as means, standard deviations, correlations, and covariance. All the data must be available and easy to traverse to simplify the calculation of these statistical variables. The LR algorithm used in machine learning is represented in Eq. (1).

Because only one feature is used for prediction, the output (y) is given as follows:

$$y = w_0 + w_1 x_1 \tag{1}$$

The gradient descent algorithm can be written as follows:

$$J(w) = 1/2 \sum_{i=1}^{M} (y_i - y1_i)^2 \tag{2}$$

where w_0, w_1, and x are the augmented vectors, $y1$ is the predicted value, and w is the coefficient used in the LR equation. The value of w influences the learning rate (i.e., the gradient descent) and controls the performance of the LR equation.

3.2. SVM

SVM is a supervised learning model associated with learning algorithms, and it analyzes data used for classification and regression analysis. For a given dataset, each data point is classified into one category. The SVM training algorithm formulates a model that assigns the data to one category with the help of a probabilistic binary linear classifier. In an SVM model, data points are represented in space and mapped into categories with a suitable wide gap. The test data is mapped into space, and they are classified into a category with a suitable wide gap. SVM also performs efficiently in nonlinear classification when using a kernel trick. In nonlinear classification, the points are implicitly mapped into high-dimensional feature space [26,27].

The SVM algorithm is applied and activated by a kernel. The hyperplane formed due to this activation, especially for the linear-based kernel in the SVM, is created through a transformation of any type of problem by using some linear equations with a polynomial degree of 1. A very interesting part of this insight is that the linear SVM algorithm can be reused or retransformed using the inner product of any two given observations rather than the observations themselves. The inner product between any two vectors or observations is the sum of the dot product of each of these pairs of input values.

For a linear kernel, the dot product is used. The dot product, especially between two pairs, is a measure of similarity used in a linear-based SVM algorithm or linear-based kernel because the distance defines a linear combination of these inputs or pairs. Therefore, in the case of curved lines, where the degree of the polynomial is greater than 1, instead of using a direct dot product between the two pairs, a polynomial kernel is applied. For a gamma polynomial, which is the most commonly used polynomial, gamma is a crucial parameter specified during learning by the SVM algorithm. For a curved line, the polynomial kernel used is known as a radial kernel. The radial kernel is localized and has the ability to create many complex regions within the feature space, such as closed polygons in two-dimensional space. The equation for the linear kernel of the SVM algorithm is given as follows:

$$K(x, y) = \left(x^t y + c\right)^d S \tag{3}$$

The equation for the radial kernel of the SVM algorithm is as follows:

$$K_{rbf}\left(x, x'\right) = e^{-\gamma x - x'^2} \tag{4}$$

where d is taken as 0 and $\gamma = \frac{1}{2\sigma^2}$ is a free and crucial parameter that must be optimized. One might roughly state that this parameter can be written on the basis of the concept of similarity between a pair of samples or observations. The minus sign in Eq. (4) has the sole purpose of inverting the distance measured into a similarity score. Due to the exponential factor,

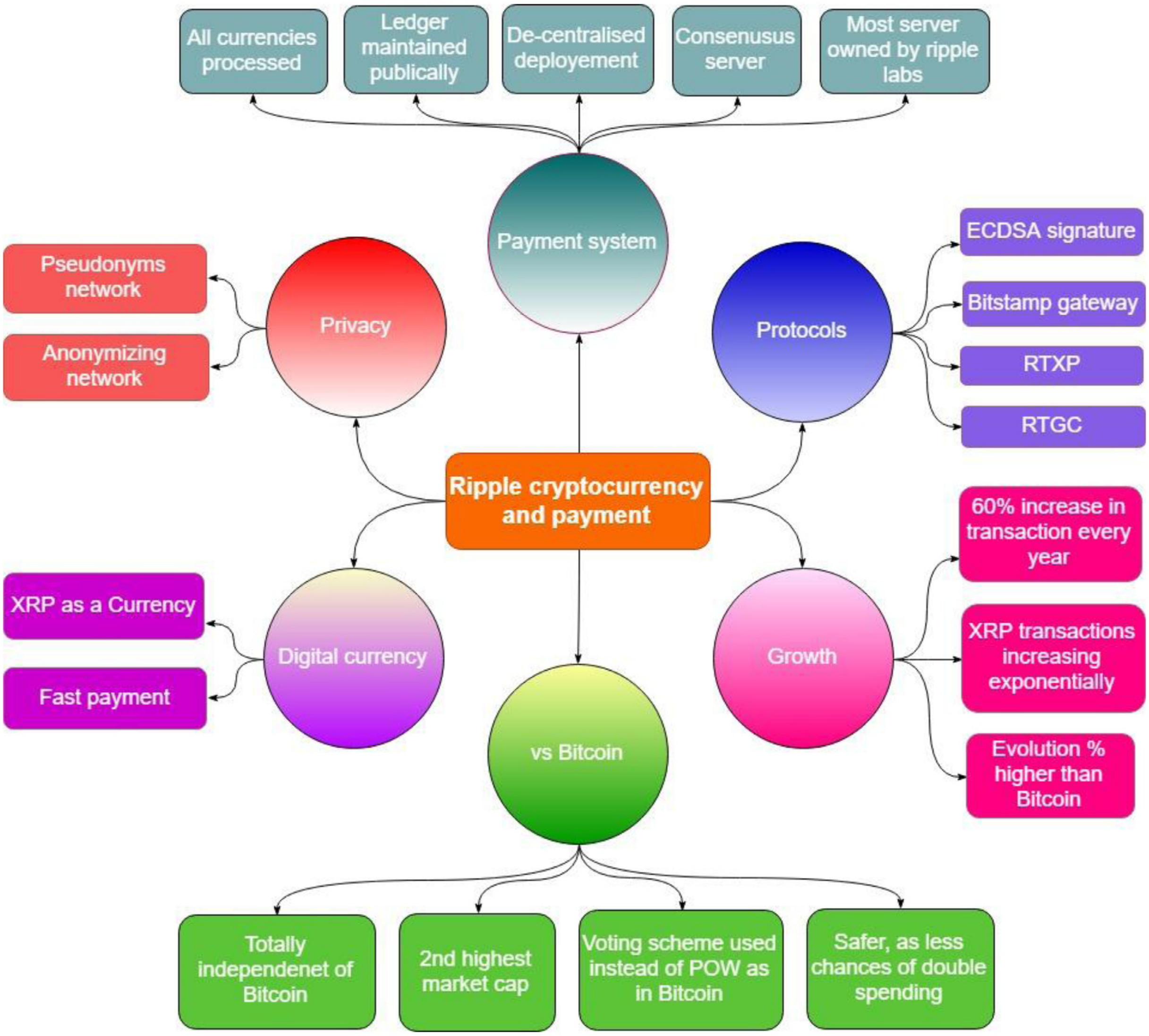

Fig. 4. Mind map for cryptocurrency payments.

the resulting similarity score falls between 1 (for exactly similar samples) and 0 (for very dissimilar samples). Fig. 4 uses protocols such as the real-time gross settlement system and Ripple transaction protocol (RTXP).

Ether has seen a 60% growth in transactions every year and has an evolution percentage higher than that of Bitcoin. Ether is completely independent of Bitcoin and has the second highest market cap. Moreover, ether is safer than Bitcoin because the possibility of double spending is lower in ether.

4. Dataset and processing

4.1. Data collection

Information from the website https://etherchain.org/programminginterface/measurements/cost, which contains each swell exchange preceding the present date, was used in the price prediction experiment. This type of information is an expanding content record, and substantial arrangements of diagrams containing the data about the exchanges are made through swell installments. Every element of a record incorporates the exchange id, sender, beneficiary incentive in Ripple (XRP), and a timestamp. Exchanges including numerous senders and different collectors are referred to by various lines with a similar exchange id. These documents are used to change our information collection into a rearranged collection listed by “client” substance as opposed to addressing the union-find calculation because for some random exchange, only a single element could be the sender of that exchange regardless of whether various sending accounts were utilized. This representation of the data allows the model to explore various properties of the graphs and to use them as features in price prediction. After the extraction of the prices of the ether cryptocurrency, the extractions are stored in the form of a JSON file, which is later converted into a CSV file for making the applications easier.

4.2. Feature extraction

The experiment aggregates a few system-based highlights to build up our administered machine learning calculations. The component families selected the following accompanying methodology: Features were produced for time allotments of 60 min at some point 1 week and 1 month before the pre-phrasing time. The objective was to foresee the cost of XRP in USD 1 h ahead of time. All highlights were registered to utilize the data accessible 1 h before the expectation time and for hub highlights to utilize both the in- and out-degrees. The following accompanying highlights were extracted: (i) current XRP price; (ii) net flow per hour; (iii) number of transactions per hour; (iv) mean transaction value; (v) median transaction value; (vi) average node in- and out-degrees; (vii) median node degree; (viii) alpha constant of the power law; (ix) total amount of XRP mined; (x) number of new addresses; (xi) mean initial deposit amount among new addresses; and (xii) number of transactions performed by new addresses. A considerable number of these highlights were to some degree associated with the cost of XRP; however, most were only inexactly related.

4.3. Feature selection

Selecting the feature with the highest contribution, which highlights the usage, is an essential part of any relapse or arrangement improvement. To prune highlights, the common data between each component must be determined and each irrelevant element must be eliminated from the model. The highlights that were demonstrated to be the most instructive and whose incorporation enhanced our outcomes were (i) current XRP price; (ii) net XRP flow for A, B, and C; (iii) closeness centrality for A, B, and C; (iv) mined XRP in the last hour; (v) number of transactions among new addresses in the last hour; (vi) mean node degree in the last hour; and (vii) net flow in the last hour.

Interestingly, none of the highlights that looked back more than the most recent hour were instructive. By a considerable margin, the most useful element for forecasting the price was the current cost.

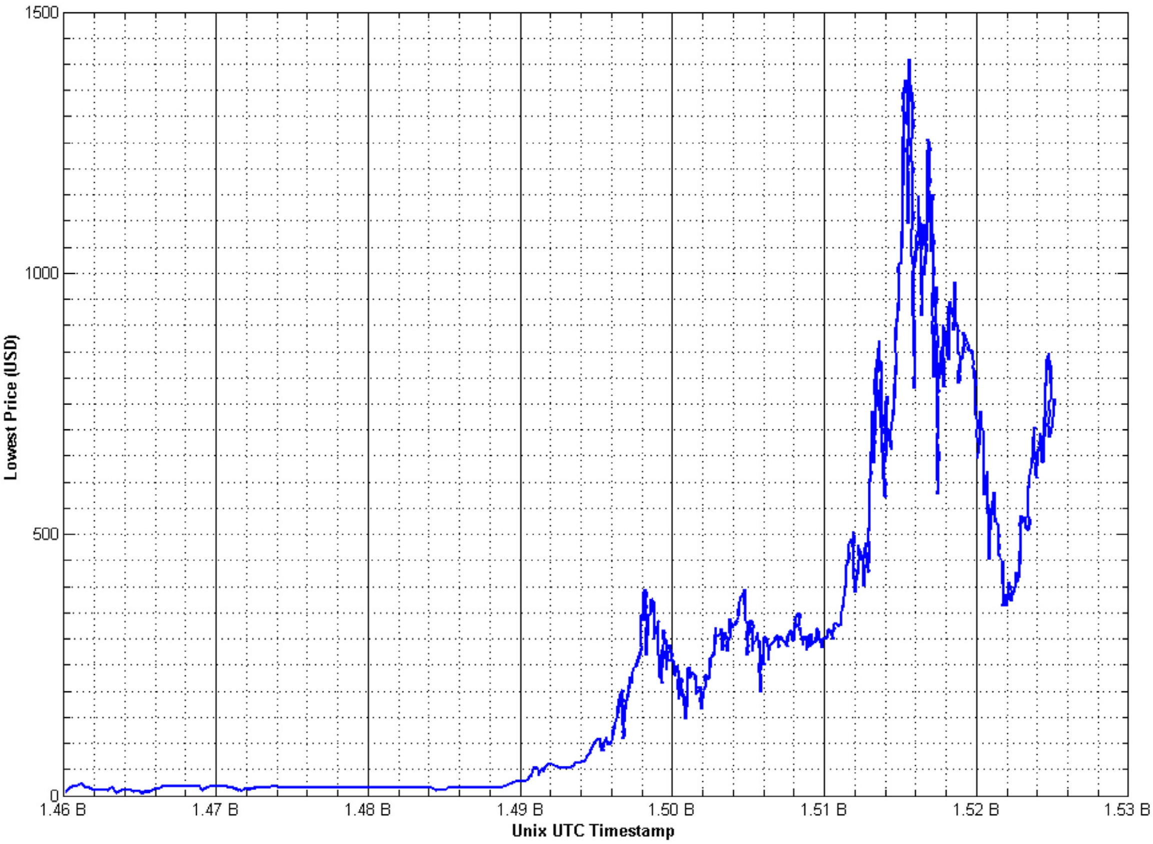

Fig. 5. Ether price variation.

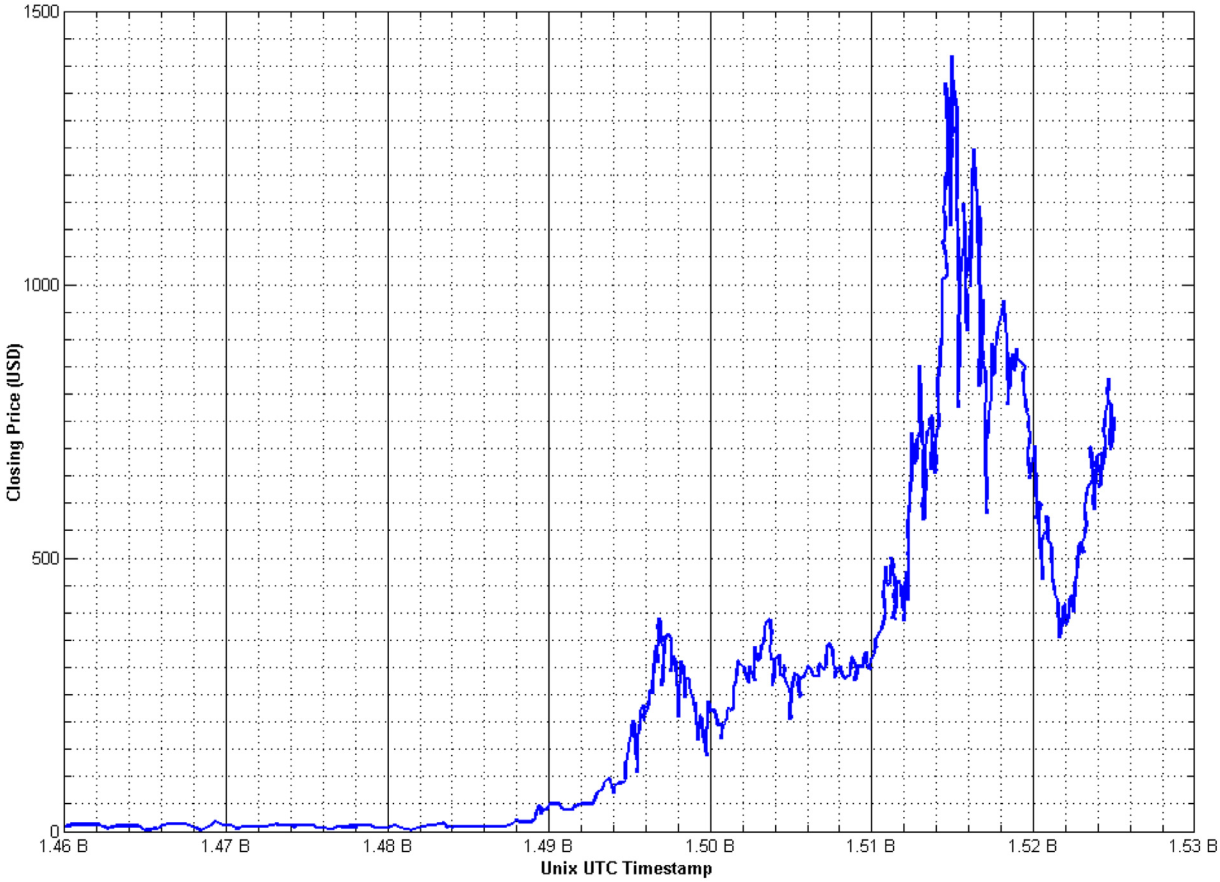

Fig. 6. Prediction of the ether cryptocurrency price by using the SVM algorithm.

5. Experimental evaluations and results

5.1. Experiment setup

An experiment was conducted using a system running on a Hexa-Core i7 Processor clocked at 4.1 GHz and equipped with 16 GB DDR4 RAM and a solid state device (SSD). The experiment was performed using a Jupiter notebook, and the overall run time of the script was approximately 10 min. The algorithm processed a multivariate dataset with 34,428 records, and graphs were generated.

5.2. Simulation results

The experimental procedure and the simulation results for the proposed algorithms are presented in the following text.

STEP 1. All the data available on https://etherchain.org/, which contains every ether transaction made prior to the present date, were extracted. After this extraction, the data were saved in the JSON file format.

STEP 2. The JSON file was converted into a CSV file for the convenience of data processing. This dataset had attributes or parameters such as close, date, open, high, low, quote volume, and the weighted average. Here, our aim was to predict the closing price of the ether cryptocurrency with respect to time or date.

STEP 3. The data were pair-plotted according to the attributes of the dataset. Pair-plotting is conducted to determine what attributes may have a large effect on the price prediction for the ether cryptocurrency. Subsequently, a graph was plotted using the attributes of close and date because the other attributes were able to give a clear picture of the dataset.

STEP 4. As displayed in Fig. 5, the peak occurred around January and a sharp dip occurred in October. Therefore, if you have any plans to invest in the ether cryptocurrency, you should buy ether in October and if you want to make a profit, sell it again in January.

STEP 5. The correlation matrix, which is presented in the form of a heatmap, was checked. The correlation factor should lie between 0.6 and 0.8. A correlation factor lower than this range would possibly lead to underfitting, and a corre-

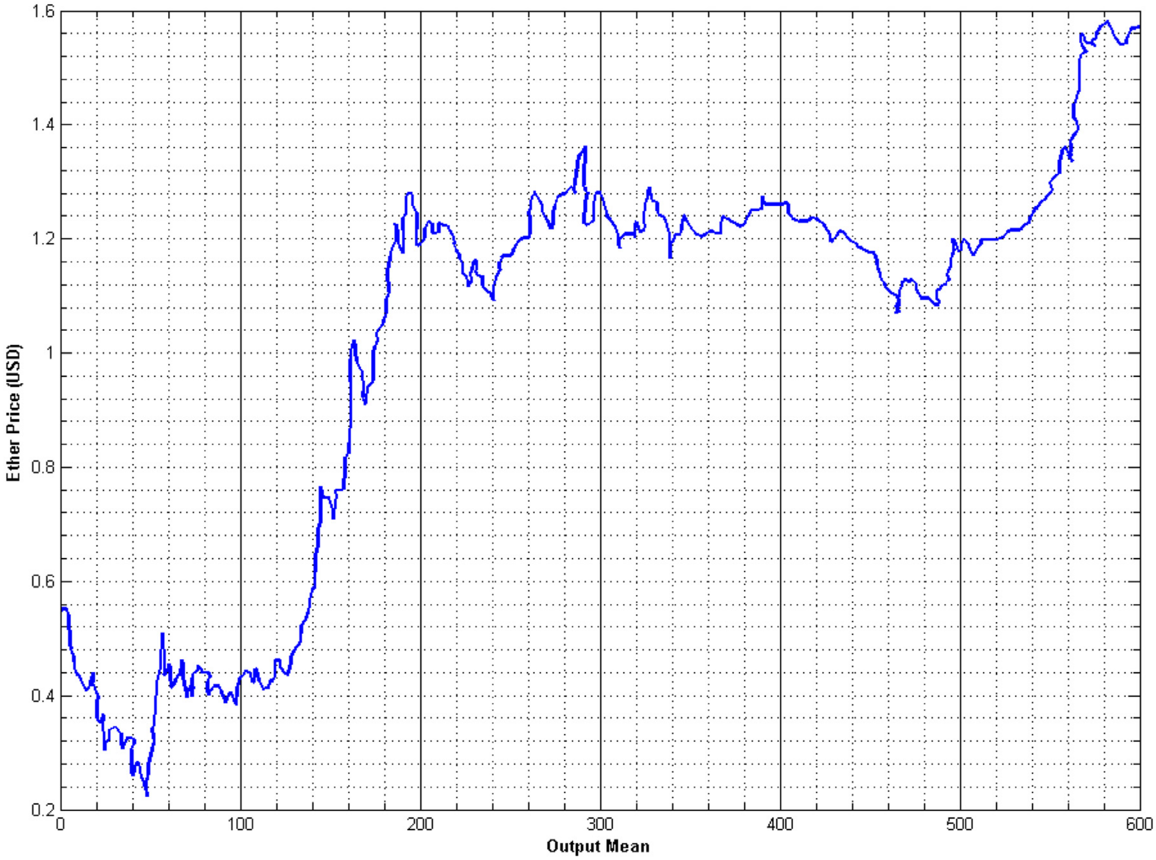

Fig. 7. Prediction of the prices of the ether cryptocurrency by using LR.

lation factor higher than this range would lead to overfitting. The close and date attributes or features had the ideal correlation factor (lying between 0.6 and 0.8). These attributes provided superior prediction results. Price prediction was conducted by performing training over 38,428 data points.

STEP 6. The prediction was initially performed using the LR algorithm; however, the accuracy score was unsatisfactory. Therefore, the SVM algorithm was implemented before the LR method was improved further.

STEP 7. The SVM algorithm used a radial basis filter kernel with values ranging from 0.01 to 0.00001. This machine learning algorithm was assisted by the GridSearchCV function to achieve high accuracy and efficiency.

STEP 8. After running the SVM algorithm, the resultant equation or resultant weights learned during the process of running the SVM algorithm were plotted in the form of a graph, as displayed in Fig. 6.

Fig. 6 clearly depicts that during the previous year, the monetary value of Bitcoin spiked, which clearly reflects that people are accepting cryptocurrency as modes of payments. The *x*-axis represents the exchange value of ether coin in USD, and the *y*-axis represents the UNIX timestamp recorded during the recording of the data.

STEP 9. To improve the performance of the LR algorithm, 17,834 additional data points were added. In addition to adding extra data, feature engineering was applied and the mean, median, and mode of the prices were recalculated. These three parameters of the prices were grouped together for the remaining points in the dataset.

STEP 10. Instead of using only the two attributes of close and date, the mean, median, and mode were included to improve the accuracy of the LR algorithm for the prediction of the prices of the ether cryptocurrency.

STEP 11. The dataset was split into 80% training samples and 20% test samples for cross-validation. Then, the LR algorithm was run on the dataset, and the resultant equation or resultant weight coefficients were obtained. With the help of these coefficients, the price of the ether cryptocurrency could be predicted. A graph was plotted with the equation of the LR algorithm, as displayed in Fig. 7.

5.3. Comparison

In the study, the LR and SVM models were used to provide an easy and computationally feasible approach for price prediction. Initially, when using LR directly over the collected dataset, an unsatisfactory result was obtained (accuracy score

Table 1
Comparison of the accuracy percentage of the proposed algorithms.

Sr. No.	Proposed algorithm	Accuracy percentage (%)
1	Linear regression	85.46
2	Support vector machine (without features)	96.06
3	Support vector machine (with features)	99

of 85.46%). The use of the SVM algorithm with the radial basis kernel increased the accuracy to 96.06%. Table 1 presents the accuracy scores obtained with the proposed algorithms.

With the addition of approximately 17,834 data points to the existing dataset and the use of feature engineering, a very high accuracy score of 99% was achieved. Hence, with the given selected features and SVM model, the price of the ether cryptocurrency can be accurately predicted.

6. Conclusion and future directions

Price prediction can provide stock brokers and cryptocurrency traders with an upper hand in the market. It can also help miners mine coins for maximum profit. The algorithms used yielded highly accurate results. Therefore, the trained model can be deployed and implemented in real time. Compared with the SVM algorithm, LR exhibits a higher accuracy on the window, especially for a large dataset. The LR algorithm works so well because the last three mean values come with sufficient information for the prediction of the next mean value. The results indicate that proposed SVM model without additional features had a 10.62% higher accuracy (96.06%) than the LR method (85.46%). The accuracy difference further increased to 13.54% when adding additional features in the SVM algorithm. Even during a price rise, the proposed models would fail only once when the price suddenly rises. After that, the last three values would reflect a rise in price and the fourth value can be easily inferred. This accuracy is extremely suitable because only the next mean value must be predicted.

In the future, the proposed algorithms may be used to predict the prices of other types of cryptocurrencies. Moreover, with help of the proposed algorithms, blockchain's function can be further improved, which can enhance cybersecurity in the digital world. Such an enhancement in cybersecurity would considerably reduce cyber threats or attacks. Thus, with suitable and sufficient data, one can easily predict the price of cryptocurrencies.

Declaration of Competing Interest

I declare that I have no Conflict of Interest.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compeleceng.2019.106527.

References

[1] Rüßmann M, Lorenz M, Gerbert P, Waldner M, Justus J, Engel P, et al.. Industry 4.0: the future of productivity and growth in manufacturing industries, 9. Boston Consulting Group; 2015. p. 54–89.
[2] Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S. Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press; 2016.
[3] Tschorsch F, Scheuermann B. Bitcoin and beyond: a technical survey on decentralized digital currencies. IEEE Commun Surv Tutorials 2016;18:2084–123.
[4] Grinberg R. Bitcoin: an innovative alternative digital currency. Hastings Sci Tech Law J 2012;4:159.
[5] S. Nakamoto, Bitcoin: a peer-to-peer electronic cash system, 2008.
[6] Wang L, Fu X. Data mining with computational intelligence. Springer Science & Business Media; 2006.
[7] Aitzhan NZ, Svetinovic D. Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. IEEE Trans Dependable Secure Comput 2018;15:840–52.
[8] Khan MA, Salah K. IoT security: review, blockchain solutions, and open challenges. Future Gener Comput Syst 2018;82:395–411.
[9] Niranjanamurthy M, Nithya B, Jagannatha S. Analysis of Blockchain technology: pros, cons and SWOT. Cluster Comput 2018:1–15.
[10] Yuan Y, Wang F-Y. Towards blockchain-based intelligent transportation systems. In: 2016 IEEE 19th international conference on intelligent transportation systems (ITSC); 2016. p. 2663–8.
[11] Garcia D, Tessone CJ, Mavrodiev P, Perony N. The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. J R Soc Interface 2014;11:20140623.
[12] Ciaian P, Rajcaniova M, Kancs dA. The economics of Bitcoin price formation. Appl Econ 2016;48:1799–815.
[13] Zurada J. Data mining with computational intelligence. IEEE Trans Neural Netw 2006;17 826-826.
[14] Yli-Huumo J, Ko D, Choi S, Park S, Smolander K. Where is current research on blockchain technology?—a systematic review. PLoS ONE 2016;11:e0163477.
[15] Beck R, Avital M, Rossi M, Thatcher JB. Blockchain technology in business and information systems research. Springer; 2017.
[16] Sin E, Wang L. Bitcoin price prediction using ensembles of neural networks. In: 2017 13th international conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD); 2017. p. 666–71.
[17] Karasu S, Altan A, Saraç Z, Hacioğlu R. Prediction of Bitcoin prices with machine learning methods using time series data. In: 2018 26th signal processing and communications applications conference (SIU); 2018. p. 1–4.

[18] Phillips RC, Gorse D. Predicting cryptocurrency price bubbles using social media data and epidemic modelling. In: 2017 IEEE symposium series on computational intelligence (SSCI); 2017. p. 1–7.
[19] Rizzo AS, Kim GJ. A SWOT analysis of the field of virtual reality rehabilitation and therapy. Presence Teleoperat Virtual Environ 2005;14:119–46.
[20] Salimitari M, Chatterjee M, Yuksel M, Pasiliao E. Profit maximization for bitcoin pool mining: a prospect theoretic approach. In: 2017 IEEE 3rd international conference on collaboration and internet computing (CIC); 2017. p. 267–74.
[21] Lahmiri S. A comparative study of backpropagation algorithms in financial prediction. Int J Comput Sci Eng Appl (IJCSEA) 2011;1:15–21.
[22] Buterin V. A next-generation smart contract and decentralized application platform. white paper; 2014.
[23] Yang R, Yu FR, Si P, Yang Z, Zhang Y. Integrated Blockchain and edge computing systems: a survey, some research issues and challenges. IEEE Commun Surv Tutorials 2019.
[24] Yeow K, Gani A, Ahmad RW, Rodrigues JJ, Ko K. Decentralized consensus for edge-centric internet of things: a review, taxonomy, and research issues. IEEE Access 2018;6:1513–24.
[25] Sagi O, Rokach L. Ensemble learning: a survey. Wiley Interdiscip Rev 2018;8:e1249.
[26] Jang H, Lee J. An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. IEEE Access 2018;6:5427–37.
[27] Peng Y, Albuquerque PHM, de Sá JMC, Padula AJA, Montenegro MR. The best of two worlds: forecasting high frequency volatility for cryptocurrencies and traditional currencies with support vector regression. Expert Syst Appl 2018;97:177–92.

M. Poongodi worked as Post-Doc fellow at Hamad Bin Khalifa University, Qatar. She received her B.Tech (I.T) from Anna University and M.E (CSE) from St.Peters University. She has obtained her PhD in the area of information security from Anna University, Chennai. Her research interest is also extended in the areas of network analysis using Cloud Computing, and Information Security.

Ashutosh Sharma is currently working as Assistant Prof. in School of Electronics and Communication Engineering at Lovely Professional University, India. He has completed his Ph.D. and M.Tech. in ECE from JUIT, India in 2019 and 2016, respectively. His topics of interest in research are SLA and QoS in communication, risk analysis, VANET, UAV, cloud computing, and machine learning.

Vijayakumar V. is currently a Prof. and an Asso Dean for School of Computing Science and Engineering at VIT University, India. He has completed Diploma with First Class Honors. He has completed BE CSE and MBA HRD with First Class. He has also completed ME CSE with First Rank Award. He has completed his PhD from Anna University in 2012.

Vaibhav Bhardwaj has received their B.Tech from VIT, Chennai, India in 2019 and working in MNCs nowadays and dedicating his time for research in security domain.

Abhinav Parkash Sharma has received their B.Tech from VIT, Chennai, India in 2019 and working in MNCs nowadays as a senior engineer. His area of research interest is industrial management system and security in communication networks.

Razi Iqbal is an Associate Prof at the College of Computer Information Technology at the American University in the Emirates and a Senior IEEE Member. He earned his PhD and Master's degree in Electrical, Electronics and Computer Systems Engineering from Akita University, Akita, Japan. His current research areas are use of short range wireless technologies in Precision Transportation and Education.

Rajiv Kumar received B.Tech. from G.B. Pant Uni. of Agriculture and Technology, India in 1994. He received M.Tech. and Ph.D. in computer network reliability from NITKU in the year 2001 and 2010, respectively. He is working as an Associate Prof in ECE Dept. at JUIT India. His areas of interest in research are performance and reliability of computer network.