International Journal of Advanced Engineering **ISSN: 2457-0397**

Vol. 5, Issue IV, Oct-Dec, 2022 http://bharatpublication.com/current-issue.php?jID=29/IJAE

USING MACHINE LEARNING TECHNIQUE TO PREDICT DIGITAL CURRENCY PRICE

Swastik Rout

Sai International School, Odisha

ABSTRACT

Coming into the 21st century, the perspectives towards money and speculations have radically moved from customary resources like gold, land, and property to resources in the advanced area, such as putting resources into different stocks, shared reserves, monetary standards, and so on.

Innovative turns of events and more current speculation choices have brought different block-tied advanced monetary standards to the front. Computerized money choices offer gigantic gains however are exceptionally powerful and fast-moving ventures, making anticipating what's to come values for most extreme additions troublesome, consequently lessening the certainty of even prepared financial backers.

While advanced monetary standards saw huge additions in the previous ten years, research in profound learning has seen identical development in proficiency, required calculation, and expectation rate. In this undertaking, we investigated different computerized monetary forms, utilizing various AI models from stretched-out trees to time series examination. The Long Short-Term Memory (LSTM) model was the only one we used to accurately predict the tested parameters. While there is a gamble of overfitting the Dataset, considering that this undertaking utilizes APIs to get continuous information, the gamble is moderated by the way that it gives exact results for inconspicuous information.

INTRODUCTION

The coming of Bitcoin in 2008 as a singular digital money that gradually moved to a worldwide space-driving means for cryptographic money exchanges prompted a spike of improvement in the field of blockchain. At first, digital currencies conveyed set esteem and were viewed as a prize for a movement properly named mining. Nonetheless, as the web turned into the primary wellspring of correspondents around the world, and independent ventures and contests began to obtain a worldwide crowd, the limitation of utilizing a particular managed money, i.e., the US dollar, should have been overhauled. This prodded fast development in the blockchain and crypto market, which made cryptographic money the primary method for exchanges between individuals over the web without the lumbering nature of banking past worldwide lines, procuring particular money, and the expenses related to the interaction. The Unexpected explosion in the crypto market additionally made it one of the chief ventures to open doors, with the admonition that it was likewise exceptionally unsteady, given it was decentralized and offered no security for money securities. The arrangement of determining future costs utilizing AI calculations was proposed to offset this unpredictability and absence of financial backer trust in the cash contributed. For the ongoing task, the group intends to involve the mathematical information for the computerized money got from the Programming interface to prepare a repetitive brain organization, to be specific Long Momentary Memory (LSTM), to foresee computerized cash costs. LSTM is a repetitive brain network in which associations are made along the hubs in a transient succession to frame consecutive charts where the singular hubs process both particular pieces of information and recall information throughout erratic time spans to such an extent that the model can deal with both individual data of interest furthermore, the all-out consecutive information to represent the impact of individual places and the patterns in the complete information for figures for a required length of time.

METHODOLOGY

API for the dataset: - The fundamental Dataset for this task is coordinated with a Programming interface that offers information from 4 months before the point at which the code is run. This Dataset was acquired from the Programming interface and was at first for USDT in two distinct periods, to be specific, minute-by-moment and day-to-day. For the ongoing cycle of model preparation, the Dataset separated was from 21 November 2021 to 28 November 2021 for the moment-by-minute Dataset, while for the everyday Dataset, the Dataset removed was from 4 May 2021 to 28 November 2021. There are a total of 209 entries in the daily Dataset, whereas the minute-by-minute Dataset has 11752 entries. Utilizing a similar Programming interface, information was accumulated for six more digital currencies: Bitcoin (BTC), Ethereum (ETH), Polkadot (Spot), Litecoin (LTE), Run Coin (Run), and Dogecoin (DOGE) for time frames seconds, 21600 seconds, 3600 seconds, 300 seconds, and 60 seconds each.

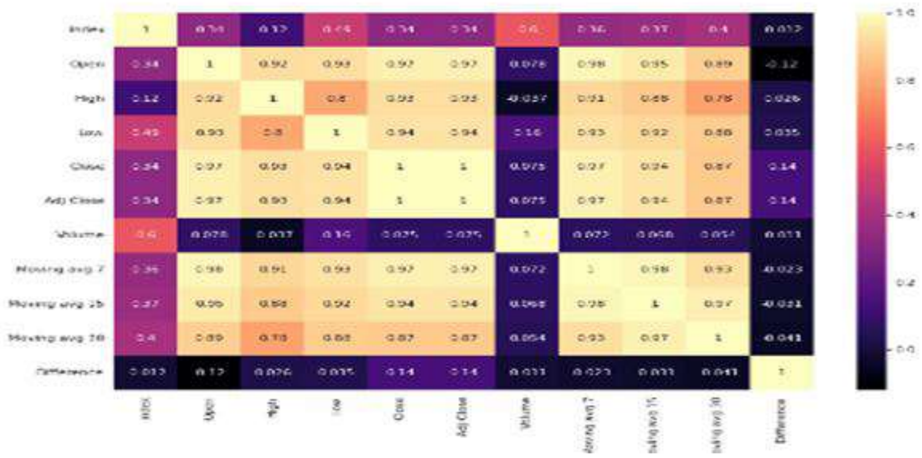


Fig. 1: Passive USDT dataset heatmap

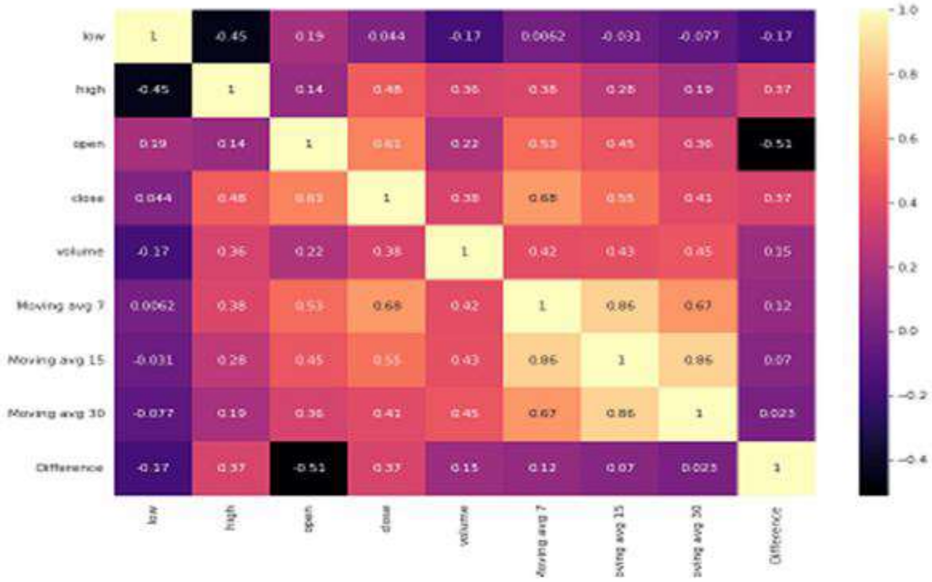


Fig. 2: API USDT dataset heatmap

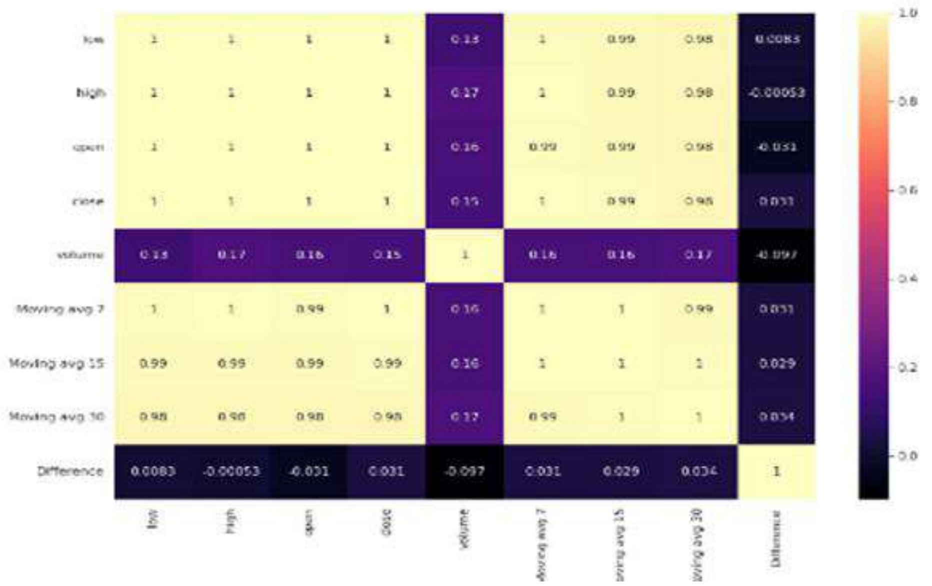


Fig. 3: API BTC dataset heatmap

These datasets have 6 primary elements, which are DATE-TIME, LOW, HIGH, OPEN, CLOSE, and VOLUME, out of which DATETIME is utilized as the file.

A. Latent Dataset: - For the model preparation to yield results with higher exactness, an inactive dataset comprising verifiable information of the cryptographic money [USDT] was used. This Dataset had 288562 passages for the moment-by-minute development of the cryptographic money from 4 May 2021 to 22 November 2021[11:14 AM]. Likewise, the Dataset for the day-to-day esteems has a sum of 784 sections from 18 September 2019 to 9 November 2021. This Dataset has similar elements as the Programming interface dataset except an extra section includes the day-to-day Dataset, specifically, Changed CLOSE, which represents any reseller's exchange exchanges and changes in the worth of the digital currency.

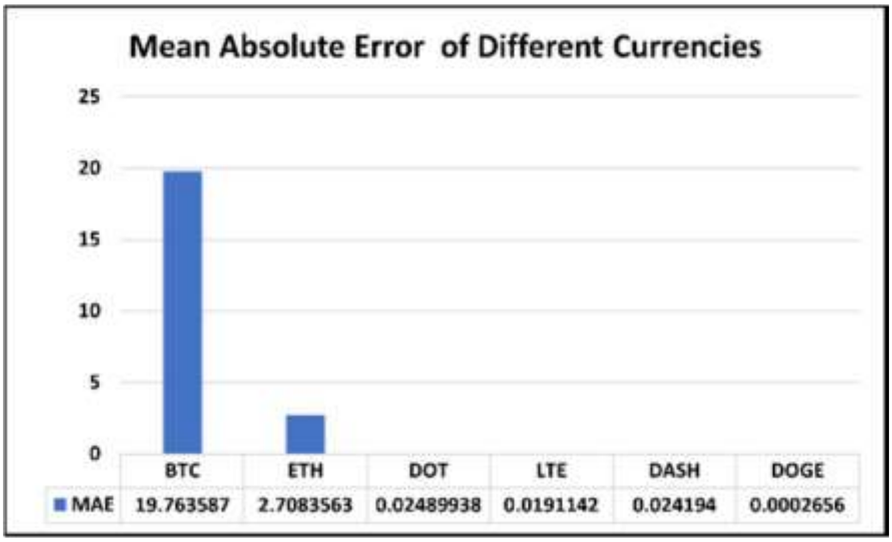
B. Information Pre-processing

Information cleaning: - After the underlying getting of information, the information should be handled for some variables before giving it to the AI pipeline. After cautious thought, the Programming interface information was gathered for the greatest productivity and least time misfortune in the information-gathering stage. Since the Programming interface just gives the essential information, no sections were dropped. During each training cycle, the entire Dataset is checked for missing or corrupted values. The idea of the Dataset gained from the Programming interface is dynamic and dependent upon debasement, making missing qualities show up in the Dataset, which will make the model break down and foresee measures incorrectly. To keep away from this issue, the Python library capability of forward-fitting is used, where the library takes the normal of the previous and succeeding worth and fits it instead of the missing qualities, which assists the Dataset with staying away from any 'NaN' values from showing up in the Dataset to be utilized as the preparation set for the AI model.

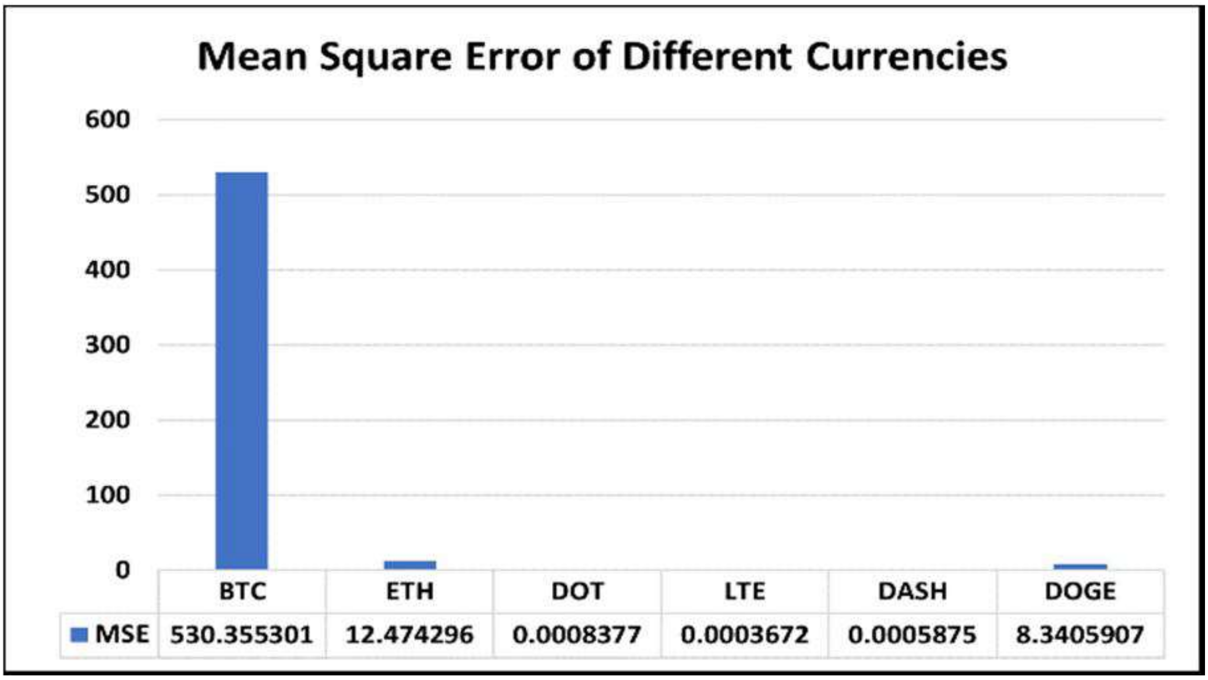
RESULTS

For this examination, numerous datasets were used for the preparation cycle to lay out a near investigation for computational effectiveness and exactness. Given the sheer size of the datasets, the time expected for the underlying preparation interaction might have been exceptional for dynamic changes of the estimate according to necessities. Hence, the information was managed for day-to-day esteems until just 300 sections were thought of. Likewise, the information was approved for future estimates rather than a preparation and testing split of the generally accessible information, building up that the model performed well in certifiable applications. The approval boundaries chosen for checking the precision of the estimates are as per the following:

1) Mean Absolute Error (MAE): - MAE gives the outright error between two pieces of information simultaneously example and midpoints the all-out error of the time series determined by the model versus the unaffected time series information.



2) Mean Square Mistake (MSE): - The mean of the squared error is calculated by MSE, which squares off all error values and makes them strictly positive.



3) Root Mean Square Error (RMSE): - The absolute positive error of the forecast at each time interval is calculated by RMSE by taking the squared error at each data point and finding its root. This fills in as the approval boundary, which gives the mathematical exactness of the model.

4) R-Squared score: - The R-squared score gives how precisely the model notices and duplicates the patterns and results of the given Dataset. This approval design is the extent of the variety of the reliant data of interest from the autonomous piece of information.

5) Mean Absolute Percent Error (MAPE): - The accuracy of the forecasted values generated by the model in comparison to the actual values is measured using MAPE.

Table 1:

Comparative Analysis of Validation Parameters for Various Datasets

Dataset	Validation Parameters			
	MAE	MSE	RMSE	R-Squared
Passive – Daily	0.026232	0.0010377	0.0322144	0.9903814
Passive - Min by Min	2.243226e-05	5.032575e-10	2.243340e-05	0.34119011
API – Daily	0.028197	0.0009952	0.03154797	0.9907753
API - Min by Min	1.935327e-06	7.1587463e-12	2.675583e-06	0.9971194
BTC	40.42163	2646.81303	51.447186	0.9995420

Table 2:

Analysis of Validation Parameters for Different Datasets with 86400 Seconds of Interval

Dataset (1 day interval data)	Validation Parameters 84600 sec			
	MAE	MSE	RMSE	R-Squared
BTC	19.763587	530.355301	23.029444	0.9976625
ETH	2.7083563	12.474296	3.5318970	0.9987444
DOT	0.02489938	0.0008377	0.0289442	0.9911232
LTC	0.0191142	0.0003672	0.0191631	0.9997705
DASH	0.0241940	0.0005875	0.0242392	0.9997586
DOGE	0.0002656	8.3405907	0.0002888	0.9978963

Based on the project's progress report results, the model overfits the Dataset. There is a contention to be made that the model overfits the gauging boundaries, with a precision of almost 99.46 per cent for various time frames utilizing the Dataset procured from the Programming interface for BTC, 98.80 per cent for ETH, 99.21 per cent for Spot, 98.61 per cent for LTE, 99.54 per cent for Run, and 99.02 per cent for DOGE. Notwithstanding, on the off chance that we notice the underlying Dataset itself, it vacillates inside a fine cross-section of values, showing a negligible change in pattern and worth over a drawn-out period. This can highlight the way that the Dataset is a lot in range for negligible development. As a result, the periodic changes in values are easy for the designed LSTM model to recognize.

CONCLUSION

After a durable investigation of the outcomes, it is laid out that the conveyed LSTM model gives a very serious level of precision while being effective as far as computational asset prerequisite and time for continuous figures when conveyed for constant information examination and ensuing gauges. The low level of error noticed for expectations with regards to Mean Outright Mistake, Mean Square Mistake, Root Mean Square error, and the serious level of precision laid out by the R-square score of 99.97 per cent for Tie (USDT), 99.63 per cent for Bitcoin (BTC), 99.91 per cent for Ethereum (ETH), 98.51 per cent for Polkadot (Dab), 99.96 per cent for Litecoin (LTE), 98.62 per cent for Run

Coin (Run), 99.65 per cent for Dogecoin (DOGE) for day-to-day stretches implies that the chosen calculation and AI model can be involved with alterations for ongoing applications according to the prerequisites of the end client.

REFERENCES

[1] Shefali Arora, Ruchi Mittal, and M P S Bhatia. Automated cryptocurrencies prices prediction using machine learning collaborative approach for trend analysis using clustering mechanisms and big data technologies. Article in International Journal of Soft Computing, page 4, 2018.

[2] Temesgen Awoke, Minakhi Rout, Lipika Mohanty, and Suresh Chandra Satapathy. Bitcoin price prediction and analysis using deep learning models. Volume 134, pages 631–640. Springer Science and Business Media Deutschland GmbH, 2021.

[3] Seema A Bhalegaonkar et al. Automated metaphase chromosome image selection techniques for karyotyping: Current status and future prospects. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(6):3258–3266, 2021.

[4] Krishna Chakravarty, Manjusha Pandey, and Siddharth Routaray. Bitcoin prediction and time series analysis, 2020.

[5] Jaiswal Rupesh Chandrakant and D Lokhande. Comparative analysis using bagging, logitboost and rotation forest machine learning algorithms for real time internet traffic classification.

[6] Jaiswal Rupesh Chandrakant and D Lokhande. Statistical features processing based real time internet traffic recognition and comparative study of six machine learning techniques.

[7] Jaiswal Rupesh Chandrakant and D. Lokhande Shashikant. Analysis of early traffic processing and comparison of machine learning algorithms for real time internet traffic identification using statistical approach. volume 28, pages 577–587. Springer Science and Business Media Deutschland GmbH, 2014.

[8] Reshma Sundari Gadey, Nikita Thakur, Naveen Charan, and R Obulakonda Reddy. Price prediction of bitcoin using machine learning, 2020.

[9] Guus Van Heijningen. Making predictions in highly volatile cryptocurrency markets using web scraping, 2017.

[10] Alvin Ho, Ramesh Vatambeti, and Sathish Kumar Ravichandran. Indian journal of science and technology bitcoin price prediction using machine learning and artificial neural network model. Indian Journal of Science and Technology, 14:2300, 2021.

[11] Rupesh Jaiswal and Shashikant Lokhande. Rupesh jaiswal and shashikant lokhande: A novel approach for real time internet traffic classification a novel approach for real time internet traffic classification.

[12] Rupesh Jaiswal, Shashikant Lokhande, Aashiq Ahmed, and Prateek Mahajan. Performance evaluation of clustering algorithms for ip traffic recognition, 2013.

[13] Rupesh Jaiswal, Shashikant Lokhande, and Aditya Gulavani. Implementation and analysis of dos attack detection algorithms, 2013.

[14] Patrick Jaquart, David Dann, and Christof Weinhardt. Short-term bitcoin market prediction via machine learning. The Journal of Finance and Data Science,7:45–66, 11 2021.

[15] Danish Khan and Rupesh C Jaiswal. Issue 11 www.jetir.org (issn-2349-5162), 2020.

[16] Shreya Mondhe, Mayank Mukundam, and R C Jaiswal. Issue 6 www.jetir.org (issn-2349- 5162), 2019.

[17] M Munot, M Joshi, and Nikhil Sharma. Automated karyotyping of metaphase cells with touching chromosomes. Int J Comput Appl, 29(12):14–20, 2011.

[18] Mousami V Munot, Jayanta Mukherjee, and Madhuri Joshi. A novel approach for efficient extrication of overlapping chromosomes in automated karyotyping. Medical & biological engineering & computing, 51(12):1325–1338, 2013.

[19] Shilpa Nair. Cryptocurrencies price movement prediction using machine learning, 2021.

[20] Institute of Electrical, Electronics Engineers. India Council, Institute of Electrical, Electronics Engineers. Bombay Section, Annual IEEE India Conference 10 2013.12.13-15 Mumbai, Annual Conference of the IEEE India Council 10 2013.12.13-15 Mumbai, and INDICON 10 2013.12.13-15 Mumbai. Annual IEEE India conference (INDICON), 2013 13-15 Dec. 2013, Mumbai, India.

[21] Sarika A Panwar, Mousami V Munot, Suraj Gawande, and Pallavi S Deshpande. A reliable and an efficient approach for diagnosis of brain tumor using transfer learning. Biomed Pharmacol J, 14:283–294, 2021.

[22] Aashay Pawar and R C Jaiswal. Stock market study using supervised machine learning, 2020.

[23] Chen Peng and Guo Yichao. Isss608 visual analytics and applications cryptocurrency price analysis and time series forecasting group 7, 2020.

[24] Prajwal Pitlehra, R C Jaiswal, and Associate Professor. Credit analysis using k-nearest neighbours model, 2021.

[25] Ajith Premarathne, Malka N., R. Samarakody, and Ampalavanapillai Nirmalathas. Real-time cryptocurrency price prediction by exploiting iot concept and beyond: Cloud computing, data parallelism and deep learning. International Journal of Advanced Computer Science and Applications, 11, 2020.

[26] Vojtech Pulec. Cryptocurrency returns: short-term forecast using google trends, 2019.

[27] Vladimir Puzyrev. Deep convolutional autoencoder for cryptocurrency market analysis. 10 2019.

[28] Lekkala Sreekanth Reddy and DrP Sriramya. A research on bitcoin price prediction using machine learning algorithms.

[29] Jaiswal Rupesh and Lokhande D Shashikant. Measurement, modeling and analysis of http web traffic.

[30] Jacopo De Stefani, Olivier Caelen, Dalila Hattab, Yann Aël LeÂ Borgne, and Gianluca Bontempi. A multivariate and multi-step ahead machine learning approach to traditional and cryptocurrencies volatility forecasting. volume 11054 LNAI, pages 7–22. Springer Verlag, 2019.

[31] Franco Valencia, Alfonso Gómez-Espinosa, and Benjamín Valdés-Aguirre. Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. Entropy, 21:589, 6 2019.