

Cryptocurrency Price Forecasting

Shubham Bhattad, Stefin Sunnymon, Dallas Vaz, Prof. Chhaya Dhavale
Dept. of Information Technology, Xavier Institute of Engineering, Mumbai, India

Abstract: Cryptocurrency is a class of digital asset that is very challenging to monitor and forecast. Predicting cryptocurrency price action and its locus is difficult because it does not coincide with market movements. Our objective is to analyze the machine learning algorithms used in various researches and find out the best model which can be used to forecast the prices of time series models. In this work, we have compared and analyzed earlier methodologies in which several machine learning models were applied to forecast the trend of cryptocurrency time series data. The outcomes support the machine learning models' ability to predict trends reasonably well. In total, there are approximately 3000 cryptocurrencies. People are frequently baffled about which coin to invest in for a profitable future. This paper analyzes the trends of the various crypto currencies and predicts the closing price of a cryptocurrency.

Index Terms - Cryptocurrency, Bitcoin, Ethereum, Prophet, XG-Boost, Auto Regressive Integrated Moving Average (ARIMA), Long Short-term Memory model (LSTM), Yfinance, Time Series data, Machine Learning.

I. INTRODUCTION

Machine learning is a technology which enables computers to learn from past data. Machine learning uses various algorithms for building models with mathematical calculations and making predictions using known data or information. ML is used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system, and many more.

Cryptocurrency is a digital currency form that uses cryptography for security and verifies transactions on the network. The market capitalization of all cryptocurrencies is over \$870.81 billion. People are constantly looking for long-term investments that will yield a decent return on their investment. A Blockchain.com poll estimates that 420 million individuals possess cryptocurrency worldwide, which is a staggering amount. 150 million people, or around

11.5% of India's total population possess cryptocurrencies.

Investors would benefit greatly from having knowledge about the cryptocurrency they are investing in as well as the future trends of coins in order to make better decisions. There are several websites and newsletters that advise cryptocurrency investors on the best investments to make money, but none of them forecast prices. Using ML algorithms would make it easier to analyze cryptocurrencies and make the predictions of closing prices of cryptocurrencies with higher volume. In this approach, the users or investors may stop being terrified of the cryptocurrency market due to its volatility and make smarter judgments about when and how much to invest in a certain crypto coin. This research paper introduces machine learning models like the Prophet Model, ARIMA, LSTM, XGBOOST that are used to predict the closing price of a specific cryptocurrency in the future with accuracy and simple implementation.

The prophet model was chosen for the final model predictions because the results are simple to interpret and the prophet model can detect seasonal trends with simple parameters. This model attribute will be useful when the technique is scaled to additional crypto coins with erratic volume and closing rates. This part provides a brief overview of the project, followed by a review of the literature, which includes examples of similar work by other researchers.

II. RELATED WORK

Blockchain technology, which is the underlying framework of cryptocurrencies, has gained a lot of attention and trust because it provides secure transactions and fast data transfer.

According to the study in [1], the volatility in the crypto market is very high and it is tough to forecast the prices of the coins like bitcoin, ethereum etc. Paper [2] compares various deep learning-based price prediction models using Bitcoin blockchain

information. Deep learning models such as deep neural networks (DNN), long short-term memory (LSTM) models, convolutional neural networks (CNN) and their combinations are used.

[3] proposed a price prediction system of different cryptocurrencies using different technical trade indicators and neural network algorithms. [4] depicts the use of Recurrent Neural Network which uses the Long Short-Term Memory algorithm and the Root Mean Square Error was found to be 3.38%. Paper [5] mentioned two deep learning techniques such as LSTM and GRU as prediction models.

Paper [7] focuses on Bitcoin’s future values using the PROPHET and ARIMA methods. According to the results in the paper, the PROPHET model makes predictions quite closer to reality than the ARIMA model. The paper [8] proposes two time series ML models: ARIMA and Prophet model. Both the ARIMA model and Prophet model perform similarly, but the ARIMA model has a slightly higher R-square score.

Paper [9] used four different models, namely LSTM, ARIMA, XGBoost and Facebook Prophet were used and the ARIMA Model came out to be the best model.

III. DATASET DESCRIPTION

While searching for an appropriate dataset for this project, we came across various sites and python libraries providing data on the topic. Websites like kaggle, google public datasets, etc and libraries like pyDataset, Yahoo Finance. So we decided to fetch data from Yahoo Finance library which is a library for all Stock and Crypto related datasets.

The historical cryptocurrency data was obtained from <https://finance.yahoo.com/> which gathered information from the start of cryptocurrency to till date. The research was based on the cryptocurrency’s closing price, which refers to the number of coins traded on that particular day.

This dataset contains seven attributes, including the opening, high, low, and closing prices, as well as the market capitalization of publicly traded outstanding shares.

```
[ ] data.info()
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 1935 entries, 2017-11-09 00:00:00+00:00 to 2023-02-26 00:00:00+00:00
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  ---      -
0   Open        1935 non-null   float64
1   High        1935 non-null   float64
2   Low         1935 non-null   float64
3   Close       1935 non-null   float64
4   Volume      1935 non-null   int64
5   Dividends   1935 non-null   float64
6   Stock Splits 1935 non-null   float64
dtypes: float64(6), int64(1)
memory usage: 120.9 KB

[ ] print(data.isnull().sum())
Open      0
High      0
Low       0
Close     0
Volume    0
Dividends 0
Stock Splits 0
dtype: int64
```

Fig 1: Description of Data

IV. METHODOLOGY

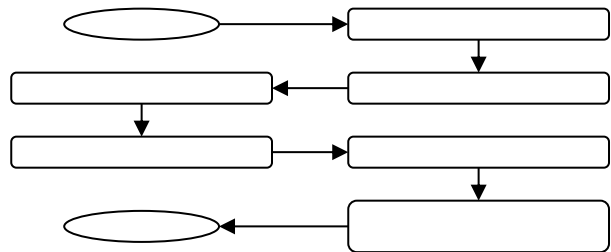


Fig 2: Implementation Process

The implementation process consisted of 6 major parts.

A. Fetching Live Data

Since the cryptocurrency data gets updated regularly, a dynamic data source was chosen to fetch data. So data was collected from the yahoo finance library.

B. Analyzing the Data

After collecting the data some basic exploratory data analysis is performed on the data. Here if any null values or missing values are found then they are discarded.

C. Training Models

In this step different models are created using different algorithms. PROPHET, ARIMA, LSTM and XGBoost algorithms are used to train our models.

D. Testing Models

After training the data, we need to check if the models are working properly or not. Test data is used to test the model predictions.

E. Evaluating Models

For this there are various methods and parameters like Accuracy, MAE, MSE, etc. We have plotted the

graphs of the actual price and the predicted price of our models.

F. Forecasting Future Prices

The prices for coming 10 to 15 days are forecasted using the trained models and are checked with their actual prices later.

V. ALGORITHMS

A. Prophet Model

Prophet is a Facebook ASCII text file tool for forecasting statistical knowledge that helps businesses perceive and possibly predict the market. There are two vital terms in Prophet Model:

Trend - It shows the tendency of the info to increase or decrease over a period of time and it filters out the differences due to the season.

Seasonality - Seasonality is the variations that occur over a short period of time.

The “Prophet Equation” fits, as mentioned above, trend, seasonality and errors. This is given by,

$$y(t) = g(t) * s(t) * e(t)$$

B. Auto-Regressive Integrated Moving Average (ARIMA)

ARIMA is specific by 3 order parameters: (p, d, q) AR(p) Autoregression - a regression model that makes use of the dependent connection between a present observation and observations from the past.

I(d) Integration - This involves the subtraction of the values of a series with its previous values d range of times.

MA(q) Moving Average - a model that applies the dependence between an observation and a residual error from a moving average model to lagged data.

C. Long Short Term Memory (LSTM)

LSTM networks are an extension of continuous neural networks (RNNs) that were designed to handle problems where RNNs failed. It is a network that works on the current input by considering the previous output (feedback) and storing it in its memory for a short period of time (short-term memory). The LSTM has been designed in such a way that the vanishing gradient disadvantage is nearly completely removed, whereas the coaching model is left unchanged. In bound issues, there are significant time lags. square measure bridged LSTMs are used wherever noise, distributed representations, and continuous values are also handled.

D. XGBoost

XGBoost is a C++ implementation of Gradient Boosted call trees. It is a package library designed to boost model speed and performance. This algorithmic rule generates call trees serially. Weights play an important role in XGBoost. XGBRegressor employs a number of gradient boosted trees to predict the value of a variable quantity. This is typically accomplished by combining call trees to form a combined strong learner.

VI. OUTPUTS AND RESULTS

In this case study we have trained four different models with past cryptocurrency data for future price prediction of various cryptocurrencies. The models are PROPHET MODEL, ARIMA, LSTM and XGBOOST. From our results, we can see that the PROPHET model outperforms all the other models by some margin.

TABLE I COMPARISON OF DIFFERENT MODELS

MODELS	ACCURACY
PROPHET	98.62316472574063
ARIMA	96.59130980478658
XG-BOOST	95.7221363205208
LSTM	96.53501113674446

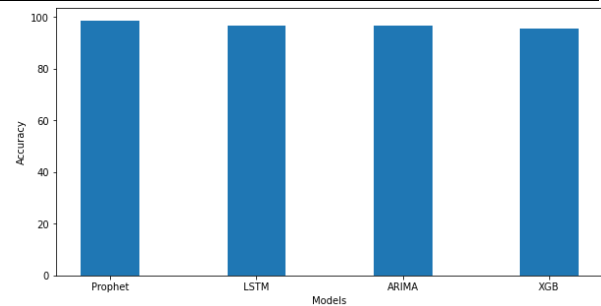


FIG 3: Accuracy Comparison Graph of Models

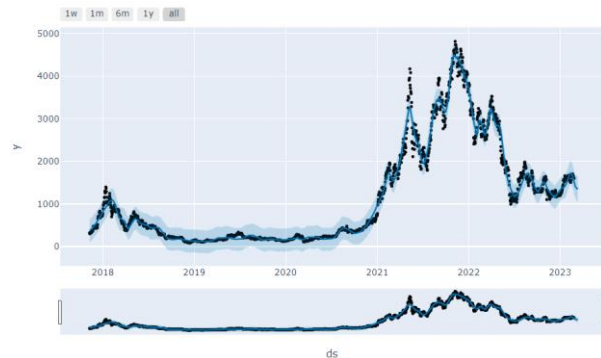


FIG 4: Prediction Graph of PROPHET

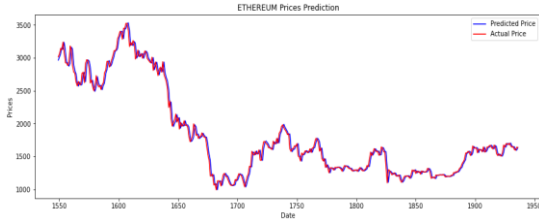


FIG 5: Prediction Graph of ARIMA

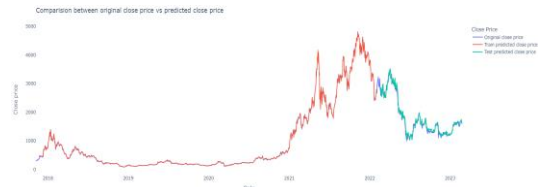


FIG 6: Prediction Graph of XGBOOST

VII. CONCLUSION

Cryptocurrencies are a form of digital asset that are built on a network that is dispersed across many computers. They are able to exist independently of governments and other central authority thanks to their decentralized nature.

We developed and compared various deep learning-based cryptocurrency price prediction models in this study using historical data for cryptocurrencies obtained from <https://finance.yahoo.com/>, which collected data from the start of cryptocurrency to till date. The study was based on the crypto currency's closing price and volume, which is the number of coins traded on that particular day. We specifically tested cutting-edge deep learning models like the Prophet Model, Auto Regressive Integrated Moving Average (ARIMA), Long Short-term Memory model, XGBoost, and their combinations. The ARIMA model and the Prophet model perform very similarly, with the PROPHET model having a slightly higher R-square score. ARIMA has an R-square score of 96.59%, while Prophet has a score of 98.62%.

In further advancements to this research, we got to know that only price trends don't affect the future prices of cryptocurrencies. So the value depends on various other social, environmental, political and other elements. Capturing the governing elements would improve prediction accuracy of the various coins. For instance, accessing information from Twitter's trending page would provide us insights into how the general public feels about the different coins, which would immediately affect demand for it and drive up its price globally. Analyzing the posts on the

Reddit site would be helpful in a similar way since it would enable us to go further and understand how the public and investors feel about the market.

To improve the accuracy of the prediction models, considering other factors such as stock market indices and major currency exchange rates would be beneficial.

By incorporating these insights into the training of models, efficiency and accuracy may be further improved and the investors risk can be mitigated to a great extent.

REFERENCES

- [1] Matthew Chen, Neha Narwal and Mila Schultz, (2017) "Predicting Price Changes in Ethereum", Stanford University, Stanford, CA 94305
- [2] Ji, Suhwan & Kim, Jongmin & Im, Hyeonseung, (2019) "A Comparative Study of Bitcoin Price Prediction Using Deep Learning.", Mathematics. 7. 898. 10.3390/math7100898.
- [3] Mohammed Khalid Salman and Abdullahi Abdu Ibrahim, (2020) "Price Prediction Of Different Cryptocurrencies Using Technical Trade Indicators And Machine Learning" IOP Conf. Ser.:Mater. Sci. Eng. 928 032007
- [4] Samiksha Marne, Delisa Correia, et al, (2020) "Predicting Price Of Cryptocurrency - A Deep Learning Approach", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181
- [5] Muniye, Temesgen & Rout, Minakhi & Mohanty, et al, (2020) "Bitcoin Price Prediction and Analysis Using Deep Learning Models.", Springer 10.1007/978-981-15-5397-4_63.
- [6] Lekkala Sreekanth Reddy, Dr. P. Sriramya, (2020) "A Research On Bitcoin Price Prediction Using Machine Learning Algorithms", International Journal Of Scientific & Technology Research Volume 9, Issue 04, April 2020 ISSN: 22778616
- [7] I. Yenidoğan, A. Çayır, O. Kozan, T. Dağ and Ç. Arslan, (2018) "Bitcoin Forecasting Using ARIMA and PROPHET," 3rd International Conference on Computer Science and Engineering (UBMK), 2018, pp. 621-624, doi: 10.1109/UBMK.2018.8566476.
- [8] Gowtham Saini, Dr. M. Shobana, (2022) "Cryptocurrency Price Prediction Using Prophet And Arima Time Series", International Research Journal of Modernization in Engineering Technology and Science e-ISSN: 2582-5208

- [9] Yash Wadalkar, Yellamraju V H Sai Tarun, et al, (2021) “A Comparative Analysis Based Approach For Bitcoin Price Forecasting”, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181