Predicting the Trends of Price for Ethereum Using Deep Learning Techniques

Deepak Kumar
Department of Computer Science and Engineering
National Institute of Technology
Rourkela, India
deepak.acumen@gmail.com

Dr S. K. Rath
Department of Computer Science and Engineering
National Institute of Technology
Rourkela, India
skrath@nitrkl.ac.in

ABSTRACT

This study intends to predict the trends of price for a cryptocurrency, i.e. Ethereum based on Deep Learning techniques considering its trends on time series particularly. This study analyses as how Deep Learning techniques such as Multi-layer perceptron (MLP) and Long Short-Term Memory (LSTM) help in predicting the price trends of Ethereum. These techniques have been applied based on historical data that were computed per day, hour, and minute wise. The data set is sourced from the CoinDesk repository. The performance of the obtained models are critically assessed using statistical indicators like Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).

Keywords—Deep Learning; Ethereum; MLP; LSTM; Cryptocurrency.

I. INTRODUCTION

Cryptocurrency is new to the financial market but considered as an alternative to existing paper currency. Cryptocurrencies like Bitcoin, Litecoin, and Ethereum etc. are attracting significant investment at present. Ethereum is a cryptocurrency that is gaining prominence after Bitcoin. Unlike Bitcoin, Ethereum offers unique features like Smart Contract and decentralized application platform to its users [1]. The prediction of the cryptocurrency is a topic of interest, in particular for those who intend to invest in it. Prediction of price for cryptocurrency is a time series prediction problem which uses historical data to predict on time series scale [2].

In time series analysis, the purpose is to estimate the potential future value with the help of the past data. However, Machine learning techniques are also considered as a better alternative applied to predict and classify on the basis of the accuracy of the time series problem. The available methodologies for time series forecasting are Moving Average, Auto Regression, Autoregressive Moving Average, Autoregressive Integrated Moving Average and Artificial Neural Network (ANN) etc.

In this paper, two different Deep Learning models: Multi-Layer Perceptron(MLP) and Long Short-Term Memory (LSTM) have been implemented and their performances are critically assessed for the Ethereum price prediction.

II. LITERATURE SURVEY

Cryptocurrency like Ethereum is very new to the financial world with a highly volatile price. Analysis of price prediction of cryptocurrency using deep learning techniques especially for Ethereum can be considered as a potential one, because of its good number of advantages. Shah, D and Zhang apply Bayesian regression to Bitcoin price prediction [3].

M Matta et al. have made a study on the relationship between Bitcoin cost, the volume of tweets and views for Bitcoin. They have compared Bitcoin price trends with Google Trends and concluded that positive tweets may predict the price trends [4].

Greaves, Alex, and Benjamin Au. have computed that the techniques like support vector machines (SVM) and ANN are successfully predicting the price of Bitcoins using Bitcoin Blockchain data and achieved an accuracy of 55% for price trends. The authors have investigated that using only Blockchain data is insufficient in price prediction [5].

There has been work on price prediction by sentiment analysis using Support Vector Machines(SVM). I Madan, S Saluja and Aojia Zhao have implemented SVM, Generalised Linear Model(GLM) and Random Forrest on Bitcoin Blockchain data and achieved the accuracy of 99% but without any validation of their models [6].

McNally Sean, Jason Roche, and Simon Caton have compared Recurrent Neural Network (RNN) and LSTM for predicting the price trends of Bitcoin. They have achieved the accuracy of 52% for the LSTM model and 50% for the RNN model. They have applied ARIMA with an accuracy of 50% and compared with RNN and LSTM models [7].

Earlier ANN has been used for stock market prediction (which is also a time prediction problem). A comparative study of different deep learning models of stock market prediction is presented in Reference [8] They have applied RNN, LSTM on stock price data of different companies and compared all three techniques.

Hiransha M., E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman have used MLP, RNN, LSTM model to predict price trends of the stock price of different companies stock price. They compared the performance of the models on the basis of the Mean Absolute Percentage Error (MAPE) [9].

III. METHODOLOGY ADOPTED

A. **Dataset Considered**

Data was collected from the coindesk and coinmarket repositories. Data set collected in a Comma-separated values(CSV) file. Models were compared on daily, hourly and minute basis data. The data on a daily basis from August 2015 to August 2018 which contain open price, closed price, high and low price and volume per day and hourly data from July 2017 to August 2018 have been considered. Daily data consist of 1000 data samples and hourly data consist of 1500 data samples and minute data consist of 400000 rows of data.

B. **Deep Learnings Models**

Neural networks can be implemented for price prediction problem due to a few distinctive characteristics. To begin with, Neural Networks have been a self-adjusting model based on training data, and they have the capability to look after the issue with only a little knowledge about its design and without any compulsion to the prediction model with the addition of any extra assumption. Secondly, Neural Networks have generalization capacity which implies that after training they can identify the new patterns even if they haven't any training data set. Since in most of the pattern recognition issues, predicting future events is based on past data. Due to the high volatility in cryptocurrency ANN is considered as a useful technique for this type of problem.

B.1 **Multi-Layer-Perceptron**

Multi-layer perceptron (MLP) is a feedforward Artificial Neural Network (ANN) with more than one hidden layer. In feedforward network information flows from the input layer to the output layer and model is trained through the Backpropagation technique. The Backpropagation technique is used to solve the non-linear separable problem. MLP can be used for pattern classification, prediction and approximation.

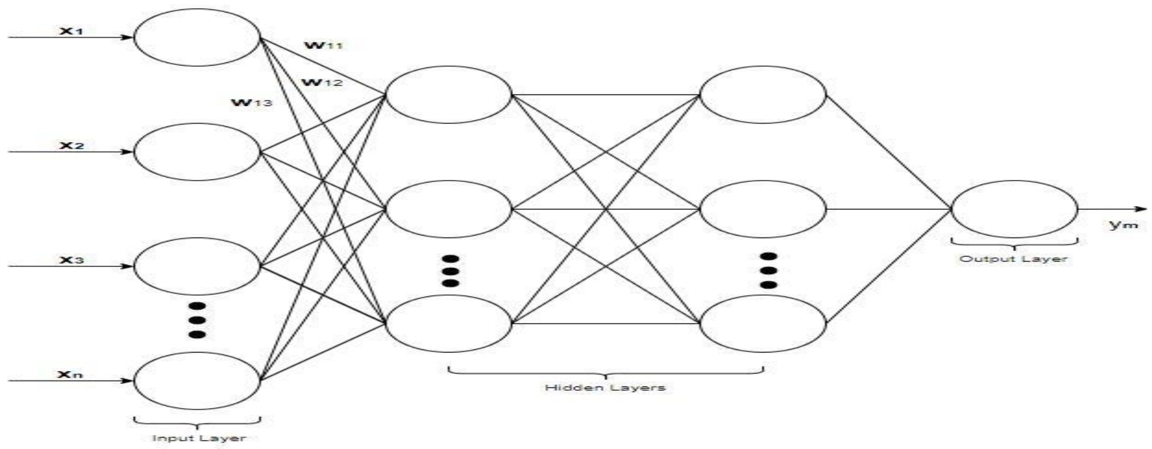

Figure 1: Multi-Layer Perceptron Model

The above figure shows an MLP network inspired by biological neuron [10] [11]. The neuron has n inputs and every neuron is connected through weight links. MLP has three sections of layers, the first input layer second hidden layer and the last one is the output layer [10]. Neurons in the networks are connected through neurons of the next subsequent layers.

B.2 **Long Short-Term Memory**

Long short-term memory [12] is a special kind of recurrent neural network [13]. RNN is a powerful and robust neural network as they have internal memory. Since they have internal memory, RNN can recall important concerning the input signal. The preciseness in prediction is the motivation behind why they are the favoured techniques for consecutive information like time series, speech, sound, video, and climate. RNN works fine for short-term dependencies, but also suffers from Exploding Gradients [14] and Vanishing Gradients [15].

LSTM is specifically designed to prevent the long-term dependency issue. It has the ability to remember the information for longer time period. The structure of LSTM is different to RNN as in RNN there is a basic structure of neural network with the feedback loop between neurons of each hidden layers, whereas in LSTM it has memory cell that propagates the information through network and control as well. The key to LSTM structure is that it uses four Gates to process the input information in the networks.

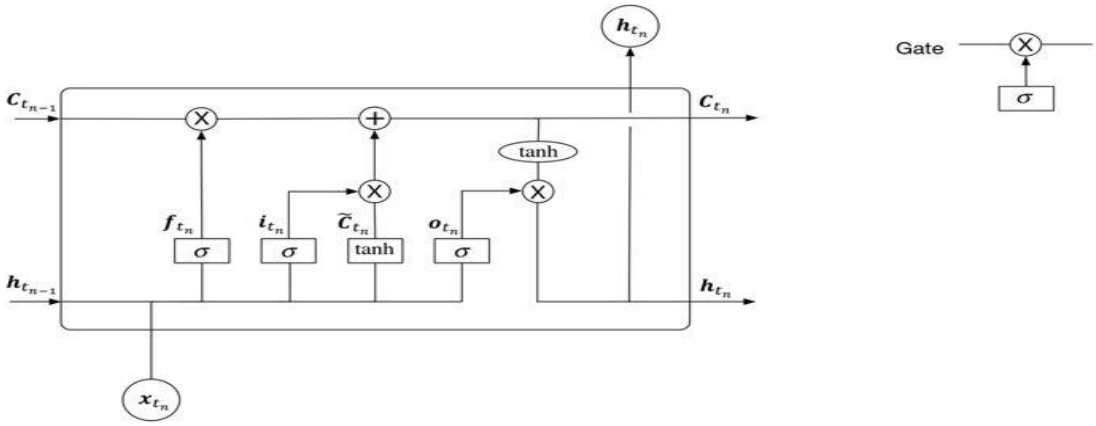

Figure 2: LSTM Architecture

Here,

- C_{tn-1} represents the old memory cell state.
- $h_{tn\text{-}1}$ for the output of the previous cell.
- C_{tn} for present cell state which runs through the entire network and has the capacity to remove or add information with the assistance of gates.
- h_{tn} output of present cell.
- f_{tn} forget gate layer that decides portion of information to be permitted.
- i_{tn} is the input gate layer.
- $\tilde{C}_{tn}$ is the new values created by a tanh layer that generates a new vector, which will be added to the state.
- o_{tn} is an output of sigmoid gate layer this layer creates numbers between zero and one, describing just how much of each element should be permitted through.

The cell state is updated based on the outputs from the gates that can be represented mathematically as follows.

$$ft_n = \sigma(W_f.[h_{tn-1},x_t] + b_f) \quad (1)$$

$$i_{tn} = \sigma(W_i.[h_{tn-1},x_t] + b_i) \quad (2)$$

$$C_{tn} = tanh(Wc.[h_{tn-1},x_t] + b_c) \quad (3)$$

$$o_{tn} = \sigma(Wo[h_{tn-1},xt] + bo) \quad (4)$$

$$h_{tn} = o_{tn} * tanh(C_{tn}) \quad (5)$$

Where:

- x_t: input vector
- h_{tn-1}: output vector
- C_{tn}: cell state vector
- ft_n: forget gate vector
- i_{tn}: input gate vector
- o_{tn}: output gate vector
- and W,b are the parameter matrix and vector.

IV. IMPLEMENTATION

A. Data Preparation

In this paper, three data sets are used for model evaluation, first daily ethereum price, secondly hourly price data and the third one is minute wise price data. From this, the data has been extracted only day-wise, open price, hourly-wise open price and minute wise open price, because the open price is preferred by the investor to decide whether to buy particular cryptocurrencies or not. The input data is usually normalized in neural networks applications in the range of (0, 1) and it depends on the activation function. So in this research, values of the open price have been normalized in the range of (0, 1) using the equation (6) and then models were used for evaluation using MLP and LSTM algorithms.

$$x(norm) = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{6}$$

Where x (norm) is the normalized value, x (min) is minimum and x (max) is the maximum value in the training dataset. This normalized data was given as the input to the network in a window size of 60 to predict days, hours and minutes based trend for prediction.

B. Evaluation criteria

The price prediction problems are judged upon the values of performance measure such as Mean Absolute Percentage Error (MAPE), since it shows the error in the predicted model. In addition to the MAPE, three other measures are used to compare prediction methods which are Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The equation for calculating these measures are given below:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{(Actual_i - Predicted_i)}{Actual_i} \right| \tag{7}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Actual_i - Predicted_i)^2 \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Actual_i - Predicted_i)^2} \tag{9}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Actual_i - Predicted_i| \tag{10}$$

V. SIMULATION RESULTS

The result of MLP and LSTM models are compared over Daily, Hourly, and Minute wise data.

A. Dataset: 1 (Per Day)

Dataset 1 consists of ethereum price on the basis of daily wise. Approximately 1200 days of price data has been considered for prediction of ethereum price by using of MLP and LSTM. The performance measurements are compared through various error function. The results quantified are shown in table 1.

Table 1: Error Function Incurred During The Prediction

Parameters	MLP	LSTM
MSE	0.021	0.018
MAE	0.114	0.013
MAPE	32.29	3.67
RMSE	21.3	20.53

Table 1 shows the various loss functions value for predicting the open price for 1000 days by using MLP and LSTM.

B. **Dataset: 2 (Per Hour)**

This dataset consists of ethereum price on an hourly basis (Price per hour). It contains data from 01-07-2017 to 02-8-2018 that comprise of 1500 data sample. The results quantified are shown in table 2.

Table 2: Error Function Incurred During The Prediction(Per Hour)

Parameters	MLP	LSTM
MSE	0.0120	0.0130
MAE	0.0048	0.0072
MAPE	8.856	1.38
RMSE	17.29	7.12

Table 2 shows the result obtained from over 1500 hours of data by using MLP and LSTM. Results show that both MLP and LSTM produce nearly the same result, but LSTM slightly performs better, since it has a better mechanism for long term dependency.

C. **Dataset: 3 (Per Minute)**

The minute wise dataset contains 400000 data sample with per minute price from 01-01-2018 to 30-10-2018. The results quantified are shown in table 3.

Table 3: Error Function Incurred During the Prediction (Per Minute)

Parameters	MLP	LSTM
MSE	0.011	0.0024
MAE	0.0028	0.0014
MAPE	3.25	2.21
RMSE	90.06	18.16

Table 3 shows the result over one-minute data. Comparative analysis of the results that are shown in table 1, table 2, and table 3. It is observed that the LSTM model shows better prediction over the MLP. Figure 3 and 4 show the trend analysis on daily price.

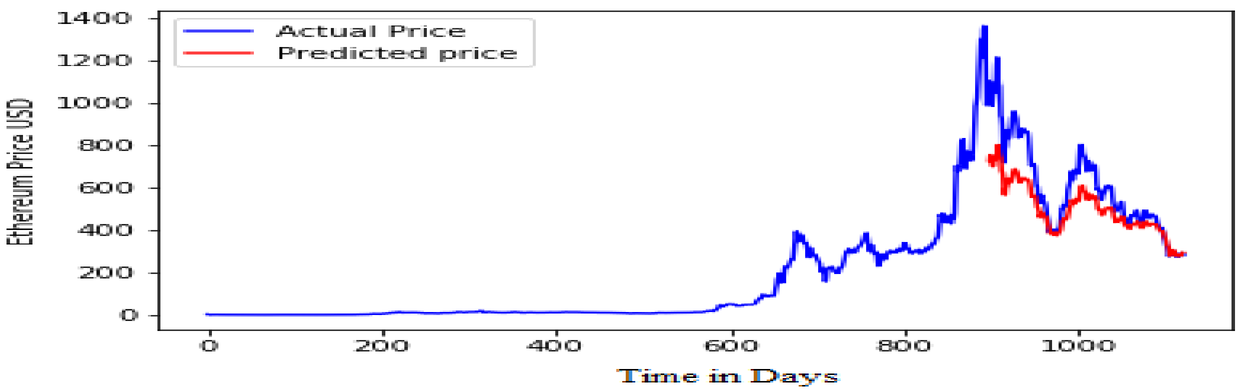

Figure 3: Daily wise Actual Vs Predicted Price by using MLP

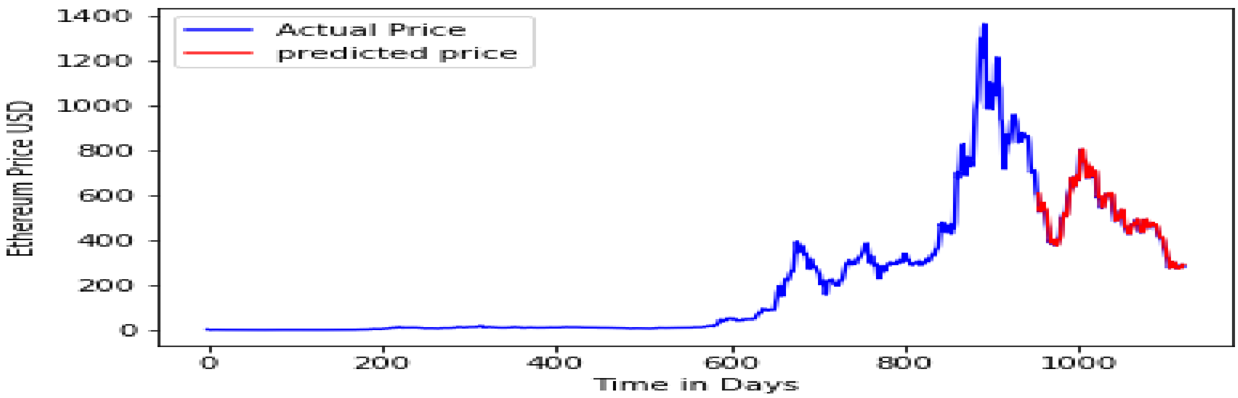

Figure 4: Daily wise Actual Vs Predicted Price by using LSTM

Figure 3 and 4 show the trend analysis of daily price by using the MLP and LSTM. As figures show, MLP is not that sufficient to capture the sudden and non-linearity in the price between 800 and 1000 days where LSTM has proven slightly better than MLP, but not completely satisfactory. Trends analysis obtained for hourly data are shown in figure 5 and 6 as mentioned below.

Figure 5: Hourly wise Actual Vs Predicted Price by using MLP

Figure 6: Hourly wise Actual Vs Predicted Price by using LSTM

Figure 5 and 6 shows hourly trends and it is shown that LSTM works better than MLP. From table 2 LSTM has better mse than MLP. Trends analysis obtained for minute wise data are shown in figure 7 and 8 as mentioned below.

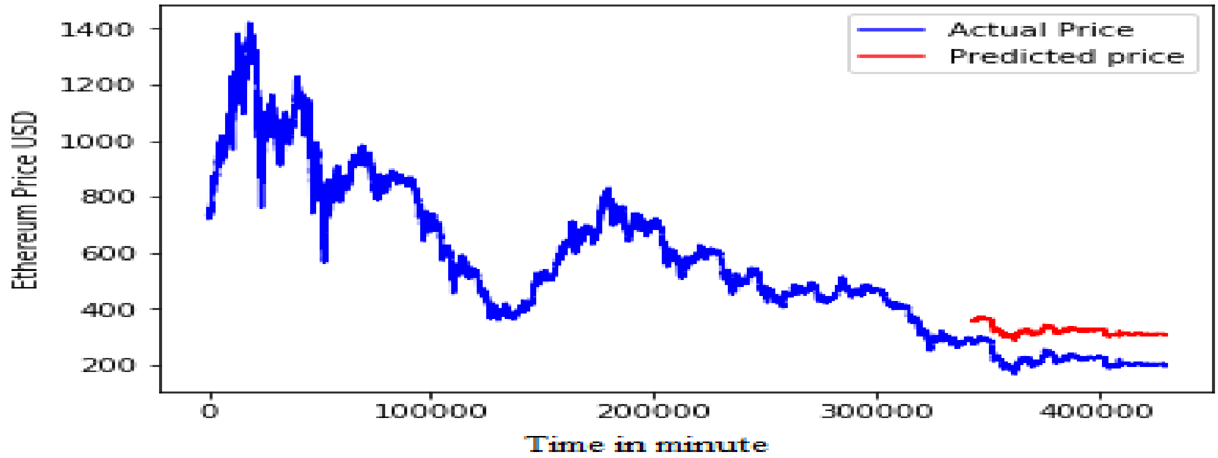

Figure 7: Minute wise Actual Vs Predicted Price by using MLP

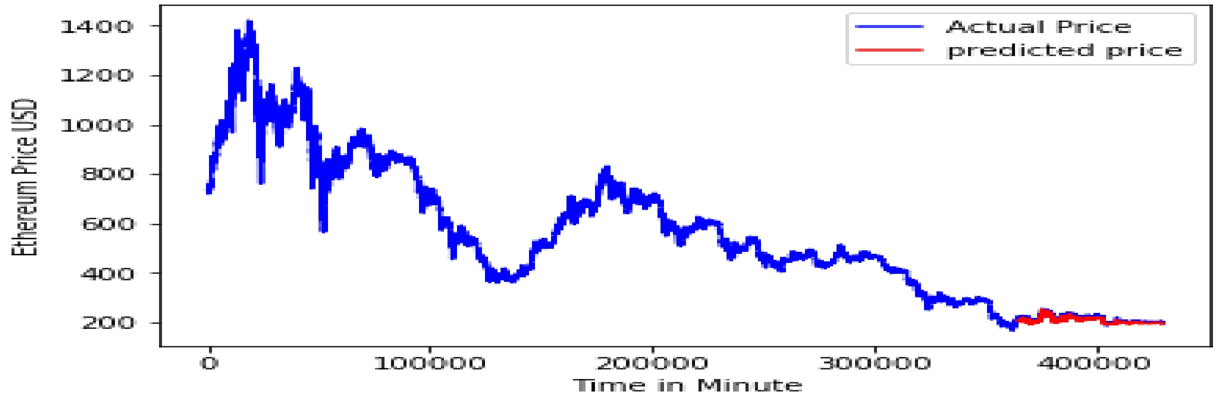

Figure 8: Minute wise Actual Vs Predicted Price by using LSTM

Figure 7 and 8 shows the price trends by the minute. Here MLP and LSTM models perform better than previous data because the deep learning network performs better with a large amount of data. By evaluating both figure 7 and 8 and table 3, it is observed that LSTM works far better than MLP with minimum losses.

VI. CONCLUSION AND FUTURE WORK

In this paper, it may be concluded that Deep Learning Models are suitable techniques for capturing the price trends of cryptocurrencies for larger datasets. MLP and LSTM both can predict the price trends, but results show that the LSTM model is more robust and precise for long term dependency as compared to MLP. Here both MLP and LSTM models are applied on Daily, Hourly and Minute wise data, and from the result, it is observed that neural network is able to predict more precisely if there is less deviation between two subsequent values in prices. From the results, it is clear that LSTM outperformed the MLP marginally but not very significantly. In future work, it is planned to study the performance of prediction using other prediction models under deep learning techniques. Also, impacts of parameters such as hash-rate and difficulty level may be considered to study the impact on performance.

REFERENCES

[1] Christidis, Konstantinos, and Michael Devetsikiotis. "Blockchains and smart contracts for the internet of things." *Ieee Access* 4 (2016): 2292-2303.

[2] Georgoula, Ifigeneia, Demitrios Pournarakis, Christos Bilanakos, Dionisios Sotiropoulos, and George M. Giaglis. "Using time-series and sentiment analysis to detect the determinants of bitcoin prices." (2015).

[3] Shah, Devavrat, and Kang Zhang. "Bayesian regression and Bitcoin." In *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*, pp. 409-414. IEEE, 2014.

[4] Matta, Martina, Ilaria Lunesu, and Michele Marchesi. "Bitcoin Spread Prediction Using Social and Web Search Media." In *UMAP Workshops*. 2015.LSTMTM

[5] Greaves, Alex, and Benjamin Au. "Using the bitcoin transaction graph to predict the price of bitcoin." *No Data*(2015).

[6] Madan, Isaac, Shaurya Saluja, and Aojia Zhao. "Automated bitcoin trading via machine learning algorithms." *URL: http://cs229. stanford. edu/proj2014/Isaac% 20Madan* 20 (2015).
[7] McNally, Sean, Jason Roche, and Simon Caton. "Predicting the price of Bitcoin using Machine Learning." In *Parallel, Distributed and Network-based Processing (PDP), 2018 26th Euromicro International Conference on*, pp. 339-343. IEEE, 2018.
[8] Selvin, Sreelekshmy, R. Vinayakumar, E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. "Stock price prediction using LSTM, RNN and CNN-sliding window model." In *Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on*, pp. 1643-1647. IEEE, 2017.
[9] Hiransha, M., E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. " NSE Stock Market Prediction Using Deep-Learning Models." *Procedia Computer Science* 132 (2018): 1351-1362.
[10] Moghaddam, Amin Hedayati, Moein Hedayati Moghaddam, and Morteza Esfandyari. "Stock market index prediction using artificial neural network." *Journal of Economics, Finance and Administrative Science* 21, no. 41 (2016): 89-93.
[11] Menon, Vijay Krishna, Nithin Chekravarthi Vasireddy, Sai Aswin Jami, Viswa Teja Naveen Pedamallu, Varsha Sureshkumar, and K. P. Soman. "Bulk price forecasting using spark over nse data set." In *International Conference on Data Mining and Big Data*, pp. 137-146. Springer, Cham, 2016.
[12] Guresen, Erkam, Gulgun Kayakutlu, and Tugrul U. Daim. "Using artificial neural network models in stock market index prediction." *Expert Systems with Applications* 38, no. 8 (2011): 10389-10397.
[13] Jia, Hengjian. "Investigation into the effectiveness of long short term memory networks for stock price prediction." *arXiv preprint arXiv:1603.07893* (2016).
[14] Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "Understanding the exploding gradient problem." *CoRR, abs/1211.5063* (2012).
[15] Hochreiter, Sepp. "The vanishing gradient problem during learning recurrent neural nets and problem solutions." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6, no. 02 (1998): 107-116.