Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2022)
IEEE Xplore Part Number: CFP22AV8-ART; ISBN: 978-1-6654-8425-1

A Novel Approach for Analytical Comparison of Cryptocurrencies using Time Series

Syed Zishan Ali,Pooja Bhargava,Ranjana Soni,Shivanshi Tiwari

2022 International Conference on Electronics and Renewable Systems (ICEARS) | 978-1-6654-8425-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICEARS53579.2022.9752071

Abstract— **Cryptocurrencies gaining popularity as a digital currency in today's world. The market is altering dramatically as a result of the rapid increase in investment in digital currency. The ability to accurately forecast and change the behavior of digital currency is becoming a critical aspect in today's digital environment. This prompted us to investigate bitcoin trading and investor adoption in the market. We began by examining and analyzing the costs of currencies, and after reviewing numerous methods, one of them proved to be the ideal technique for our data representation, taking into account various criteria such as date, time, and year. A time series forecasting technique has been implemented for forecast the future profit or loss in currencies.**

IndexTerms—**Cryptocurrency, Forecasting, Xgboost, Algorithm, Prophet Function, PolkaDot, Price.**

I. INTRODUCTION

There are several People in the twenty-first century are always coming up with new ideas to improve their sources of income, and after years of study and development, a new form of cash known as virtual currency has emerged, the reason for the increase in sources of revenue is due to substantial increases in the prices of goods as well as consumer desires. These factors have put peopleunder initiatives to raise their income because human beings are always in desire to make their lives more pleasant.

Virtual currency is a type of digital asset that is dispersed across a large number of computers via a network. They are able to exist beyond the control of governments and central authorities because of their decentralized structure. Users can exchange value digitally without the involvement of a third party when they use a cryptocurrency, thanks to the concept of virtual money.

We explored one of the most recent cryptocurrencies, Polkadot, in our work, which is aimed to combine private and cooperative chains, public and anonymous networks, oracles, and other technology. This supports an internet where various blockchains can communicate information and transactions in a trust less manner through the Polka Dot relay chain. This currency solves these issues with Para chains. Rather than deploying all of the apps on the same network, each one has its own tiny blockchain (or parachain) that connects to the main chain. DOT was first traded at $2.76 per coin when it was released in August 2020. Market cap valued it at $20.21 at the moment of writing, representing a 630 percent gain in less than a year. It is being built by the Web3 Foundation in collaboration with best-in-class organizations, who will also help to design the services and applications that will operate on it. Researchers from Inria Paris and ETH Zurich, as well as developers from Parity Technologies and capital partners from crypto-funds like Poly chain Capital, are collaborating to create a superior Web3 implementation with PolkaDot at its core. The Web3 Foundation offers grants to help the ecosystem grow.[4]

We analyzed and examined the data of PolkaDot prices in past years in this article, and based on that, a prediction graph of future years at which rate mining will be halved had been shown. Also, a comparison has been done with two major cryptocurrencies that are Bitcoin and dogecoin.

1. XGBoost Algorithm

The beauty of this efficient algorithm is its scalability, which allows for rapid learning via parallel and distributed computing while still managing memory efficiently. It's really no surprise that CERN chose it as the most effective method for classifying signals from the Large Hadron Collider. CERN's challenge demanded a system that could process 3 petabytes per year and effectively identify an exceptionally rare signal from background disturbances in a complex physical process. XGBoost proved to be the most practical, easy, and reliable approach.

XGBoost is a method of supervised learning. It may not always be enough to rely on the outcomes of a single machine learning model. Ensemble learning is a method for combining the predictive abilities of numerous learners in a systematic way. The end result is a single model that integrates the outputs of numerous models. The foundation learners, or models that make up the ensemble, could be from the same learning algorithm or from distinct learning algorithms. **Bagging** and **Boosting** are two types of ensemble learners that are frequently practiced. Though these two strategies can be applied to a variety of statistical models, decision trees have been the most popular.

1.2Bagging

Although decision trees are one of the simplest models to grasp, their functionality is highly varied. Develop a unique training dataset that we divided into two sections at random. Let's now use each component to train a decision tree and get two models. Both of these models would produce different results if we fitted them together. Due of this propensity, decision trees are believed to be associated with large variance. Bagging or boosting aggregation can assist any learner to reduce variation. The foundation learners of the bagging technique are several decision trees that are produced concurrently. These learners are taught using data that has been sampled using replacement. The mean output from all of the learners is used to make the final forecast.

1.3 Boosting

The trees are built successively in boosting, with each consecutive tree aiming to reduce the previous tree's mistakes. Each tree builds on the knowledge of its ancestors and corrects any lingering faults. As a result, the following tree in the sequence will learn from an updated set of residuals.

In boosting, the base learners are weak learners with a massive bias and predictive power that is just slightly better than random guessing. Each of these weak learners gives some crucial information for prediction, allowing the boosting strategy to efficiently combine these weak learners to build a strong learner. Both the bias and the variance are reduced by the final strong learner.

Boosting uses trees with fewer splits than bagging techniques like Random Forest, which uses trees that have been developed to their full potential. The interpretation of such little trees, which are not very deep, is extremely easy. Through validation approaches like k-fold cross-validation, parameters like the number of trees or iterations, the pace at which the gradient boosting learns, and the depth of the tree can be appropriately adjusted. The presence of a significant number of trees may result in overfitting. As a result, the boosting stopping conditions must be deliberately crafted.

Three simple steps make up the boosting ensemble technique:

- To predict the target variable y, an initial model F0 is defined. A surplus (y – F0) will then be linked with this model.
- The residuals from the previous phase are fitted to a new model h1.
- F0 and h1 are now merged to form F1, which is a boosted version of F0. The F1 mean square error will be smaller than the

- F0 mean squared error:

1. $F_1(x) <- F_0(x) + h_1(x)$

To optimize the productivity of F1, we might develop a new model F2 based on F1's residuals:

2. $F_2(x) <- F_1(x) + h_2(x)$

This can be repeated for 'm' iterations till the residuals are as low as possible:

3. $F_m(x) <- F_{m-1}(x) + h_m(x)$

Concluding the explanation of XGBoost here are some of its unique features:

- Regularization
- Working with sparse data
- Sketch of a weighted quantile
- Parallel learning with a block structure
- Aware of the cache
- Out-of-core computing

Coming at the end note is that's everything there is to grasp about the principles behind the popular XGBoost algorithm. If you have a firm grasp on the fundamentals, this article should have been a breeze for you. XGBoost is such a powerful algorithm that, while it has generated alternative techniques (such as CATBoost), it remains a game-changer in the machine learning community.

2. Facebook Prophet Function

Facebook Prophet is indeed an open-source time-series model generation algorithm that combines some classic principles with some novel twists. It excels at modeling time series with numerous seasonality and avoids some of the limitations associated with other techniques. Growthg(t), seasonality s(t), holidays h(t), and error e t are the sum of three-time functions plus an error term:

1. $y(t)=g(t)+s(t)+h(t)+€_t$

Talking about its growth function I. e., change points the growth function simulates the data's general trend. Anyone with a basic understanding of linear and logistic functions should be familiar with the old concepts. The new concept introduced into Facebook Prophet is that the growing trend can exist at any time in the data or can be changed at "changepoints," as Prophet refers to them. There are three primary alternatives for the growth function:

- Linear Growth
- Logistic Growth
- Flat

Prophet offers us simulations to choose from (however, newer models can be written or extended according to specific requirements). The logarithmic growth model is one, and the small bit linear model is the other. The snippet linear model is used by default in Prophet, however it can be altered by specifying the model. Choosing a paradigm is difficult since it is influenced by a number of aspects such as the size of the firm, its growth pace, and its business strategy. The logistic growth model is the ideal alternative if the data to be projected is saturating and non-linear (grows non-linearly and after reaching the theoretical saturation, shows little to no growth or decrease and only exhibits small seasonal fluctuations).Only data in a specific format is anticipated by Facebook Prophet. Columns should be recorded as ds for time series analysis and y for predicted data inside this data matrix with the data. The time series is represented by the column

Month, and the data to be projected is represented by the column.As a result, we've seen how, with only a few lines of code, we can create a prediction model that would have been extremely difficult to develop using typical machine learning techniques and probability and statistics notions alone.

II. METHODOLOGY USED

A. *Data*

We Obtain the daily prices of Polkadot(DOT1-USD), Bitcoin (BTC), Dogecoin (DOGE) from a website named finance.yahoo.com. We integrated several data frequencies to create our final dataset, which contains daily prices from September 2018 through November 2021.

B. *Methodology*

The methodology of this study is based on taking a dataset of cryptocurrencies from the previous two years and creating graphs and charts of the currency Polkadot based on the values of closing and opening prices. Then, using the algorithm XGBoost, price predictions for the next two years have been made using the prophet algorithm.
XGBoost is a gradient boosting-based decision-tree-based ensemble Machine Technique for learning. Artificial neural networks outperform all other types of networks. algorithms or frameworks in prediction issues involving unstructured data (pictures, data, etc.).However, for small to medium structured/tabular data, decision tree-based techniques now are considered the best.

Why did we utilize this algorithm in our project, and why did it work so well, is a question that arises. XGBoost and Gradient Boosting Machines (GBMs) are both ensemble tree approaches that use the gradient descent architecture to boost weak learners (CARTs in broad). XGBoost, on the other hand, enhances the fundamental GBM framework through system optimization and algorithmic improvements.[5] So, using this method, below is a graph showing the opening and closing prices month by month for the years 2020 and 2021, based on our data set.

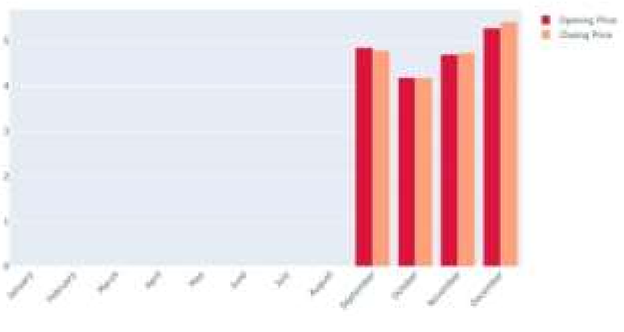

Figure 1. Graph for 2020

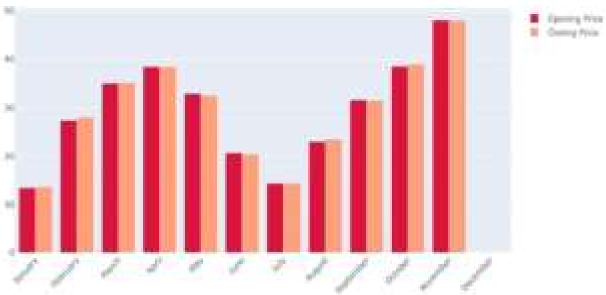

Figure 2. Graph for 2021

After visualizing the price data, we began predicting the price of the most recent digital currencies, Polkadot, and then went on to estimate the values of currencies that were a little older than PolkaDot.
Our forecasts are based on the prophet function, which is defined as a data forecasting process based on an additive model that fits non-linear seasonality on an annually, weekly, and constant basis, as well as holiday impacts. It works best with time series with substantial seasonal influences and historical data from multiple seasons. The prophet is forgiving of missing data and trend shifts, and it usually handles outliers well. The great thing about this algorithm is that it is quite adaptable when it comes to the data it is fed. You don't have to have all the dates and times planned for NAs.
It also works reasonably well without any parameters being explicitly set. You may also adjust some of the parameters to improve the model more if you have domain knowledge, but those settings are very simple to grasp.The Prophet algorithm is an additive model, which means it first analyzes the subsequent trend and seasonality in the statistics before combining them to provide the anticipated values. Seasonality Yearly, Weekly, and Daily Trends Effect of the Breaks. In a nutshell, we can acquire a lot of helpful information from the model by looking at the trend and seasonality that the Prophet recognizes.

Now, using this algorithm, we simply generate our first PolkaDot currency forecast. Below is the graph showing predictions of further two years.

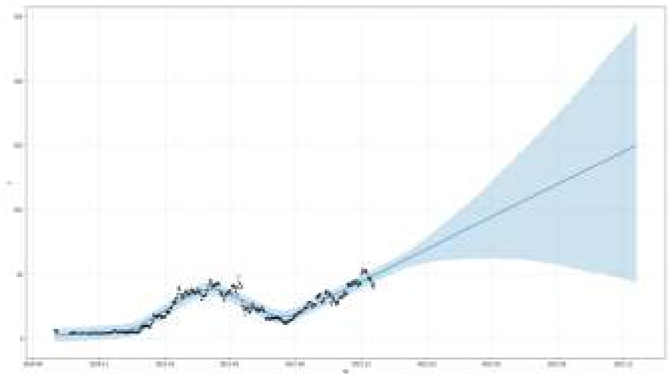

Figure 3. DOT- PolkaDot

According to this prediction graph above the graph depicts a range of low and high between which the values lie.

Another graph shows the same for cryptocurrency Bitcoin prediction for one year and also shows the range betwe

en which the price lies and according to the graph we can say that for Bitcoin there is slow increase in the price in every 4-6 months.

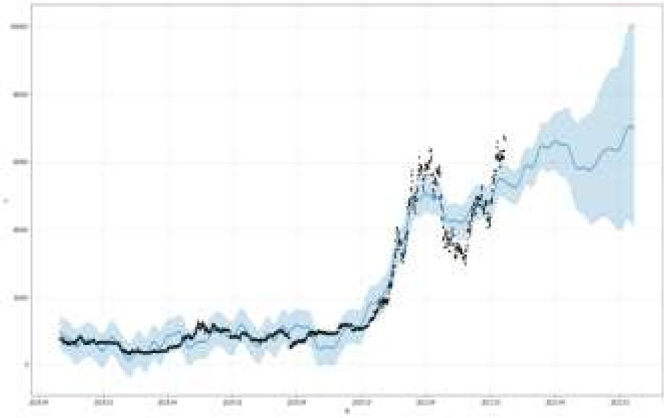

Figure 4. Bitcoin

And then at last we have predicted the values for our last cryptocurrency named Dogecoin where according to the graph we can say that there is rapid increase within a year in the price of dogecoin.These studies of prediction are made by analyzing the historical data that is inserted in our project.

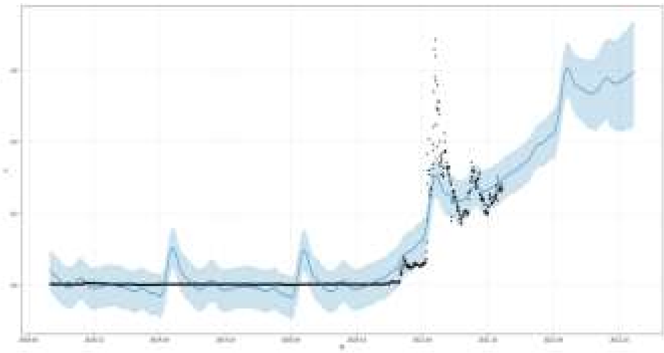

Figure 5. Dogecoin

Coming at the outcome of how values differ in with compared to the values compared here is a table designed which shows the real values of the year 2021:

TABLE: 1 Real Values of Cryptocurrency 2021

S.NO	CRYPTOCURRENCY	YEAR	MONTH	OPEN	HIGH	LOW
1	POLKADOT	2021	APRIL	35.98 USD	36.80USD	35.14 USD
			AUGUST	25.73 USD	27.56 USD	24.42 USD
			DECEMBER	25.97 USD	26.21 USD	25.17 USD
2	DOGECOIN	2021	APRIL	0.3047 USD	0.3397 USD	0.3029 USD
			AUGUST	0.2812 USD	0.2895 USD	0.2703 USD
			DECEMBER	0.1730USD	0.1765 USD	0.1715USD
3	BITCOIN	2021	APRIL	53.568USD	57.900USD	53.129 USD
			AUGUST	48.834 USD	48.925USD	46.950 USD
			DECEMBER	47.561 USD	47.999 USD	46.906 USD

And here's the graph that demonstrates the values that the graphs in this project displays, as well as the price fluctuations.

TABLE 2: Estimated Price by Project

S. NO	CRYPTO CURRENCY	YEAR	MONTH	OPEN	HIGH	LOW
1	POLKADOT	2021	APRIL	34.4556 USD	39.4078 USD	29.3226USD
			AUGUST	26.62323USD	31.7587 USD	21.9886USD
			DECEMBER	26.7779USD	30.1839USD	31.7782USD
2	DOGECOIN	2021	APRIL	0.28163USD	0.33321USD	0.318356USD
			AUGUST	0.274669USD	0.331778USD	0.220269USD
			DECEMBER	0.269964USD	0.216544USD	0.165196USD
3	BITCOIN	2021	APRIL	49697.6USD	54355.52USD	53121.86USD
			AUGUST	48368.18USD	53349.73USD	48764.3USD
			DECEMBER	43172.16USD	52243.84USD	48123.43USD

III. CONCLUSION

The aim of this paper is to present the major algorithms that will successfully forecast the prices of cryptocurrencies by examining and analyzing previous data sets and assisting us in understanding and concluding the rate at which the price will increase or decrease, as well as displaying us the range between which the price may come.

REFERENCES

[1] .LekkalaSreekanth Reddy, Dr.P. Sri Ramya Mentioned in (IJSTR) Volume 9," A Research on Bitcoin price prediction using machine learning algorithms", April 2020

[2] Dr.M.Sharmila Begum, G. Jayashree, Z. MaheboobAafia, R. Subhasri and L. Vishnu Priya Mentioned in (TJCME), "Predicting the price of Bitcoin using data analytics", April 2021.

[3] S M Raju, Ali Mohammad Tarif, "Real - Time Prediction of Bitcoin Price Using Machine Learning Techniques and Public Sentiments Analysis", in (IIUM), June (2020).

[4] https://www.investopedia.com/terms/c/cryptocurrency.asp

[5] https://towardsdatascience.com/https-medium-com-vishalmorde-xgboost-algorithm-long-she-may-rein-edd9f99be63d

[6] Mahboubeh Faghih Mohammadi Jalali and Hanif Heidari, "Predicting Changes in Bitcoin Price using gray system theorem" in (FMJHFI), 2020.

[7] Mrs. Vaidehi M, Alivia Pandit, Bhaskar Jindal, Minu Kumari, Rupali Singh," Bitcoin Price Prediction Using Machine Learning", (IJETMR) Volume 8, 2021.