Journal of Multi Disciplinary Engineering Technologies
Volume 18, Issue 01

Time-Series Forecasting and Anomaly Detection in Ethereum: A Deep Learning Approach

Aakash Sharma[1], Kunal Gulia[1], Nikunj Jindal[1], Ish Narayan Tripathi[1], Gargi Mishra[1*]

[1]Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi, 110063, India.

*Corresponding author(s). E-mail(s): gargi.mishra@bvcoend.ac.in;
Contributing authors: akashsharma19200@gmail.com; kunalgulia02@gmail.com; nikunjjindal13@gmail.com; ish.narayan.tripathi06@gmail.com;

Abstract

Cryptocurrencies markets are still heavily volatile and provide a unique activity field for more classical financial analysis. The work entails taking historical price data on Ethereum for a vast amount of time. Feature engineering is done with the intention of getting the right information that needs to be taken for the modeling process and to make the data normal for the process. The LSTM model is then used and trained on the predefined dataset because of the capability of the LSTM model for mastering complex non- linear patterns in Ethereum price fluctuations. In this paper, the accuracy of the LSTM model in the prediction of Ethereum prices is determined using the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Information derived from this model can help patients, producers, investors and other stakeholders make appropriate decision-making in this ever volatile crypto-market. The results of this research advance the literature on using deep learning approaches for predicting challenging financial dynamics, focusing on cryptoassets. The deep autoencoder applied in this work also deals with anomaly detection on the Ethereum system. In particular, autoencoder is useful for reconstructing normal patterns; therefore, deviations in transactions and prices can be detected. Such an approach is rational, preventive and based on analysis of risks and threats in the environment of cryptocurrencies, increasing their safety and dependability.

Keywords: Ethereum Price Prediction, LSTM Networks, Anomaly Detection, Deep Autoencoders, Cryptocurrency Time-Series Forecasting

1 Introduction

Cryptocurrencies may also be described in terms of a relatively new category of financial assets, which cause interest among investors, traders, and scholars. Among these, Ethereum (ETH) can be stigmatized being not only one of the most important digital currencies hosting smart contracts but also one of the most fluctuating ones. Since the cryptocurrency markets are growing to new heights

Published on: November, 2024

and innovations every day even every minutes, it is crucial to determine the price estimate for each cryptocurrency in the trading market for better decision-making in trading and investments.

Recurrent Neural Networks (RNNs) and their type, LSTM networks, solve the problem of detecting patterns into data. This paper will attempt to analyze the effectiveness of LSTM network in the prediction of price of Ethereum with the view of pointing out the direction this virtual currency can take in the future. The research focuses on gathering past price data of Ethereum to understand price patterns and trends effectively. LSTM networks should help strengthen the general understanding of the highly fluctuating cryptocurrency market in this study.

Bearing in mind different evaluation parameters and measures, this work assesses the effectiveness of LSTM networks for Ethereum price prediction. In addition to the accuracy and effectiveness of the model, the study discusses the readability of the outcomes providing insights into factors affecting Ethereum's market. Legacy financial analysis tools present one of the challenges of applying them in this dynamic digital currency environment where daily changes are typical, the currencies exposed to various influences.

In this respect, accurate technologies and computerized networks, to mention but a few, deep learning methods like LSTM are valuable assets. Incorporating the LSTM networks among other predictors make the prediction of price of cryptocurrencies more accurate. In addition, the analysis of future stabilization of Cryptocurrencies provide insight to the body of knowledge focusing on the efficacy of deep learning models in the field of finance.

The outcomes of this research hold potential value for various stakeholders, including investors, market analysts, and anyone looking to invest in digital assets. Scholars interested in understanding the factors influencing Ethereum's price dynamics in the context of digital assets will also find these insights valuable.

2 Related Work

A study titled *Cryptocurrency Time-Series Forecasting with Ensemble Deep Learning Models* [1] introduced a model combining LSTM, BiLSTM, and convolutional layers, which helps in building robust, steady, and reliable forecasting models. Another research, *LSTM Based Stock Price Prediction* [2], produced accurate stock price forecasts using datasets from Alcoa Corp, Carnival Corp, Tesla Corp, and Google Corp, achieving results comparable to actual stock prices.

In the study *Predicting the Price of Bitcoin Using Machine Learning* [3], the authors explored deep learning models like RNN and LSTM for Bitcoin price prediction. Although LSTM generally performs better due to its ability to identify long-term relationships, it requires more training time compared to RNN. Further, *Anomaly Detection in Time Series Data of Sensex and Nifty50 With Keras* [4] employed LSTM models to detect anomalies in the Sensex and Nifty50 indices, showing that GRU models outperformed LSTM when analyzing high, close, and open prices.

A study from Stanford University, *Predicting Price Changes in Ethereum* [5], compared various machine learning models—logistic regression, naive Bayes, SVM, random forest, and ARIMA—to predict Ethereum prices. ARIMA was found to provide the most accurate predictions. Similarly, the work *Ethereum Analysis and Predictions Utilizing Deep Learning* [6] employed six types of neural networks and found that LSTM and GRU achieved the highest accuracy at 71%.

In the paper *Cryptocurrency Forecasting with Deep Learning Chaotic Neural Networks* [7], RMSE was calculated using Lyapunov's nearest neighbor rating for predicting Bitcoin, Digital Cash, and Ripple values. A separate study, *Using GRU for Predicting Ethereum Price* [8], developed a GRU-based model for cryptocurrency price prediction. GRU, being simpler and capable of retaining long-term memory, produced favorable results compared to other models and proposed ways to reduce information leakage for improved accuracy.

In *LSTM Based Crypto Prediction* [9], the researchers used both LSTM and MLP models to forecast Ethereum prices, analyzing 1000 data points daily, 1500 per hour, and 400,000 per minute. The study *Bitcoin Price Prediction Using Machine Learning* [10] applied random forest and Bayesian

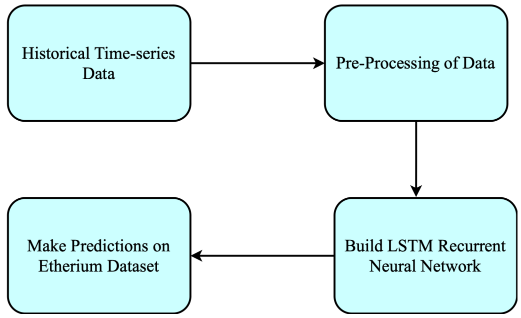

Fig. 1 Basic Flow of Proposed Model

regression algorithms, predicting Bitcoin prices with high accuracy after developing a framework to identify daily market patterns.

The work *Enhancing Bitcoin Price Fluctuation Prediction Using Attentive LSTM and Embedding Network* [11] presented a model that utilized autoencoders and basic features with Attentive LSTM Networks to enhance Bitcoin price predictions. The ALEN model excelled in time series forecasting. Another study, *Using Deep Learning for Predicting Bitcoin Prices: A Comparative Analysis* [12], reviewed several deep learning techniques such as DNNs, LSTMs, CNNs, ResNets, and CRNN combinations, concluding that DNNs were the most accurate.

In *Analyzing Machine Learning Algorithm's Performance to Predict Bitcoin Prices* [13], the authors compared linear regression, LSTM, and RNN for predicting Bitcoin prices, finding that the LSTM-RNN combination was more effective due to its ability to capture long-term trends. Lastly, *Neural Network Prediction of Changes in the Price of Bitcoin* [14] used RNN and MLP models for short- and long-term predictions, with MLP achieving 80% higher accuracy in long-term forecasts.

3 Methodology

Our proposed methodology consists of a LSTM network to predict cryptocurrency prices. Historical time series data is fetched from Binance, and the LSTM algorithm is built and implemented. The model then predicts the price of the cryptocurrency. The basic flow of this process is shown in Figure 1.

Recurrent Neural Networks (RNNs) with LSTM are commonly used for time series analysis and forecasting, particularly in scenarios where traditional RNNs face challenges, such as the vanishing gradient problem. This issue can also lead to inaccuracy of pattern recognition. LSTM networks manage this through the use of memory cells which are then stacked and connect through gateways. These gates regulate the in and out transportation of information within these cells. The Memberlosz input, forget, and output gates allows the network to update and output data only selectively.

LSTM networks have found application in image captioning, speech recognition and in handling natural languages. It has also successfully used in exploiting the future prices of cryptocurrency. Compared to other traditional models, LSTM network extracts features from historical price data, and then can be used for forecasting future price. To be more precise, the realized prices are input and the network returns the estimated price at some future date.

The basic structure of LSTM consists of three key components, each serving a specific purpose in data analysis and prediction:

- **Forget Gate**: This gate determines whether information from the previous cell should be retained or discarded.
- **Input Gate**: This gate processes new data to update the current cell's state.
- **Output Gate**: This gate outputs the current data, which is passed to the next time step.

These gates work together to enable the LSTM network to effectively learn and predict future cryptocurrency prices.

The equation for the forget gate is given by Equation 1.

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \tag{1}$$

The equations for the input gate are given by Equation 2, 3, 4.

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \tag{2}$$

$$\hat{C}_t = tanh(x_t U^g + h_{t-1} W^g) \tag{3}$$

$$C_t = \sigma(f_t * C_{t-1} + i_t * \hat{C}_t \tag{4}$$

The equations for the output gate are given by Equation 5, 6.

$$o_t = \sigma(x_t U^o + h_{t-1} W^o \tag{5}$$

$$h_t = tanh(C_t) * o_t \tag{6}$$

The methodology for anomaly detection in Ethereum price prediction using deep autoencoders involved several key steps. Firstly, we collected historical Ethereum price data and then preprocessed it to normalize the data and split it into training and test sets. A deep autoencoder model was designed with an encoder-decoder architecture, where the encoder reduces the dimensionality of the input data and the decoder reconstructs the original input. The model is trained using the training data to learn normal price patterns. Subsequently, the trained autoencoder is used to reconstruct the test data, and the reconstruction errors are calculated. A threshold based on the training errors flags outliers as anomalies. Data points with reconstruction errors above this threshold are flagged as anomalies. Finally, the detected anomalies are visualized on a price chart to gain insights into unusual price movements or behaviors.

4 Implementation

The programming language chosen for this work is Python 3.0, which possesses a robust and facile setting and encompasses numerous libraries to support ML/DL. A variety of methods and different tasks are used in machine learning and artificial intelligence (AI) and Python is effective, but reliable.

4.1 Data Preprocessing

Our model uses historical data to predict the closing price of Ethereum for the next day. To utilize LSTM networks, the input must be structured as a 3D tensor with dimensions (batch_size, window_size, input_dim). The number of past days that the model will consider needs to be defined. While this choice is somewhat arbitrary, we will opt for a window size of 10 days. We reorganized the data so that a sequence of values from the previous 10 days is used to predict the value at time t.

We construct small data frames, known as windows, comprising data from 10 consecutive days. For instance, the first window includes rows 0-9 of the training set, the second contains rows 1-10, and so on. Choosing a small window size allows us to input more windows into our model, but it may limit the model's ability to detect complex long-term patterns due to insufficient information. In this case, we only utilized the "close" price to make predictions, resulting in an input dimension (input_dim) of 1. However, the target variable, the "close" price of Ethereum, can be influenced by numerous other factors.

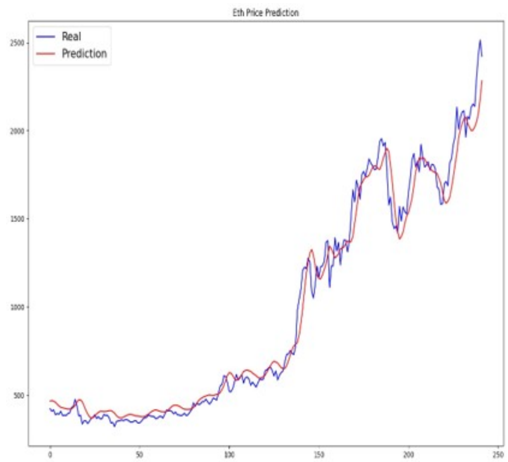

Fig. 2 Ethereum 250 days Price Prediction

Additionally, we scaled our price data to fit within a specific range. This will assist our optimization algorithm in reaching convergence more quickly. Normalization is a commonly used method in preparing data for machine learning. Its aim is to adjust the values of numerical columns in a dataset to a standard scale while ensuring that variations in the ranges of values are preserved. For this purpose, we used the MinMaxScaler from `sklearn`. The MinMaxScaler adjusts features by scaling each one to a specified range. This process scales and shifts each feature independently to fit within the desired range.

4.2 Error metrics to evaluate the results

Root Mean Squared Error (RMSE) metric calculates the square root of the mean of the squared errors between the actual and predicted values. The formula for the same is given by Equation 7.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predicted_i - Actual_i)^2}{N}} \tag{7}$$

Mean Absolute Error (MAE) calculates the difference between the actual and predicted values, and then takes the mean of these absolute values. The equation for the same is given by Equation 8.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \tag{8}$$

5 Results

The resulting graph clearly shows how well our LSTM model predicts Ethereum prices. The blue line serves as a comparison point and represents the actual historical pricing of Ethereum, while the predictions made by the model are represented by the orange line. The two lines should ideally align closely to indicate that the forecasts are accurate; deviations suggest potential areas where the model may not adequately reflect market behavior. Understanding the general trends and specific regions of varied responses or convergence can help assess the model's forecasting effectiveness. It is noteworthy that, despite the inherent challenges in forecasting cryptocurrency, the model demonstrates its ability to reflect actual prices due to its focus on value. This is visually represented, highlighting its potential effectiveness in the cryptocurrency market.

We utilized two popular assessment measures: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) which provide prediction accuracy of our LSTM model specifically. RMSE

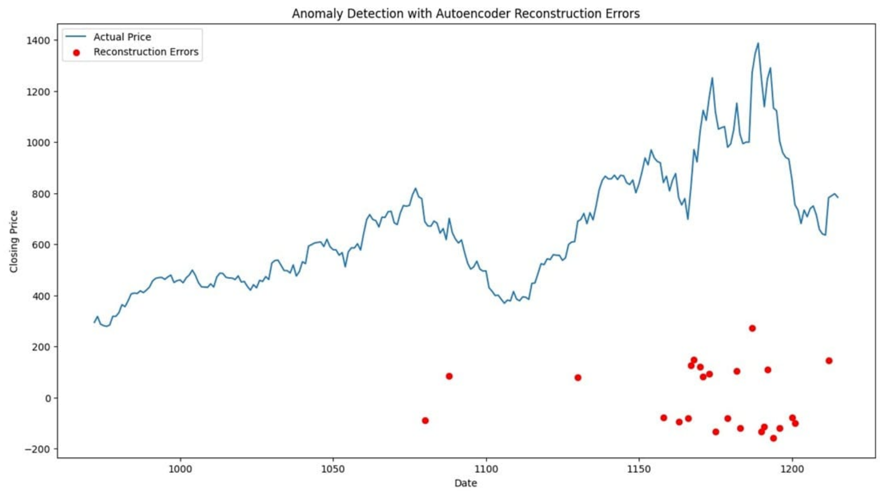

Fig. 3 Reconstruction Errors for Anomaly Detection with Autoencoder

exposes the qualities of the model which shows the average size of the errors from the actual value. On the other hand, MAE makes the representation of the absolute errors easier. The values we got for assessments were RMSE = 0.0013757152836841424 and MAE = 0.026190283157309237.

Using these indicators, effectiveness of the model can be properly instituted which gives insight of effectiveness of Ethereum price prediction. These indicators can therefore used as benchmarks when comparing the extent to which the identified model reflects the essential characteristics and changes in the cryptocurrency market. Again for anomaly detection, the result graph shows the reconstructed errors in parallel with the actual prices. The anomaly detection model is actually tested on Ethereum price data. The model applied autoencoder type with a multiple layers and it operated on initial portion of the data and tested on remaining data. Measurement errors of reconstruction for each data number were estimated. The model was able to discern trends in the Ethereum price data and is suggestive of its ability to point out any strange price bouts or fluctuations.

6 Conclusion

In other words, this study yields positive and helpful findings in predicting Ethereum prices utilizing LSTM models. Using past price movement to forecast, LSTM has shown strong analytical efficacy in showing finer fluctuations and correlation within cryptocurrencies' volatile nature. That means incorporating metrics like RMSE and MAE as a reference point suggest this model's aptness to measure and estimate changes in price over time. This data predictive capability is especially significant within the trading of cryptocurrencies since these are extremely volatile markets where trading can occur almost in the blink of an eye. However, it is essential to pay attention to the range of barriers and challenges present within the confines of natural estimate fundamental to cryptocurrencies. Market outlooks, events and acts of regulation amongst other factors are outside the realm of historical confirmation, and therefore explain future movements of price variations. Specifically, for the Ethereum cryptocurrency such as anomaly detection in the form of deep autoencoders is considered the promising approach to identifying the inconsistencies and potential fraud in the network. Since autoencoder preserves the normal patterns, the differences from these patterns can be considered as anomalies. This methodology offers a preventive/anticipative and quantitative approach toward managing risks associated with irregularities of Ethereum transactions as well as changes in their prices. Last but not the least, deep autoencoders in context of anomaly detection on Ethereum show promising results in exploring the security and utility of the crypto-currency domain. More investigations in this area, as well as future development of this field could provide a better tool for improved identification of anomalies, reinforcing the reliability and credibility of Ethereum and a variety of other Blockchain systems.

Declarations

- The authors received no specific funding for this study.
- The authors declare that they have no conflicts of interest to report regarding the present study.
- No Human subject or animals are involved in the research.
- All authors have mutually consented to participate.
- All the authors have consented the Journal to publish this paper.
- Authors declare that all the data being used in the design and production cum layout of the manuscript is declared in the manuscript.

References

[1] Livieris, I.E., Pintelas, E., Stavroyiannis, S., Pintelas, P.: Ensemble deep learning models for forecasting cryptocurrency time-series. MDPI (2020)

[2] Ahire, P., Lad, H., Parekh, S., Kabrawal, S.: Lstm based stock price prediction. International Journal of Creative Research Thoughts (IJRTI) (2021)

[3] McNally, S., Roche, J., Caton, S.: Predicting the price of bitcoin using machine learning. In: 26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (2018)

[4] Shah, D., Khade, S., Pawar, S.: Anomaly detection in time series data of sensex and nifty50 with keras. In: International Conference on Emerging Smart Computing and Informatics (ESCI) (2021)

[5] Chen, M., Narwal, N.: Predicting Price Changes in Ethereum (2017)

[6] Zoumpekas, T., Houstis, E., Vavalis, M.: Eth analysis and predictions utilizing deep learning. Expert Systems with Applications **162**, 113866 (2020)

[7] Lahmiri, S., Bekiros, S.: Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos, Solitons Fractals **118**, 35–40 (2019)

[8] Balu, G., Ali, S.: Ether price prediction using gated recurrent unit. International Science and Technology Journal (2018)

[9] Kumar, D., Rath, S.K.: Predicting the Trends of Price for Ethereum Using Deep Learning Techniques (2020)

[10] Velankar, S., Valecha, S., Maji, S.: Bitcoin price prediction using machine learning. In: 2018 20th International Conference on Advanced Communication Technology (ICACT), pp. 144–147 (2018)

[11] Albariqi, R., Winarko, E.: Prediction of bitcoin price change using neural networks. In: International Conference on Smart Technology and Applications (ICoSTA) (2020)

[12] Ji, S., Kim, J., Im, H.: A comparative study of bitcoin price prediction using deep learning. Mathematics **7**(10), 898 (2019)

[13] Rizwan, M., Narejo, S., Javed, M.: Bitcoin price prediction using deep learning algorithm. In: 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS) (2019)

[14] Li, Y., et al.: Enhancing bitcoin price fluctuation prediction using attentive lstm and embedding network. Applied Sciences **10** (2020)