Cryptocurrency Real time Data and News Fetcher and prediction

Azhar Hasan
Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India
Azharhasan013@gmail.com

P. Duhitha
Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India
Duhitha.p@gmail.com

Kaushik Bajaj
Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India
Kausheek37@gmail.com

Dr. D.S John Deva Prasanna
Professor, Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India
johndp@hindustanuniv.ac.in

Abstract

Cryptocurrency is the latest form of currency, which is used for trading and other exchanges in market. In pandemic, massive growth has been witnessed in both trading and also in returns. Cryptocurrency price prediction is very difficult and quite impossible as its highly volatile. In this paper, an algorithm involving recurring neural networks and Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) is proposed to forecast the cryptocurrency's price. The unsupervised learning technique is seeming to be used as there is more accuracy in prediction in spite of highly volatile dataset. The price prediction of the cryptocurrency is done by taking bitcoin price as the input to the machine learning algorithm. A complete web app is also implemented using various backend APIs to display detailed information about the various cryptocurrencies and its market valuer predictions. Data like market Cap, Volume, Last traded, 24Hours Volume has been used for prediction. Real time graphs have been plotted for the given crypto currency using Chart-JS. APIs like coin ranking api and bing news api specifically target cryptocurrency as it's subject. The performance of the model is compared with the existing algorithms and found to provide better accuracy.

Keywords: *cryptocurrency, Deep learning, Machine learning, Application Program Interface.*

I. Introduction

Bitcoin is the foremost digital currency which was first introduced in early 2009. Digital Currencies have become one of the most important components of the international financial market after several years. Nevertheless, the prices of crypto changes drastically, and price of the cryptocurrency changes not only depends on market but it is also volatile to the other factors which can cause changes to the price of the currency based on the behavior of traders. There are shallow market complications in addition to the market risks. These factors make it difficult to have a thorough grasp of the cryptocurrency industry. On such issue, several researches have been undertaken. Cryptocurrencies, including a comparison of Bitcoin to other capital assets, as well as assessments of its fluctuating price and drivers.

As a result, we still lack the fundamental notions needed to completely understand cryptocurrencies. We, on the other side, have a much more expertise and understanding in traditional markets like foreign currency exchanges and the stock exchanges, which may help us better grasp the cryptocurrency market and also help us to understand about the market and how it works. We generally chose asset tree and correlation matrix as two of the most extensively used methodologies, based on price dataset recorded from January 2012 to October 2017, to analyze the features of the crypto market and place it in context with other traditional markets through comparisons. The key contributions and conclusions of this study can be stated as follows as a continuation of effort.

First, we are going to study about the cryptocurrency and its influence in the market at the financial market level and in contrast with the traditional markets to check similarities and differences between them. Second, we discovered several fascinating facts throughout our inquiry. Despite the fact that cryptocurrencies were created as a blockchain based digital currency, however the market for the cryptocurrency is dynamic is similar to our stock exchange but with few differences, thus it is a high risk, high return market because of its high volatility. Our research demonstrates the nature of cryptocurrency price fluctuations and give suggestions for investing, regulation, and law. We also plan to implement server-side rendering to the application as it takes the load of the client side machines off and rather performs all the computations and prediction and heavy processing over at server side. Piispanen,M. et. al. [10] explains the importance of modern architecture through his approach. As a result, the client needn't have a very high-end device to use the application and can enjoy highly powerful computational results on any device.

As a result, the client needn't have a very high-end device to use the application and can enjoy highly powerful computational results on any device. To achieve this, we will be implementing Next-JS which is another framework of ReactJS and is known mostly for server-side rendering and its server supported functionalities. We are using a huge dataset to achieve the task of price prediction here forth, consisting of more than 4 lakh entries with their timestamps.

II. LITERATURE REVIEW

Huang et. al. proposes the method where the research is directed for predicting the price of the cryptocurrencies using sentimental analysis of the social media posts and using unsupervised algorithm finding correlation between them. The author has taken posts from the Chinese media app called sina-weibo and created a pipeline for capturing the posts and creating a sentiment crypto dictionary. The proposed approach by the model shows 18.5% in precision and 15.4% in recall.

Rebane et. al. proposes the method where they are comparing the model performance between

ARIMA (Autoregressive Integrated Moving Average Models) and seq2seq using recurrent deep multi-layered neural network by considering various input types selection. Cryptocurrency price prediction has been an attractive topic, which draws attention for media and interests of investors. proposed model compares the performance of Arima and sequence to sequence model where the results showed that seq2seq was better in performance over ARIMA

Velankar et. al. proposes the method where they are predicting the cryptocurrency price by using Bayesian regression and generalized linear regression. Bitcoin is one of the latest technology and current trend. Hence there are many Prices prediction models. The author has taken dataset considering features of the bitcoin price over the time period of five years. Dealing with Bayesian regression where the crypto currency data is break into one third of consecutive intervals of different sizes 180s, 360s and 720s and by applying the k means clustering to narrow down the best 20 or most varied and effective clusters.

McNally et. al. proposes model which shows the implementation of Bayesian optimized RNN (recurring neural networks) and LSTM and comparing it with ARIMA where LSTM performance was outstanding marginally. The ARIMA model which is popularly known for its forecasting using time series performance was poor. The proposed model shows 52% accuracy and RMSE of 8%.

Derbentsev et. al. proposed model where the 90 days horizon for the three currencies namely Bitcoin, Ethereum, and ripple were predicted as they are most capitalized currencies. by using, Neural Networks, Binary Autoregressive Tree Model (BART) and random forest. A comparison of the created model's and the prediction ability revealed that all of them accurately explain the dynamics of cryptocurrencies. The models BART and MLP gives MAPE of 3.5% and random forests gives error within 5%.

Chen et. al. proposed method was using Machine Learning approaches to predict Bitcoin price at various timestamps, we first divide Bitcoin price into two: high frequency price and daily price. For Bitcoin daily price forecast, a collection of high-dimension factors which includes market attention, trading and market, property and network, and gold price are employed. Author uses statistical method and also machine learning models where the accuracy is 66% for statistical method and 67.2% for machine learning models respectively.

Awoke et. al. proposed method where the goal is to use deep learning-based models for predicting price, by using Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM), to deal with cryptocurrency market volatility and achieve high accuracy and high efficiency. Their research compares deep learning with these two-part series approaches and demonstrated their usefulness for forecasting bitcoin prices.

III. PROBLEM STATEMENT

In current world we all are heading towards latest technologies and trends. Cryptocurrency is also similar technology which is in form of digital currency and having slight differences when compared to physically available currencies. Cryptocurrency is highly volatile and so the market. It means the prices goes high and low which states the risk in investing money It has similar work functions for example you can buy a Car or similar things with digital money. But there are 50 (fifty) or more than fifty digital currencies in the market. As its highly volatile and seeing the risk one may find difficulty in finding best digital currency and invest in it. This aim of this system is to Develop a user-friendly web application and news fetching for cryptocurrencies which by gives user real-time

information about the currencies and we are also implementing deep learning algorithms.

IV. MATERIALS AND METHODOLOGY

4.1. Model Selection:

In this paper we are going to train our model with two different models by taking bitcoin as example and using historical prices for bitcoin. Then for evaluation of the model we are going to compare to the current models in following stages

1.) Historical data for cryptocurrency

2.) Exploring the cryptocurrency data and visualizing it

3.) Training and testing of the models by giving cryptocurrency data as input.

4.) Getting output and comparing it

Here we are going to compare between Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) and based on which the model will predict the price for cryptocurrency using historical data. It starts with data preparation.

There are few steps in data preparation:

1. Collection of the data: The very first step in data preparation is information gathering. This is very important step as choosing right and standard data indirectly affects the prediction of your model.

2. Data pre-processing: In this step it is going to split the dataset into the two parts the training data and testing data respectively.

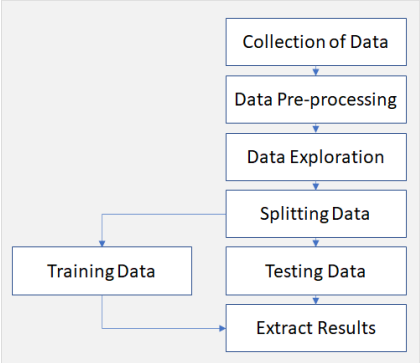
Fig. 1. Train set and Test state

3. Selecting the model: The next step is model selection.

4. Training the dataset: Training the dataset is one of the most important tasks in training the model. To improve the ability of model We will apply the data progressively i.e., the result should be approx. to predicted price.

5. Evaluation: The accuracy which we calculated or obtained gives us brief picture about the best algorithms amongst all.

Fig. 2. Methodology of data processing and model selection.



4.2. Dataset Description:

The data for this study consists of one .csv file for Bitcoin (BTC), The prices which were recorded in the dataset is collected with data counting to approximately 31 Lakhs. From 1 January 2012 to 12 December 2017, data points were taken tenth of a second

updates by considering Open, High, Low, Close, and volume, which denoted currency, and its proportional bitcoin price. Figure 3 shows BTC dataset sample and it demonstrates the dataset and gives brief idea for the targeted digital asset.

4.3. Auto-correlation and partial correlation:

Autocorrelation means the degree of similarity between time series and the delayed or lagged version of itself over particular time intervals. It measures the relationship between present value to the past value.

Partial correlation is a strategy for determining the relationship between the dependent and independent variables while controlling for additional variables.

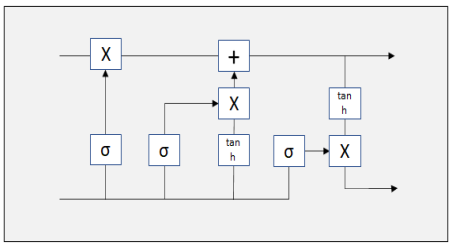
4.4. Deep learning models:

In this section we demonstrate the models used which is Gated recurrent network (GRU), Long Short Term Memory (LSTM).

4.4.1 Long Short Term Memory:

In LSTM and recurrent neural network, they both have one thing in common which is it has a similar control flow mechanism. While working with LSTM's, the data flow through the system which is known as cell states. In this way, LSTM's can remember or forget values to avoid long- term dependency problem.

Fig. 3. Architecture diagram for LSTM



Forget gate layer: Forget get layer also known as sigmoid layer. It is to decide whether the data is to be ignored. While performing the test the input ht-1 and xt, the outcome is a number between forget and keep where forget denotes o and keep denotes 1 representing each number present in the cell state ct-1.

Newer information to store: A sigmoid input gate layer is the deciding factor responsible for which values should be updated and a Tan-h activation function is responsible for the creation of a vector of the new value, ct for the candidate.

Updating the old cell state: Updating the old cell state ct-1 into the new cell state ct. And then multiplying it with the old state ct-1 by f1 and add it* ~ct.

Output: Firstly, we run the sigmoid layer, which is the deciding factor for the particular segments of the cell state that we take into consideration. Then using tan-h activation function values is fed in cell state and to get final output multiplying it with the result of the sigmoid gate.

4.4.2 Gated Recurrent unit (GRU):

Gated Recurrent unit is a part of recurrent neural network that demonstrate its effectiveness in many of the applications which is need of sequential. GRU's are the gated mechanism which is part of recurrent neural networks. The Gated Recurrent Unit is similar to Long Short Term memory with a forget gate, but has fewer inputs when compared to LSTM that's why it is also known as lighter version of LSTM. According to [8] and [9] the deep learning models always shown positive impact in outperforming the poor models for example ARIMA.

Update gate: It is similar to the forget gate and the parameter gate of the Long Short Term memory. It's also one of the deciding factors which decides what value is to be ignored and what value should be added.

Reset gate: It is another gate which decides how much history of information can be forgotten.

4.5. Technologies used in creating web app:

React JS: ReactJS is an open-sourced library made by Facebook for JavaScript used in developing UI for single page applications or SPAs. Handling the view layer for web and mobile apps is performed by react JS.

Ant design: It is a library that has a numerous of components that are easy to use for developing magnificent user interfaces. We are using it for the UI of our application.

React-Redux: It is a state container especially for JavaScript apps. It helps in developing apps that are consistent in behavior and which is platform independent. We are using Redux for the state management of our application.

Chart js: Chart.js is an open-source library that makes it simple to visualise data/datasets using JavaScript. We are using it to plot graphs and prediction graphs in our web app.

React-Moment: React-moment is a standalone open-source JS framework wrapper for date objects that replicates the clunky native JavaScript dates. Using a clean API, MomentJS makes it simple to display,parse, format, validate, and manipulate dates and various timezones.

React-Millify: Millify is an NPM package for react which converts long numbers into pretty, human-readable and convenient strings. Since we are dealing with crypto this is a great tool to have.

4.6. Accuracy and Evaluation matrix:

Performance assessments are criteria that aid in the comparison of various machine learning models i.e., it shows the best one among all other models or method which may be used to predict the price of cryptocurrency. To predict results, we used a variety of methods. These measures are root mean squared and let us understand the performance evaluation measures. The evaluated result of the model is done using the Root mean square error (RMSE) and symmetric mean absolute Percentage error (SMAPE).

RMSE: The Root mean squared Error is used often and straightforward statistic. Mean Squared Error is the square difference between the actual value and the predicted value.

SMAPE: It is error-based measuring system for calculating accuracy for the model. The full form for SMAPE is symmetric mean absolute percentage error.

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(actual - forecasted)^2}{No\ of\ samples}}$$

$$\text{SMAPE} = \frac{100\%}{no\ of\ samples} \sum_{t=1}^{n} \frac{|forecasted - actual|}{|actual| + |forecasted|}$$

Where Ai and fi are Actual price and Forecasting price in order, and n denotes the total number of values/samples.

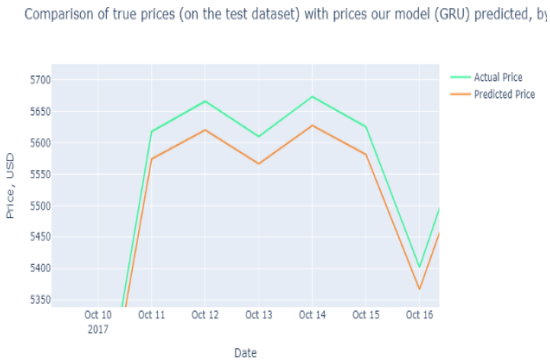
V. RESULTS AND GRAPH DISCUSSIONS

In this research, the obtained results using LSTM and GRU Algorithms using bitcoin (BTC) as it is popular cryptocurrency. The model with the lowest RMSE score and SMAPE is regarded as the best one. Although all are solid models, the GRU was proven to be the best of the all.

In above figures 4 and 5 the x axis is set for date and time and the y axis is set for the price in us dollars. Figure 4 and Figure 4 are maximized image.

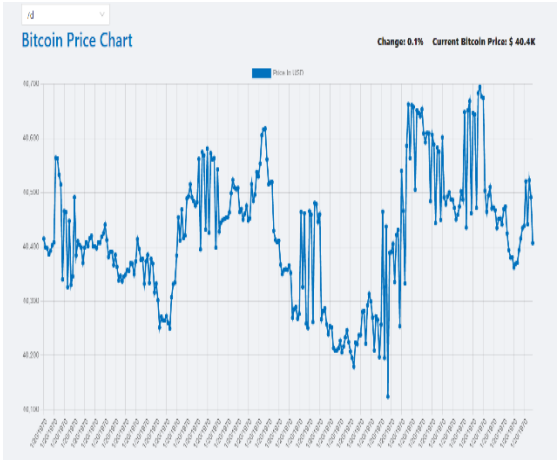
Fig. 4. Graph showing maximized image for LSTM.

Fig. 5. Graph showing maximized image for GRU.



In graph it follows the trend. It notes the points and observes the places where the point is deviating from the trend. At this point when it shows the deviation, identify the factors or variable points which is responsible for that deviation. Then, when making future projections based on the graph, take those variables into account.

Fig. 6. Graph representing price in web app using chart js



VI. CONCLUSION AND FUTURE ENHANCEMENTS

We used Bitcoin historical data to train the Gated Recurrent Unit and Long Short Term Memory the in this study. These models might be used to estimate bitcoin price movements in the future. The models' performance is found to be pretty good. On average, both of the models studied here produce error in few dollars. For GRU the RMSE is 21.265 and SMAPE is 0.236 and for LSTM the RMSE is 54.942, SMAPE is 0.57 respectively. In future to incorporate the work advancing the field from the present state of knowledge and suggest future experiments we are going to Add SSR support and next js for the web app. For prediction model, we can consider external factors and try to integrate with our model so our prediction comes out without giving error. We also plan to incorporate SSR using Next JS and Nest JS.

References

[1] Huang, X., Zhang, W., Tang, X., Zhang, M., Surbiryala, J., Iosifidis, V., ... & Zhang, J. (2021, April). Lstm based sentiment analysis for cryptocurrency prediction. In International Conference on Database Systems for Advanced Applications (pp. 617-621). Springer, Cham.

[2] Rebane, J., Karlsson, I., Papapetrou, P., & Denic, S. (2018). Seq2Seq RNNs and ARIMA models for cryptocurrency prediction: A comparative study. In SIGKDD Fintech'18, London, UK, August 19-23, 2018.

[3] Velankar, S., Valecha, S., & Maji, S. (2018, February). Bitcoin price prediction using machine learning. In 2018 20th International Conference on Advanced Communication Technology (ICACT) (pp. 144-147). IEEE.

[4] McNally, S., Roche, J., & Caton, S. (2018, March). Predicting the price of bitcoin using machine learning. In 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP) (pp. 339-343). IEEE.

[5] Derbentsev, V., Matviychuk, A., & Soloviev, V. N. (2020). Forecasting of cryptocurrency prices using machine learning. In Advanced Studies of Financial Technologies and Cryptocurrency Markets (pp. 211-231). Springer, Singapore.

[6] Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 365, 112395.

[7] Awoke, T., Rout, M., Mohanty, L., & Satapathy, S. C. (2021). Bitcoin price prediction and analysis using deep learning models. In Communication Software and Networks (pp. 631-640). Springer, Singapore.

[8] Rizwan, M., Narejo, S., & Javed, M. (2019, December). Bitcoin price prediction using deep learning algorithm. In 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS) (pp. 1-7). IEEE.

[9] Dutta, A., Kumar, S., & Basu, M. (2020). A gated recurrent unit approach to bitcoin price prediction. Journal of Risk and Financial Management, 13(2), 23.

[10] Piispanen,M. (2017). Modern architecture for large web applications.