BITCOIN PRICE PREDICTION USING MACHINE LEARNING

[1]G.Naga Jyothi, [2]Dr.M.Jahir Pasha

[1]M.Tech Student, [2]Professor & HOD
Department of CSE,
Ashoka women's Engineering College

***Abstract*: Crypto currencies, such as Bit coin, are one of the most controversial and complex technological innovations in today's financial system. This study aims to forecast the movements of Bit coin prices at a high degree of accuracy. In order to test these algorithms, besides existing continuous dataset, discrete dataset was also created and used. For the evaluations of algorithm performances, the F statistic, accuracy statistic, the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE) and the Root Absolute Error (RAE) metrics were used. The test was used to compare the performances of the SVM, ANN, NB and RF with the performance of the LR. Empirical findings reveal that, while the RF has the highest forecasting performance in the continuous dataset, the NB has the lowest. On the other hand, while the ANN has the highest and the NB the lowest performance in the discrete dataset. Furthermore, the discrete dataset improves the overall forecasting performance in all algorithms (models) estimated.**

INTRODUCTION:

Some statistical tests can be carried out on the close price of financial instruments (to ascertain key metrics which can further be used to better understand market behavior); one of which is the Augmented Dickey-Fuller test. This test is used to check if a particular asset or instrument will revert to its rolling mean after a market swing (in upwards or downwards direction).
The Bitcoin stock-to-flow model makes it possible to trade it against a base currency on the foreign exchange market. This means that as with the Volatility 10 index above, bitcoin can be represented (on the charts) by its open, high, low and close prices and consequently traded with leverage at varying trade volumes.

IMPLEMENTATION:

LSTM (Long Short-Term Memory) is a deep learning model that helps withprediction of sequential data. LSTM models prevail significantly where there is a need to make predictions on a sequence of data. The daily OHLC (Open, High, Low and Close) price of any financial asset constitutes a good example of a sequential data.
LSTM's are an extension of the classic recurrent networks, which address the (the gradient tends to zero as the error propagates through many layers recursively). The long-short term memory cell uses an input, a forget and an output gate. Those gates help the network learns what to save, what to forget, what to remember, what to pay attention and what to output. Pretty neat right? Remember that a gate is nothing more than a simple multilayer perceptron, but a smart combination of them can provide amazing results. Each LSTM cell has its cell state (c) and has the ability to add or remove information to it. The forget gate decides what to remove from the cell state(f), while the input gate (i) decides which values it will update.

Output Screens

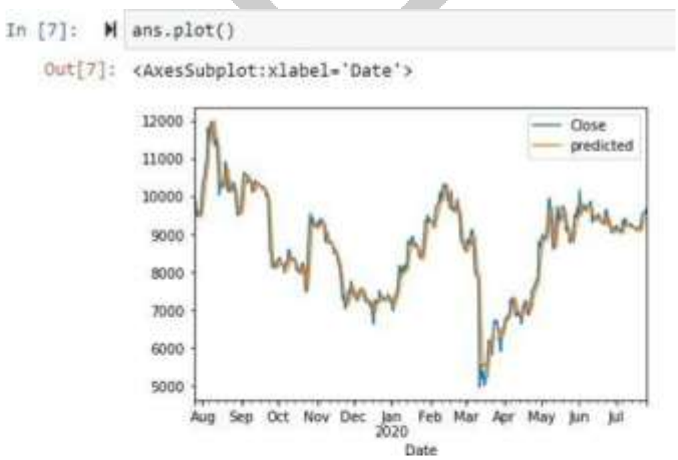


Fig.1: Graph of 2020

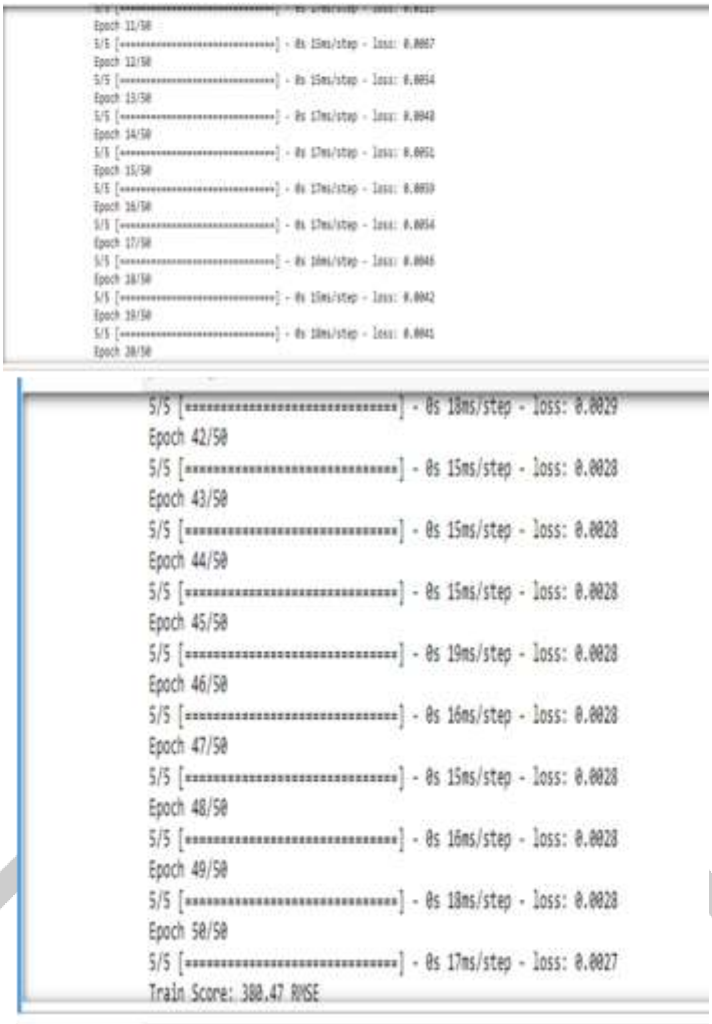


Fig.2: Epoch Graph

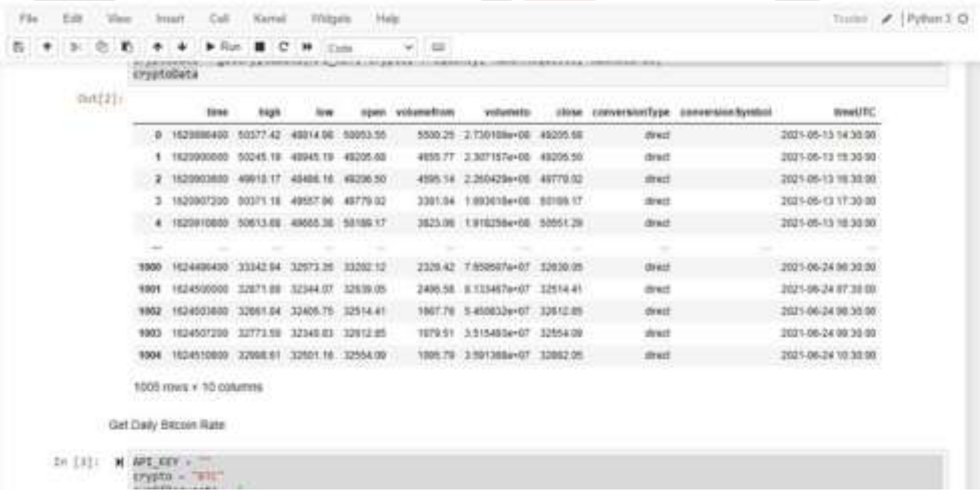
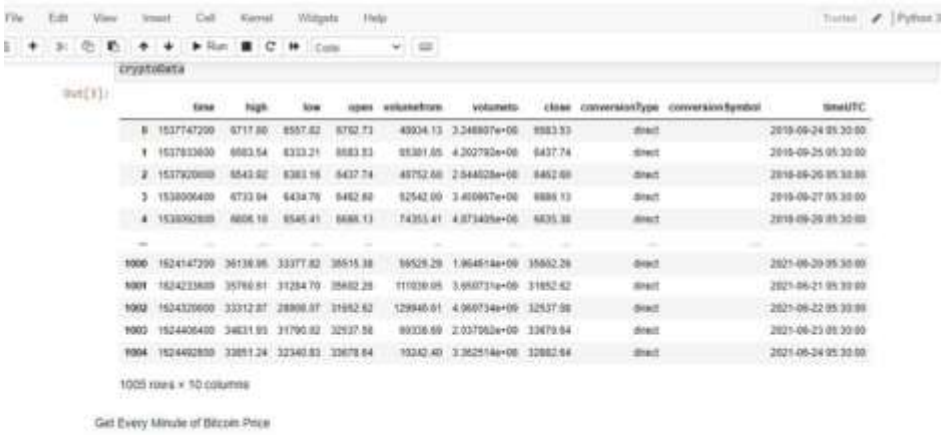


Fig.3: Daily Price Rate



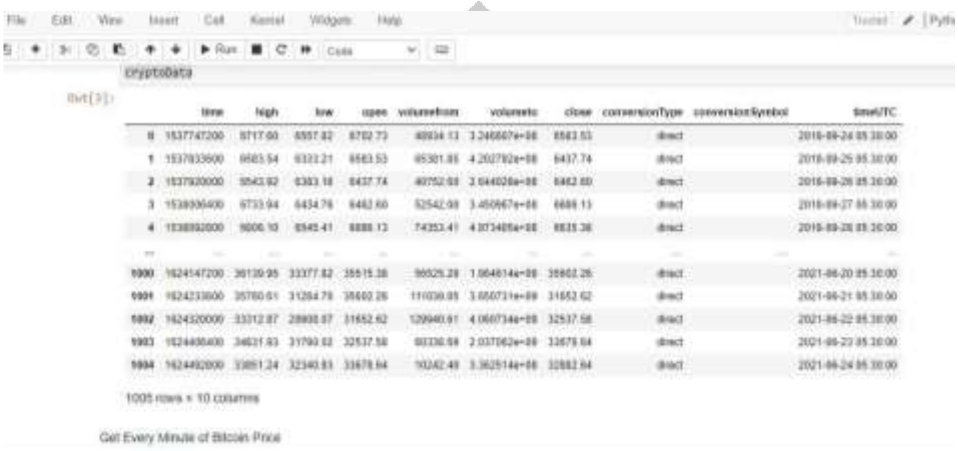
cryptoData

	time	high	low	open	volumefrom	volumeto	close	conversionType	conversionSymbol	timeUTC
0	1537747200	6717.00	6557.82	6702.73	49004.13	3.248807e+08	6583.53	direct		2018-09-24 05:30:00
1	1537833600	6683.54	6333.21	6583.53	95301.05	4.202792e+08	6437.74	direct		2018-09-25 05:30:00
2	1537920000	6543.02	6383.16	6437.74	49752.60	2.544028e+08	6462.60	direct		2018-09-26 05:30:00
3	1538006400	6733.04	6434.76	6462.60	52542.00	3.400867e+08	6686.13	direct		2018-09-27 05:30:00
4	1538092800	6806.10	6545.41	6686.13	74353.41	4.873405e+08	6635.38	direct		2018-09-28 05:30:00
...	...	...	...	...	...	...	...	...	...	...
1000	1624147200	36138.95	33377.82	35515.38	59529.28	1.964614e+09	35602.26	direct		2021-06-20 05:30:00
1001	1624233600	35760.61	31284.70	35602.26	111039.05	3.650731e+09	31652.62	direct		2021-06-21 05:30:00
1002	1624320000	33312.87	28808.07	31652.62	129946.01	4.060734e+09	32537.58	direct		2021-06-22 05:30:00
1003	1624406400	34631.93	31790.02	32537.58	89336.69	2.037062e+09	33679.64	direct		2021-06-23 05:30:00
1004	1624492800	33851.24	32340.83	33678.84	10242.40	3.362514e+08	32882.64	direct		2021-06-24 05:30:00

1005 rows × 10 columns

Get Every Minute of Bitcoin Price

Fig.4: Minute Price Rate



cryptoData

	time	high	low	open	volumefrom	volumeto	close	conversionType	conversionSymbol	timeUTC
0	1537747200	6717.00	6557.82	6702.73	49004.13	3.248807e+08	6583.53	direct		2018-09-24 05:30:00
1	1537833600	6683.54	6333.21	6583.53	95301.05	4.202792e+08	6437.74	direct		2018-09-25 05:30:00
2	1537920000	6543.02	6383.16	6437.74	49752.60	2.544028e+08	6462.60	direct		2018-09-26 05:30:00
3	1538006400	6733.04	6434.76	6462.60	52542.00	3.400867e+08	6686.13	direct		2018-09-27 05:30:00
4	1538092800	6806.10	6545.41	6686.13	74353.41	4.873405e+08	6635.38	direct		2018-09-28 05:30:00
...	...	...	...	...	...	...	...	...	...	...
1000	1624147200	36138.95	33377.82	35515.38	59529.28	1.964614e+09	35602.26	direct		2021-06-20 05:30:00
1001	1624233600	35760.61	31284.70	35602.26	111039.05	3.650731e+09	31652.62	direct		2021-06-21 05:30:00
1002	1624320000	33312.87	28808.07	31652.62	129946.01	4.060734e+09	32537.58	direct		2021-06-22 05:30:00
1003	1624406400	34631.93	31790.02	32537.58	89336.69	2.037062e+09	33679.64	direct		2021-06-23 05:30:00
1004	1624492800	33851.24	32340.83	33678.84	10242.40	3.362514e+08	32882.64	direct		2021-06-24 05:30:00

1005 rows × 10 columns

Get Every Minute of Bitcoin Price

Fig.5: Minute Price Rate

TESTING STRATEGIES:

A Strategy for software testing integrates software test cases into a series of well-planned steps that result in the successful construction of software. Software testing is a broader topic for what is referred to as Verification and Validation. Verification refers to the set of activities that ensure that the software correctly implements a specific function. Validation refers the set of activities that ensure that the software that has been built is traceable to customer's requirements

Unit Testing
Unit testing focuses verification effort on the smallest unit of software design that is the module. Using procedural design description as a guide, important control paths are tested to uncover errors within the boundaries of the module. The unit test is normally white box testing oriented and the step can be conducted in parallel for multiple modules.

Integration Testing
Integration testing is a systematic technique for constructing the program structure, while conducting test to uncover errors associated with the interface. The objective is to take unit tested methods and build a program structure that has been dictated by design.

Top-down Integration
Top down integrations is an incremental approach for construction of program structure. Modules are integrated by moving downward through the control hierarchy, beginning with the main control program. Modules subordinate to the main program are incorporated in the structure either in the breath-first or depth-first manner.

Bottom-up Integration
This method as the name suggests, begins construction and testing with atomic modules i.e., modules at the lowest level. Because the modules are integrated in the bottom up manner the processing required for the modules subordinate to a given level is always available and the need for stubs is eliminated.

Validation Testing
At the end of integration testing software is completely assembled as a package. Validation testing is the next stage, which can be defined as successful when the software functions in the manner reasonably expected by the customer. Reasonable expectations are those defined in the software requirements specifications. Information contained in those sections form a basis for validation testing approach.

System Testing
System testing is actually a series of different tests whose primary purpose is to fully exercise the computer-based system. Although each test has a different purpose, all work to verify that all system elements have been properly integrated to perform allocated functions.
Security Testing
Attempts to verify the protection mechanisms built into the system.
Performance Testing
This method is designed to test runtime performance of software within the context of an integrated system.

CONCLUSION:
In this project we conclude that survey report will be just introducing modules of Bitcoin price prediction and machine algorithms. Hear the Comparison table of ML algorithm model accuracy which tells that the linear regression model will have most accuracy then the other algorithms. In this project we conclude that the linear regression algorithm is more efficient than the other algorithms. By taking help from that linear regression algorithm, we can implement the LASSO also. The time complexity reduction in bit coin price prediction using LASSO algorithm is tested by referring all other algorithms and came to a conclusion that LASSO is the best among all. The machine learning algorithms will improve that feature idea of crypto currencies. That will improve the market price of globule investments. In this paper we proposed the new algorithm to find the feature price accuracy. That helps the customer increments and profits.

FUTURE ENHANCEMENT:
- With the help of real time data it can further be predicted for upcoming years.
- It can be predicted accurately with help of API key this project can be improvised to predict next years price.

REFERENCES:
[1] Bitcoin Price Prediction using Machine Learning, Siddhi Velankar, Sakshi Valecha, Shreya Department of Electronics & Telecommunication, Pune Institute of Computer Technology, Pune, Maharashtra, India.
[2] D. Shah and K. Zhang, ―Bayesian regression and Bitcoin,in 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2015, pp.
[3] Huisu Jang and Jaewook Lee, ―An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information, IEEE Early Access Articles, 2017.
[4] ―Project Based Learning: Predicting Bitcoin Prices using Deep Learning S. Yogeshwaran, Piyush Maheshwari, Maninder Jeet Kaur, 2019.
[5] ―Predicting the Price of Bitcoin Using Machine Learning Sean McNally, Jason Roche ;Simon Caton; Ireland, Dublin, IEEE 2018.
[6] ―Bitcoin Volatility Forecasting with a Glimpse into Buy and Sell Orders Tian Guo, Albert Bifet , Nino Antulov Fantulin , IEEE 2018.
[7] F. Andrade de Oliveira, L. Enrique and M. de Azevedo Reis; C. NeriNobre , The use of artificial neural networks in the analysis and prediction of stock prices,‖ in IEEE International Conference on Systems, Man, and Cybernetics, 2020.
[8] M. Daniela and A. BUTOI, ―Data mining on Romanian stock market using neural networks for price prediction in 2019.
[9] ―Bitcoin Cost Prediction using Deep Neural Network Technique Kalpanasonika ,Sayasri S , Vinothini , SugaPriya , IEEE 2020.
[10] ―An improved K-Means clustering algorithm Juntao Wang, Xiaolong Su, IEEE 2017.
[11] ―Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression‖ ;Jitendra Kumar Jaiswal, Rita Samikannu IEEE 2017
[12] ―Improved Random Forest for Classification‖ ; Angshuman Paul, Dipti Prasad Mukherjee, Senior Member, IEEE, Prasun Das, AbhinandanGangopadhyay, Appa Rao Chintha and SaurabhKundu IEEE2018.