Knowledge and Information Systems (2023) 65:1179–1197
https://doi.org/10.1007/s10115-022-01796-0

REGULAR PAPER

A new method of ensemble learning: case of cryptocurrency price prediction

Akhter Mohiuddin Rather[1]

Received: 26 July 2022 / Revised: 29 August 2022 / Accepted: 14 November 2022 /
Published online: 1 December 2022

Abstract
This work proposes a novel method of ensemble learning for time series prediction. Different machine learning-based models have been integrated, and a combined prediction model has been created. The objective of the ensemble model is that it must outperform all other individual models that are used to construct the ensemble model in terms of producing excellent predictions. The field of cryptocurrencies has been selected as the domain of this work where the focus is to predict the cryptocurrency prices using the proposed model. A new regression model is proposed and implemented in this work. Different machine learning techniques have been adopted and integrated to form a combined prediction model. The machine learning models include deep neural networks, support vector regression, and decision trees. The regression scheme has to be implemented on each machine learning model separately as well as their performance is also to be improved. The combined prediction model requires optimal weights generation for integration, and therefore, time complexity is a concern. A large set of experiments have been carried out on various cryptocurrencies and the results are displayed. Real-world data has been used here and a comparison is also performed. It is observed that the combined prediction model outperforms other models resulting in excellent predictions capturing most of the nonstationary movements in the data.

Keywords Cryptocurrency · Deep neural networks · Support vector regression · Decision tree · Machine learning

1 Introduction

A cryptocurrency or simply crypto is a kind of network type having a collection of binary data which serves as a digital exchange medium. The transactions are secured using blockchain technology which uses algorithms of cryptography where data travels in the form of ciphertext, and therefore there is no way to decipher it. The idea behind cryptocurrency was first proposed by [23]. Despite the hurdles in accepting cryptocurrency as valid by many nations, it still gained a lot of attention from researchers as well as from industries and investors.

✉ Akhter Mohiuddin Rather
akhter.m@greatlakes.edu.in

1 Great Lakes Institute of Management, Gurgaon 122413, Haryana, India

Some of the famous cryptocurrencies are Bitcoin, Ethereum, Dogecoin, ADA, etc. Bitcoin was the first cryptocurrency created in 2009 which achieved huge popularity and with a result, hundreds of cryptocurrencies were created later on. The actual bitcoin trading happened in 2013 and there is still a lot to be explored in this field [7]. It is reported that by mid-2020, more than 7000 cryptocurrencies were active and their market cap was more than 300 billion US dollars [38].

The area of predictive analytics has seen a lot of presence in academia as well as in industries for decades. Predictive analytics can have two major sub-categories – classification-based problems and regression-based problems. There are two major methods in solving predictive analytics problems – the statistical approach and the artificial intelligence (AI) or machine learning (ML)-based approach. The statistical models for time-series-based prediction include the famous autoregressive integrated moving average (ARIMA) models [2], smoothing models [5], autoregressive conditional heteroscedasticity (ARCH) models [12], generalized autoregressive conditional heteroscedasticity (GARCH) models [3], etc. AI or ML-based models include various models such as artificial neural networks (ANNs) [31], which have given rise to deep neural networks (DNN) [22, 24, 25, 36], support vector machines (SVM) [6, 33], etc. Both statistical and ML models have their limitations and advantages. For instance, the prediction of time-series data requires a lot of statistical tests to be performed beforehand if solved with statistical regression-based models. Moreover, statistical models are unable to capture the nonstationary movements of the data therefore their use is limited at present. AI models have also come with a cost such as the high computation time required to reach the solution and the cost of implementation, deployment, and retraining of the model with newer data. However, by witnessing the complexity and nature of data in the present era, the AI-based approach seems to be the optimal approach to tackle any problem.

The work presented here is based on the regression-based problem, and therefore, the focus here will be on regression-based problems. Depending on the data available and the statement of the problem, an appropriate regression-based model is chosen. If the data is cross-sectional, then we have multiple regression available. However, if the data is in a time-series fashion, in that case, autoregressive models will be more appropriate to use. This work proposes a combined prediction model for cryptocurrency prices. The integration is done of three different models – DNN, decision trees (DT), and support vector regression (SVR). The idea is that the performance of the proposed model must outperform the other three models in terms of robustness and efficiency with the least prediction error. To test the robustness and efficiency of the proposed model, a large set of experiments were carried out on different cryptocurrency prices and the results were found to be phenomenal.

The remainder of this paper is organized as follows:

In Sect. 2 literature review is discussed; the proposed model is discussed in detail in Sect. 3; materials and methods are presented in Sect. 4; Sect. 5 shows the results obtained; in Sect. 6 the discussion is presented and finally conclusions are presented in Sect. 7.

2 Literature review

Cryptocurrencies are the most happening digital currencies or digital transformations that have attracted many researchers, investors, individuals, and industrialists worldwide. Within just twelve years since their inception, cryptos are now the fastest growing industry, also considered a threat to the current financial system by some experts [19]. There are various cryptocurrencies trading currently and the number is growing. Various studies have appeared that highlight intercorrelations between them and it was observed that Bitcoin is dominant among all [18]. Similar work on connectedness between cryptos was done by (Q. [20, 21]. Some researchers believe that there's some explosivity between various cryptocurrencies, a study on this was also performed [4]. Similarly, competition among cryptocurrencies was performed and it was found that Bitcoin is more valuable than the US dollar [16]. Since 2013, Bitcoin witnessed an amazing growth of 8000% of its value leaving other cryptos behind [10]. Due to the popularity and increased users of cryptocurrencies, problems started to appear such as "double spending" or using the same digital currency in two transactions at a time. This also led cryptocurrencies to become an easy target for hackers [11, 35]. To solve this problem and other similar problems, a robust protection system was created known as 'blockchain' which is an anti-fraud system [13]. Investors have found cryptocurrency as a big asset of investing which gives absolute returns. At the same time, they are worried about the risks involved too, thus mitigating risk while investing is another factor that any risk-averse investor would be interested in [8].

Research is going to predict cryptocurrency prices, being a difficult task, researchers take advantage of robust models of AI. ANN, RF, and multiple regression techniques were applied in predicting Bitcoin prices and the results have shown that RF has outperformed other models; however, there is still a lot of improvement to be made [30]. A new method showing the prediction of cryptocurrency exchange rates appeared where researchers have used ANN with convolutional components [1]. A classification-based work for Australian listed cryptocurrency was performed based on various factors such as investors of blockchain technology and published information on the company website [14]. A time-series-based prediction was performed using linear regression and SVM of daily closing prices using different window sizes and it was observed that SVM outperformed linear regression [26]. A recurrent neural network was also used for a similar purpose, and the results were satisfactory [17]. Similarly, a hybrid prediction model having a hybridization of DNN and gated recurrent units was proposed for the prediction of two cryptocurrencies – Litecoin and Zcash [34]. [28] proposed a hybrid model for stock prediction by integrating recurrent neural network with two other statistical models. Similarly for Bitcoin prediction, DNN, and long short-term memory (LSTM) were used and the results showed that the LSTM outperformed DNN with excellent performance [20, 21]. LSTM was also used to predict stock returns using a new regression model and finally, a prediction-based portfolio was constructed for investment purposes [27].

3 Proposed model

This section discusses the proposed model in detail which includes the mathematical foundations behind the model as well as its implementation in the AI-ML model. Henceforth, the final model is named a combined prediction model (CPM).

3.1 The autoregression model

The autoregression model (ARM) is inspired by the early work of [15]. ARM treats time-series data $C_t = \left(c_{t-(T-1)}, \ldots, c_{t-2}, c_{t-1}, c_t\right)$ where T is the number of observations in the series, c_t being the latest observation and $c_{t-(T-1)}$ being the last observation in the series. Let $\wp$ be the order of regression, a new equation can be built shown in Eq. 1 which can be treated as a moving reference.

$$\mathcal{T} = c_{t-(\wp-1)-1} \tag{1}$$

An autoregression scheme can be thus formed as shown in Eq. 2.

$$\widehat{\hat{C}_{t+1} - \mathcal{T}} = \left(c_{t-(\wp-1)} - \mathcal{T}, \ldots, c_{t-2} - \mathcal{T}, c_{t-1} - \mathcal{T}, c_t - \mathcal{T}\right) \tag{2}$$

where $\hat{C}_{t+1} - \mathcal{T}$ is the future prediction.

The predictions will be obtained using the ML-based approach by implementing the above regression model on it. $\hat{C}_{t+1} - \mathcal{T}$ is by default the dependent variable, independent variables will be created by choosing the regression order $\wp$. By applying Eq. 1 in Eq. 2, independent variables and a dependent variable can be created. The dimension of independent variables is equal to chosen regression order $\wp$. The prepared data is thus a differenced series where moving reference $\mathcal{T}$ is subtracted from each value. By using the data thus prepared, an ML model can be trained and tested. Once the output is collected from the model, the final prediction values can be calculated using Eq. 3.

$$\hat{C}_{t+1} = \widehat{C_{t+1}} - \mathcal{T} + \mathcal{T} \tag{3}$$

3.2 ARM implemented on ML models

An ARM can be implemented on various ML models with the common goal of producing excellent output compared to statistical models. After defining ARM, the next step is to implement it on ML models. In this work, three such models have been selected – DNN, SVR, and DT. The method of implementation of ARM on the three ML models is shown below.

3.2.1 Implementation of ARM on deep neural network

DNNs are advanced and enhanced ANNs whose function is to process and predict information. An ANN is a data processing unit inspired by the human brain to make the machine as intelligent as a human brain. ANN is composed of several layers and these layers are composed of units known as neurons (S. [37]. There is always one input layer, one output layer, and one or more hidden layers between these two layers. The input layer determines the dimensionality of the input or independent variables of the dataset where one neuron represents one feature. Similarly, output neuron(s) in the output layer represents the dependent variable. This allows the network to learn in a supervised manner. Figure 1 shows an ANN or even a DNN on which ARM can be implemented in a supervised fashion. There are n number of neurons in the input layer where each component on the right-hand side of Eq. 2 can be fed to the input layer such that $\left(x_1 = c_t - \mathcal{T}, x_2 = c_{t-1} - \mathcal{T}, \ldots, x_n = c_{t-(\wp-1)} - \mathcal{T}\right)$ and $\widehat{\hat{C}_{t+1} - \mathcal{T}}$ is shown to the output layer so that the network learns in a supervised manner. The network can have any number of hidden layers and the number of neurons in each layer is chosen by trial and error. The connection links between the layers are associated with some random weights in the first iteration of weight generation. To minimize the prediction error, these random weights get optimized from the second iteration onwards using the delta learning rule [29]. A set of such iterations till the entire data is fed to the network constitutes one epoch. Depending upon the length of data, noise present, nonstationarity, or nonlinearity in data, a network can take thousands or even millions of epochs to converge.

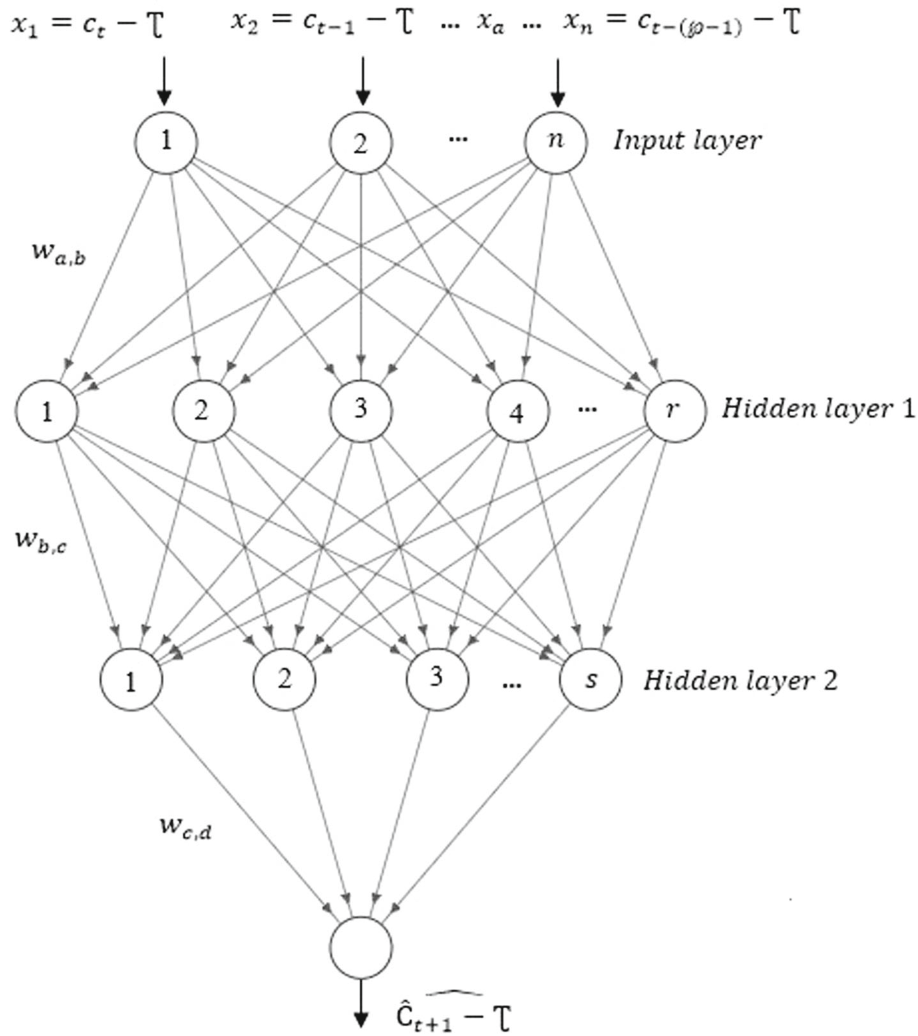


Fig. 1 Deep neural network for ARM

ANNs or DNNs possess some nonlinear activation functions in hidden layers which let the network capture and thereby learn complex patterns in the data. As shown in the figure, the data is passed from the input layer to the first hidden layer which has r number of neurons. The weighted sum is calculated at each hidden neuron (Eq. 4) followed by passing the weighted sum through a nonlinear activation function (Eq. 5).

$$h_r = \sum_{a=1}^{n} x_i w_{a,b} \tag{4}$$

$$y_r = \xi(\gamma + h_r) \tag{5}$$

where ξ is a nonlinear activation function chosen and γ is a bias unit. A user has a choice of choosing any activation functions from several functions. These activation functions include sigmoid, hyperbolic tangent, rectified linear unit (ReLU), etc.

Similarly, in the second layer similar process takes place as shown in the following equations. Equation 6 shows the weighted sum taking place at the second hidden layer.

$$h_s = \sum_{b=1}^{r} y_r w_{b,c} \tag{6}$$

The weighted sum is thereby passed through a nonlinear activation function ξ as shown by Eq. 7.

$$y_s = \xi(\gamma + h_s) \tag{7}$$

Similarly, in the second layer similar process takes place as shown in the following equations. Finally, the output is calculated at the output layer as shown in Eq. 8.

$$\widehat{\hat{C}_{t+1} - \mathcal{T}} = \sum_{c=1}^{s} y_s w_{c,d} \tag{8}$$

The nonlinearity present in the data is captured by the hidden layers and it is the activation functions that do this job. Therefore, hidden layers are the most important components of ANN. Figure 2 shows one hidden neuron and its functionality, as shown there are n inputs $(x_1, x_2, \ldots, x_n)$ coming from the input layer to hidden neuron; these inputs are associated with random weights $(w_1, w_2, \ldots, w_n)$; bias input is shown by $x_0 = 1$ which is also associated with its weight w_0. The output y is calculated by Eq. 9.

$$y = \left(\sum_{i=1}^{n} x_i w_i + w_0 \right) \tag{9}$$

The limitation of DNN is that it required huge amount of data to train and extensive use of GPUs that increase the cost.

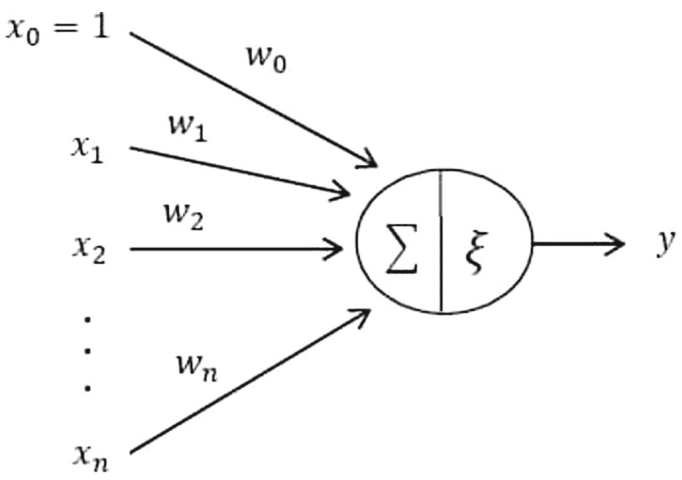


Fig. 2 Artificial neuron in a hidden layer

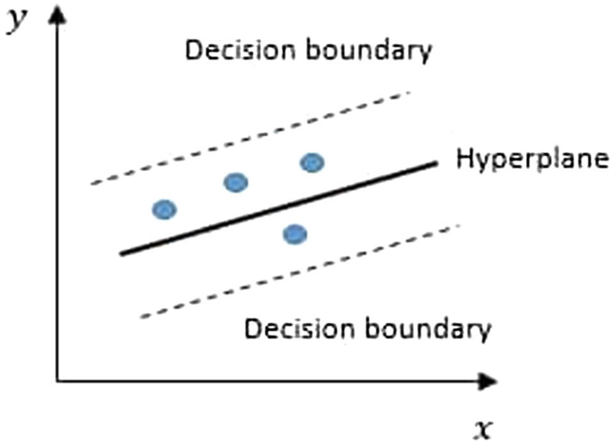


Fig. 3 Support vector machine

3.2.2 Implementation of ARM on a support vector machine

The set of inputs and desired output was already created using ARM. The same regression technique can be now implemented on SVR too in a supervised manner. SVR can be used for regression problems while retaining the characteristics of SVM which states finding the optimal hyperplane. SVM regression or more specifically SVR uses a nonparametric method and completely relies on various kernel functions. Figure 3 shows the typical SVM in two-dimensional space, where there is a hyperplane with decision boundaries. In the classification task, the optimal hyperplane would be the one where the distance between the two classes is maximum. However, a regression problem has a dependent variable which is a continuous number, and it is difficult to estimate due to infinite possibilities. The hyperplane has decision boundaries that are used to estimate or predict the output. The data points closest to the hyperplane are known as support vectors.

The support vectors play a crucial role in both classification as well as in regression. SVM uses mathematical functions known as kernel functions that take input data in the required form. The role of any kernel function is to find a hyperplane in a higher-dimensional space. Various available kernel functions are radial basis function (RBF), sigmoid, linear, and polynomial. The main idea of choosing SVM is that it is robust to outliers and can be easily updated and implemented. The limitation of SVM is its time complexity especially while finding optimal parameters during hyperparameter optimization.

3.2.3 Implementation of ARM on a decision tree

Similarly, ARM can be implemented on DT in a supervised manner. The input–output pairs are already created using ARM, and therefore, the data is ready to be fed to DT. Figure 4 shows a typical DT with different nodes each performing some function. The root node starts with any of the independent variables created earlier $(c_{t-(\wp-1)} - \mathcal{T}, \ldots, c_{t-2} - \mathcal{T}, c_{t-1} - \mathcal{T}, c_t - \mathcal{T})$ and the dependent variable being $\widehat{\hat{C}_{t+1} - T}$. The splitting starts from the root node by checking the condition using entropy or Gini index. The algorithm chooses a variable and its value for splitting so that the two groups are different from each other. A good split is

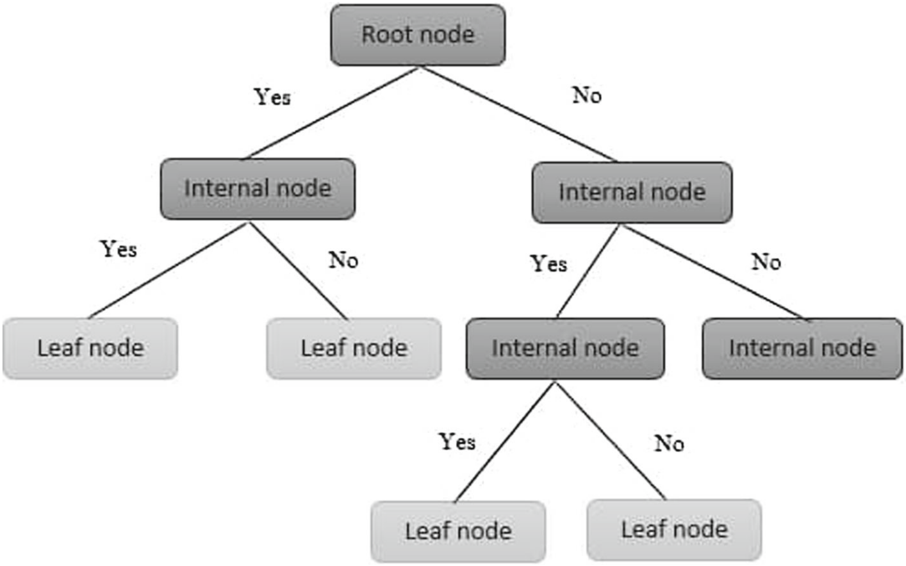


Fig. 4 Decision Tree

decided based on the weighted mean of the mean squared error (MSE) of the two groups. The final decision happens at terminal or leaf nodes. To find the best parameters, hyperparameter optimization can be used which results in the optimal results with the least prediction error. Depending upon the complexity of noise in the data, this process can involve high computation time which is a limitation of the model.

3.3 Combined prediction model

CPM is constructed by the combination of different predictive models with the aim that its predictive performance should be superior to the models that it is constructed. To combine different models optimal weights will be required for each model. These weights have to be generated using an optimization technique. CPM can be a mix of statistical and AI-ML-based models. Suppose we have a choice of N different predictive models $(hv_1, hv_2, \ldots, hv_N)$ as shown in Fig. 5. Each of the models produces N different prediction series $\left(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_N\right)$, respectively.

These predictions are multiplied by the corresponding weights $(v_1, v_2, \ldots, v_N)$ which must be optimal weights such that there is the least minimum prediction error. Lastly, the final prediction $¥_{t+1}$ is generated. The weights are generated using the following optimization model.

$$\text{Min } C_t = \frac{(C_t - ¥_{t+1})^2}{T} \tag{10}$$

$$s.t.$$

$$0 \le v_i \le 1 \tag{11}$$

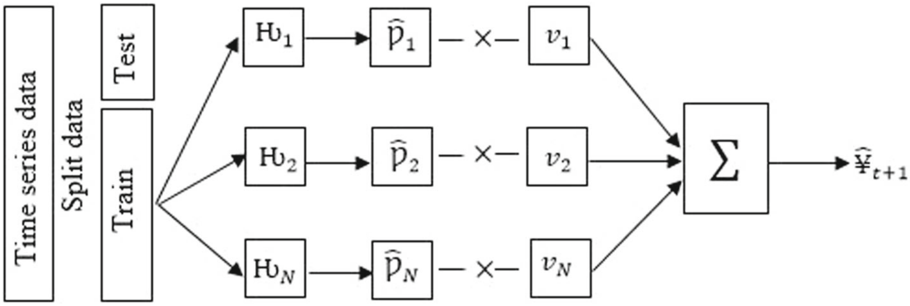


Fig. 5 Combined prediction

$$\sum_{i=1}^{N} v_i = 1 \tag{12}$$

Equation 10 minimizes the mean squared error (MSE) between target data and predicted data; the first constraint shown by Eq. 11 ensures that the weights range between 0 and 1 and Eq. 12 makes a sum of weights equal to 1.

Since CPM is the hybridization of multiple individual predictive models, it is obvious that the models which perform better attain better weights and vice versa.

4 Materials and methods

To prove the robustness of the proposed model, the domain of cryptocurrency prices was selected. The idea behind choosing such a domain is that the historical prices of cryptocurrencies are nonstationary. Forecasting such nonstationary can be challenging for most predictive models. Table 1 shows the list of top ten cryptos (according to Forbes) that have been considered for experimentation in this work. The name, symbol, and market cap of each of the

Table 1 List of top ten cryptos. Source: https://www.forbes.com/

S.No	Crypto	Symbol	Market cap ($ billion)	Date: since (dd:mm:yyyy)	Number of observations
1	Bitcoin	BTC	723	01–01-2020	863
2	Ethereum	ETH	333	01–01-2020	863
3	Tether	USDT	83	01–01-2020	863
4	Binance Coin	BNB	62	01–01-2020	863
5	USD Coin	USDC	49	01–01-2020	863
6	Solana	SOL	29	11–08-2020	640
7	XRP	XRP	29	01–01-2020	863
8	Terra	LUNA	28	21–08-2020	630
9	Cardano	ADA	26	01–01-2020	863
10	BinanceUSD	BUSD	19	01–01-2020	863

Table 2 Mean and standard deviation

S.No	Crypto	μ	σ
1	Bitcoin	30,986.0	18,611.0
2	Ethereum	1754.0	1420.0
3	Tether	1856.0	1530.0
4	Binance coin	231.0	212.0
5	USD coin	242.0	220.0
6	Solana	68.0	69.0
7	XRP	1.0	0.0
8	Terra	30.0	32.0
9	Cardano	1.0	1.0
10	BinanceUSD	1.0	0.0

ten cryptos are shown in columns 2, 3, and 4, respectively. Daily closing prices of each cryptocurrency exchanged for US dollars have been considered here. The fifth column shows the past or beginning date of trading for each crypto, since then the data is considered here. The data has been considered from 01–01-2020 for most of the cryptos (except for Solana and Terra as data) till 12–05-2022.

The mean (μ) and standard deviation (σ) of each crypto has been calculated as shown in Table 2. Some cryptos have a high mean and high standard deviation; for instance, Bitcoin, Ethereum, and Tether while others have a low mean and low standard deviation, for instance XRP, Cardano, and BinanceUSD.

The following are the core steps that have been followed:

- *Data extraction* The data for the above-listed cryptocurrencies have been obtained using 'alpha vantage' (https://www.alphavantage.co). Alpha vantage provides a free API key that can be used to extract the stock data directly to a python program. τ
- *Exploratory data analysis* This step involves visualizing the data to check its nonstationary behavior.
- *Data preparation* Using Eq. 1 and Eq. 2 of ARM prepare input and output variables. The order of regression ($\wp$) has been chosen as equal to 4 initially. To further assess the regression scheme and its impact on the final model, $\} = 5$ and $\wp = 6$ have also been tested. The length of the dataset thus prepared varies with the chosen regression order. If the length of a variable is L, then after performing ARM the length reduces to $L - \}$.
- *Training & testing models* This step involves training and testing the chosen AI-based models with the data prepared in the previous step. Data has to be split into training and test sets sequentially, i.e., the major part of past data can be used to train the model and the recent part can be used to test it. Since different models have to be trained and tested several times (as different regression orders have been chosen) and also depending on the nonstationary behavior of the data, this step requires heavy computation time.
- *Calculating forecasts for each model* As discussed earlier, the output collected from each model is not the final prediction. According to Eq. 3, the final prediction is calculated by adding $\mathcal{T}$ values back to the prediction collected from each model.
- *Constructing CPM* CPM is finally built by integrating the predicted data of each model. To integrate the data, optimal weights are generated using an optimization model discussed earlier.

Table 3 Model parameters

Models	Parameters
DNN	***Neurons***:$\wp$: 200 : 150 : 100 : 50 : 1
	($\wp$: Number of neurons in an input layer and the rest are the number of neurons in each of the four hidden layers and one neuron in the output layer, respectively)
	Optimizer: Adams
	Epochs: Set to 1000 initially
	Dropout rate: 20% in each hidden layer
	Early stopping: Yes
	Loss function: MSE
SVR	***Kernel***: Radial basis function (rbf)
	Gamma**:** Scale
	C*(regularization parameter)*: 1
	Method used: Grid search cross-validation
DT	***Error metric***: Squared error
	Max_depth*(maximum depth of a tree)*: 5
	Min_samples_split*(minimum samples required to split a node of a tree)*: 40
	Min_samples_leaf*(minimum samples required in a leaf node)*: 30
	Criterion*:* Gini
	Method used: Grid search cross-validation

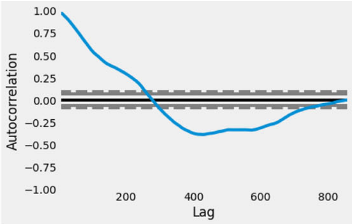
4.1 Optimizing model performance

AI-ML-based models are prone to overfit and the initial results need not be always true. There's a lot of scope for improving the performance of models. K-fold cross-validation method is a very famous technique for avoiding overfitting and is used by many researchers [39]. There are other methods too such as ensembling learning [9], bagging & boosting [32], or including more samples in training data. Some of the methods are common to various AI-ML models while some are restricted to particular models. For instance, K-fold cross-validation is a method common for all models while early stopping and dropout are methods available to DNNs only. For the given problem, DNN, SVR, and DT were finally run and results were obtained. To avoid overfitting and achieve genuine results, hyperparameter optimization was performed for SVR and DT and DNN was also reconfigured and some parameters were changed. Table 3 shows the best parameters obtained for each model.

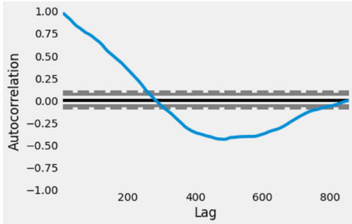
The models were implemented in python on google colab. Scikit-learn (https://scikit-learn.org) libraries of SVR and DT were for the purpose. For implementing DNN, Keras (https://keras.io) which is an API for DNN, and TensorFlow (https://www.tensorflow.org) which is an AI-ML platform was used.

5 Results

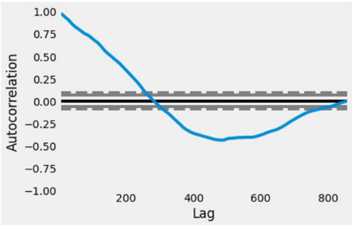
The set of experiments was carried out for each of the selected cryptocurrencies. To check how the observations in the time series are related autocorrelation function (ACF) was calculated for each cryptocurrency. Figure 6 shows the correlogram (an ACF plot) for all six



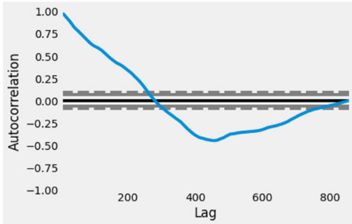
ACF of BTC



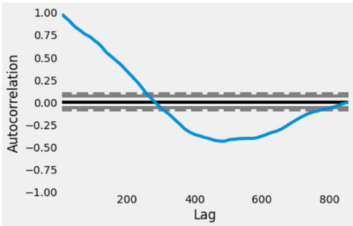
ACF of ETH



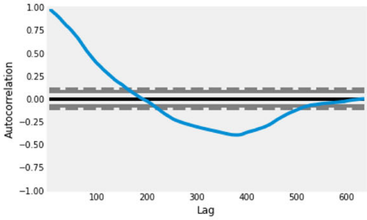
ACF of USDT



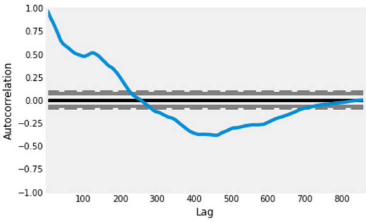
ACF of BNB



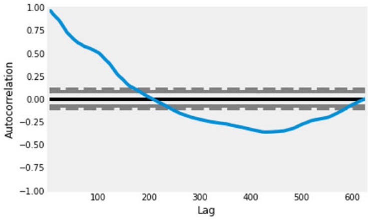
ACF of USDC



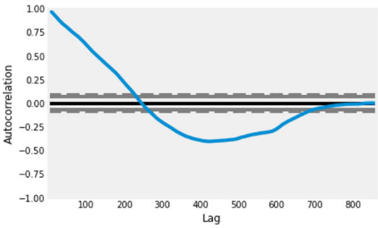
ACF of SOL



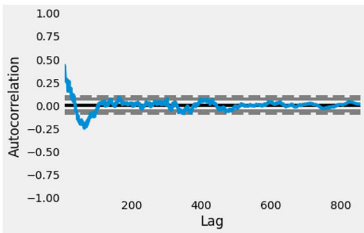
ACF of XRP



ACF of LUNA



ACF of ADA



ACF of BUSD

Fig. 6 Correlogram of top ten cryptos cryptocurrencies. Depending on the number of observations in each crypto, the correlogram is drawn accordingly. As seen, some correlograms are drawn up to 600 lags while others are drawn up to 800 lags. The purpose of calculating ACF for each crypto is that it helps to know the similarity of the time series and its lagged form. This allows comparing the current price of any crypto to its past value. For instance, Figure 6.1 shows that the correlation of BTC is positive with itself till 300 past time lags, beyond which the correlation is negative till almost 800 lags and it is 0 at 800th lag. Similar insights can be drawn for the rest of the cryptos.

The results obtained from DNN, SVR, and DT were collected and corresponding $\mathcal{T}$ values were added back which constituted final predictions. The next step was to construct CPM by integrating the predictions obtained from the three models. To do so optimal weights were generated as discussed earlier. Once the models generated predictions, it was observed that the predictive performance of DNN, SVR, and DT was satisfactory. The models were able to capture the nonlinear movements of the data properly. DNN outperformed SVR and DT; as expected the performance of CPM exceeded the expectations as its predictive performance has remained extraordinary for all the cryptos. For the sake of convenience, the output of CPM is shown below for three cryptos BTC, SOL, and BUSD in Figs. 7, 8, and 9, respectively. Each graph shows training data, test data, and prediction. As seen the nonlinear or nonstationary movements are captured by the CPM properly. Similar observations were found for the rest of the cryptos also.

Depending on the nonstationarity in the data, different cryptos had taken a different number of epochs to reach minimum prediction error. As stated earlier, the early stopping and dropout method was adopted, thereby the optimal number of epochs required by DNN for each crypto is shown in Table 4. Bitcoin consumed more epochs followed by Ethereum, Tether, and Binance Coin. The rest of the cryptos required less epochs.

Table 5 shows the R^2 of the test set for each model. It is observed that DNN outperforms SVR and DT and as expected CPM outperforms all the three selected models.

Similarly, Table 6 shows the root mean squared error (RMSE) of the test set for each model. As expected, the CPM outperforms other models having the least prediction error.

Fig. 7 Prediction for BTC

Fig. 8 Prediction for Solana

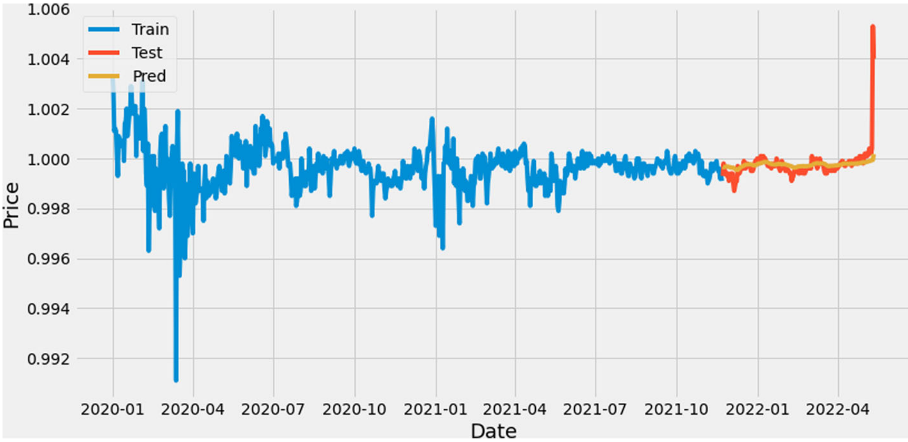


Fig. 9 Prediction for BUSD

Table 4 Epochs for each crypto

S.No	Name	Epochs #
1	Bitcoin	180
2	Ethereum	120
3	Tether	110
4	Binance Coin	110
5	USD Coin	100
6	Solana	100
7	XRP	100
8	Terra	100
9	Cardano	100
10	BinanceUSD	90

Table 5 R-squared values of all models

S.No	Name	R^2			
		DNN	SVR	DT	CPM
1	Bitcoin	0.88	0.76	0.74	0.94
2	Ethereum	0.84	0.71	0.70	0.91
3	Tether	0.81	0.72	0.68	0.90
4	Binance Coin	0.82	0.72	0.70	0.94
5	USD Coin	0.82	0.71	0.73	0.93
6	Solana	0.80	0.69	0.71	0.91
7	XRP	0.79	0.65	0.68	0.91
8	Terra	0.76	0.70	0.71	0.92
9	Cardano	0.78	0.71	0.71	0.90
10	BinanceUSD	0.82	0.70	0.66	0.92

Table 6 RMSE of all models

S.No	Name	RMSE			
		DNN	SVR	DT	CPM
1	Bitcoin	700.0	720.0	725.0	673.0
2	Ethereum	1701.0	1800.0	1820.0	1650.0
3	Tether	1800.0	1870.0	1840.0	1765.0
4	Binance Coin	72.0	80.0	88.0	42.0
5	USD Coin	1840.0	1900.0	1950.0	1645.0
6	Solana	18.0	42.0	68.0	13.30
7	XRP	0.45	0.60	0.80	0.25
8	Terra	0.78	0.90	1.40	0.59
9	Cardano	1.0	1.4	1.3	0.08
10	BinanceUSD	0.8	0.85	0.70	0.01

6 Discussion

The cryptocurrency market is very new, unexplored as well as highly volatile. Lack of information and the trading barriers on cryptocurrency still make its presence among selected groups and individuals. However, there is a growing interest among various sections of society who are keen to know, learn and invest in crypto. Most of banks and financial institutions do not have the authorization to operate on cryptocurrency and there are some valid reasons for this. For instance, in India, all the banks and financial institutions whether private sector or public sector follow the guidelines of the Reserve Bank of India (RBI) therefore RBI is treated as the governing bank. RBI does not recognize cryptocurrency for various reasons, one such reason is that cryptocurrency is decentralized and it is not issued by any authentic agency, or by any government. Most countries follow similar rules for cryptocurrency trading, therefore its reach is up to individual traders or investors. After the arrival of Bitcoin, other players also followed the suit. The number of cryptocurrencies that arrived in the market became too high, therefore building a predictive model that captures patterns of all the cryptocurrencies is very challenging. Another factor that makes cryptocurrency prices difficult to predict is that the prices are independent of time and these prices are affected by some external factors.

Technology has advanced a lot, such as increased processing that includes GPU processing on a local system or cloud space such as Amazon Web Services, Google Cloud, and Microsoft Azure. Such advancements in technology affect various cryptocurrency prices differently. This makes the crypto prices more volatile and thus difficult to predict. Public opinion or their perception also affects the prices. Research has revealed that popular currencies are expected to have higher prices, the best example is Bitcoin. To predict cryptocurrency prices based on sentiment analysis and opinion polls is challenging. Many countries do not accept cryptocurrency as a legal currency and those who accept they do with some legal aspects and terms and conditions. Yet many countries have banned cryptocurrencies. Such factors have a lot of effect on the perception of people which as a result impacts crypto prices directly. Even being a very young area, the prediction and analysis of cryptocurrencies is still one of the most debated areas and also a very elusive task. The researchers and scholars are putting massive efforts into understanding the complexity of cryptocurrency prices and various methods associated with it. The pace at which the crypto industry is growing poses a threat to the existing monetary and financial system and many organizations have already started to feel about the risks associated with it. Cryptocurrencies are very different than traditional currencies and as already discussed it is challenging to capture the volatility of such currency, therefore this area deserves further study and research.

7 Conclusions

This work proposed a new autoregression scheme for cryptocurrency prediction and implemented it on three models that include DNN, SVR, and DT. A combined prediction model was proposed by integrating these three models. The aim of introducing a combined prediction model is to produce excellent predictions and it must outperform the three individual predictive models, i.e., DNN, SVR, and DT. The integration of the models requires optimal weights to be generated corresponding to each of the three models. The optimal weights were generated using an optimization technique. Different cryptocurrencies were included for experimentation that tested the robustness of the proposed model. All three predictive models were able to produce excellent predictions and as expected, the integrated model was able to outperform the other three models. The goal of this work is to capture the volatility of the cryptocurrencies and produce up-to-date predictions that will help an investor to select the better cryptocurrencies and invest accordingly. The proposed model may not always capture complexity in the data, therefore this is a limitation of this work. Therefore, there is still scope for refinement and improvement in this work. The future work includes researching new methods to further improve the prediction performance of the proposed model and to know the public opinion on cryptocurrencies by performing sentiment analysis. This is certainly a promising approach for future research.

Acknowledgements The author highly acknowledges the research support received from the Great Lakes Institute of Management, Gurgaon.

Author contributions I am the sole author.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Alonso-Monsalve S, Suárez-Cetrulo AL, Cervantes A, Quintana D (2020) Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators. Expert Syst Appl 149:113250. https://doi.org/10.1016/J.ESWA.2020.113250
2. Bartholomew DJ, Box GEP, Jenkins GM (1971) Time Series Analysis Forecasting and Control. Operat Res Q (1970-1977) 22(2):199. https://doi.org/10.2307/3008255
3. Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31(3):307–327. https://doi.org/10.1016/0304-4076(86)90063-1
4. Bouri E, Shahzad SJH, Roubaud D (2019) Co-explosivity in the cryptocurrency market. Financ Res Lett 29:178–183. https://doi.org/10.1016/J.FRL.2018.07.005
5. Brown, R. G. (2004). Smoothing, Forecasting and Prediction of Discrete Time Series (Phoenix Edition). https://woqylenipo.leftebook.icu/smoothing-forecasting-and-prediction-of-discrete-time-series-book-8208wv.php
6. Burges CJC. (1998) A tutorial on support vector machines for pattern recognition. Data Mining Knowl Discov 2(2):121–167. https://doi.org/10.1023/A:1009715923555
7. Caporale GM, Gil-Alana L, Plastun A (2018) Persistence in the cryptocurrency market. Res Int Bus Financ 46:141–148. https://doi.org/10.1016/J.RIBAF.2018.01.002
8. Corbet S, Meegan A, Larkin C, Lucey B, Yarovaya L (2018) Exploring the dynamic relationships between cryptocurrencies and other financial assets. Econ Lett 165:28–34. https://doi.org/10.1016/J.ECONLET.2018.01.004
9. Cui S, Wang Y, Wang D, Sai Q, Huang Z, Cheng TCE (2021) A two-layer nested heterogeneous ensemble learning predictive method for COVID-19 mortality. Appl Soft Comput 113:107946. https://doi.org/10.1016/J.ASOC.2021.107946
10. David SA, Inacio CMC, Nunes R, Machado JAT (2021) Fractional and fractal processes applied to cryptocurrencies price series. J Adv Res 32:85–98. https://doi.org/10.1016/J.JARE.2020.12.012
11. de Albuquerque BS, de Castro Callado M (2015) Understanding Bitcoins: Facts and Questions. Rev Bras Econ 69(1):3–16. https://doi.org/10.5935/0034-7140.20150001
12. Engle RF (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50(4):987. https://doi.org/10.2307/1912773
13. Extance A (2015) The future of cryptocurrencies: Bitcoin and beyond. Nature 526(7571):21–23. https://doi.org/10.1038/526021A
14. Frankovic J, Liu B, Suardi S (2021) On spillover effects between cryptocurrency-linked stocks and the cryptocurrency market: Evidence from Australia. Glob Financ J. https://doi.org/10.1016/J.GFJ.2021.100642
15. Freitas FD, de Souza AF, de Almeida AR (2009) Prediction-based portfolio optimization model using neural networks. Neurocomputing 72(10–12):2155–2170. https://doi.org/10.1016/J.NEUCOM.2008.08.019
16. Gandal N, Halaburda H (2014) Competition in the Cryptocurrency Market. SSRN Electron J. https://doi.org/10.2139/SSRN.2506577
17. Hamayel M J, Owda AY (2021) A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. AI 2(4):477–496. https://doi.org/10.3390/ai2040030
18. Jalan A, Matkovskyy R, Yarovaya L (2021) "Shiny" crypto assets: A systemic look at gold-backed cryptocurrencies during the COVID-19 pandemic. Int Rev Financ Anal 78:101958. https://doi.org/10.1016/J.IRFA.2021.101958
19. Jiménez I, Mora-Valencia A, Perote J (2022) Semi-nonparametric risk assessment with cryptocurrencies. Res Int Bus Financ 59:101567. https://doi.org/10.1016/J.RIBAF.2021.101567
20. Ji Q, Bouri E, Lau CKM, Roubaud D (2019) Dynamic connectedness and integration in cryptocurrency markets. Int Rev Financ Anal 63:257–272. https://doi.org/10.1016/J.IRFA.2018.12.002
21. Ji S, Kim J, Im H (2019) A comparative study of bitcoin price prediction using deep learning. Mathematics 7(10):898. https://doi.org/10.3390/math7100898
22. Lauriola I, Gallicchio C, Aiolli F (2020) Enhancing deep neural networks via multiple kernel learning. Pattern Recogn 101:107194. https://doi.org/10.1016/J.PATCOG.2020.107194
23. Nakamoto, S. (n.d.). Bitcoin: A Peer-to-Peer Electronic Cash System. Retrieved November 22, 2021, from www.bitcoin.org

24. Nazir A, Mir RN, Qureshi S (2020) Exploring compression and parallelization techniques for distribution of deep neural networks over Edge-Fog continuum – a review. Inter J Intell Comput Cyber 13(3):331–364. https://doi.org/10.1108/IJICC-04-2020-0038/FULL/XML
25. Pahič R, Ridge B, Gams A, Morimoto J, Ude A (2020) Training of deep neural networks for the generation of dynamic movement primitives. Neural Netw 127:121–131. https://doi.org/10.1016/J.NEUNET.2020.04.010
26. Poongodi M, Sharma A, Vijayakumar V, Bhardwaj V, Sharma AP, Iqbal R, Kumar R (2020) Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system. Comput Electr Eng 81:106527. https://doi.org/10.1016/J.COMPELECENG.2019.106527
27. Rather AM (2021) LSTM-based Deep Learning Model for Stock Prediction and Predictive Optimization Model. EURO J Decision Process 9:100001. https://doi.org/10.1016/J.EJDP.2021.100001
28. Rather AM, Agarwal A, Sastry VN (2015) Recurrent neural network and a hybrid model for prediction of stock returns. Expert Syst Appl 42(6):3234–3241. https://doi.org/10.1016/J.ESWA.2014.12.003
29. Samarasinghe, Sandhya. (2007). Neural networks for applied sciences and engineering : from fundamentals to complex pattern recognition. Auerbach. https://www.routledge.com/Neural-Networks-for-Applied-Sciences-and-Engineering-From-Fundamentals/Samarasinghe/p/book/9780849333750
30. Shakri, IH (2021) Time series prediction using machine learning: a case of Bitcoin returns. Studies in Economics and Finance, ahead-of-print(ahead-of-print). https://doi.org/10.1108/SEF-06-2021-0217
31. Shin SG, Jin S (1998) Optical neural network using fractional Fourier transform, log-likelihood, and parallelism. Opt Commun 153(4–6):218–222. https://doi.org/10.1016/S0030-4018(98)00231-4
32. Simidjievski N, Todorovski L, Džeroski S (2015) Predicting long-term population dynamics with bagging and boosting of process-based models. Expert Syst Appl 42(22):8484–8496. https://doi.org/10.1016/J.ESWA.2015.07.004
33. Syam N, Kaul R (2021) Support Vector Machines in Marketing and Sales. Mach Learn Artif Intell Market Sales. https://doi.org/10.1108/978-1-80043-880-420211005
34. Tanwar S, Patel NP, Patel SN, Patel JR, Sharma G, Davidson IE (2021) Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. IEEE Access 9:138633–138646. https://doi.org/10.1109/ACCESS.2021.3117848
35. Tschorsch F, Scheuermann B (2016) Bitcoin and beyond: A technical survey on decentralized digital currencies. IEEE Commun Surv Tutorials 18(3):2084–2123. https://doi.org/10.1109/COMST.2016.2535718
36. Wang J, Liu L (2020) A multi-attention deep neural network model base on embedding and matrix factorization for recommendation. Inter J Cogn Comput Eng 1:70–77. https://doi.org/10.1016/J.IJCCE.2020.11.002
37. Wang S, Zhang Y, Zhang C, Yang M (2020) Improved artificial neural network method for predicting photovoltaic output performance. Global Energy Interconnect 3(6):553–561. https://doi.org/10.1016/J.GLOEI.2021.01.005
38. Wu J, Liu J, Zhao Y, Zheng Z (2021) Analysis of cryptocurrency transactions from a network perspective: An overview. J Netw Comput Appl 190:103139. https://doi.org/10.1016/J.JNCA.2021.103139
39. Yan T, Shen S-L, Zhou A, Chen X-S (2022) Prediction of geological characteristics from shield operational parameters using integrating grid search and K-fold cross validation into stacking classification algorithm. J Rock Mech Geotechn Eng. https://doi.org/10.1016/J.JRMGE.2022.03.002

Akhter Mohiuddin Rather Dr. Akhter Mohiuddin works as an Associate Professor, Business Analytics at Great Lakes Institute of Management, Gurgaon, India. He holds a Ph.D. degree in Computer Science with a specialization in Machine learning & Data Sciences under joint guidance from the Institute for Development & Research in Banking Technology, Hyderabad (*established by the Reserve Bank of India*), and the University of Hyderabad in 2016. His research interest includes Artificial neural networks and deep learning models in computer vision, object detection, and image classification. Dr. Akhter is also an Editor of GLIMS Journal of Management Review and Transformation. In 2019, Dr. Akhter Mohiuddin was nominated for the award of "*Best young scientist below 40 years*" by GISR India.