iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** II **Month of publication:** February 2022

DOI: https://doi.org/10.22214/ijraset.2022.40521

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

Crypto Currency Market Price Prediction Using Data Science Process

Karthik Vikram[1], Nikhil Sivaraman[2], Dr. P. Balamurugan[3]
[1, 2, 3]*Computer Science Engineering, SRM Institute of Science and Technology, Chennai, India*

Abstract: ***Cryptocurrency is commonly called as digital currency where the coin ownership records are stored in an electronic ledger existing in a form of a computerized database using strong cryptography approach which is used to maintain the creation and updation of an addition coin in a market, it is also used to check and maintain the current ownerships of the coins. Nowadays cryptocurrency is used in largescale and there is a sudden rise or decrease in their share and it is difficult to predict the price of the crypto currency. In this project a machine learning model is built to predict the price of crypto currency. The application of data science process is applied for getting the better model for predicting the outcome. Variable identification and data understanding is the main process in building the successful model. Different machine learning algorithms are applied on the pre-processed data and the accuracy are compared to see which algorithm performed better other performance metrics like precision, recall, scores are also taken in consideration for evaluating the model. The machine learning model is used to predict the crypto currency outcome***

Keywords: crypto currency, Decision Tree, Prediction, Random Forest, Logistic regression, Linear Regression, LASSO Regression, Ridge Regression.

I. INTRODUCTION

A cryptocurrency is a type digital currency which can be used a mode of online payment or make an individual to sell/purchase cryptocurrency to utilizing various monetary forms. The Cryptocurrency transaction are stored and maintained in an electronic ledger and even supports blockchain transactional databases. Bitcoin is the earliest invented cryptocurrency and has been in trading since 2009. There are around 7577 different Cryptocurrencies available in the market. We can use straightforward utilization methods like Machine learning or Data science to predict its values. Machine learning has a huge amount of regression algorithms. Which can make our task easier. Data science techniques like Random Forest regression, Ridge regression, LASSO regression, Linear regression can be used to predict the future price of a Cryptocurrency and predict future outcomes. In this paper we are going to use 6 different algorithms like Random Forest regression, Ridge regression, LASSO regression, Linear regression, Logistic regression and compare each algorithms accuracy and root mean square value. In this paper we are also going to use 4 different and popular Cryptocurrency available in the current marker like Bitcoin, Ethereum, Cardano (ADA), Binance Coin (BNB). Find which algorithm has the best possible accuracy and implement it in our Machine learning model. Finally, this model can be used to predict the market price of each cryptocurrency. This paper can be useful for Several traders who are scared about investing in cryptocurrency because of the fear of frequent drop in price. But with data science and constant monitoring, it will be easier to identify the trend in stock price. It also becomes easier to find the right time to invest and to draw the money.

II. LITERATURE SURVEY

They presented a computational approach for identifying and characterizing cryptocurrency pump and dump operations that are carried out in social media. They had used financial and Twitter data pertaining to a particular coin, the method used was able to detect, with reasonable accuracy, whether there is an unfolding attack on that coin on Telegram, Twitter, and whether or not the resulting pump operation will succeed in terms of meeting the anticipated price targets. They also analyzed activities of users involved in pump operations, and observe a prevalence of Twitter bots in cryptocurrency-related tweets in close proximity to the attack. To analyze such activities. “Bitcoin price prediction using machine learning” [1]. In this paper they used algorithms like RNN, ARIMA, SVM, Logistic Regression all the algorithms performed equally good but had a very low accuracy of around 50 %. **“**A Research on Bitcoin Price Prediction Using Machine Learning Algorithms” [2]. In this paper they used an algorithm linked to artificial intelligence named LASSO (least absolute shrinkage selection operator. Their main aim was to reduce the time and efficiency in calculating a Crypto currency market trend. The other papers used different algorithms like SVM, Linear regression, Logistic regression, RNN, CNN, ANN, ARIMA. Which did not have a great time management. So, our main project goal was to increase the algorithms accuracy and have a better time management.

©IJRASET: All Rights are Reserved

III. PREDICTION TECHNIQUES

1) Logistic regression is mainly used in classification problems. Therefore, the outcome will not be correct for our problem. It mainly gives output based on probability either the value will decrease or increase, win or lose,1 or 0. So it is not possible to use this algorithm and predict the market price. It is generally used to predict if the market price will increase or decrease.

2) Linear regression is a regression-based machine learning algorithm which uses supervised data learning which means that it makes use of ladled data set. While predicting a dependant variable (y) based on a given independent variable (x), we use Linear Regression. This technique is also used to find out the linear relationship between x(input) and y(output).

3) Decision tree is also a supervised machine learning algorithm which is used for either classification or regression-based problems. It is a like tree-based structure, where the root nodes represent the entire sample, the Interior Nodes represent the different features of the data set like name, symbol, date, market price, volume, etc... and finally the leaf node is used to represent the outcome of the problem.

4) Lasso regression is like linear regression but the main difference is that in LASSO regression it uses a simple method called “shrinkage” where the coefficient is shrunk towards zero. This allows us to regularize the coefficients to avoid overfitting problem and make them work better for our datasets

5) Ridge regression is a machine learning algorithm that is used in analysing a data when it suffers multicollinearity and it also makes use of coefficient shrinkage. It's a regularization technique that reduces the complexity of the model. It is popularly used for models that have multicollinearity. This technique uses regularization to increase the "Bias" and reduce "Variance" thereby improving the overall fit, and also adds a penalty scaling with the complexity.

6) Random Forest Regression is basically a number of classifying decision tree on various samples of the dataset. It is similar to concept of averaging; it is used to improve the predictive accuracy of the Decision tree and prevent the model from overfitting

IV. DATASET DESCRIPTION

For this project we took the cryptocurrencies price from a popularly known dataset website called Kaggle. The dataset (.CSV) is of size 100KB and contains historical prices of 4 different cryptocurrencies namely Ethereum, Bitcoin, Cardano and Binance-coin. The data consists of 1079 rows and Thirteen columns out which 1-144 is for Ethereum ,145-532 is for Bitcoin,533-773 is for Cardano and 774-1079 is for Binance-coin. Each column has a parameter of the data such as slug, symbol, name, date, rank, open, high, low, close, volume, market price, close ratio and spread. We are choosing this dataset for analyzing and predicting the market price of these 4 different cryptocurrencies. A brief description of the dataset is given in Table.2

Table.1. Sample dataset

	slug	symbol	name	date	ranknow	open	high	low	close	volume	market_price	close_ratio	spread
0	ethereum	ETH	Ethereum	8/7/2015	3	2.830000	3.540000	2.520000	2.770000	164329	166610555	0.2451	1.02
1	ethereum	ETH	Ethereum	8/8/2015	3	2.790000	2.800000	0.714725	0.753325	674188	45486894	0.0185	2.09
2	ethereum	ETH	Ethereum	8/9/2015	3	0.706136	0.879810	0.629191	0.701897	532170	42399573	0.2901	0.25
3	ethereum	ETH	Ethereum	8/10/2015	3	0.713989	0.729854	0.636546	0.708448	405283	42818364	0.7706	0.09
4	ethereum	ETH	Ethereum	8/11/2015	3	0.708087	1.130000	0.663235	1.070000	1463100	64569288	0.8715	0.47

Table.2. Dataset description

Parameter	Description
slug	Name of the coin
symbol	Type of the coin
name	Name of the coin
date	Date at which data is taken
rank now	Current rank of the coin in market
open	Opening price of the day
high	Highest price of the day
low	Lowest price of the day
close	Closing price of the day
volume	Total Volume of the currency
Market price	Market capital of the currency
Close ratio	Close ratio = (Closed deals) / (Total sales leads) x 100
Spread	Difference between high and low value

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

V. PROPOSED METHODOLOGY

A. *Block Diagram*

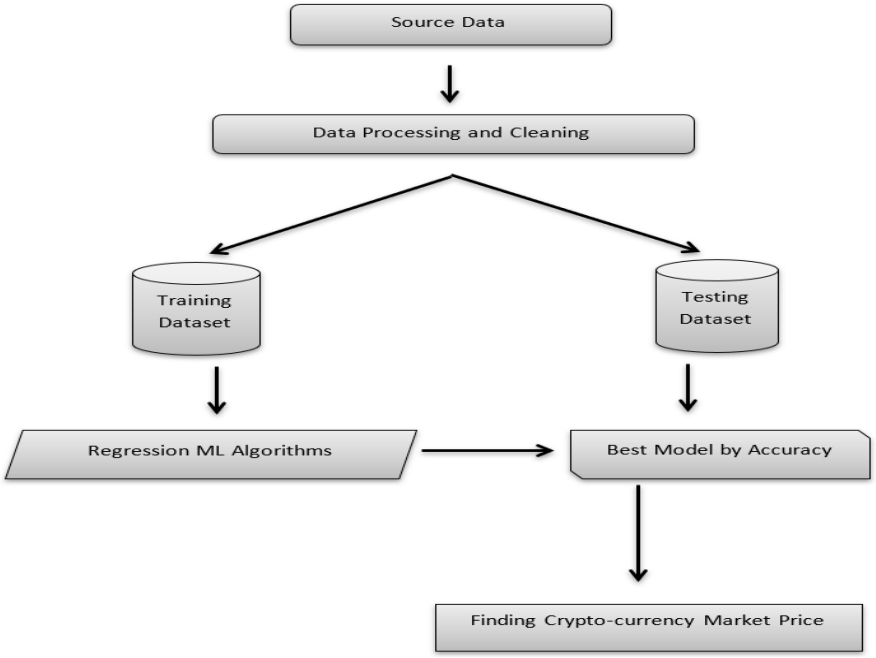


Fig.1.Cryptocurrency price prediction

B. *List of Modules*

1) Data Pre-processing
2) Data Analysis of Visualization
3) Comparing the Algorithm
4) Deployment Using Flask

VI. RESULT AND DISCUSSION

We considered previous cryptocurrencies transaction in which market price and timestamps are the attributes used to predict the cryptocurrencies price for future. We used six methods for price predictions such as Logistic regression, Linear Regression Prediction, Decision Tree, Lasso Regression and Ridge regression and calculated the accuracies for these six methods. Then we have to compare the algorithm which has the best accuracy and implement in the machine learning model. For our project Random Forest regression had best possible result of 97.49 % and this model was implemented in our flask framework. On an average the linear regression algorithm had the highest possible mean square error rate of 9.014 and it even had the lowest accuracy of around 60.85%. Decision tree regression and the random forest regression performed equally good both of them had an accuracy of 97 % and their mean square error were quite low up to some extent. But Random Forest regression was the best possible fit for the given data set.

Table.3. Accuracy description

Model Test	**Logistic Regression**	Linear Regression	Decision Tree Regression	Random Forest Regression	Lasso Regression	Ridge Regression
Accuracy	-1.525	60.85%	97.31%	97.49%	60.86%	70.73%

Table.4. Mean square error description

Model Test	**Logistic Regression**	Linear Regression	Decision Tree Regression	Random Forest Regression	Lasso Regression	Ridge Regression
Mean square error	4.821	9.014	6.860	6.459	8.717	5.057

©IJRASET: All Rights are Reserved

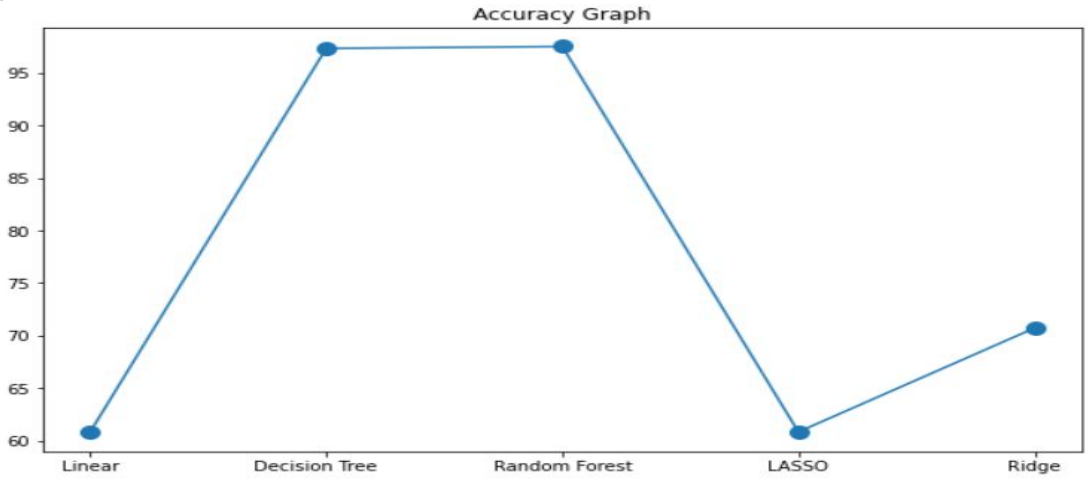


Fig.2.Visualization of accuracy

VII. CONCLUSION AND FUTURE WORK

The systematic process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test data is taken into account and higher accuracy score for the given test data is implemented in the machine learning model. This application can help to find the Cryptocurrency Market Price. This project will improve that future idea of crypto currencies and it can even improve the market price of the cryptocurrency. In this paper we proposed a machine learning algorithm to find the market price of the given Cryptocurrency and even calculated the accuracy of various different Machine learning model. This application will help the users get more profit and improve their trading ability.

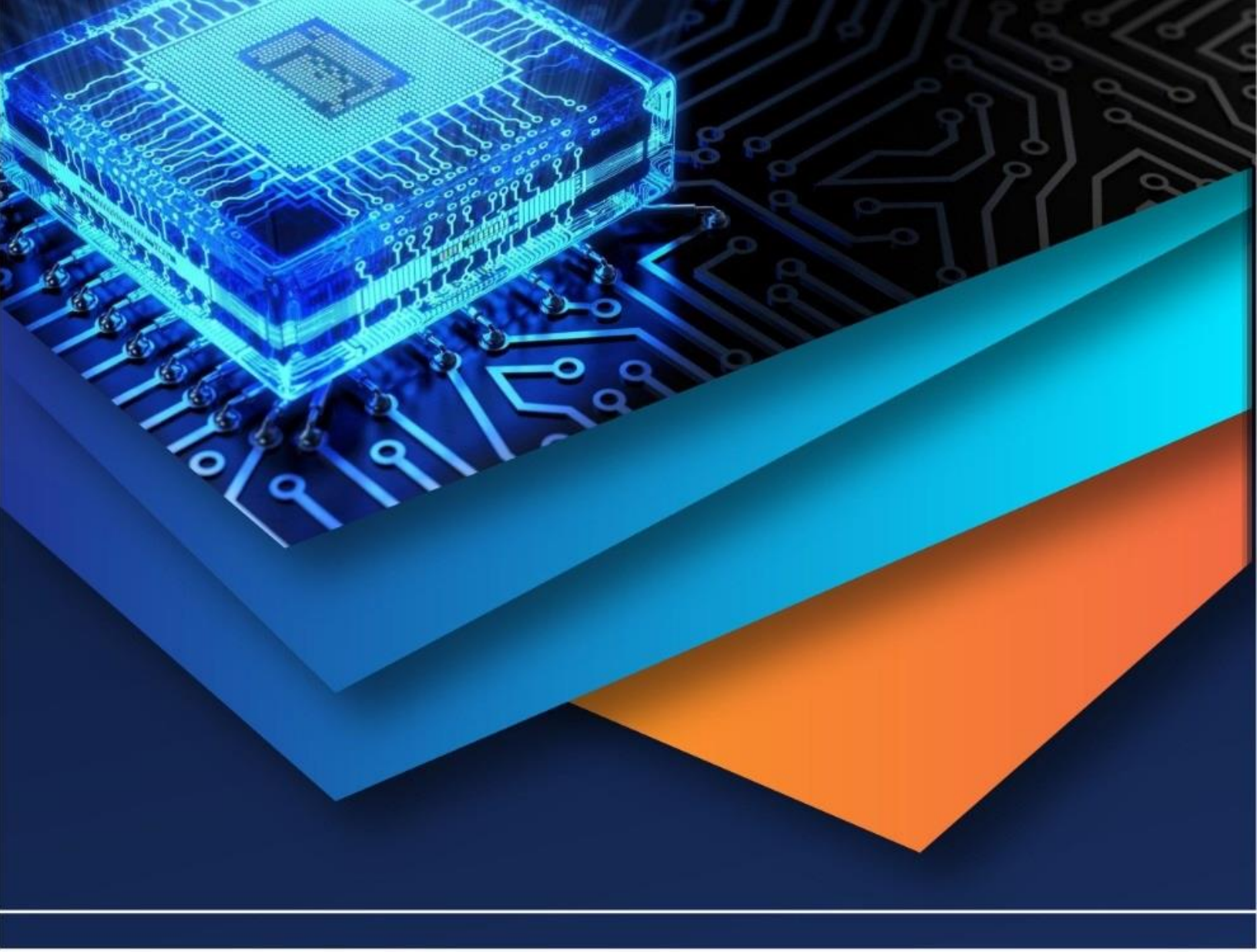
A. Future Work

1) Show the statistical trend in cryptocurrency for 1 week.
2) Automate this process and show the prediction result in android application or IOS application.
3) Even improve the result using Artificial Intelligence environment.
4) Add more coins and predict its market value.

REFERENCES

[1] Neha Mangla,Akshay Bhat, Ganesh Avabratha, Narayana Bhat ”bitcoin price prediction using machine learning ” INTERNATIONAL JOURNAL OF INFORMATION AND COMPUTING SCIENCE, Volume 6, Issue 5, May 2019
[2] Lekkala Sreekanth Reddy, Dr.P. Sriramya ”A Research On Bitcoin Price Prediction Using Machine Learning Algorithms” International Journal of Scientific & Technology Research, Volume 9, Issue 04, April 2020
[3] RuchiMittal,ShefaliArora andM.P.SBhati ”AUTOMATED CRYPTOCURRENCIES PRICES PREDICTION USING MACHINE LEARNING” International Journal of Soft Computing Volume: 08, Issue: 04 ,July 2018
[4] Alireza Ashayer “MODELING AND PREDICTION OF CRYPTOCURRENCY PRICES USING MACHINE LEARNING TECHNIQUES” East Carolina University (ECU),May 2019
[5] Franco Valencia, Alfonso Gómez-Espinosa and Benjamín Valdés-Aguirre “Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning” Entropy 2019, 21, 589, June 2019
[6] Python Data Analysis: Perform Data Collection, Data Processing, Wrangling, Visualization,and Model Building Using Python Book by Armando Fandango, Avinash Navlani, and Ivan Idris
[7] Python for data analysis by Wes McKinney ,published by O'Reilly Media, 2nd Edition, Sep 2017
[8] Laura Alessandretti,Abeer ElBahrawy,Luca Maria Aielloand Andrea Baronchelli **“**Anticipating Cryptocurrency Prices Using Machine Learning” Volume 2018, 4 Nov 2018
[9] Steve Hyun, Jimin Lee, Jong-Min Kim and Chulhee Jun "What Coins Lead in the Cryptocurrency Market: Using Copula and Neural Networks Models" Volume 3 , 8 August 2019
[10] Jiang Z, Liang J "Cryptocurrency portfolio management with deep reinforcement learning" Volume 7 11 May 2017
[11] Kim YB, Kim JG, Kim W, Im JH, Kim TH, Kang SJ, Kim CH "Predicting Fluctuations in Cryptocurrency Transactions Based on User Comments and Replies" Volume 2; August 17, 2016
[12] Li X, Wang CA "The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin" Volume 3; Dec'2018
[13] S. Yogeshwaran ; PiyushMaheshwari; ManinderJeet Kaur “Project Based Learning: Predicting Bitcoin Prices using Deep Learning‖ Amity University Dubai Dubai” ,UAE; IEEE 2019

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)