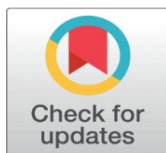# BITCOIN PRICE PREDICTION WITH COVID-19 SENTIMENT USING LSTM NEURAL NETWORK

Bhavsar Shachi [1]✉ , Ravi Gor [2]✉

[1] Research Scholar, Department of Mathematics, Gujarat University, Ahmedabad, India
[2] Department of Mathematics, Gujarat University, Ahmedabad, India

ABSTRACT

Cryptocurrencies are nowadays getting popular for investment due to its various benefits such as low transaction cost, blockchain secured platform, profit, etc. Bitcoin being top of the market capitalization currency, gained more popularity during covid-19 pandemic. This study focuses on bitcoin price prediction with covid-19 sentiment. Here Long Short-Term Memory Deep learning model based on machine learning is used for price prediction. At the end both results i.e., with covid-19 sentiment and without it are compared which shows model performs better by adding sentiments.

Keywords: Supervised Learning, Neural Networks, Back Propagation, Long Short-Term Memory, Gated Recurrent Unit

1. INTRODUCTION

Habitually described as cryptocurrency, Bitcoin is a digital money working without any central control, bank, or government administrations. It is based on peer-to-peer blockchain technology which makes it more secure and trustable by public. Bitcoin is used for investment, buying goods and services, etc. As bitcoin trading is allowed in many developed countries, it is also used for international transfers. As of now, the market cap of Bitcoin has a staggering of $1.1 TN, comprising half of the cryptocurrency market, which is over $2 TN according to the free press journal. Ethereum Races Clock to Collect Enough Coins for Big Upgrade (2020)

In recent years of pandemic, cryptocurrencies have gained more attention especially bitcoin. The covid-19 outbreak sentiments boosted the bitcoin prices sky-high. During pandemic bitcoin has gained attention from investors, researchers, government, media. Also, people without any experience started investing in bitcoin. Bitcoin is based on blockchain technology which makes it more efficient and make people feel safe to use. Many research and start-ups on block chain system makes it more competent. Due to pandemic when other firms were losing their worth, at the other side prices of cryptocurrencies such as bitcoin raised excessively. This instability in stock markets also made more investors move towards bitcoin investment. Government also promoted online money transfers to avoid the spread of virus. Due to such reasons bitcoin got more attention day by day relating to pandemic situations resulting in sudden hike in price. At the same time many cryptocurrency exchange firms were expanded in India too. Impact-of-covid-19-on-cryptocurrencies (2021)

Due to highly volatile nature, forecasting bitcoin price is a complex process. At the same time, it is a demanding topic. There are many regression techniques in Machine learning which are helpful in prediction systems. In which Deep learning (DL) helps when complexity and abstraction increase. Deep learning is basically made up of Artificial Neural Networks (ANN). Artificial Neural Networks imitates the working behaviour of human brain using Back Propagation. ANN is made up of three neurons layers i.e., Input, Hidden and Output layer. Neurons are weighted nodes from which data flows. They compute and decide whether output should be passed to next layer or not using an activation function. Recurrent Neural Network (RNN), Convolutional Neural Network, Deep Boltzmann Machine, etc. are models of Deep Learning. Long Short-Term Memory (LSTM) is subset of RNN. It has vast applications in areas such as image recognition, speech recognition, recommender system, fraud detection, etc. LSTM is also made up of three layers which are input layer, hidden layer, and output layer. DL is a part of supervised learning in machine learning, where inputs and outputs are fed to system to predict the output for unseen data. Here, the bitcoin price is predicted with covid impact using LSTM approach. Some factors behind bitcoin price fluctuations are neglected in this paper.

2. LITERATURE REVIEW

Pengfei and Yan (2019) combined PSR method with LSTM neural network to predict prices for six stock indices. Also, other prediction methods namely ARIMA, SVR, MLP, LSTM were used to compare the best fit model. The comparison shows that results for stock indices, combination of LSTM and PSR model gives better predictions as compared to others. Pengfei and Yan (2019)

Li et al. (2020) predicted the price fluctuations i.e., sudden rise or fall of bitcoin using Attentive LSTM with Embedding network (ALEN). Some of the features were selected manually such as change in price, change in volatility, etc. Also, some traditional technical indicators such as Simple moving average, Bollinger band, moving average convergence divergence were selected as features with Denoising Autoencoders features. They showed that the novel proposed model named ALEN can capture the representation of Bitcoin. Yang et al. (2020)

Livieris et al. (2021) proposed advanced CNN and LSTM models based on multi-input architecture for price fluctuations and predictions of cryptocurrencies: Bitcoin, Ethereum and Ripple due to its highest market capitalization. The experimental analysis showed that the Multiple-Input Cryptocurrency Deep Learning Model (MICDL) was efficient to exploit mixed cryptocurrency data. Ioannis et al. (2021)

Shen (2021) predicted trend of stock market prices using comprehensive Deep Learning. Hybrid feature algorithm is used with feature extraction, recursive feature elimination and randomized Principal Component Analysis. LSTM model with 29 features was reduced to 5 principal features using hybrid feature extraction algorithm. The compared result with extracted features shows that reduced principal features gives better results. Shen (2021)

Ping et al. (2021) examined the psychological state analysis relationship between Bitcoin prices and covid-19. Long term positive significant state of public and investors psychology was observed due to various reasons such as cashless transactions, secured, unbanked, easy payments, decentralised, etc. Also, unidirectional relation between bitcoin prices and cumulative covid deaths is observed. Ping et al. (2021)

Ghosh and Gor (2022) used K-means clustering and Random Forest Regression algorithms for sales prediction. They used clustering methods for ad campaigning analysis. First, ad groups are created using the K-Means clustering algorithm then Random Forest Regressor algorithm is used to optimize sales conversion and predict future sales. Impressions, clicks, and spent are used as independent variables to predict total number of people that asked about the product after viewing the ad on Facebook. They also calculated Mean Absolute Error and Root Mean Square Error. The integration of two algorithms K-means clustering and Random Forest regressor gives permissive result with 75% accuracy. Ghosh and Gor (2022) "Ad-Campaign Analysis and Sales prediction using K-means Clustering and Random Forest Regressor", 2022) Bhavsar and Gor (2022)

Ghosh and Gor (2022) used the Nave Bayes and Decision Tree algorithms to detect whether a Short Message Service (SMS) is spam or not. Count Vectorizer tool was used to transform a given text into a vector on the basis of the frequency of each word that occurs in the entire text. Each message and each word were labelled by number of occurrences in a text. 80% of data was used for training and 20% of data was used for testing purpose. Results of both algorithms were compared, where Naïve Bayes gives the better result with accuracy 98.4%. Also, SMS classifier application was created by using Flask and Naïve Bayes Algorithm. Ghosh and Gor (2022)

Ghosh and Gor (2022) used Block chain Technology and different supervised learning algorithm such as Support Vector Regression, Lasso Regression, Ridge Regression, Multiple Linear Regression and Random Forest Regression algorithms to predict insurance premium. The data was stored using a user interface made by blockchain technology. In this study age, sex, BMI, and children taken as an independent variable to predict charges of health insurance. Results of algorithms were compared with each other and conclude that Random Forest Regression gives the better result. Ghosh and Gor (2022)

2.1. DEEP LEARNING RECURRENT NEURAL NETWORK

Recurrent Neural Network is a Deep Learning algorithm. It is a type of Artificial Neural Network in which time series or sequential data is fed. It is made up of three layers input layer, hidden layer, and output layer. It takes inputs from previous time steps. RNNs can be both unidirectional and bidirectional. There are many variations of RNN such as vanilla RNN, LSTM, Bi-directional LSTM, GRU, etc in which according to the output/application one to one, one to many, many to many, etc models are used. RNN can take multiple time steps to predict the output. Following Figure 1 is the diagram of folded and unfolded RNN with multiple n time steps. Here x

represents input state, h represents hidden state and y represents output state. Initially parameters stored in input state was computed by hidden layer and result is obtained in output state. Then information collected from initial time step was passed to the next time step and results for the same were calculated using previous information. In unfolded RNN x_t represents t th time step of input state, h_t represents t th time step of hidden state which carries information from h_{t-1} and y_t represents t th time step of output state which is passed to h_{t+1} time step. This process continues till n time steps and then again, the model is back propagated to update the weights as shown in Figure 1 RNN and its other variants such as LSTM, BiLSTM, GRU, etc. uses backpropagation through time to update the weights of neurons and repeats back propagating until it best fits the model. Bhavsar and Gor (2022)
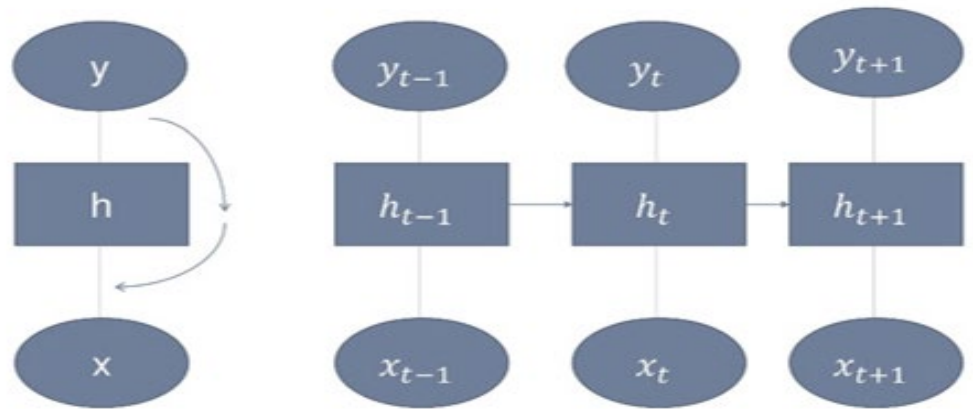
Figure 1

Figure 1 Recurrent Neural Network with t+1-time steps Bhavsar and Gor (2022)

2.2. BACK PROPAGATION

Backpropagation Through Time (BPTT) also involves a repeated application of chain rule same as a standard back propagation. Time step in BPTT includes input time step, copy of network and output. It works unfolding each input time step. Then from obtained results errors are considered and gathered for each time step. The network is folded again, and weights of each neuron are updated. The process of back propagating is repeated to make the model learn.

Learning through backpropagation, RNN faces vanishing gradient problem due to long data sequences. While calculating loss or updating weights suppose if each state has 0.01 gradient value with 100 states, then 0.01/ (100) ~0 value would change in next updated weight. Which shows that after computing each and every input time step the change in updated weights would be almost negligible. That will increase computing time and increase number of epochs, making it unaffordable to work. To avoid such vanishing gradient problem LSTM neural network was made for long term dependencies. Rajpurohit et al. (2021) Bhavsar and Gor (2022)

2.3. LONG SHORT-TERM MEMORY NEURAL NETWORK [LSTM NN]

LSTM NN stores information in memory cell over time using feedback loop in its recurrent layer maintaining sequence information. Information from previous output is carried to next input. RNN has ability to learn short term dependencies,

but LSTM has ability to learn short as well as long term dependencies. In LSTM memory cell has ability to forget or keep the relevant information. This cell is made up of input, forget and output gates. Forget gate decides whether the information from previous output is relevant or not and accordingly it discards or keeps the information by setting threshold. Input gate computes the new data to enter in cell. The information from previous to next time step was conveyed by Cell state often called as Memory cell. Output gate regulates the output of that particular cell. The gates are connected with activation functions such as sigmoid or tanh according to the need of outputs of any gate. In this model forget and input gates are attached with sigmoid function and input node and output gates are attached with hyperbolic tangent function. Hochreiter and Schmidhuber (1997)
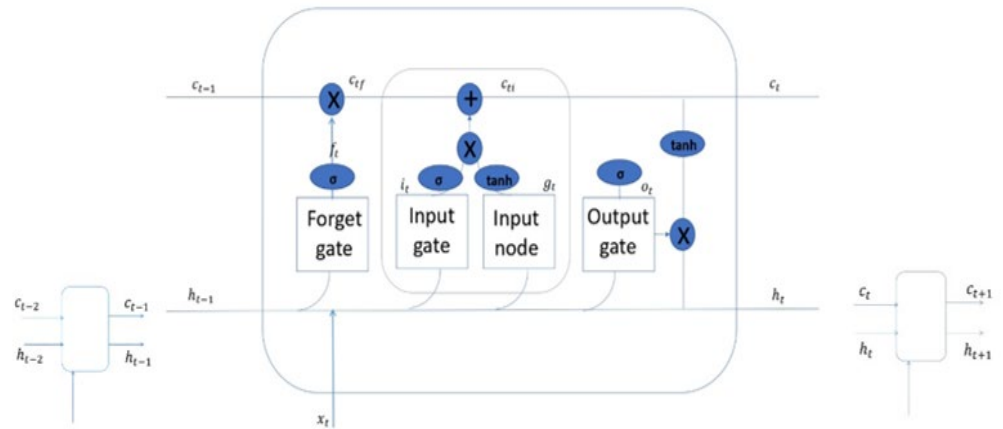
Figure 2

Figure 2 Working of Long Short-Term Memory

Figure 2 shows the cell structure of LSTM neural network. Where c_{t-1} shows the information passing from previous cell to next cell. c_1 collects computed information from forget gate and input gate and decides whether to pass information or not. Similarly, h_t collects information from output gate. All gates are connected with activation function. Also, brief view of previous and next cell states is shown in Figure 2 in left and right of LSTM cell. Here, the information passed by previous state i.e., h_{t-1} is passed to h_t cell along with x_t to compute further. Hochreiter and Schmidhuber (1997)

- Output of forget gate: $f_t = \sigma[w_{fh}h_{t-1} + w_{fx}x_t + b_f]$ is passed to cell $c_{tf} = c_{t-1}f_t$; where b_f is bias of forget gate, w_{fh} and w_{fx} are weights for forget gate with respect to hidden state and input state. Here forget gate decides which information from previous state should be continued and which to be dropped by threshold value.

- Input gate is computed and output of input gate is $i_t = \sigma[w_{ih}h_{t-1} + w_{ix}x_t + b_i]$; where b_i is bias of forget gate, w_{ih} and w_{ix} are weights for output gate with respect to hidden state and input state.

- Output of output gate is $o_t = \sigma[w_{oh}h_{t-1} + w_{ox}x_t + b_o]$; where b_o is bias of forget gate, w_{oh} and w_{ox} are weights for output gate with respect to hidden state and input state.

- Output of input node is $g_t = tanh[w_{gh}h_{t-1} + w_{gx}x_t + b_g]$; where b_g is bias of forget gate, w_{gh} and w_{gx} are weights for output gate with respect to hidden state and input state.

- Output of input gate and input node makes $c_{ti} = i_t g_t$. So, $c_t = c_{ti} + c_{tf}$ and $h_t = tanh(c_t) + o_t$ is passed to the next cell state of time step $t + 1$. Hochreiter and Schmidhuber (1997)

LSTM Neural Network model learns from the different features fed. Using the data with high dimension or more features can be helpful as well as problematic too. It might affect increase computing time, increase storage space, difficult to visualize data in higher dimensions, etc. so most of the correlated features are redundant. Here feature selection technique is used for covid-19 dataset variables.

2.4. DIMENSION REDUCTION OR FEATURE SELECTION

For features selection simple baseline approach is used known as Backward feature elimination. This feature selection technique in machine learning is opposite to Forward feature selection where features get selected turn by turn. Here features get eliminated. Initially, dataset with all expected variables is trained. Then by dropping each variable, model is trained. The variable with low or almost no change in performance can be eliminated. The process is repeated till no variables are left. The parameters with minor difference will not make many changes in results of given model. It instead increases the computing time giving similar result. So, such parameters are removed from the model to improve the interpretation of the parameters and to avoid over fitting problems.

3. DATA COLLECTION AND METHODOLOGY

First data includes open, high, low prices (in USD) of bitcoin between 22-01-2020 to 21-09-2021 total 609 days. [Data was collected from investing.com] (inv). The consideration of bitcoin over other cryptocurrencies was that it has highest market capitalization and volatile behaviour.

Second data includes covid-19 number of positive cases, recovered cases and deaths per day between 22-01-2020 to 21-09-2021. [Data was collected from data.humdata.org] (dat). The data includes covid-19 information from different countries over the globe. So, the final merged data contains seven attributes. The model is built to predict the price of bitcoin.

3.1. DATA PRE-PROCESSING

Data cleaning: was done to fill some missing values found in variable of recovered cases by mean. Noisy data: Here noisy data was not removed as the data was volatile in nature and removing it is not reasonable. Data normalization: was used to scale the data values using minmax scaler.

For training and testing dataset 80:20 ratio was taken. Training dataset was from 22-01-2020 to 22-05-2021 and testing dataset was from 23-05-2021 to 21-09-2021which is considerable amount of unseen dataset. So, from 609 days 487 were taken as train set and 122 as test set. Reason of such time period taken was that it considers covid-19 crisis, volatile and deviated behaviour with structural breaks.

Initially LSTM neural network was fed with all features. In LSTM neural network sigmoid was taken for forget gate and tanh for input node and output gate as an activation function, Adam optimizer was used with 0.2 dropout rate.

To check relation between all attributes, feature reduction technique was applied. Here, backward elimination method was used for variables selection. Here LSTM model was fed with different combinations of variables to check accuracy of

proposed model using backward elimination method by comparing results. After applying backward elimination method, variable of recovered cases showed low performance as compared to other two variables of cases and deaths. So, the variable of recovered cases was dropped, and ultimate dataset is shown in Table 1

Table 1

	Table 1 Dataset with reduced parameters					
	Date	**Price**	**Open**	**High**	**Low**	**Positive Cases Per Day**	**Deaths Per Day**
0	2020-01-22	8678.5	8733.0	8805.4	8610.8	557	17
1	2020-01-23	8405.1	8678.5	8687.3	8309.6	655	18
2	2020-01-24	8439.9	8404.9	8522	8242.6	941	26
3	2020-01-25	8341.6	8439.9	8447.6	8277.2	1433	42
4	2020-01-26	8607.8	8341.6	8607.8	8304.9	2118	56

The loss function for train set and test set for predicted price were calculated by Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). r2 score was used to check accuracy of model, higher the score better the accuracy. Python programming is used for computing LSTM neural network.

4. RESULTS AND DISCUSSION

For the experimental results of the model, different datasets were considered,

Dataset-a contains reduced dataset by feature selection technique i.e., open, high, low prices with number of covid-19 positive cases and deaths.

Dataset-b contains bitcoin dataset with open, high, low prices.

Dataset-c contains bitcoin and covid-19 merged dataset i.e., open, high, low prices with number of covid-19 positive cases, recovered cases and deaths.

The following is the table of loss function for train set and error for test set of different datasets.

$$\text{MSE} = \frac{\sum_{t=1}^{n}|y_t - \bar{y}_t|}{n}, \text{ RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \text{ and } r^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y}_t)^2} \text{ where } \bar{y}_t = \frac{\sum_{t=1}^{n}y_t}{n}.$$ Bhavsar and Gor (2022)

As observed in Table 2 dataset-c shows the highest error as compared to other datasets. (Hence, recovered cases are removed by backward elimination method.) Now dataset-a have less error as compared to dataset-b. Note that, here dataset-a includes covid-19 dataset and dataset-b does not contain. Comparative study shows that, the prediction of dataset-a is better than dataset-b, which proves that addition of covid-19 sentiments provides better results in price prediction.

Also, r^2 score was calculated in Table 3 to check accuracy of model for different dataset. In r2 score higher the values better the result. The score is observed and compared with all datasets. Here it shows that dataset-a shows better score as compared to dataset-b and dataset-c.

To check accuracy, the purposed LSTM model was compared with Gated Recurrent Unit model for dataset-a. The compared results showed that MAE for train set is 0.0292 and test set is 0.0323. Also, RMSE for train set is 0.044 and 0.041 for test set. Which shows that purposed LSTM model performed better.

Table 2

Table 2 MAE and RMSE comparison for dataset-a, dataset-b, and dataset-c				
	MAE		RMSE	
	Train set	Test set	Train set	Test set
Dataset -a	0.0265401	0.0322269	0.04175	0.041121
Dataset -b	0.0280616	0.0323449	0.042	0.0415218
Dataset -c	0.0301871	0.0346946	0.043109	0.0444608

Table 3

Table 3 r² score comparison for dataset-a, dataset-b and dataset-c		
r^2 score		
	Train set	Test set
Dataset -a	0.8370507	0.79773
Dataset -b	0.8410582	0.791684
Dataset -c	0.8377789	0.777424

This shows that LSTM with reduced feature in dataset shows higher accuracy in test set. Hence, keeping both errors and accuracy score in mind, we consider features of dataset-a as best fit features. In the Figure 3 and Figure 4 predicted price of training and test set are shown respectively.

Figure 3

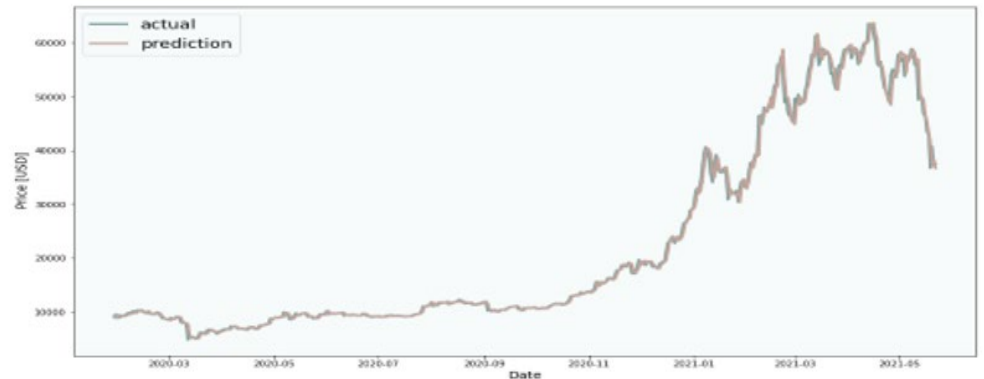**Figure 3** Dataset-a price prediction of train set
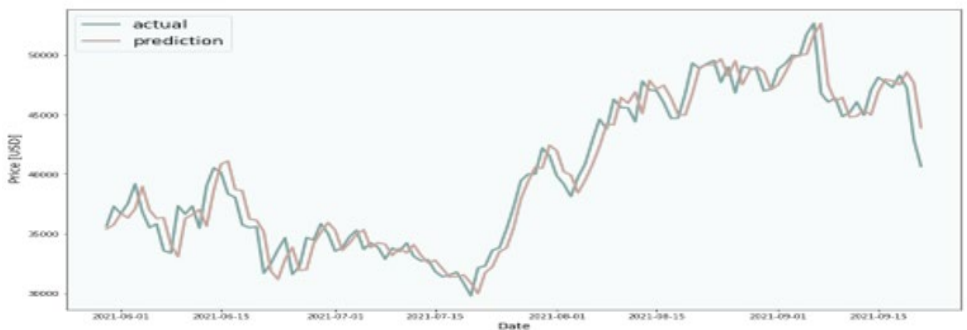
Figure 4

Figure 4 Dataset-a price prediction of test set

5. CONCLUSION

Deep learning-based LSTM model was designed and optimized by suitable activation functions and backward feature elimination. Impact of covid sentiments was calculated using various parameters which shows positive results compared to model without covid sentiments. Moreover, final results shows that model achieves good accuracy of 0.797 for test set and 0.837 for train set. This study can be helpful in understanding relation between bitcoin price movements with covid sentiments. Deep learning models works better with more data but here as covid cases are considered we are limited with dataset that is one drawback of the model. Future work can be done by including other sentiments which affects bitcoin prices movements.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Achyut, G. Soumik, B. Giridhar, M. Narayan, C. & Somya, S. (2019). Stock Price Prediction Using LSTM on Indian Share Market. 32nd International Conference on Computer Applications in Industry and Engineering, 63, 101-110.

Alvin, H. Ramesh, V. & Kumar, R. S. (2021). Bitcoin Price Prediction Using Machine Learning and Artificial Neural Network Model. Indian Journal Of Science And Technology, 2300-2308. https://doi.org/10.17485/IJST/v14i27.878

Aniruddha, D. Kumar, S. & Meheli, B. (2020). A gated recurrent unit approach to bitcoin price prediction. Journal of Risk and Financial Management, MDPI.

Asian, C. W. (n.d.). shares end quarter in sombre mood, dollar on high Asian shares end quarter in sombre mood, dollar on high.

Bhavsar, S. & Gor, R. (2022). Comparison of Back propagation algorithms: Bidirectional GRU and Genetic Deep Neural Network for Churn Customer. International Organization of Scientific Research Journal of Computer Engineering (IOSR-JCE).

Bhavsar, S. & Gor, R. (2022). Predicting Restaurant Ratings using Back Propagation Algorithm. International Organization of Scientific Research journal of Applied Mathematics (IOSR-JM), 18(2), 5-9.

Ethereum Races Clock to Collect Enough Coins for Big Upgrade. (2020).

Ghosh, M. & Gor, R. (2022). Ad-Campaign Analysis and Sales prediction using K-means Clustering and Random Forest Regressor. International Organization of Scientific Research Journal of Applied Mathematics, 18(2), 10-14.

Ghosh, M. & Gor, R. (2022). Health Insurance Premium Prediction Using BlockChain Technology and Random Forest Resression Algorithm (In press). International Journal of Engineering Science Technologies (IJOEST). https://doi.org/10.29121/ijoest.v6.i3.2022.346

Ghosh, M. & Gor, R. (2022). Short Message service Classifier Application using Naive Bayes algorithm. IOSR Journal of Computer Engineering (IOSR-JCE), 24(3), 1-6.

Hochreiter, S. & Schmidhuber, J. (1997). Long Short Term Memory. Neural Computation, 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Ioannis, L. Kiriakidou, N. Stavroyiais, S. & Pitelas, P. (2021). An advanced CNN LSTM model for cryptocurrency forecasting. Electronics MDPI, 287 (1-16).

Impact-of-covid-19-on-cryptocurrencies. (2021).

Pengfei, Y. & Yan, X. (2019). Stock price predictions based on deep neural networks. Neural computing and applications, 1609-1628. https://doi.org/10.1007/s00521-019-04212-x

Ping, H. J. Liu, J. & Jie, Y. (2021). Examaning the psychological state analysis relationship between bitcoin prices and COVID-19. Frontier in Psychology, 1-7.

Rajpurohit, V. Bhavsar, S. & Gor, R. (2021). A comparision of GRU-based ETH price prediction. Proceeding of International Conference on Mathemaitcal Modelling and Simulation in Physical Sciences (MMSPS-2021), 424-431.

Salman, K. A. & Peter, A. (2019). Predictive Analytics in Cryptocurrency Using Neural networks : A Comparative Study. IJRTE, 7(6).

Shen, J. (2021). Short term stock market price trend prediction using comprehensive deep learning system. Journal of big data, 1-33. https://doi.org/10.1186/s40537-020-00333-6

Sheng, C. & Hongxiang, H. (2018). Stock Prediction Using Convolutional Neural Network. AIAAT Materials Science and Engineering, 435. https://doi.org/10.1088/1757-899X/435/1/012026

The Humanitarian Data Exchange (n.d.). data.humdata.org

Yang, L. Zibin, Z. & Dai, H. N. (2020). Enhancing bitcoin price fluctuation predictionn using attentive LSTM and Embedding network. Applied sciences MDPI, 4872. https://doi.org/10.3390/app10144872