JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Bitcoin Price Prediction Using Deep Learning

Sai Mouna Bogireddy[1], Sai Mithilesh Reddy Bogireddy[2]

[1] Department of Computer Science and Engineering, Osmania University, Hyderabad, India
[2]Department of Information Technology, Osmania University, Hyderabad, India

Abstract: - Bitcoin has recently become a household name due to its incredible growth, unpredictable volatility, and interesting applications. Bitcoin is currently a thriving open-source community and payment network, which is currently used by approximately 10 million people. As the value of Bitcoin in US Dollar fluctuates every day, it would be very interesting for investors to forecast the Bitcoin value but at the same time making it difficult to predict. This work focuses on predicting Bitcoin prices using a Long Short Term Memory (LSTM) and Gated Recurrent Units (GRU) algorithm. The goal is to ascertain with what accuracy the direction of Bitcoin price in USD can be predicted.

Index Terms – Deep Leaning, Long Short Term Memory, Gated Recurrent Units, Bitcoin Price.

I. Introduction

Bitcoin is a crypto currency, a form of electronic cash. It is a decentralized digital currency without a central bank or single administrator that can be sent from user to user on the peer-to-peer bitcoin network without the need for intermediaries. Transactions are verified by network nodes through cryptography and recorded in a public distributed ledger called a blockchain. Bitcoin was invented by an unknown person or group of people using the name Satoshi Nakamoto and released as open-source software in 2009. Bitcoins are created as a reward for a process known as mining. They can be exchanged for other currencies, products, and services. Research produced by the University of Cambridge estimates that in 2017, there were 2.9 to 5.8 million unique users using a crypto currency wallet, most of them using bitcoin. Bitcoin has been criticized for its use in illegal transactions, its high electricity consumption, price volatility, thefts from exchanges, and the possibility that bitcoin is an economic bubble. Bitcoin has also been used as an investment; hence the need to predict the prices is very much needed at the hour.

Bitcoin is more accessible, with more exchanges, more merchants, more software and more hardware that support it. Bitcoin has two things going for it that help significantly in this respect — Stability and entrepreneurship. It has the most entrepreneurs creating companies around it with a lot of intellect, dedication and creativity going toward making it more useful. As Bitcoin evolves, we can expect Bitcoin to grow in unexpected ways as new utility is found. Bitcoin owners can expect that its usefulness will only increase over time, hence creating a huge opportunity for investment and make huge profits. But when to invest and how much to invest is questionable and hence we have built this model to help predict the best time to invest.

Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised. Huisu Jang and Jaewook Lee's [1] Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases superior to human experts.

II. Literature Review

Various approaches have been used in the past to carry out the price prediction task. There are mainly two sets of literature that are highly relevant to this work. One is financial data analysis; the other, time series data analysis.

A.Financial Data Analysis:

Several approaches are describes in the literature including, one called technical analysis also known as "charting" that forecasts future prices. According to it, stock market prices do not follow random walk, that is the price movements follow asset of patterns. These price movements can be used to predict the future price. There exist some other empirically designed patterns such as head-and-shoulders, double-top-and-bottom that can be used to predict future prices. We refer the interested reader to the work of Lo et al (2000).In this paper, authors have used kernel based regression techniques to find out the patterns in historical data, that is-price is predicted based on past data. This work is theoretically close to the current project work. However, it does not employ the same strategies followed in the current project.

B. Time Series Data Analysis :

In the context of future price predictions, classical methods are quite popular. Autoregressive integrated moving average (ARIMA) models are a popular choice for forecasting over a short term. It works very well when the data exhibits consistent or stable pattern over time with least possible outliers. The ARIMA methodology works well only when that data is “stationary”, which means that the series remains almost constant. But this is not always possible in the real time scenario, where the data fluctuates drastically, and it is highly volatile.

Ediger and Akar used the seasonal ARIMA model to estimate the future fuel energy demand in Turkey over certain years. However, the similar scenario is not guaranteed to work for unseasonal or non-linear data. To solve real time prediction problems, artificial neural networks are very much useful to increase the speed of computation due to its ability to handle nonlinearities in the data. In one of our research papers, by Amjad, M.J. & Shah [2], the prediction is done using ARIMA model, and that could predict prices with only small datasets and the predictions are not that accurate because of the choice of the model, for the highly volatile bitcoin data. Referring to another research paper, by Tian Guo and Nino Antulov Fantulin [3], the prediction is done using ARIMA model and also by using Random Forests algorithm. The limitation of this work is the choice of the algorithms, where the data is prone to over fitting, collecting more noise and this had a serious effect on prediction.

Random Forest algorithms works very well for classification models rather than prediction. For time series data prediction, it is always good to choose, Recurrent Neural Networks (RNN) or variants of RNN. In another of our research papers, the work presents an application of artificial neural networks for making one day ahead prediction of highest and closing price of crypto currency Bitcoin. Two temporal neural network architectures have been considered: a time-delay neural network (TDNN) and a recurrent neural network (RNN). Her results indicate that TDNN models predicted the values closer to the actual price compared to RNN. The limitation of this work is that her work focuses only on one-day ahead and partly on three-day ahead prediction (2016-17 dataset) of Bitcoin price. The number of hidden layers in these models had been restricted to only one and hence the results may not be accurate.

In research paper of, Siddhi Velankar’s [4] thesis contributes to the relatively new field of Bitcoin price prediction in several ways. He used ARIMA and Artificial Neural Network models on real Bitcoin price data to predict the Bitcoin price. His model takes transaction costs in algorithmic Bitcoin trading into consideration. This has given him more realistic return rates and a better grasp of how well the applied machine learning techniques would actually work when trading Bitcoin. The limitations are that his neural network model was not able to manage to get any positive return rates considering he took into account transaction costs when creating and evaluating our models. He even did not consider transaction latency. In his prediction experiments, the prediction models only predict one step into the future which results in missing out on profits.

In another research paper by P. Katsiampa [5], they have tested the performance of three forecasting models on daily crypto currency prices for 1, 681 currencies. Two of them (Method 1 and Method 2) were based on gradient boosting decision trees and one is based on long short-term memory recurrent neural networks (Method 3). In Method 1, the same model was used to predict the return on investment of all currencies; in Method 2, they built a different model for each currency, that uses information on the behaviour of the whole market to make a prediction on that single currency; in Method 3, they used a different model for each currency, where the prediction is based on previous prices of the currency. It is important to stress that their study has limitations. First, they did not attempt to exploit the existence of different prices on different exchanges, the consideration of which could open the way to significantly higher returns on investment. Second, they ignored intra-day price fluctuations and considered an average daily price. Finally, and crucially, we run a theoretical test in which the available supply of Bitcoin is unlimited and none of our trades influence the market.

III. IMPLEMENTATION

A. Data Sets

The Quandl API provides the Bitcoin price data set, starting from January 2014 – 2019(present). This API gives access to Bitcoin exchanges and daily Bitcoin values. It allows users to customize the query while using interface to download the historical Bitcoin prices. The data is available in three different formats i.e JSON, XML and CSV. Data is downloaded in CSV format.

B. Data Pre-processing

We start with collection of the data from Quandl API and this data is sent for pre-processing and we reduce the dimensionality by selecting the features that are actually required for the prediction. Then we start labelling the data and also segregate training and testing data. We also start building our neural networks model, where we are trying to find out the best model, which yields more accuracy. We build a LSTM model and also a GRU model and will find out which model is more ideal. After building the neural network model, we train the model, with previous or historical prices and then test the model with our testing data for calculating or evaluating the performance using performance metrics like, Root Mean Squared Error (RMSE) and we plot graphs to understand it better. We use the best model for real-time purposes as well.

C. Design

The need system architecture diagrams to understand, clarify, and communicate ideas about the system structure and the user requirements that the system must support. It's a basic framework that can be used at the system planning phase to help partners understand the architecture, discuss changes, and communicate intentions clearly.

UML (Unified Modelling Language) is a graphical language for modelling the structure and behaviour of object-oriented systems. UML is widely used in industry to design, develop and document complex software. The sequence diagram is yet another important UML diagram which shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality.

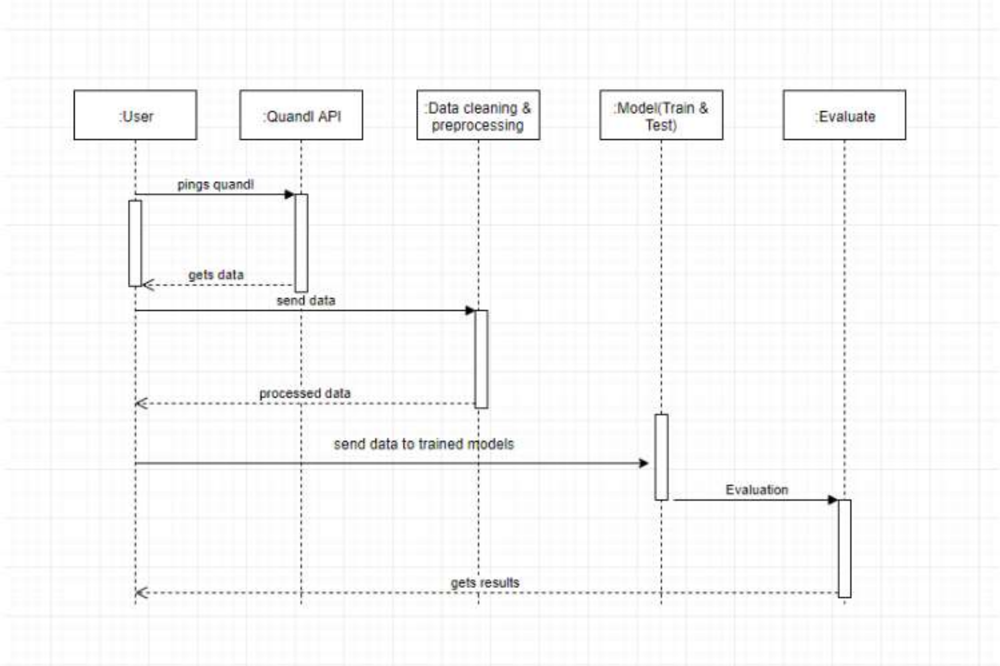

Fig 1. Sequence Diagram for Bitcoin Price Prediction

The sequence diagram is as follows:

- The user pings the Quandl API requesting data.
- The data is sent by the Quandl API and is ready to be used for pre-processing.
- The data is then cleaned by replacing the NANs with last valid observation and pre-processed.
- And then we build the model and train it with training data.
- The model is tested with testing data and the results are evaluated.

The Design of the model, takes the following steps as follows:

We import Day-to-Day data of the Bitcoin price from Quandl API. The Next step is Data Cleansing, which is filling out the values which are null or zero with Nan. We have seven features out of which, we have selected one attribute, weighted Price. We then label the dataset into Train and Test on the basis of 7:3 ratios. We create our neural network models, LSTM (Long Short Term Memory) and GRU (Gated Recurrent Units) and we train these models with the Training data. We now have the Final neural network model, and test the final model with testing data and then we plot the graphs, with predicted and actual prices with blue and green labels for predicted and actual prices. We will follow the same process by taking multiple features which are suitable for this project and we compare the accuracies of both the models, LSTM and GRU against each other with single feature and multiple features. The parameters that define the accuracy are evaluated, that is RMSE, and the graphs are plotted against Actual and predicted prices, for all cases, LSTM with single and multiple attributes and similarly for GRU. So the best model will be most suitable for time series data prediction.

D. Deep Learning Models

i. Long Short Term Memory Networks

Long short-term memory (LSTM) units (or blocks) are a building unit for layers of a recurrent neural network (RNN). A RNN composed of LSTM units is often called an LSTM network. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell is responsible for remembering values over arbitrary time intervals. The expression long short-term refers to the fact that LSTM is a model for the short-term memory which can last for a long period of time. An LSTM is well-suited to classify process and predict time series given the time lags of unimportant series and duration between important events.

ii. Gated Recurrent Units

Gated Recurrent Units (GRU) is another variation of RNN. It's network structure is less sophisticated than LSTM with one reset and forget gate but getting rid of the memory unit. It is claimed that GRU's performance is on par with LSTM but more efficient. They are a gating mechanism in recurrent neural networks, introduced in 2014 by Kyunghyun Cho et al. Their performance on polyphonic music modeling and speech signal modeling was found to be similar to that of long short-term memory (LSTM). However, GRUs has been shown to exhibit better performance on smaller datasets. They have fewer parameters than LSTM, as they lack an output gate.

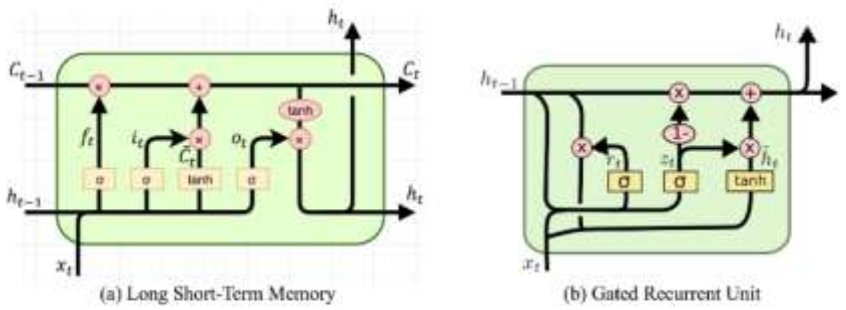

Fig 2. LSTM and GRU Architecture

IV. RESULTS AND DISCUSSION

The results are plotted against actual and predicted price for single and multiple attributes with timestamp being on X-Axis. The results and discussions are a crucial part of the project. The results discuss a lot about the project, analysis of the models used. Here the RMSE of the models, LSTM and GRU are depicted using histograms and scatter plots for prices.

A. LSTM Model:

Fig 3. Graph depicting prices using LSTM

The above is by using a single attribute. We need to predict the prices using multiple attributes as well. To find out which attribute to be considered, we use a heat map to see the correlation between attributes.

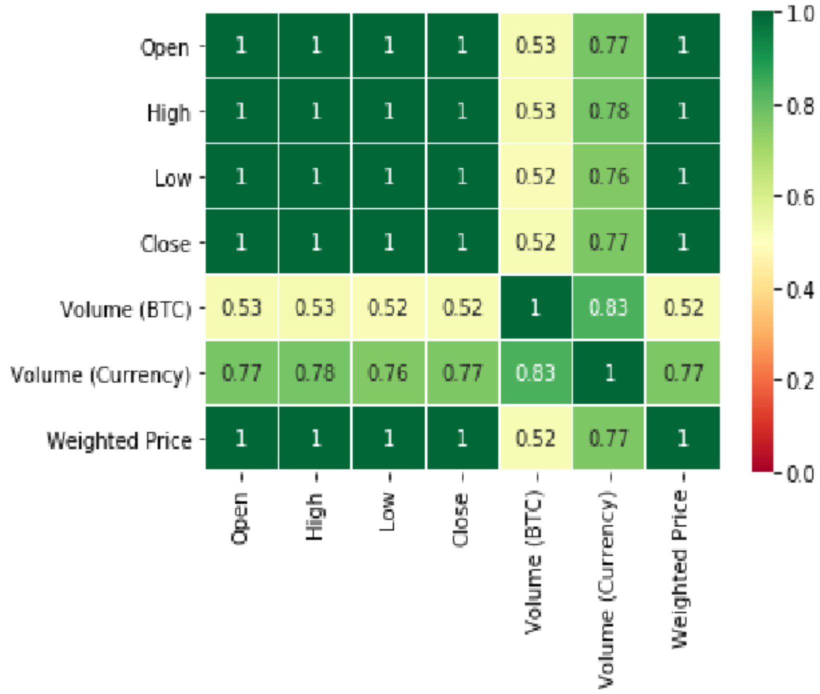

Fig 4. Heat map depicting correlation between attributes

The high level of correlation is between Currency, Weighted price and Volume (BTC). In this step, all the unnecessary columns are dropped and features which are only useful for prediction are retained.

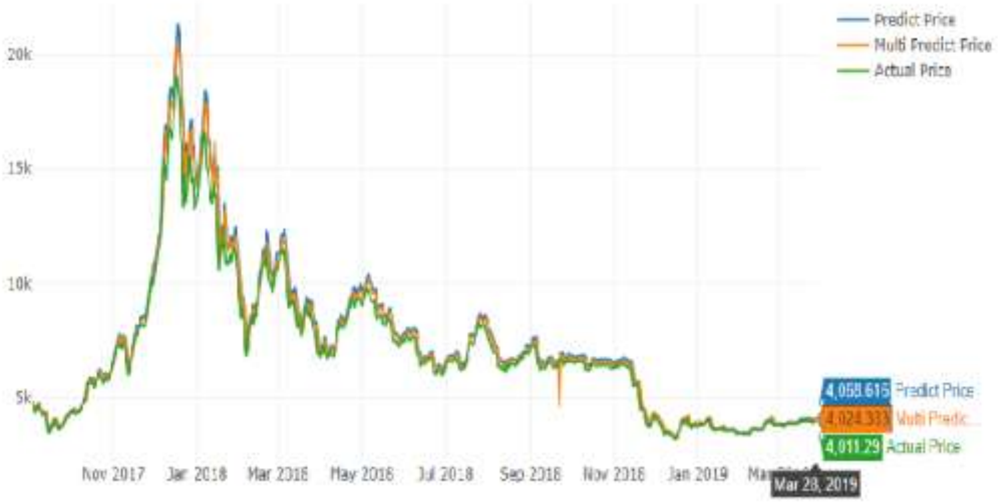

Fig 5. Graph depicting single and multiple attributes using LSTM model

B. GRU Model:

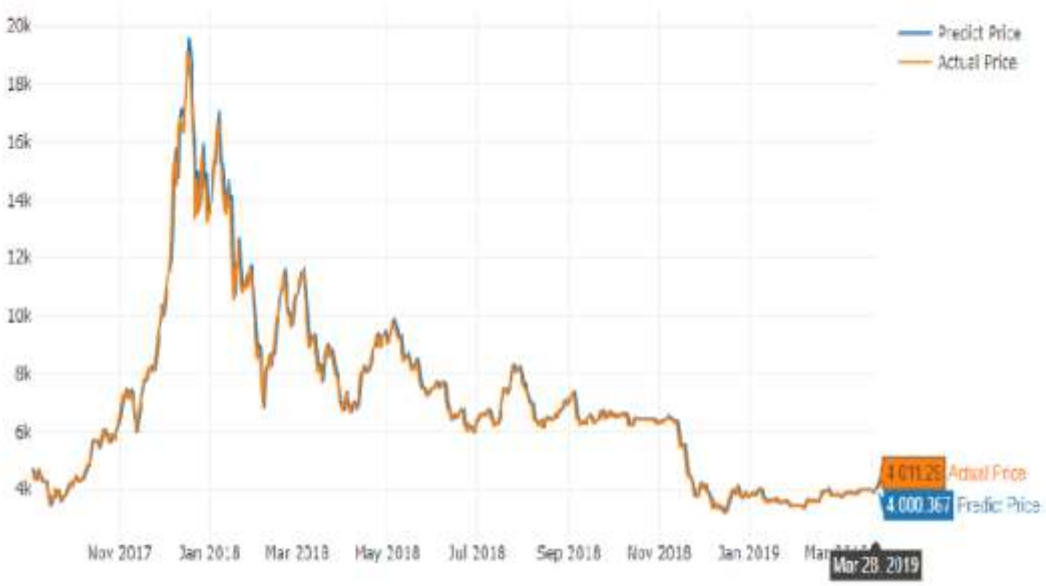

Fig 6. Graph depicting single attribute bitcoin prices using GRU model

The above is by using a single attribute. We need to predict the prices using multiple attributes as well. To find out which attribute to be considered, we use a heat map to see the correlation between attributes.

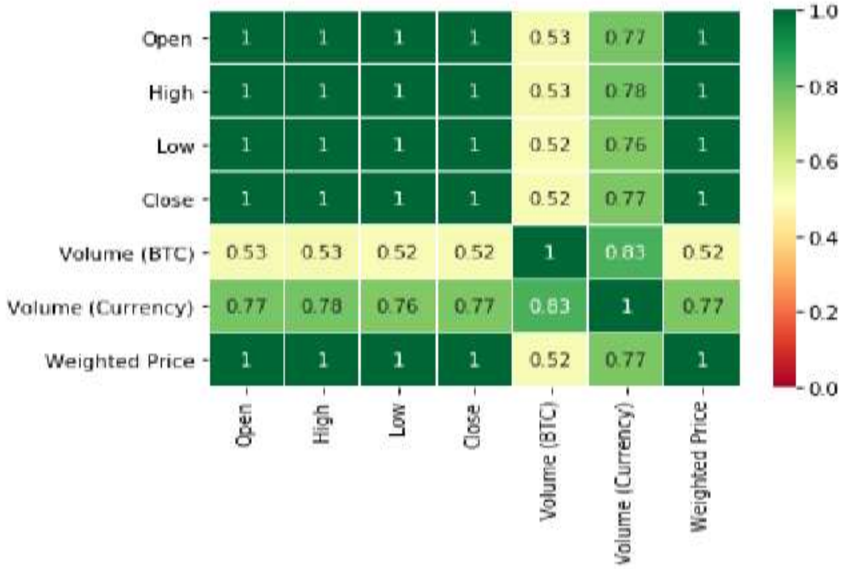

Fig 7. Heat map depicting correlation between attributes

The high level of correlation is between Currency, Weighted price and Volume (BTC). In this step, all the unnecessary columns are dropped and features which are only useful for prediction are retained.

Fig 8. Graph depicting single and multiple attributes using GRU attributes

The LSTM model is good at receiving more attributes and it predicts more accurately with multiple attributes .In the above graph, S indicates single attribute and M indicates multiple attributes. The LSTM model's RMSE for a single attribute model is 396 and the same model with multiple attributes showed 380.

From the above histogram, it can be understood that the LSTM model with single attribute gives more error compared to LSTM model with multiple attributes. Thus, we come to a conclusion that LSTM model with multiple attributes is a better model to use for bitcoin price prediction in the future.

V. Conclusion

This work presents an application of artificial neural networks for predicting bitcoin prices. Two neural network architectures have been considered: Long Short Term Memory (LSTM) and Grated Recurrent Neural Networks (GRU). Also, comparisons between LSTM and GRU have been presented. It is observed that LSTM model works better for model with multiple attributes than single attribute. The RMSE is lower for LSTM in case of multiple attributes. It is observed that GRU model works better for single attribute than for multiple attributes. Comparing the overall performances, it is always proved that RMSE for GRU models is considerably low when compared to LSTM models. After training, both the network models provided satisfactory results.

VI. REFERENCES

[1] Huisu Jang and Jaewook Lee, "An Empirical Study on Modeling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information",28 February,2018.

[2] Amjad, M.J. & Shah, D, "Trading Bitcoin and Online Time Series Prediction", Proceedings of the Time Series Workshop at NIPS, 2017.

[3] Tian Guo and Nino AntulovFantulin, "An experimental study of Bitcoin fluctuation using machine learning methods", proceedings of ACM conference 2018, New York, USA.

[4] Siddhi Velankar, Sakshi Valecha, Shreya Maji, "Bitcoin Price Prediction using Machine Learning", 20th International Conference on Advanced Communication Technology(ICACT), 2018.

[5] P. Katsiampa, ''Volatility estimation for Bitcoin: A comparison of GARCH models,'' Econ. Lett. , vol. 158, pp. 3–6, Sep. 2017.

[6] S. Nakamoto, ''Bitcoin: A peer-to-peer electronic cash system,'' Tech. Rep., 2008.

[7] Takuya Shintate, Lukas Pichi, "Trend Prediction Classification for High Frequency Bitcoin Time Series with Deep Learning", 21 January, 2019.