Cryptocurrency Price Prediction Using Deep Learning and Machine Learning

D Siddharth and Jitendra Kaushik

January 30, 2022

Cryptocurrency price prediction using Deep learning and Machine learning

Siddharth.D
CHRIST University, India
d.siddharth@science.christuniversity.in

Dr. Jitendra Kaushik
CHRIST University, India
jitendra.kaushik@christuniversity.in

ABSTRACT:-
A cryptocurrencies is a network-based computerized exchange that makes imitation and double-spending pretty much impossible. Many cryptocurrency are built on distributed networks based on blockchain technology, which is a distributed ledger enforced by a network of computers. Thanks to blockchain technology, transactions are secure, transparent, traceable, and immutable. As a result of these traits, cryptocurrency has increased in popularity, especially in the financial industry. This research looks at a few of the most popular and successful deep learning algorithms for predicting bitcoin prices. LSTM and random forest outperform our generalized regression neural architecture benchmarking system in terms of prediction. Bitcoin and Ethereum are the only cryptocurrencies supported. The approach can be used to calculate the value of a number of different cryptocurrencies.

Keywords:- Bitcoin, Ethereum, long, short-term memory, random forest, time series, cryptocurrency, and price prediction.

I. INTRODUCTION

Nowadays, there are over 5000 cryptocurrencies available over the world. However, there are several issues to deal with regarding scientific research. Similarly, they are not highly ranked for market capitalization as market drivers. Another top-ranked pre-mined currency has a third characteristic that cannot be quickly asserted for Society open - sourced non-mineable coins. A controlled blockchain also supports Non-mineable coin transactions. We will now see the top-rated two cryptocurrencies up to the market date: bitcoin (BTC) and ethereum(ETH). These currencies they used for trading and got a lot of profit. That currency they can buy or sell in the coin switch platform. They are network-based exchange mediums, and they use cryptography algorithms to secure the transaction. They have their wallet to exchange all their currency and have their savings. And cryptocurrency is relevant to the stock market in that they buy and sell the shareholders. In this also similar, they will buy and sell currency. But its price fluctuates a lot depending on the parameter we discussed before. Many academics have worked with machine learning and deep learning techniques as well as well-known cryptocurrencies such as bitcoin and ethereum, as per the literature. Many digital currencies, including Tron, Bitcoin, Stellar, and others, have a wide range of applications in banking firms.
The trading companies demand cryptocurrency price prediction because they have to fix the target. So it will get floated like ups and downs. So by a projection of the price, it would be helpful to them to know the strategy and move according to that. So this literature would be beneficial to them.

II. BACKGROUND WORK

Sci-kit-learn and Keras were utilized in this study for data analysis, deep learning, and machine learning models. The Tensorflow software is also employed to create data flow diagrams in this study.

Scikit-learn

Scikit-learn is a free and open-source data mining analysis package. Python is used to analyse and build machine learning algorithms such as classification, regression, and clustering. In addition to normalisation, standardisation, and cleaning aberrant or missing data, Scikit-learn may be used to analyze data in a variety of ways.

Tensorflow

Tensorflow is a deep learning framework designed by Google as an open-source project. Dealing with a large number of Graphical Processing Units (GPUs) may be used to train and forecast Neural Network (NN) architectures, allowing for the adoption of intense deep learning and NN methodologies. Speech recognition, computer vision, robotics, and other sectors might benefit from this design. When graphs are made up of node groups, Tensorflow may construct data flow graphs for processing.

Keras

Keras is a high-level NN open-source library. It is a Python API for neural network programming. Tensorflow, CNTK, and Theano libraries are also supported. Keras can generate models for machine learning, NN, and deep learning. Keras is easy to develop and comprehend, thanks to the division of programs into pieces. Neural layers, cost functions, optimizers, and activation functions are the most prevalent creation model components. Python may be used to quickly construct new specified functions or classes.

III. RELATED WORK

Patel, M. M., Tanwar,[1] In their work, they have highlighted how time series models such as ARIMA, SARIMA, ARCH, and GARCH are widely used to investigate various benefits of implementing. They did, however, rely heavily on the time series forecasting method, which has apparent limitations to assumptions. As per the researchers, neural networks have shown promising results in time-series data prediction. They've presented a technique for predicting bitcoin prices incorporating GRU and LSTM models. Ji, S., Kim, J. [2] Only bitcoin estimation techniques based upon Bitcoin blockchain data were expected in this study. DNN, LSTM models, CNN, and deep residual networks are just a few of the deep learning techniques they've researched (ResNet). By a bit of margin, LSTM has outpaced models. In respect of logistic regression, all deep learning approaches tend to predict the Price movement quite well. Lahiri, S., & Bekiros, S.[3] In their work, they use deep learning to predict virtual currency prices. The research of unpredictable and irregular movements is vital to the predictability of non-linear systems. They have added to the tectonophysics literature by exploring the complicated properties of the three most widely traded digital currencies. Cryptocurrencies showed substantial self-similarity in training and testing subsamples when this segmentation was performed during deep learning processing. Livieris, I. E., Pintelas, E.[4] This study explains how to estimate cryptocurrency prices and movement using advanced deep learning models and ensemble learning strategies. Apart from that, it resembles deep learning and neural network models. Phaladisailoed, T., & Numnonda, T[5], In this investigation, the GRU findings show the best accuracy, although they require more time to calculate than Huber regression. Open, Close, High, and Low may not be sufficient to predict Bitcoin values since different aspects, including social sites reactions, rules, and legislation, require them to make accurate predictions. Aggarwal, A., Gupta, I.[6] This paper used RMSE values to compare the various deep learning models on a comprehensive analysis of multiple parameters affecting bitcoin price prediction. The findings reveal that the different deep learning models can adequately predict bitcoin prices. When a good tweet in bitcoin is written, the data of bitcoin price prediction using Twitter sentimental analysis demonstrate a positive link and are presented.

IV. PROPOSED WORK

A. Data

All data used in this research are everyday records in US dollars for Bitcoin, Ethereum, and Ripple, the three most valuable cryptocurrencies by market valuation. From the 1st of January 2017 until the 31st of October 2020, all bitcoin data was obtained from https://www.cryptodatadownload.com. The bitcoin data was divided into three sets for assessment: training, validation, and testing. The training set contained everyday period of January 1, 2017, to February 1, 2018. 18, 2020 (1381 data points), the validation set included data from March 1, 2020, to May 31, 2020 (94 data points). Data from June 1, 2020, to October 31, 2020 (152 data points) was included in the testing set, assuring a significant number of observed and in pieces of data during testing.

- Price: The day's average cryptocurrency price.
- Close: The price of cryptocurrencies at the end of the day.
- Open: the day's starting price of cryptocurrencies
- Low: Today's bitcoin price is the lowest it's ever been.
- High: Cryptocurrency's highest price of the day.
- Volume: The number of cryptocurrencies that were exchanged in a single day.

B. Model Description

Random Forest

Random Forest is a standard bagging method that involves training decision trees in parallel and aggregating the results. Decision trees are very dependent on the data used to train them, and the results can be drastically different when the training data is changed. As a result, the projections are constantly other. Furthermore, training decision trees is computationally expensive. Because they can't go backward after a split, it has a great danger of overfitting, and they tend to find local optimization rather than global optimization.

Lstm

The RNN version of Long Short-Term Memory (LSTM) is smart enough to learn long-term connections. The structure of LSTMs is very comparable to those of RNNs. However, the repeating unit is quite different. They feature four neural network layers that interact with each other rather than just one.

V. Model Training and validation

The validation split option in the Keras library's fit function is used to validate the model. The validation split is set at 0.2, a fraction between 0 and 1. The model will separate this portion of the training data from the rest and not be trained on the same data. The model will evaluate the loss and any model metrics on the data set aside at the end of each epoch. Epochs are fine-tuned using various models and values from the SK learn library's grid search cv, with the best value picked. Hyper-parameters like as the number of hidden layers, batch size, drop-out rates, and learning rate are all changed on different architectures to produce the best results. The random searching and grid research functions in the SK learn library are used to find the best values.

The error between the output and the supplied goal value is calculated using the loss function. It calculates our distance from the goal value. Many loss functions are dependent on the problem; thus, we must choose them carefully. The binary pass loss function is utilized in binary classification issues. It is based on the entropy approach, which demonstrates chaos or uncertainty. It is calculated using a probability distribution for a random variable X. A higher entropy value for a probability distribution indicates a more extended time until it is issued. On the other hand, a smaller value indicates a more certain distribution.

The first layer has 40 neurons in the first model, glorot uniform as the kernel initializer and tanh as the activation function. It is reduced to 20 neurons in the second layer, with tanh as an activation function. Drop-out is added at a rate of 0.3 between these two layers to add regularisation to the model and prevent overfitting. The model is built using the binary cross-entropy loss function and Adam optimizer, with an ideal learning rate of 0.005.

The second model contains 40 neurons in the first layer, with the uniform as the kernel initializer and relu as the activation function. It is reduced to 20 neurons in the second layer, and relu is used as an activation function. Drop-out is added with 0.2 between these two layers to regularise the model and prevent overfitting. It is reduced to 10 neurons in the third layer, and relu is used as an activation function. As the research objective is the classification of consumers into churners and non-churners, the output layer contains a component of sigmoid activation. The model was created using the binary cross-entropy gradient descent and the Adams algorithm, and it has an optimum learning rate of 0.01.

VI. RESULTS AND DISCUSSIONS

All comparison ensemble methods were tested for machine learning and data mining issues, specifically for forecasting the cryptocurrency price for the next hour (regression) and whether the cost will rise or decrease for the next hour (classification). According to our findings, including deep learning models into an ensemble learning framework improved predictive performance more often than using a single deep learning model. Bagging was the most accurate classification, followed by average and stacking(kNN); stacking(LR) had the best regression accuracy. The confusion matrices showed that the primary learner, stacking(LR), was biassed because the majority of the cases were mistakenly categorized as "Low," whereas bagging and stacking(CNN) exhibited a balanced forecast dispersion among "Lower" and "High," correspondingly. Predictions of "down" or "up." Since bagging can be understood as a perturbation approach to boost resilience, particularly against outliers and very volatile pricing, it's worth noting. The average classifier model based on a modified training dataset promotes normalization to such disturbances, and better reflects the direction movements of a randomness process represented, according to the numerical analyses.

The model should be tested on new data after it has been trained. Following the model's performance, it is evaluated using performance assessment criteria such as accuracy, ROC curve, etc. The plots are drawn between training and testing data, and the measure utilized is the accuracy of the model's predictions. The loss of training and testing data is depicted in figure 1.1. The cost of training begins at 40000 and steadily reduces to 8000 as the number of epochs increases from 0 to 100.

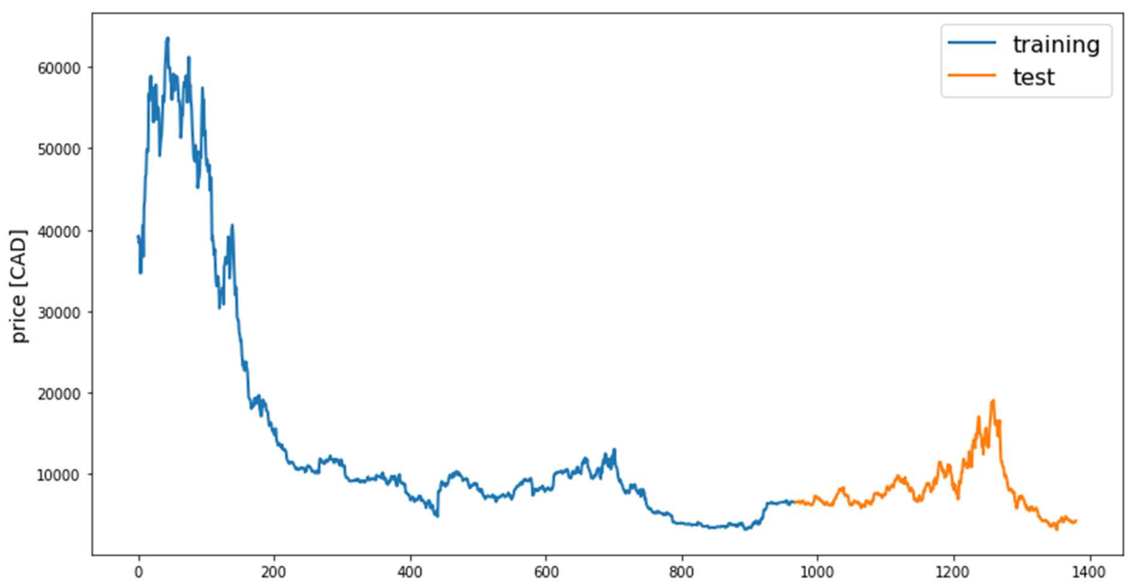

Fig.1. Training and testing data

The ROC curve of the classifier model is represented in figure.2. The ROC curve is a tool for evaluating binary classification issues. The mean square error is plotted on a probability curve. The Epochs values are shown on the ROC curve.

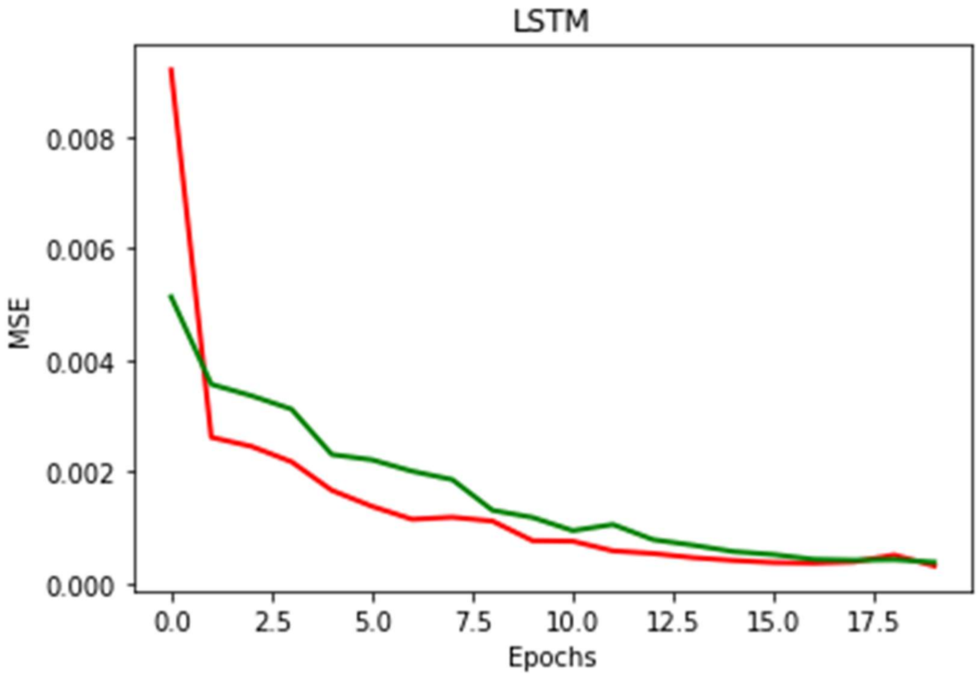

Fig.2. ROC Curve

For the first model, after training, the model is evaluated on unseen data. Both actual and prediction look similar. It shows models have predicted well. Will show the actual price and predicted price inline plot graph

Fig.3. Actual price and predicted price of LSTM

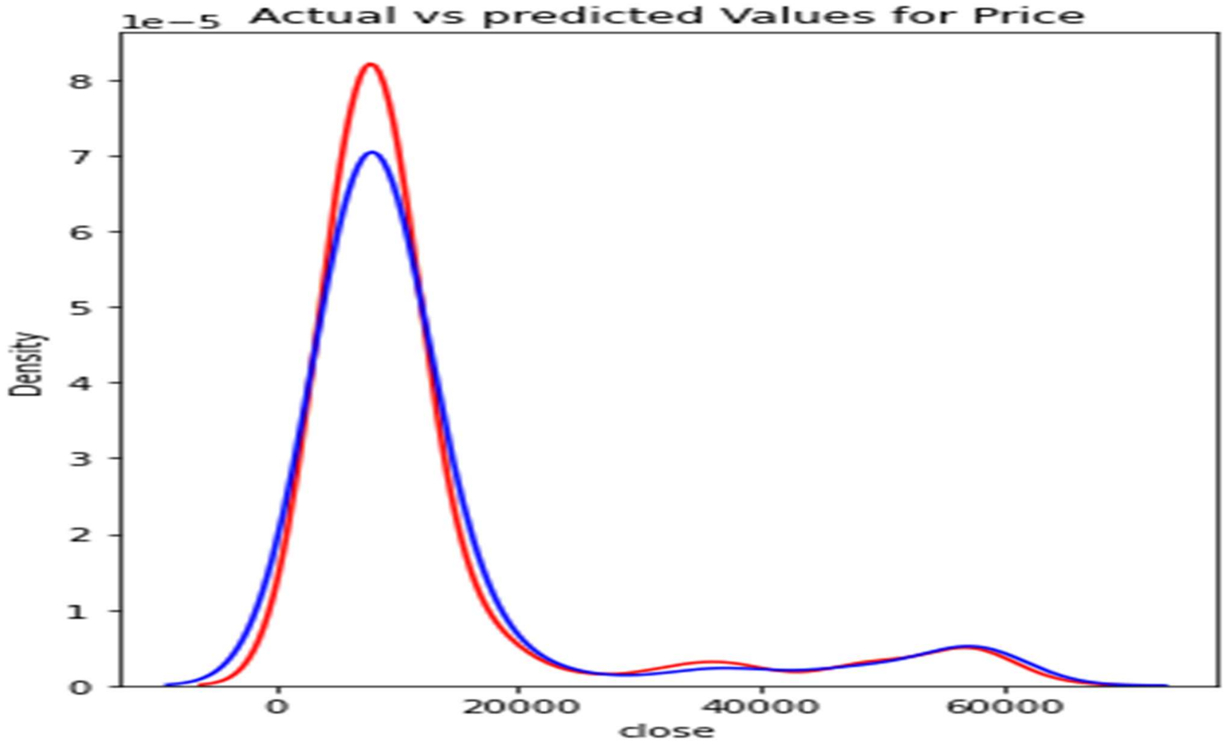

Fig.4. Actual price and predicted price of Random forest

	Actual	Predicted
654	11396.08	11418.75343
974	6465.12	6485.88480
445	8033.31	8424.61699
609	8063.73	8036.42211
459	9936.40	9845.11827
...	...	...
358	9666.24	9563.78123
724	8115.82	8509.32872
263	10256.20	10165.33374
511	6965.71	7038.67325
359	9518.04	9652.20406

415 rows × 2 columns

Fig.5. Actual price and predicted price of model -2

COMPARISON OF MODELS

models	Mean absolute error	Mean squared error	Root mean squared error	Accuracy
Lstm	**0.012922516470067806**	**0.00048548876771886**	**0.9722279097949948**	**97%**
Random forest	**242.38503812047594**	**322996.65109717037**	**568.3279432661835**	**98.43%**

Table. 1. COMPARISON OF MODELS

Mean Squared Error (MSE) and R-Square are two of the most often used metrics to measure the correctness of data set (R2). Table 1 shows the MSE and R2 of all of the algorithms we used, whereas table 2 shows the time we computed. The results indicate that regression models based on deep learning are effective. Such as Random Forest and LSTM outperform Theil-Sen and Huber regression. MSE is 0.01292, and R2 is 0.9722, or 97.2 percent, with the best results coming from LSTM. Compared to LSTM and random forest, Huber regression takes substantially less time to calculate. Random forest models are overfitting after cross-validation; also, accuracy is the same.

VII. Conclusion

This paper predicting bitcoin price and movement, this article developed a deep convolutional neural network based on the multi structure.. The proposed prediction model takes cryptocurrency data as an input and processes it separately, allowing each coin to be exploited and treated independently at first. Each cryptocurrency's data was compiled of inputs to numerous convolutional and LSTM layers, that are used to train the internal structure and discover short- and long-term relationships for each coin. The algorithm then integrates and analyses the recorded information from the LSTM layers' input vectors to produce the final forecast. It's worth noting that, to satisfy the stationarity property, all bitcoin time-series were transformed using the returns transformation. Mean Squared Error (MSE) and R-Square are two of the most common measures for measuring data set accuracy (R2). Table 1 shows the MSE and R2 for each of the models we used, while table 2 shows the computed time. Deep learning-based regression models such as Random Forest and LSTM, according to the findings, are more effective than traditional regression models. outperform Theil-Sen and Huber regression models. MSE is 0.01292, and R2 is 0.9722, or 97.2 percent, with the best results coming from LSTM. Compared to LSTM and random forest, Huber regression takes substantially less time to calculate.

References:-

1. Patel, M. M., Tanwar, S., Gupta, R., & Kumar, N. (2020). A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions. Journal of Information Security and Applications,

2. Ji, S., Kim, J., & Im, H. (2019). *A Comparative Study of Bitcoin Price Prediction Using Deep Learning. Mathematics, 7(10), 898.*

3. Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos, Solitons & Fractals, 118, 35–40. doi:10.1016/j.chaos.2018.11.014

4. Livieris, I. E., Pintelas, E., Stavroyiannis, S., & Pintelas, P. (2020). *Ensemble Deep Learning Models for Forecasting Cryptocurrency Time-Series. Algorithms,*

5. Phaladisailoed, T., & Numnonda, T. (2018). *Machine Learning Models Comparison for Bitcoin Price Prediction. 2018 10th International Conference on Information Technology and Electrical Engineering*

6. Aggarwal, A., Gupta, I., Garg, N., & Goel, A. (2019). Deep Learning Approach to Determine the Impact of Socio-Economic

7. D. Shah, K. Zhang, "Bayesian regression and Bitcoin," Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference. IEEE, 2014.

8. I. Madan, S. Saluja, and A. Zhao. "Automated Bitcoin Trading via Machine Learning Algorithms," 2015.

9. A. Greaves, B. Au, "Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin," 2015.

10. J. Almeida, S. Tata, A. Moser, and V. Smit, "Bitcoin prediction using ANN," June 2015

11. S. McNally, "Predicting the price of Bitcoin using Machine Learning" Diss. Dublin, National College of Ireland, 2016

12. S. Patro, K. Sahu. "Normalization: A preprocessing stage," arXiv preprint arXiv:1503.06462, 2015.

13. X. Dang, H. Peng, X. Wang, and H. Zhang "Theil-Sen Estimators in a Multiple Linear Regression Model," Olemiss Edu, 2008.

14. P. Huber. "Robust statistics." International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, 2011, pp. 1248-1251.

15. A. Bernal, S. Fok and R. Pidaparthi. "Financial Market Time Series Prediction with Recurrent Neural Networks." 2012.

16. S. Hochreiter, J. Schmidhuber. "Long short-term memory." Neural computation vol.9 no.8, 1997, pp. 1735-1780.

17. K. Cho, B. Van Merrienboer, C. Gulcehle, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio "Learning phrase representations using RNN encoder-decoder for statistical machine translation." arXiv preprint arXiv:1406.1078, June 2014.