Cryptocurrency Price Prediction Using Machine Learning

Dhruva Parag Kawli
Department of Computer Engineering
St. John College of Engineering and Management,
Palghar
dhruvakawli123@gmail.com

Aditi Sunil Chaudhari
Department of Computer Engineering
St. John College of Engineering and Management,
Palghar
aditi1d2801@gmail.com

Prajakta Dnyaneshwar Ingale
Department of Computer Engineering
St. John College of Engineering and Management,
Palghar
prajaktaingale641@gmail.com

Gaurav Arun Telange
Department of Computer Engineering
St. John College of Engineering and Management,
Palghar
gauravtelange786@gmail.com

Abira Banik
Department of AI/ML
St. John College of Engineering and Management,
Palghar
abirab@sjcem.edu.in

***Abstract*—The application of machine learning algorithms in predicting cryptocurrency prices has gained significant attention in recent years. Researchers have explored various approaches such as recurrent neural networks, deep learning neural networks, Bayesian regression, k-nearest neighbor, support vector machine, and other algorithms to forecast the prices of cryptocurrencies like Bitcoin, Ethereum, Dogecoin and Litecoin. This paper will draw on established literature on price prediction using machine learning, including studies on NFT sales predictability, NFT sale price fluctuations prediction, gold price prediction, and silver price forecasting. The research paper has focused on utilizing high-dimensional features, time-series analysis, as well as the comparison of different statistical models and machine learning algorithms. Additionally, the prediction models have incorporated factors such as market liquidity, exchange market dynamics. While the literature acknowledges the potential of machine learning in cryptocurrency price prediction, gold, silver and NFT's there is a recognized gap in the application of these techniques across a broader range of cryptocurrencies. The proposed methodology will integrate various machine learning models and statistical methods to predict the prices of cryptocurrencies, gold, silver, and NFTs, taking into account factors such as market trends, trade networks and visual features. Furthermore, the studies emphasize the importance of feature engineering, sample dimension engineering, and the use of various machine learning techniques to enhance the accuracy and stability of cryptocurrency price predictions. As the cryptocurrency market continues to expand, there is a need for further research to develop robust machine learning models that can effectively forecast the prices of diverse cryptocurrencies, contributing to the advancement of this field.**

Keywords: Price Prediction, Random Forest, Long Short-Term Memory (LSTM)

I. INTRODUCTION

The use of machine learning for price prediction in various asset classes, including cryptocurrencies, gold, silver, and non-fungible tokens, has gained significant attention in recent years. The challenges in real-time price prediction for cryptocurrencies due to their deterministic nature emphasized the use of machine learning algorithms to predict and forecast cryptocurrency prices, aiming to facilitate trading activities acknowledged the difficulty in predicting cryptocurrency prices due to their high volatility.

Cryptocurrency, gold, silver and Non-Fungible Token (NFT) are distinct assets which having unique features [15]. Gold is often seen as a hedge against economic uncertainty, while silver serves both investment and industrial purposes. Non- Fungible Tokens (NFTs) are indivisible and have gained popularity in the digital world for proving ownership. Cryptocurrency was created in 2009, that was Bitcoin and since then, thousands of another cryptocurrency was created such as Ethereum, Litecoin, Dogecoin, etc. cryptocurrency are relatively unpredictable compared to traditional financial instruments. This technology is indeed having challenges to overcome [26]. There have been several studies on using machine learning algorithms to predict prices including, Bitcoin, Litecoin, Ethereum, Gold and Silver. This Studies have various data sources to make predictions about future prices.

Machine learning algorithms analyze historical price data, market trends, and various indicators to make predictions about the future value of these digital assets. The later section gives more knowledge on how to predict the price of these Crypto currencies with Machine Learning [26].

II. LITERATURE SURVEY

This paper focuses on the limited exploration of alternative machine learning algorithms beyond LSTM and ARMA for Bitcoin price prediction. While the paper introduces random forest regression as a potential method with superior prediction errors, there is a need for further investigation into other innovative algorithms and their effectiveness in enhancing prediction accuracy [1].

The specific machine learning algorithms employed for Bitcoin price prediction and the absence of detailed exploration into the nature of biases directly affecting the model's predictions. Additionally, there is a need for further elucidation on how domain expert feedback has specifically contributed to refining the proposed Bitcoin price prediction model [2].

This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 International License

There is a lack of clarity on how the selected factors, such as the volume of Bitcoin and other currencies, contribute individually or interactively to the accuracy of the predictive model. Further investigation into the rationale behind feature selection and their respective impact on forecasting could enhance the understanding of the model's robustness and effectiveness [3].

The challenge of accurately predicting the price action and locus of cryptocurrency due to its unique characteristics and lack of alignment with traditional market movements [4].

In machine learning for gold price prediction has primarily focused on historical data analysis, pattern recognition, and model evaluation. However, there is a notable gap in understanding the specific limitations of current machine learning models for gold price forecasting, and further investigation is needed to address these limitations and enhance the accuracy and reliability of predictions. [5].

The limited exploration of hyperparameter tuning specifically for forecasting silver prices using the XGBoost machine learning method. While XGBoost is acknowledged for its computational efficiency, there is a need to systematically identify and optimize hyperparameter combinations to enhance the accuracy of silver price predictions over a short-term horizon. [6].

The limited progress made in developing robust prediction models for long-term fluctuations in cryptocurrency prices, particularly Bitcoin. While various methods, including traditional statistical approaches and machine learning algorithms, have been employed for short-term forecasting, there is a lack of comprehensive models that consider the interplay of multiple cryptocurrencies in predicting Bitcoin prices over an extended period [8].

The underlying factors contributing to the observed predictive performance, and further investigation is needed to identify and address potential limitations or biases in the models that could impact their reliability in real-world trading scenarios [9].

The individual and interactive contributions of selected factors, like Bitcoin and other currency volumes, to the accuracy of predictive models. Further research is needed to delve into the rationale behind feature selection and understand how these factors impact forecasting, thereby enhancing the comprehension of the model's robustness and effectiveness [10].

There is a need for a more comprehensive analysis that incorporates a wider array of economic variables, considering the dynamic nature of financial markets and potential external factors impacting gold prices. Additionally, the sustainability of high gold prices and the potential reversal of trends remain underexplored aspects in the current research landscape [11].

The utilization of machine learning for gold price prediction through historical data analysis and model creation. However, it lacks specificity regarding the identification of specific gaps in existing research or the novel contributions made by the study, leaving room for a clearer delineation of the research gap and the unique aspects addressed by the proposed research [13].

A systematic review on Non-Fungible Tokens (NFTs) delves into their influence across industries, particularly in art and gaming. The analysis underscores NFTs' transformative impact on digital ownership, presenting both challenges and opportunities while reshaping economic paradigms [14].

The research by Kin-Hon Ho, Haoyuan Pan, and Tse-tin Chan employs machine learning to analyze factors influencing Non-Fungible Token (NFT) pricing, offering valuable insights into the determinants shaping the value of NFTs in various markets. The study contributes to understanding the complex dynamics and predicting pricing trends within the NFT ecosystem [15].

The study "Price Determinants of Non-fungible Tokens in the Digital Art Market" by Florian Horky, Carolina Rachel, and Jarko Fidrmuc investigates factors influencing NFT prices, providing insights into the dynamics of the digital art market. The research offers a comprehensive analysis of the determinants shaping the value of non-fungible tokens in this evolving ecosystem [16].

Perry Sadorsky's study utilizes tree-based classifiers to predict gold and silver price directions, offering valuable insights for predictive modeling in the precious metals market and advancing understanding of classification techniques for anticipating trends in their prices [17].

The research by Mrs. Vaidehi M, Alivia Pandit, Bhaskar Jindal, Minu Kumari, and Rupali Singh focuses on predicting Bitcoin prices using machine learning, contributing insights into the application of advanced algorithms for forecasting cryptocurrency trends. Their study aims to enhance understanding and accuracy in anticipating Bitcoin price movements [18].

The research on gold price prediction by M. Sravani, Ch. Abhilash, T. Divya, Ch. Vasthav, and D. Priyanka likely employs predictive models to forecast gold prices, providing insights into potential trends and factors influencing the precious metal's value. The study contributes to the understanding of forecasting methodologies in the context of gold pricing [19].

Ilan Alon's study on predictors of NFT prices, utilizing an automated machine learning approach, investigates key factors influencing non-fungible token valuations, providing valuable insights into the pricing dynamics within the NFT market [20].

Nandini Tripurana, Binodini Kar, Sujata Chakravarty, Bijay K. Paikaray, and Suneeta Satpathy's research on gold price prediction utilizes machine learning techniques, aiming to forecast gold prices and provide insights into the factors influencing their fluctuations. The study contributes to understanding predictive modeling applications in anticipating trends within the gold market [21].

Lekkala Sreekanth Reddy and Dr. P. Sriramya conduct research on Bitcoin price prediction employing machine learning algorithms, providing insights into the application of advanced techniques for forecasting cryptocurrency trends. The study contributes to enhancing understanding and accuracy in predicting Bitcoin price movements [22].

Vaddi, Neelisetty, Vallabhaneni, and Prakash employ machine learning and deep learning techniques to predict cryptocurrency prices, aiming to enhance accuracy and understanding of market trends. Their research contributes advanced methodologies for forecasting in the dynamic cryptocurrency space [23].

Maleki, Nikoubin, Rabbani, and Zeinali use machine learning and time series analysis to predict Bitcoin prices, providing insights into its interactions with other cryptocurrencies. The study seeks to improve predictive modeling by considering relationships within the broader cryptocurrency market [24].

Jodeiri Shokri, Shamsi, and Dehghani employ a novel approach by combining multiple linear regression and imperialist competitive algorithm to predict silver prices. This integrated machine learning and optimization technique enhances forecasting methodologies, contributing to advanced understanding of silver price trends [25].

III. METHODOLOGY

Crypto currency price prediction using machine learning has got some real time prediction power by integrating machine learning algorithms. The process of prediction of prices using machine learning as follows:

Data Collection: In the data collection process, massive amounts of data are gathered and stored to conduct data filtering operations. The data collection phase is initiated with the downloading of data from various online sources. Thus, a substantial file containing numerous columns and rows is collected. Here, data is assembled into CSV files, encompassing different parameters such as Open, Close, High, Low, Volume, and Market Cap [26].

Data Preprocessing: Technical steps need to be performed in data preprocessing, including the normalization method applied to the data after splitting it for training, validation, and testing [26]. Variations in the range of values are observed, and 30 percent of the data is utilized for testing, while the remaining 70 percent is available for training. The data is loaded as dataset matrices, normalization is performed, and then the data is split and improved through training. A problem-solving approach is employed, and Random Forest and LSTM models are utilized.

Random Forest: An ensemble form of multiple regression trees is represented by the Random Forest [9]. Its advantages include high explicability, but the predicted results are constrained by the training samples. The principle of the regression tree involves dividing the parent group into subgroups using an indicator of a specific variable, with the classification being determined by minimizing the average sum of squared residuals for each group.

Long Short-Term Memory (LSTM): The Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) designed to overcome the limitations of traditional RNNs in capturing long-term dependencies in sequential data [4][10]. LSTMs is pivotal for time-series analysis and sequential data modelling. LSTMs utilize memory cells and gating mechanisms to selectively retain and forget information over extended sequences. The architecture includes input, forget, and output gates, enabling the network to store and retrieve information effectively. This design helps LSTMs address the vanishing gradient problem, allowing for the modelling of intricate temporal relationships, showcasing their proficiency in capturing and learning complex sequential patterns.

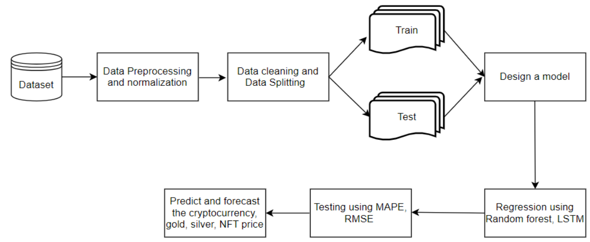
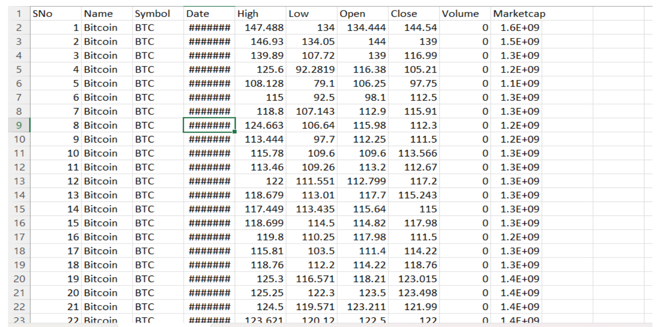


Fig. 1. Block Diagram



1	SNo	Name	Symbol	Date	High	Low	Open	Close	Volume	Marketcap
2	1	Bitcoin	BTC	#######	147.488	134	134.444	144.54	0	1.6E+09
3	2	Bitcoin	BTC	#######	146.93	134.05	144	139	0	1.5E+09
4	3	Bitcoin	BTC	#######	139.89	107.72	139	116.99	0	1.3E+09
5	4	Bitcoin	BTC	#######	125.6	92.2819	116.38	105.21	0	1.2E+09
6	5	Bitcoin	BTC	#######	108.128	79.1	106.25	97.75	0	1.1E+09
7	6	Bitcoin	BTC	#######	115	92.5	98.1	112.5	0	1.3E+09
8	7	Bitcoin	BTC	#######	118.8	107.143	112.9	115.91	0	1.3E+09
9	8	Bitcoin	BTC	#######	124.663	106.64	115.98	112.3	0	1.2E+09
10	9	Bitcoin	BTC	#######	113.444	97.7	112.25	111.5	0	1.2E+09
11	10	Bitcoin	BTC	#######	115.78	109.6	109.6	113.566	0	1.3E+09
12	11	Bitcoin	BTC	#######	113.46	109.26	113.2	112.67	0	1.3E+09
13	12	Bitcoin	BTC	#######	122	111.551	112.799	117.2	0	1.3E+09
14	13	Bitcoin	BTC	#######	118.679	113.01	117.7	115.243	0	1.3E+09
15	14	Bitcoin	BTC	#######	117.449	113.435	115.64	115	0	1.3E+09
16	15	Bitcoin	BTC	#######	118.699	114.5	114.82	117.98	0	1.3E+09
17	16	Bitcoin	BTC	#######	119.8	110.25	117.98	111.5	0	1.2E+09
18	17	Bitcoin	BTC	#######	115.81	103.5	111.4	114.22	0	1.3E+09
19	18	Bitcoin	BTC	#######	118.76	112.2	114.22	118.76	0	1.3E+09
20	19	Bitcoin	BTC	#######	125.3	116.571	118.21	123.015	0	1.4E+09
21	20	Bitcoin	BTC	#######	125.25	122.3	123.5	123.498	0	1.4E+09
22	21	Bitcoin	BTC	#######	124.5	119.571	123.211	121.99	0	1.4E+09
23	22	Bitcoin	BTC	#######	123.621	120.12	122.5	122	0	1.4E+09

Fig. 2. Dataset

TABLE I. COMPARISON OF RANDOM FOREST ALGORITHM AND LSTM ALGORITHM

Parameters	Random Forest Algorithm	LSTM Algorithm
Accuracy	The proportion of correctly classified instances among the total instances. Formula: (True Positives + True Negatives) / Total Instances	Similar to the Random Forest accuracy, measuring the proportion of correctly classified instances.
Precision	Representing the ability of the Random Forest to correctly identify positive instances among all instances predicted as positive. Formula: Precision = True Positives / (True Positives + False Positives)	Similar to Random Forest, evaluating the precision for positive and negative classes.
Handling Missing Data	Random Forest can handle missing data well by imputing missing values based on other features.	LSTM, being a neural network, may require more preprocessing to handle missing data effectively.
Mean Absolute Error	For regression tasks, calculate the Mean Absolute Error, representing the average absolute difference between predicted and actual values. Formula: MAE = (1 / N) * Σ |Actual - Predicted|	Similarly, for regression tasks, calculate the Mean Absolute Error for the LSTM model. Formula: MAE = (1 / N) * Σ |Actual - Predicted|

Ensemble Nature	Random Forest is an ensemble method, combining multiple decision trees, which helps in reducing overfitting and improving generalization.	LSTM is a type of recurrent neural network (RNN), which has a more complex architecture designed for capturing temporal dependencies in sequential data.

IV. RESULT AND DISCUSSION

LSTM outperforms Random Forest due to its inherent capacity to model sequential dependencies, beneficial for time-series data. As a type of recurrent neural network, LSTMs excel in learning temporal patterns, enabling accurate predictions by comprehending complex relationships. In contrast, Random Forest's ensemble of decision trees may struggle with intricate temporal patterns, underscoring the superior accuracy of LSTMs in scenarios where sequential information is pivotal.

The below Fig.3 gives the comparison between original close price and predicted close price, which helps to achieve the prediction result using LSTM algorithm.

Fig. 3. comparison between original close price and predicted close price using LSTM Algorithm

The Fig. 4 gives the comparison between the actual close price and predicted close price using Random Forest algorithm.

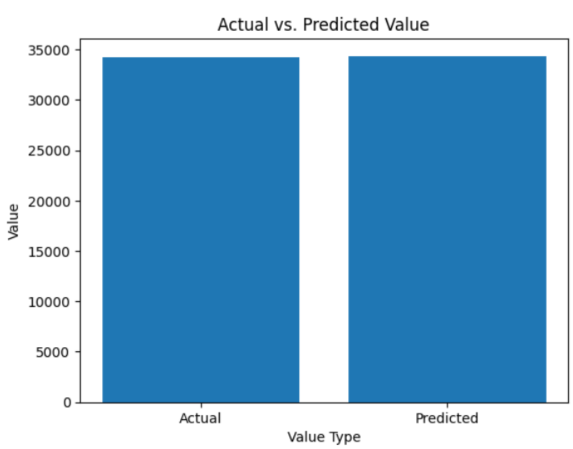


Fig. 4. comparison between actual close price and predicted close price using Random Forest Algorithm

LSTM (Long Short-Term Memory) over random forests due to LSTM's ability to recognize sequential dependencies in time-series data. Price movements are predicted with greater accuracy and flexibility by LSTMs.

TABLE II. ERROR CHECKING PARAMETERS

Parameters		Cryptocurrency	Gold- Silver
RMSE	Train	622.042	15.60
	Test	1193.89	41.44
MSE	Train	386936	243.49
	Test	142539	1717.36
MAE	Train	412.87	12.67
	Test	854.591	36.31
Variance Regression Score	Train	0.9824	0.7699
	Test	0.9710	0.7596
R Square Score	Train	0.9814	0.76939
	Test	0.9707	0.4234

V. CONCLUSION

Advanced machine learning algorithms are utilized in this paper to forecast price trends in cryptocurrency, gold, silver, and non-fungible tokens (NFTs). Historical data is leveraged, and factors such as market sentiment, trading volume, and macroeconomic indicators are integrated by the models. Promising predictive accuracy is indicated by our findings, demonstrating the potential of machine learning in forecasting asset prices. Higher volatility is exhibited by cryptocurrency and NFTs, while more stability is shown by gold and silver. In the dynamic landscape of diverse asset classes, valuable insights for investors seeking informed decision-making strategies are offered by this research as the digital and traditional markets evolve.

VI. FUTURE WORK

Promising accuracy was demonstrated by our model, utilizing historical data and relevant features. Moving forward, the focus of our future work is on refining the models by incorporating real-time market sentiments, macroeconomic indicators, and technological advancements. Additionally, ensemble methods and neural network architectures are aimed to be explored to enhance predictive capabilities. Furthermore, the importance of continually adapting the models to dynamic market conditions and incorporating ethical considerations in algorithmic trading is emphasized. This research lays the groundwork for more robust and adaptive forecasting in the evolving landscape of digital assets.

REFERENCES

[1] Junwei Chen, “Analysis Of Bitcoin Price Prediction Using Machine Learning”,Greduate School Of Economics, Kobe University, Japan ,13 January 2023.

[2] Snega . S, Nivedha . B, Ramachandran . C. A , “ Bitcoin Price Prediction using ML”,B.S Abdur Rahman Crescent Institute Of Science and Technology, Chennai, India, 2023

[3] Gurupradip G, Harishvaran m,Amsavalli, “Cryptocurrency Price Prediction Using Machine Learning”, International Journal Of Engineering and Technology and Communication Engineering, Chennai India, VoL 12, Issue 4, April 2023

[4] Shubham Bhattad, Stefin Sunnymon, “Review Of Machine Learning for cryptocurrency Price Prediction,” Dept. Of Information Technology,Xavier Institute Of Engineering ,Mumbai India, 17 May 2023

[5] Ranganath Maruti Bhajantri, Rohith F Puttappanavar, Shashank, Vishal, Prof. Bhaskar Rao, “Gold Price Prediction Using Machine Learning”, Department of Computer Science and Engineering, Bangalore, Karnataka, India, Vol-9, Issue-3, 2023

[6] Dylan Norbert Gono, Herlina Napitupulu, Firdaniza, “Silver Price Forecasting Using Extreme Gradient Boosting (XGBoost) Method”, Department of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia, 5 September 2023

[7] Ram Prasad S K, Dr. Vibha M B, “A Study On Gold Price Prediction Using Machine Learning”, Department of MCA, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India, Vol-11, Issue 06 2023

[8] N. Maleki, A. Nikoubin, M. Rabbani, Y. Zeinali, “Bitcoin Price Prediction based on Other cryptocurrencies using Machine learning and Time series analysis”, University of Technology, Tehran, Iran, 2023

[9] Pratik Jaquart, Sven Kopke, Christof Weinhardt , “Machine Learning For cryptocurrency and market Trading,” Institute of Information Systems and Marketing Germany, pp. 331–352, 14 December 2022.

[10] Ashmit K. Khobragade, Omkar C. Keskar, “Forecasting of Cryptocurrency Values Using Machine Learning,” International Journal Of Scientific & Technology, Marathwada Mitra Mandals College of Engineering, Pune, VOL 10 Issue 5 May 2022

[11] Adena Wahyu Gumelar, Tacbir Hendro Pudjiantoro, Puspita Nurual Sabrina “NFT Coin Price Prediction (Non-Fungible Token) Using K-Nearest Neighbors Method ”, Informatics Department Faculty of Science and Informatics Universitas Jenderal Achmad Yani Cimahi, Indonesia, International Conference on Industrial Engineering and Operations Management, Johor Bahru, Malaysia, September 13-15, 2022

[12] Chodavrapu Pragna, Bhavan Purra, Adari Viharika, Garabhapu Roshni, Prof. B. Prajna, “Gold Price Prediction Using Machine Learning”, Andhra University College Of Engineering For Women, Visakhapatnam, Andhra Pradesh, India, Vol- 4, Issue 5, May 2022

[13] Rushikesh Ghule, Abhijeet Gadhave, Prof. Manish Dubey, Dr. Jyoti Kharade, “Gold Price Prediction Using Machine Learning”, Bharati Vidyapeeth’s Institute of Management and Information Technology, Navi Mumbai, Vol-06, Issue-06, June 2022

[14] Hamed Taherdoost, “Non-Fungible Tokens(NFT): A Systematic Review”, Department of Arts, Communication and Social Science, University Canada West, Vancouver, 31 December 2022

[15] Kin-Hon Ho, Tse-Tin Chan, Haoyuan Pan, “Analysis of Non-Fungible Token Pricing Factors with Machine Learning”, University of Hong Kong, Octomber 2022

[16] Florian Horky, Carolina Rachel, Jarko Fidrmuc, “Price Determinants of Non-Fungible Tokens in Digital Art Market”, May 2022

[17] Perry Sadorsky, “Predicting Gold and Silver Price Dieextion Using Tree-Based Classifiers”, Schulich School of Business, York University, Toronto, 29 April 2021

[18] Mrs Vaidehi M, Alivia Pandit, Bhaskar Jindal, Minu Kumari, Rupali Singh , “Bitcoin Price Prediction Using Machine Learning”,International Journal Of Engineering and Technology and Management Research, Bangalore India III, pp. 20–28,13 May 2021.

[19] M. Sravani, Ch. Abhilash, T. Divya, Ch. Vasthav, D.Priyanka, “Gold Price Prediction”, Department of CSE, Gudlavalleru Engineering College, Gudlavalleru, India, Vol- 9, Issue 6 June 2021

[20] Ilan Alon, “Predictors of NFT Prices: An Automated Machine Learning Approach”, University of Ariel, Israel, Vol- 31, Issue 1, 2021

[21] Nandini Tripurana, Binodini Kar, Sujata Chakravarty, Bijay K. Paikaray, Suneeta Saptpathy, “Gold Price Prediction Using Machine Learning Techniques”, University of Technology and management, Odisha, India, 2021

[22] Lekkala Sreekanth Reddy ,Dr . P. Sriramya , “A research On Bitcoin Price Prediction Using Machine Learning Algorithms”, International Journal Of Scientific & Technology Research, Vol 9, 4 April 2020.

[23] Lokesh Vaddi, Vaishnavi Neelisetty, “Predicting Crypto Currency Prices Using Machine Learning And deep Learning”,International Journal Of Advanced Trends in Computer Science and Engineering, VOL 9,4 July 2020

[24] N. Maleki, A . Nikoubin , M. Rabbani, “Bitcoin Price Prediction Based on other Cryptocurrencies Using machine Learning and tinme Series Analysis., Sharif University Of Technology, Iran, pp.285-301, 2020

[25] Behshad Jodeiri Shokri, Hesam Dehghani, Reza Shamsi, “Predicting silver price by applying a copled multiple linear regression (MLR) and imperialist competitive algorithm (ICA)”, Department of mining Engineering, Hamedan University of Technology, Hamedan, Iran, Vol. 1, No. 1, June 2020

[26] K. Ramya Laxmi, Marri Abhinandhan Reddy, CH. Shivasai, P. Sandeep Reddy, “Cryptocurrency Price Prediction Using Machine Learning”, Department of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad, India, SAMRIDDHI Volume 12, Special Issue 3,2020