Bitcoin Price Prediction Using Machine Learning

Prabhat Pundir[1], Saksham Phadke[2], Saurabh Gajbhiye[3], Preethi Y.R[4]
[1, 2, 3] Dept of ISE
4Assistant Professor, Dept of ISE
[1, 2, 3, 4] Sapthagiri College of Engineering, VTU, Bangalore, India

Abstract- *Cryptocurrency such as Bitcoin are more popular these days among investors. In the proposed work, it is attempted to predict the Bitcoin price accurately taking into consideration various parameters that affect the Bitcoin value. For the first phase of investigation, it is aimed to understand and analyze daily trends in the Bitcoin market while obtaining insight into optimal features surrounding Bitcoin price. The data set consists of numerous features relating to the Bitcoin price and payment network over the course of time, recorded daily. For the second phase of investigation, using the available information, we will predict the sign of the daily price change with highest possible accuracy.*

I. INTRODUCTION

A. **Bitcoin:**

In October 2008, Bitcoin was first introduced by Satoshi Nakamoto through his white paper entitled "Bitcoin: peer-to- peer Electronic Cash System" [1]. Bitcoin is the first decentralized cryptocurrency while other digital currencies are created by cloning or adjusting the mechanism of Bitcoin [2]. All transactions controlled by cryptography make them secure, validated, and stored in "blockchain" by a decentralized network [3]. With the concept based on the new electronic cash system, online payment transactions can be done directly between any two willing parties without the need for a trusted third party such as a financial institution. Users send payments by reporting digitally signed messages to the network. Members known as miners verify and timestamp transactions into a common public database called the blockchain, for which they are repaid with transaction fees and newly minted bitcoins. Bitcoin was the largest and most popular in cryptocurrency market measured by market capitalization in March 2017. Bitcoin accounts engrossed 72% of the total cryptocurrency in market and number of transactions were 286,419 in January – February 2017 which are more than all other cryptocurrencies [2]. In 2007, the value of Bitcoin was at 1,000 USD and went up to 16,000 USD in December 2017. This makes Bitcoin's prices extremely tough to predict.

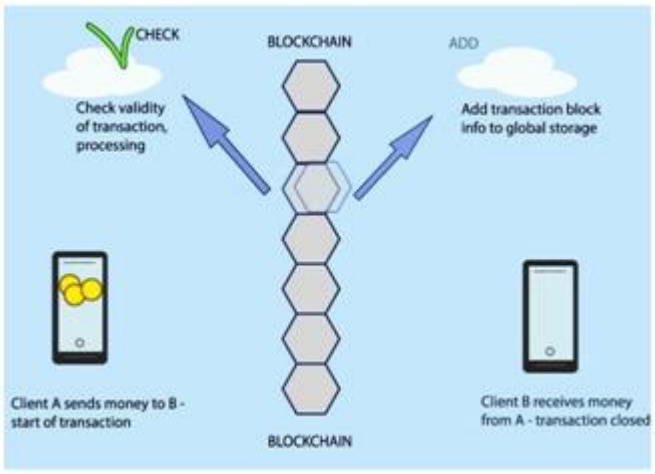

Fig.1. Cryptocurrency workflow using blockchain mechanism

B. Prediction :

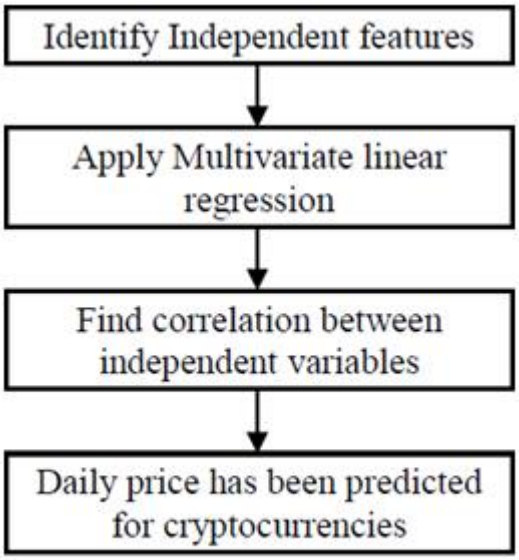

Fig.2. Working of bitcoins

The Bitcoin's value varies just like any other stock . There are many algorithms used on stock market data for price forecast. However, the parameters affecting Bitcoin are different. Therefore it is necessary to foretelling the price of Bitcoin so that appropriate investment decisions can be made. The price of Bitcoin does not depend on the business events or intervening government authorities, unlike the stock market. Thus, to forecast the value we feel it is necessary to leverage machine learning[6][7] technology to predict the

price of Bitcoin. In our given dataset, we have made use of the following independent features to determine the highest price of Bitcoin. These features are:

- Open: Price at which bitcoin opened.
- Low: Lowest price achieved.
- Close: Closing price of bitcoin.

II. PREDICTION TECHNIQUES

A. Data Collection

Machine learning models in this research use Bitcoin transaction data from the bitstamp website with publishing on the Kaggle, Blockchai.info website. However, in this research, 1-minuteinterval trading exchange data rate in USD from January 1,2012 to January 8, 2018 is focused. The datasets are in CSV files. Features of the datasets from *A* are as follows:

Features	Definition
Close	latest trade
Open	opening trade
High	highest trade during day
Low	lowest trade during day
Weighted price	mean Bitcoin price
Volume_(BTC)	total trade volume of day in BTC
Volume_(Currency)	total trade volume of day in USD
Timestamp	data recorded time

B. Data Cleansing

From exchange data we consider relevant only the Volume, Close, Open, High prices and Market capitalization. For all data sets if NaN values are found to be existent, they are replaced with the mean of the respective attribute. After this, all datasets are merged into one, along the time dimension. Judging from Bitcoin price movements during the period from which it started until the present date.Hence the data which will be passed to the neural network lies from the starting date till the present date.

C. Data Preparation

In our study, the basic dataset consisted of 2590 days Bitcoin data (from 29 November 2011 to 31 December 2018). Specific data point, p, is an 18-dimensional real-valued vector where each dimension stores a daily value of one of the Bitcoin blockchain features discussed. To predict Bitcoin prices, sequences of past data were exploited. When the size of each sequence was m, a total of 2590 m + 1 sequence data were used to train and test various deep learning-based prediction models, that is, from S[1 : m] to S[2590 m + 1 : 2590], as shown in Figure 4. The sequence size was experimentally determined as explained , and the first 80% of the sequence data were used as the training data and the rest as the test data. Given a sequence S[i : i + m 1], for a regression problem, we predicted the Bitcoin price of the (i + m)-th day, and for a classification problem, we predicted whether the price would go up or down in the (i + m)-th day with respect to the (i + m 1)-th day. In other words, by analyzing past Bitcoin prices including today's price, we predict tomorrow's price and predict if tomorrow's price will go up or down with respect to today's price for the analysis case, as it is what most investors or traders would be interested in.

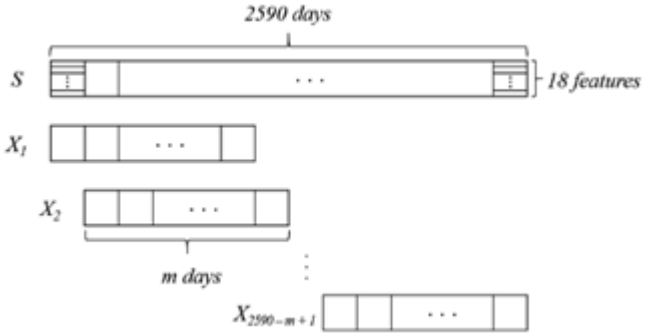

Fig.4. Data preprocessing. A total of 2590 m + 1 sequence data are generated from the whole dataset of 2590 days if m consecutive days are analyzed to predict the next Bitcoin price.

D. Block diagram

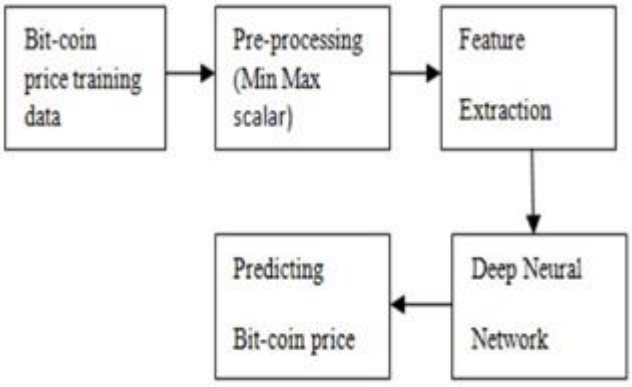

Fig.5. Block diagram

III. RESULTS AND DISCUSSION

A. Convolutional Neural Networks (CNN)

A convolutional neural network (CNN) has many applications in image analysis and classification problems. Recently, it is shown to be also effective for sequence data analysis and thus we also developed a CNN-based prediction model. Normally, a CNN consists of a series of convolution layers, ReLU layers, pooling layers, and fully-connected layers, where a convolution layer convolves the input with a

dot product. In this work, we developed a simple CNN model consisting of a single 2D convolution layer (Conv2D) as shown in Figure 7. (A more sophisticated CNN model is discussed in the next section.) The input is an m _ 18 matrix, where m is the number of days to be consulted for the prediction and 18 is the number of the features. A total of 36 2D convolution filters of size 3 _ 18 are used for convolution, where single real values are obtain from consecutive three days data through each filter. That is, 3 _ 18 feature values are translated into a single value and thus each filter produces a real-valued vector of size m 2, which is then applied to an element-wise ReLU activation function. Then, the 36 output vectors produced by the convolution filters are flattened into a single vector of size (m 2) _ 36, which is then translated into a single prediction value through a fully connected layer. The number of filters was determined experimentally. Moreover, unlike other image analysis applications, simply adding more convolution layers together with pooling layers did not improve the performance of CNN models for Bitcoin price prediction, and therefore we present only a simple CNN model.

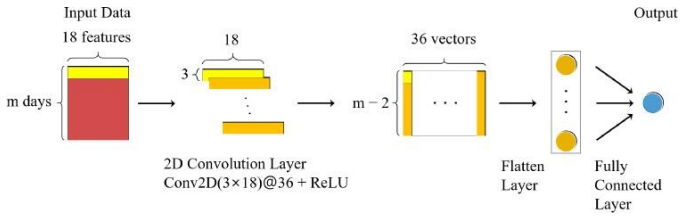

Fig.6. Our CNN model. It consists of a single 2D convolution layer where 36 filters of size 3 _ 18 are used for convolution. An m _ 18 input matrix is translated into an (m2) 36 matrix by the Conv2D layer.

The following results were obtained for CNN classification algorithm after running the process with inputs Datasets for predicting the CNN forecast.

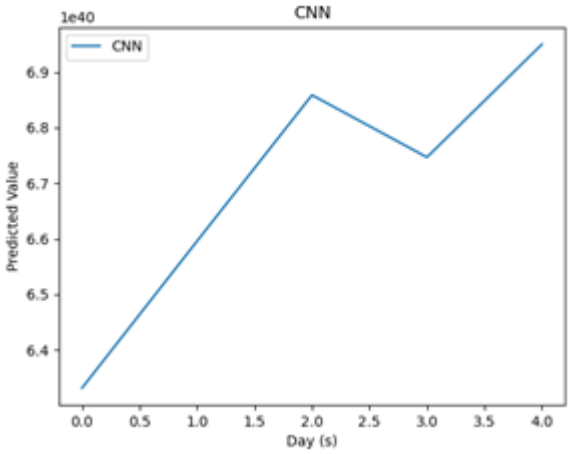

Fig.7. CNN Forecast

B. LSTM

LSTM model is a powerful type of recurrent neural network (RNN), capable of learning long-term dependencies. For time series involves autocorrelation, i.e. the presence of correlation between the time series and lagged versions of itself, LSTMs are particular useful in prediction due to their capability of maintaining the state whilst recognizing patterns over the time series. The recurrent architecture enables the states to be persisted, or communicate between updated weights as each epoch progresses.Where, ft is a sigmoid

$$f(t) = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

function to demonstrate whether to keep the previous state, Ct-1 is the old cell state, Ct is the updated cell state, Wf, Wi, and WC are the previous value in each layer, ht-1 and xt, is the input value, bf, bi,and bc are constant values, it decides which value will be used to update the state, Ct stands for the new candidate values.

LSTM for sequence prediction acts as a supervised algorithm unlike its autoencoder version. As such, the overall dataset should be split into inputs and outputs. Moreover, LSTM is great in correlation with classic statistics linear models, since it can easier handle multiple input forecasting problems. In our approach, the LSTM will use previous data to predict the next day price ahead of closing price.

The following result where obtained after processing the dataset from Quandl and extracting the features from the datasets to forecast the CNN-LSTM and Bitcoin price for prediction.

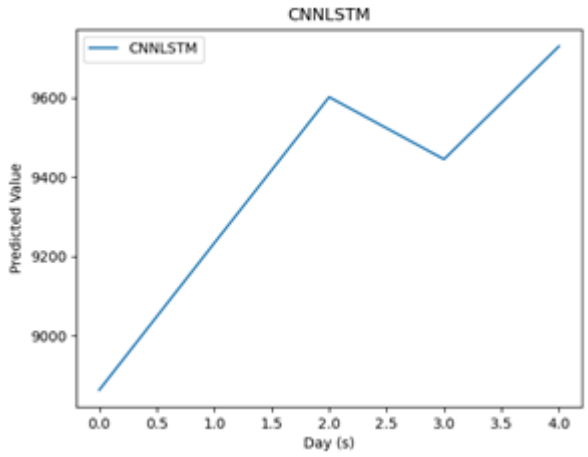

Fig.8. CNN-LSTM Forecast

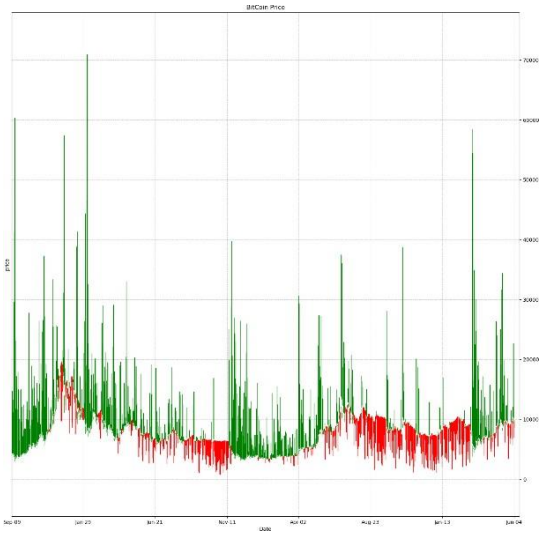

Fig.9. Graph for Bitcoin Price Prediction

IV. CONCLUSION/FUTURE WORK

In this paper, we have done few experiments to predict the price of bitcoin based on their open, high and low cost. This price prediction helps the number of users who are using cryptocurrencies for multiple types of investments. These experiments allow us to get much more in-depth knowledge about the various aspects of the cryptocurrencies. In this paper, we applied our approach to the comparatively smaller dataset. Overall, the comparison is done for the prediction of bitcoin using and CNN_LSTM algorithm using different amount and size of datasets. Taking the time and the speed also in consideration and the complexity of the database. Depending on the statistical accuracy of the Bitcoin prices where the array of datasets is taken first. We can decide the number of days by which we wish to predict the price. This is done by creating small data frames, with some days from the training set. Overall, cryptocurrencies particularly Bitcoin are still in their infancy making it challenging for users to predict how they will evolve in the future. For the sake of future work, we are planning to extend our analysis to the more significant dataset and may collaborate big data technologies with it. We believe that it will be interesting to study the impact of cryptocurrency in real life scenario and this become one of the accessible areas among researchers for researchers. That helps the customer increments and profits. The drawback of using this approach is that we might not have enough information for long-term analysis.

REFERENCES

[1] Satoshi Nakamoto.2008. Bitcoin: A peer-to-peer electronic cash system.

[2] Neha Mangla, Akshay Bhat, Ganesh Avabratha, Narayana Bhat 2019.
"Bitcoin price prediction using machine learning"
Dept. of ISE, Atria Institute Of Technology, Bengaluru, India

[3] Thearasak Phaladisailoed, Thanisa Numnonda
"Machine Learning Models Comparison for Bitcoin Price Prediction" 2018.

[4] Ruchi Mittal, Shefali Arora and M.P.S Bhatia Netaji Subhas Institute of Technology.

[5] Lekkala Sreekanth Reddy, Dr.P. Sriramya
"A Research On Bitcoin Price Prediction Using Machine Learning Algorithms".

[6] Marius Kinderis1, Marija Bezbradica1 and Martin Crane1 1School of Computing, Faculty of Engineering & Computing, Dublin City University, Glasnevin, Dublin 9, Ireland.

[7] Pratik Mehta, E. Sasikala
"Prediction of Bitcoin using Recurrent Neural Network" International Journal of Recent Technology and Engineering.

[8] Kejsi Struga, Olti Qirici, University of Tirana "Bitcoin Price Prediction with Neural Networks".