IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Research on Bitcoin Price Prediction Using Machine Learning Algorithm

GAURAV SHEKHAR

Dept. of Comp. Science Engineering.
SRM Institute of Science and Technology,
Chennai, India

BHAVIKA MOTWANI

Dept. of Comp Science Engineering.
SRM Institute of Science and Technology,
Chennai, India

DR. V ANBARASU
Associate Professor
Dept. of Comp Science Engineering.
SRM Institute of Science and Technology,
Chennai, India

***Abstract*—In this Research paper, we have tried to foresee the future value of bitcoin in the short term. We have used many trade indicators and technical analysis methods used in financial markets followed by machine learning strategies to learn from these indicators and predict the future price of bitcoin. We try to say the value of Bitcoins by paying attention towards various conditions that influence the price of Bitcoin over a certain time period. Firstly, we will focus on understanding and identifying changes that occurs in the Bitcoin market while adding value into the positive aspects of the prices. The data sets that is been used has many characteristics related to the cost of Bitcoin and the payments networking within five years, which are put down every day. Secondly, by using available info, we will be predicting the daily change in the price and indicating it with at most accuracy. We have used Long Short-term Memory (ie.LSTM) and Recurrent Neural networks with technical indicators to achieve Highest accuracy.**

***Keywords*— *Bitcoin, Bitcoin predictions, crypto currency, Arima model, Time forecasting,* Long Short-Term Memory (LSTM), RNN, Fb Prophet**

I. INTRODUCTION PART

Bitcoin can be understood as a cryptocurrency that is used globally for e-payments or just for stake. Bitcoin has been transferred to certain areas i.e. which has no ownership. The deals made in terms of Bitcoins are as simple as non-binding in any nation. Investments are done in markets called "bitcoin-exchanges". This allows anyone to trade Bitcoins for other legal tenders. Mt Gox is the largest Bitcoin trade-off in market. Bitcoins are kept in the form of e-wallet similar to a virtual bank account. For every transaction, the timestamp is put down in a location known as Blockchain. Each cell of a block chain has a cursor in the preceding data block. Statistics in blockchain is ciphered. At the time of transaction the username isn't unfolded, only the wallet ID is visible. The purpose of this paper is to investigate the accuracy of the predicted Bitcoin price using Time Series inspection and Machine Learning techniques. In machine learning techniques, we use two methods - RNN and LSTM.

II. STATE OF THE ART (LITERATURE SURVEY)

Bitcoin is a new technology so at the moment there are a few types of pricing available.

A. Bitcoin Price Prediction using Machine Learning

The Algorithm uses Bayesian Regression and GLM / random forest. Benefits- [1] It works to predict by taking a Markup cap. [2] Quandl- filter data using MAT_Lab structures. Disadvantages [1] long process of filtering data. [2] Low downtime to make predictions.

B. Project-Based Learning: Bitcoin Price Prediction, Deep Learning

The Algorithm uses CNN (Convolutional_Neural_Networks) and RNN. Benefits- [1] The Great Benefits of CNN Weight Sharing. [2] Easily calculate large amounts of data set. Disadv. - [1] Convolution works a little bit there,

say maxpool, front and back.

C. Predicting Bitcoins with glance of purchasing and Selling

The Algorithm uses LSTM (Long-Term Memory) and ARIMA (Integrated Autoregressive Rate). Benefits- [1] effortlessly handle Bitcoins. [2] Bitcoins are dealt online. [3] Easy place to do transaction. Disadvantages [[1] A single refund is not a transaction. [2] Conversion will be too late.

D. Bayesian regression and Bitcoin module

The Algorithm used Bayesian retrospect. Advantages- [1] The benefit of the decline in the Bayesian decline in the results of Bitcoin pricing is reflected in the binary options. [2] It helps to understand the results with great care. Disadvantages It takes a long time to resolve a data set.

E. Predicting Bitcoin Prices using Deep Learning

The algorithm used is SVM (Support Vector machine). Advantages- [1] Convincing for high size spaces. [2] It works well with a clear line of separation. [3] Applicable to cases where dataset has more number of samples. Disadvantages [1] not working well when data is huge. [2] Performance not that good when a data is noisy.

III. PRPOSED WORK

The LSTM model is a duplicate of a neural network that is able to learn long-term dependence. In the same way that we use previous data to inform future results, LSTM models use renewable gateways and forget gates to remember from time to time and forget pieces of historical information to inform their predictions. The database used is different from previous versions. Details are timed and recorded at intervals (1 day or so). Such a sequence of data is called the Time Series. Temporary data sets are common to performance. Energy you use and your expenses (calories in, calories expired), climate change, stock market, statistics collected from users of your product / app and even your heart (probably in love) that produces the Time Series. You may be interested in a lot of architecture in relation to your Time Series - posture, time of year and automotive automation are some of the most popular features.

Autocorrelation is the interaction of data points separated by a period of time (known as lag). Seasonal means the discovery of a particular circular pattern at a certain time. The time series is called stationarity if it always has meaning and diversity. Also, the covariance is independent of time. One obvious question you might ask yourself when looking at Time Series data is: “Does the current action value affect the next one?” eg.Time Series Forecast. There are many ways you can use this purpose. But we will build networks for short-term memory that we predict and use to predict future Bitcoin prices.

LSTM models use renewal gates and forget gates to remember from time to time and forget pieces of historical information to inform their predictions. LSTM may choose to "forget" the previous provinces. Therefore, it is better able to deal with data with a recurring tendency over time. The ability to use historical "context" thus allows these models to be particularly suitable for predictive purposes, among many other programs. LSTM solves the problem on the RNN network. There is a switch that can select specific events and note them. And it does not depend on the length of time and does not require much training.

Some useful statistics: -

Eq 1. $$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Eq 2. $$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Where, true/real +ve (T_P), true/real -ve (T_N), false +ve (FP) and false/non -ve (FN).

Eq 3. $$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j|$$

IV. IMPLEMENTATION

L.S.T.M: Long_short_term memory; R.N.N: Recurrent Neural Network comes as a part of in-depth learning or we can say deep learning. It is a repetitive algorithm due to the response cycle in its structure. Advantages over other neural networks is due to its ability to process all data sequences. Its structure consists of a cell, an i/p gate, an o/p gate and a forgotten gate.

The given cell memorizes numbers randomly, and the gates control the flowing of info inside and outside the cell. Model cell responsibly for keeps record of interdependence among consecutive input sequences. The i/p gate controls how a fresh value enters cell, the input gateway also commands the level where the values resides in the cell, and o/p gate commands rate where the cells unit is used to compute LSTM activity value.

However, there are exceptions to LSTM models like Gate Recur renting Unit (GRUs) which don’t have any o/p gateway (Fig 1.). LSTM algorithms are mainly used in timing based series for division, processes and predicting. It is popular in the app series because there are many not known time lag between imp circumstances in the timeline.

Below fig 2 refers to the architecture diagram we are expecting in the bitcoin price prediction.

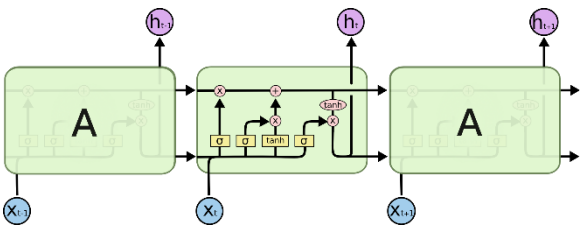


Fig 1. Gateway Diag.

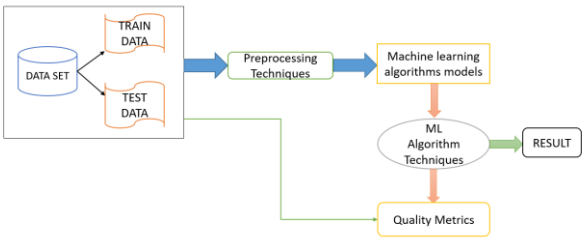


Fig 2. Architecture Diag.

A. Data Preparation

Previous data processing work can be thought of as data cleaning before using it for download to the model. In this process, we first check whether something is being extracted from the database. Database can be downloaded from Kaggle.

The database contains a total of 5 elements. The details are as follows: -

- Close price - It is the price in the closing price of the day
- Maximum value - It is the highest amount of the daily amount
- Lowest price - It is the lowest price for the day.
- Open price - It is the open market price for the day.
- Volume - The volume of sales for the day.

B. Training

The introduction of the train database seems to be one of the major players in the game. We select and provide bulk previews or data sensors as read entries. In a nutshell, the more we bring the better results we expect to get. To understand the trend, the model sets a graph with a date on the x-axis and a stock price on the y-axis. With a planned graph, it can pull out a graph the next day and be able to see the estimated value of the stock.

Therefore, finding the nearest price of an asset is very important to obtain a good database that has been processed and processed using various techniques. Therefore the training data we will provide to our model is very important in determining the level of accuracy of the proposed system. Therefore, to build an accurate model, we need to download the correct database as training data to the model.

C. Testing

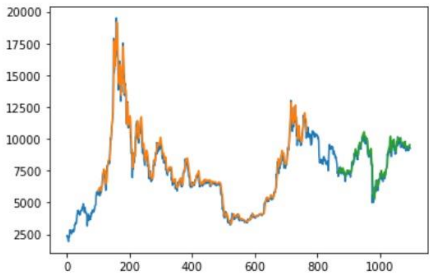
Test data is nothing but data where the actual results are already known and we test the model in that data provided to see how accurate the prediction of the model is compared to the actual result obtained with us.

Model testing can be done in two ways. One way is to provide data for an event that has already occurred in the past and to compare that result with a predictive outcome model. That is, in this case, we need to test the model by asking the model to predict the price of the daily share price that has just occurred in the past. . This process will be repeated many times and should be tested for different stocks in different areas in the past.

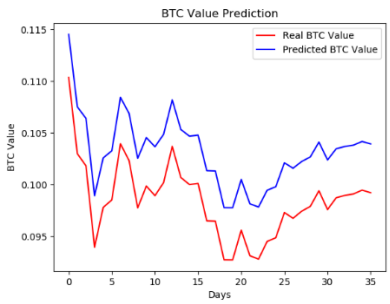
V. DISCUSSION OF RESULTS

	Date	High	Low	Open	Close	Volume	Adj Close
1092	2020-07-06	9360.6171875	9201.8154296875	9349.1611328125	9252.27734375	13839652594.0	9252.27734375
1093	2020-07-07	9450.3359375	9249.5	9253.0205078125	9428.3330078125	19702359882.0	9428.3330078125
1094	2020-07-08	9431.37890625	9234.9990234375	9427.994140625	9277.9677734375	18000702524.0	9277.9677734375
1095	2020-07-09	9287.4716796875	9118.001953125	9273.357421875	9278.8076171875	16860035604.0	9278.8076171875
1096	2020-07-10	9293.5322265625	9199.4853515625	9277.51171875	9240.3466796875	13249910444.0	9240.3466796875

Sample data



Predicting Graph received after launch



Accuracy Graph for Machine Learning Algorithm

VI. CONCLUSION

It has been a great learning and experience for me throughout the whole semester so far specially the main project. I wasn't sure if I could pull this off as it was more challenging than I was prepared for. However, with the support and guidance of all lecturers and special thanks to my Project Guide (Dr. V. Anbarasu). He is really patience and helpful person I wouldn't be able to finish my project without his guidance and support I feel it wouldn't have been possible. Classes were as productive as they could be. The whole module was set to equipped students with a great knowledge and experience when it comes to real world. Can deny it wasn't hectic or I dint feel pressure but I was sure it was for my greater good.

As the project begun, I was in touch

with supervisor once or twice in every 2 weeks. Sometime often as I was facing difficulties at some stages of the module. Although it was my idea to work on that but sure needed a guidance how I can make it work to live up to their expectations. Whenever I needed to see my Project Guide, I was given time even on short notice. I was gaining momentum each time I had meeting. I can't thank my supervisor enough to help me though my project.

It surely is one of my great learning experiences in my academic career.

VII.REFERENCES

[1] Bitcoin_Price: Prices using ML, SiddhiVelankar, SakshiValecha, ShreyMai* Dept of Tech and Comm, Pune Ins of Comp Tech, India-409-415.

[2] D. Sha noK. Zan, ressionBayesian and Bitcoin,at the 52nd annual Allert Conf on Comm, Control, and Computing, 2015.

[3] Bitcoin Predict using DNN Techniq-- Kalpana ,Saysri, Vinoth , SugPriya , publishedIEEE- 2018

[4] Bitcoin prices using gray sys theory --Mahbob Faghi Md Jalali and Haneef Hadari

[5] Huisu Jang and Jaewook Lee, — A Powerful Modeling and Prediction Study for Bitcoin with Bayesian NN based on Blockchain Info -IEEE 2017, volume 99, pages.

[6] Learning-Based Project: Bitcoin Price Predictions using Deep Learn-Yogesh ; Piyush Mahesh; Mandir Kaur ; Amity Univ Dubai; publishedIEEE 2019