Quest Journals
Journal of Software Engineering and Simulation
Volume 8 ~ Issue 3 (2022) pp: 08-12
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

A Model for Predicting Cryptocurrency Rates using hybrid Recurrent Neural Network-Long Short Term Memory

Olise, Oliseamaka Tamunobelema[1*], Anireh, V.I.E[2], Taylor, O.E[3].
[1]*(Department of Computer Science, Rivers State University, Port Harcourt, Nigeria)*
[2]*(Department of Computer Science, Rivers State University, Port Harcourt, Nigeria)*
[3]*(Department of Computer Science, Rivers State University, Port Harcourt, Nigeria)*
Corresponding Author: Olise, Oliseamaka Tamunobelema

ABSTRACT: *In this study, we assume that the network effects at work in cryptocurrency networks, play a significant role in helping to determine the rate of the cryptocurrency network. The study selected such features as the number of active addresses on the network, the transaction rates and also the hash rates associated with the network. These metrics generally can be used to measure the degree of participation by the network users. A time vector attention model was used as a way to positional encode temporal data and let the model know the position/sequence of the data points in the times series input so as to enable the model focus on the more recent and thus, more relevant data on the assumption that a data point at the beginning should affect the predicted output less than a more recent data point. The result of the attention based model developed was compared with a basic LSTM model, showcasing an improvement in forecast accuracy.*

KEYWORDS: *Recurrent Neural Network; attention; attention mechanism; Long Short Term Memory, Cryptocurrency; forecasting; prediction;*

R*eceived 03 Mar, 2022;* ***R****evised 14 Mar, 2022;* ***A****ccepted 16 Mar, 2022*
P*ublished with open access at* ***www.questjournals.org***

I. INTRODUCTION

A cryptocurrency is a network-based digital exchange medium, where the records are secured using strong cryptographic algorithms such as Secure Hash Algorithm 2 (SHA-2) and Message Digest 5 (MD5). Etc [20]. Money creation and transactions are controlled by mathematical algorithms (the so-called mining) implemented within the underlying protocol [7]. This is in contrast to a fiat currency which is any coin and/or paper money of any country that is established as legal tender by government decree or regulation and that circulates and accepted as a medium of exchange in the country of issuance. [17]. Fiat currencies do not have intrinsic value as they are not backed by physical commodities such as copper, silver or gold, but based on the creditworthiness of the issuing institution or government. Cryptocurrencies run on blockchains which are designed to be secure, transparent (except for privacy cryptocurrencies) and immutable. A blockchain is a distributed ledger (a distributed database management system/protocol) which maintains consensus about the existence and status of a shared set of facts but does not rely on this assumption of good faith by the transacting parties.[16]. Various cryptocurrencies use different consensus protocols/algorithms, which enable validation of the blockchain. This is necessary as blockchains have no centralized authority but still need to be able to determine which transactions are valid. [5][19][23]. Because of these characteristics of cryptocurrency networks and blockchains, cryptocurrencies have certain features which relate to the network and can help us understand the usage and size of the network. These features for the Etheruem network currently include, transaction rate, hash rate, number of addresses and the price of the cryptocurrency. The transaction rate tells us the number of transactions happening on the network within a set period and is a proxy for the use of the network. The more the transactions, the more in use the network. The hash rate is also a proxy for use but relates to miners which are the nodes that produce/enable the block creation in the cryptocurrency network. With miners mining on the network, the network will stop or cease to exist as no blocks will be produced meaning that no new transactions can be added to the chain. The number of addresses is a direct measurement of the growth of the network. (Please note: on cryptocurrency networks, a single user can have multiple addresses). The price of the cryptocurrency, this relates directly to the value of the network and therefore demand for the cryptocurrency. This is gotten from the supply and demand curve. As the demand for an item with a fixed or constant supply increase, the price of that item will also increase and vice versa.

Long Short Term Memory networks (LSTM's) are a type of Recurrent Neural Network which are Deep Neural Networks [11] particularly adapted to sequence data. Factors chosen as input to the LSTM model can be just as relevant in determining its predictive power. Our paper aims to use an attention-based LSTM model to predict the future value of a cryptocurrency rate, utilizing factors that act as a measure of the network growth as well as the price. These factors are transaction rate, hash rate and number of addresses.

II. RELATED WORK

Cryptocurrencies are a speculative asset and highly volatile with massive price fluctuations occurring in short time periods. Etheruem has seen a total return of around 66505.03% (*Ethereum* (ETH), 2020) since inception while bitcoin has seen a +99,999.99% since 11/01/2011 (*BTCUSD - Bitcoin - USD Cryptocurrency Performance Report - Barchart.com*, 2021). As such they present a challenge to the investor who wants to maximize his/her returns. Analysis of financial time series and prediction of future stock prices and future stock price movement patterns have been an active area of research over a considerable period of time. While some people believe that it is impossible to forecast stock prices with any degree of consistent accuracy, there exist in the literature that it is possible to predict the values of stock prices with a very high level of accuracy using optimally designed and fine-tuned models. [14][15]

Time series decomposition of stock price data is a popular approach for stock price forecasting [25], as is the use of machine learning and deep learning-based approaches are also adopted in some propositions [24][22][13]. The Recurrent Neural Network (RNN) is a Deep Neural Network (DNN) that is adapted to sequence data, and as a result the RNN is also extremely expressive. RNNs maintain a vector of activations for each timestep, which makes the RNN extremely deep. [12] Their depth, in turn, makes them difficult to train due to the exploding and the vanishing gradient problems when Recurrent neural networks with Long Short-Term Memory (LSTM) architecture where developed, which is resistant to the vanishing gradient problem. LSTM's solve the exploding and vanishing gradient problem present in simple RNNs [10]. LSTM's do this by using a set of gates which allow for a better control over the gradient flow. LSTM's have a high accuracy when it comes to predicting asset prices in forex and stock markets especially when compared to other methods such as regression methods, econometric approaches like ARIMA and other stochastic methods. LSTM's also outperform buy and hold when it comes to investing [11] making them ideal when it comes to increasing an investors performance. The central idea behind the LSTM architecture is a memory cell which can maintain its state over time, and non-linear gating units which regulate the information flow into and out of the cell. The LSTM is a recurrent neural network (RNN) architecture that remembers values over arbitrary / random lengths of time. As such it is well suited to process, classify and predict time series of unknown lengths. This Relative insensitivity to gap length gives an advantage to LSTM over alternative RNNs, other sequence learning methods or hidden Markov models. The LSTM Cell is controlled by three gates: the output, input, and forget gates. A nice metaphor for these gating function is that of a differentiable if statement, where c is the condition and v is the value. If χ is true (χ # 0), v will be passed on; if it is false (χ # 0), it will not. All the operations on the gates are differentiable, which is why gradient-based learning can be employed. Within the cell, the gradient of a state with respect to its immediate predecessor, is the only part where gradients flow through time[3]. In the functioning of an LSTM cell, the first step is, deciding what to forget and what to keep. This is done using a sigmoid in the forget gate layer.

The next step, step two, involves deciding what new data o be added back to the cell state. This involves a sigmoid gate layer deciding what to change and a Tanh gate layer creating candidates values that could be added to the cell state. The third step involves updating the cell state with the information from steps 1 and 2. and the fourth step involves the output of the modified version of the cell state. [18][26]. Various versions of the LSTM have been developed incrementally over time. the first version came only the input and output gate. The second version, introduced the forget gate[8]. Additional "peephole" connections were added by [9]. with other additional variations being introduced later by [2]. The insight that many different architectures work well inspired evolving alternative memory cells [4]. Including other modifications to the LSTM in order to increase the model performance. In their work, [1] showed that LSTM models can be improved by modification of the activation function. While [6] showed that the accuracy of an LSTM network in predicting prices can be improved by adding an attention mechanism to the LSTM network. Assigning attention weights on neural networks has been shown to achieve great success in various machine learning tasks this includes in stock prediction tasks as evidenced by the use of a dual-stage attention-based RNN namely DA-RNN using both an encoder-decoder and attention mechanism which was employed by [21] in their work, "A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction.". The LSTM model has thus emerged as a major machine learning method employed when dealing with sequential data such as time series data fore predicting rates of stocks and forex pairs and the burgeoning cryptocurrency market though different in characteristics. (It is a 24/7 market) is no exception.

III. METHODOLOGY

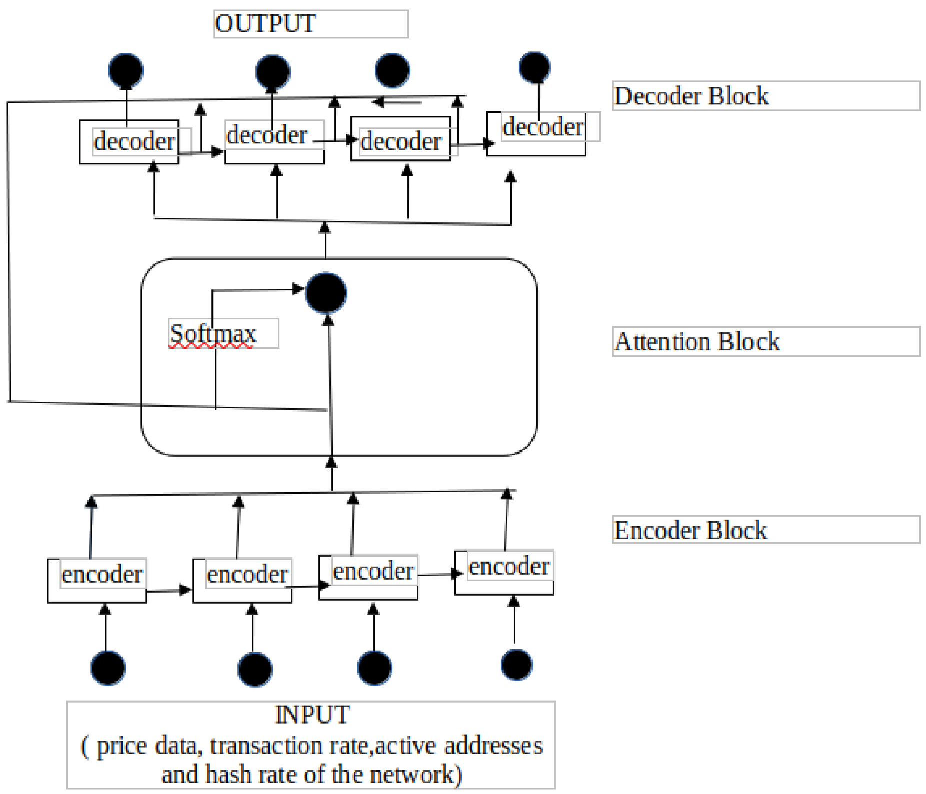
In this work, to develop the model required some parameters such as

3.1 Data Set : The dataset used is end of day price data for the cryptocurrency Ethereum. This data was gotten from Cryptoquant. This data is required to train and validate the LSTM network.

3.2 Data Pre-processing : The data was prepared in a manner that the neural network understands. Null values were removed and the data was normalized before being fed into the neural network. When the output is gotten, the output is first denormalized to get the correct value. To train the network, the data is split into 2 subsets, A set of train data and a set of test data.

3.3 Model : The model is an encoder-decoder model with an attention module between them.

Below is a top level view of the architecture of the model



The encoder-decoder model uses an LSTM with an input layer that has a timestep (sequence length) of 50 with the 4 features. The Attention layer is a time to vector layer. The time to vector layer provides a means for the transformer to understand temporal positioning of the datapoints. It has two components the periodic and non-periodic/linear components. so as to be able to represent the periodic and non-periodic patterns. This can be represented by the Mathematical equation given below:

$$\mathbf{t2v}(\tau)[i] = \begin{cases} \omega_i \tau + \varphi_i, & \text{if } i = 0. \\ \mathcal{F}(\omega_i \tau + \varphi_i), & \text{if } 1 \leq i \leq k. \end{cases}$$

t2v = Time to vector representation

$\boldsymbol{\omega_i \tau + \varphi_i}$ = non-periodic/linear component of the time vector

$F(\omega_i \tau + \varphi_i)$ = periodic/non-linear component of the time vector

ω (omega) in $\omega_i \tau + \varphi_i$ = a matrix defining the slope of the timeseries

φ_i = the intersection of the timeseries τ with the y-axis

3.4 Experiment : The model developed was used on Etheruem data and was trained on the train data. In order to provide a comparison to a simple LSTM model, a basic LSTM model was developed ands also used on the same data.

IV. RESULTS AND ANALYSIS

Graphical results of the experiment are as follows:

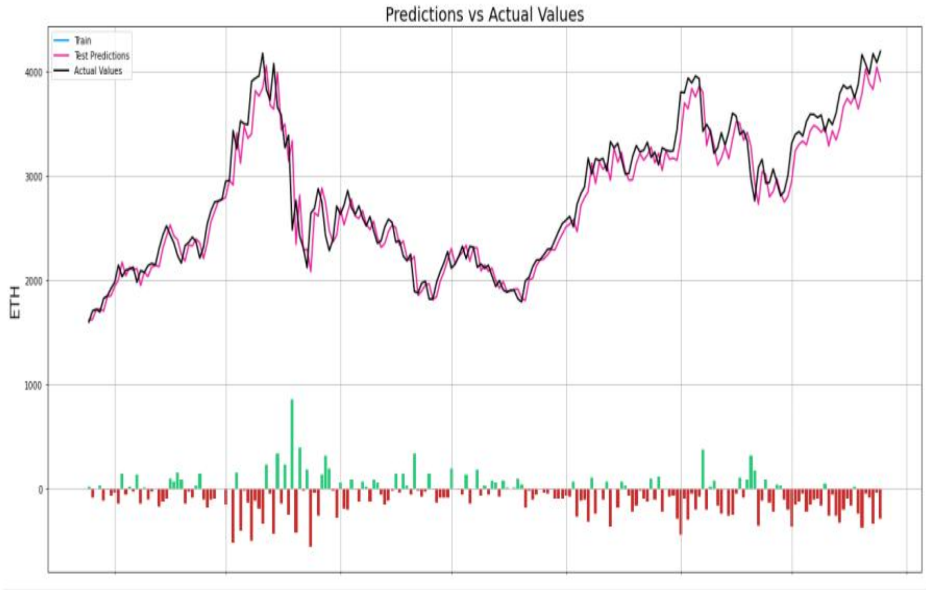


Figure 1: Result of running the model on the test portion of the data. Predicted values are in purple, actual values in black.

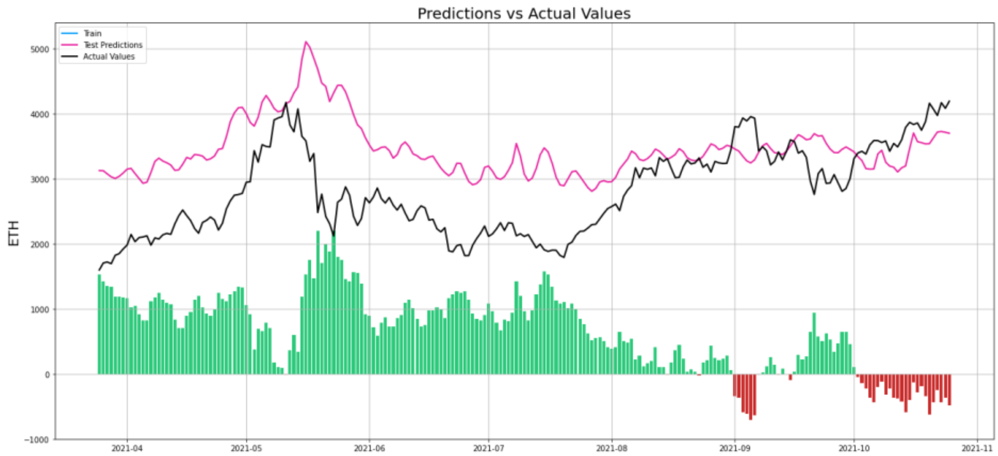


Figure 2: Result of running a standard LSTM model on the test portion of the data. Predicted values are in purple, actual values in black.

Table 1: Comparison of the error statistic between the basic LSTM system and the Model system developed (the proposed system)

Error Statistic	LSTM System	Proposed System
Mean Absolute Error (MAE)	359.26	81.87
Mean Absolute Percentage Error (MAPE)	14.12%	4.44%
Median Absolute Percentage Error (MDAPE)	11.82%	3.6%
Root Mean Square Error (RMSE)	469.88	131.02

V. CONCLUSION

This study aimed to develop and implement an attention based model for predicting cryptocurrency rates. the study used the cryptocurrency Etheruem as a case study and included as features, previous price, the number of active addresses, the transaction rates and the hash rates. Using a multi-head attention model with time-embeddings the model developed was determined to be more accurate at predicting the next price point as shown in the results above.

Further experimentation and possibly improvements can be obtained by optimizing the hyperparameters. If better resources are available, the model can be run with more data and higher epochs, thereby proving more insight into the hyperparameters and the best optimizations to use.

REFERENCES

[1]. Anireh, V. & Osegi, E. N. (2016). A Modified Activation Function with Improved Run-Times For Neural Networks. ArXiv. https://www.semanticscholar.org/paper/A-Modified-Activation-Function-with- Improved-For-Anireh-Osegi/c328656aa2de2bed162730084c7e571bdc675aa8
[2]. Bahdanau, D., Cho, K. & Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate. Https://arxiv.org/pdf/1409.0473.pdf
[3]. Bayer, J. S. (2015). Learning sequence representations.
[4]. Bayer, J., Wierstra, D., Togelius, J. & Schmidhuber, J. (2009). Evolving Memory Cell Structures for Sequence Learning. In Artificial Neural Networks – ICANN 2009. Lecture Notes in Computer Science, vol 5769. Springer, Berlin, Heidelberg.
[5]. Bouraga, S. (2021). A taxonomy of blockchain consensus protocols: A survey and classification framework. Expert Systems with Applications, 168, 114384.
[6]. Cheng, L., Huang, Y. & Wu, M. (2018). Applied attention-based LSTM neural networks in stock prediction. 2018 IEEE International Conference on Big Data (Big Data). https://doi.org/10.1109/bigdata.2018.8622541
[7]. De Filippi, P. (2015). Translating Commons-based peer production values into metrics. Handbook of Digital Currency, 463-483.
[8]. Gers, F. A., Schmidhuber, J. & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. Neural Computation, 12(10), 2451–2471.
[9]. Gers, F., Schraudolph, N. & Schmidhuber, J. (2002). Learning Precise Timing with LSTM[26] Recurrent Networks. J. Mach. Learn. Res., 3, 115-143. Journal of Machine Learning Research, 3, 115-143.
[10]. Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R. & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222–223
[11]. Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R. & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222–2232.
[12]. Hu, Z., Zhao, Y. & Khushi, M. (2021). A Survey of Forex and Stock Price Prediction Using Deep Learning. Applied System Innovation, 4(1), 9.
[13]. Jozefowlcz, R., Zaremba, W. & Sutskever, I. (2015). An Empirical Exploration of Recurrent Network Architectures. ICML.http://proceedings.mlr.press/v37/jozefowicz15.html
[14]. Kaushik, R., Jain, S., Jain, S. & Dash, T. (2020). Performance evaluation of deep neural networks for forecasting time-series with multiple structural breaks and high volatility. arXiv preprint arXiv:1911.06704, 2019 - arxiv.org
[15]. Mehtab, S., Sen, J. & Dasgupta, S. (2020). Robust analysis of stock price time series using CNN and LSTM-based deep learning models. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA).
[16]. Mehtab, S., Sen, J. & Dutta, A. (2021). Stock price prediction using machine learning and LSTM-based deep learning models. Communications in Computer and Information Science, 88-106.
[17]. Meunier, S. (2018). Blockchain 101. Transforming Climate Finance and Green Investment with Blockchains, 23-34.
[18]. Nelson, J. (2018). Cryptocurrencies as Non-Sovereign Fiat Currencies. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3308645
[19]. Olah, C. (2015). Understanding LSTM Networks -- colah's blog. Github.io. https://colah.github.io/posts/2015-08-Understanding-LSTMs/
[20]. Pahlajani, S., Kshirsagar, A. & Pachghare, V. (2019, April). Survey on Private Blockchain Consensus Algorithms. 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT). https://doi.org/10.1109/iciict1.2019.8741353
[21]. 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, 55(1), 102583. https://doi.org/10.1016/j.jisa.2020.102583
[22]. Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G. & Cottrell, G. W. (2017). A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence.
[23]. Qiu, J., & Zhou, C. (2020). Forecasting stock prices with long-short term memory neural network based on attention mechanism. PLoS ONE 15(1):e0227222.
[24]. Samy, H., Tammam, A., Fahmy, A. & Hasan, B. (2021). Enhancing the performance of the blockchain consensus algorithm using multithreading technology. Ain Shams Engineering Journal.
[25]. Sen, J. (2018). Stock price prediction using machine learning and deep learning frameworks [Paper presentation]. International Conference on Business Analytics and Intelligence, Bangalore, India.
[26]. Sen, J. & Chaudhuri, T. D. (2016). An investigation of the structural charactersictics of the Indian IT sector and the capital goods sector – an application of the R programming language in time series decomposition and forecasting. Journal of Insurance and Financial Management, 1(4), 68-132.
[27]. Staudemeyer, R. C. & Morris, E. R. (2019). Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. ArXiv: 1909.09586 [Cs].