SHORT-TERM BITCOIN PRICE FORECAST USING VECTOR REPRESENTATION OF TIME

[1]ABHISHEK SINGH KUSHWAH, [2]DEEPANSH GANDHI

[1]Department of Computer Science and Engineering, University Institute of Technology - Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India
[2]Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India
E-mail: [1]abhishek18kushwah@gmail.com, [2]deepanshgandhi1404@gmail.com

Abstract - Bitcoin is a decentralized digital cryptocurrency that can be sent from user to user on the peer-to-peer bitcoin network without the use of intermediaries. The main problem with these types of cryptocurrencies is price volatility. This paper demonstrates a high-performance model with Univariate Time Series Data and applying Regression using an invariant vector representation of data, trained on different time ranges in data (2015 to 2021 and 2019 to 2021), results show improvement over the current literature.

I. INTRODUCTION

Bitcoin is the most popular cryptocurrency in the world right now. Cryptocurrency allows users to send and receive digital currency over the Internet in a secure and anonymous manner. Because of speculation, Bitcoin's price is extremely volatile. While Bitcoin can be used to purchase tangible products in some places, the vast majority of Bitcoin transactions are still financial in nature. As a result, Bitcoins are bought and sold just like any other investment. The buy-sell cycle is what causes Bitcoin's price to fluctuate so much.

Despite its high volatility, many researchers have been working in the field of cryptocurrency price prediction and have demonstrated excellent results. These studies involve different machine learning methods for end-of-day price forecast and increase/decrease prediction.

In this paper, we present a new approach to Forecast on Univariate Time Series Data (Price) and predict Bitcoin price changes by the hour.

Our results have exceeded previous research and use a novel architecture that has never been used before.

II. RELATED WORK

The popularity of bitcoin as a digital cryptocurrency began to surge in 2014, resulting in price volatility and a high volume of daily transactions.

A variety of literature have analyzed the movement of prices of bitcoin using several indicators such as social media traffic[1] and high dimensional features of historical data[2].

There are basically two types of model-based approaches in time series forecasting: statistical and neural networks.

The statistical models for time series data, such as ARIMA[3] have shown encouraging results in generating short term forecasts but perform poorly on data with high variance, which is due to its inability to learn non-linear patterns and non-stationarity in data.

The application of neural networks to time series data forecasting has yielded promising results, particularly when RNNs are used (Recurrent Neural Networks).
[5] compared the results using ARIMA and an RNN with Bayesian Optimisation coupled with an LSTM Cell.
[7] used RNN and GRNN to obtain price predictions for a set of cryptocurrencies considering high liquidity.

The use of rapid fourier transforms and fuzzy transforms[8] as a technique to represent data into vectors and utilise that as an embedding into NN based models was the beginning of alternate representation in time series.
However, with these transformations, data is represented using a predefined set of frequencies that cannot be learned.

To learn representations of data similar to contextual vectors in computer vision and natural language processing, Time2Vec[6] was introduced. Our goal is to feed the architecture with vector representations created by Time2Vec[6] implementation.

III. METHODOLOGY

The proposed methodology describes the elements of the architecture(in figure 2).

The dataset used for our research was downloaded from CyptoDataDownload. This website provides historical cryptocurrency price data in time series format for three different time intervals: daily, hourly, and minute! Each time series is arranged by cryptocurrency exchange and includes data for the Opening price, High price, Low price, and Closing price (OHLC format) as well as volume statistics.

A. Data Preparation

We have used hourly interval data for Bitcoin prices and have limited ourselves to predict Close Prices for bitcoin in that particular hour.

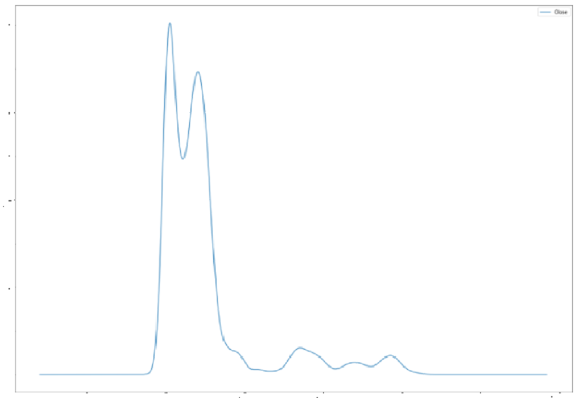


Figure 1 - The density plot of the dataset shows that the data is highly skewed.

Due to this invariability in the data, we experimented with a median based replacement based on the IOR quantile score of the data (for values of standard deviation less than 25% and more than 75%) but found that the results were almost similar and decided to drop the replacement.

We divided the data into an 80 percent train set and a 20% test set for each of our datasets (2015-21 and 2019-21).

B. Feature Engineering

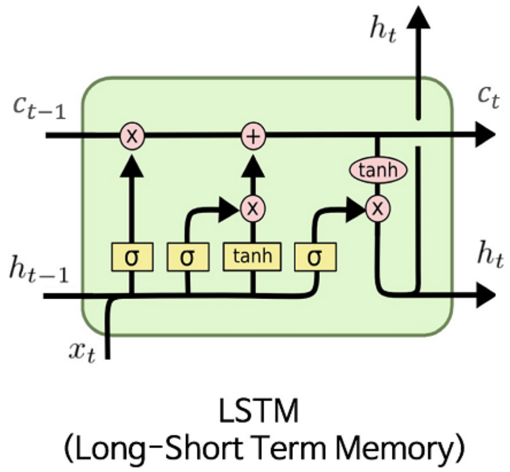
For the larger dataset (2015-21), we employed power transformation to ensure that invariability does not damage our model, whereas for the smaller dataset (2019-21), we used min-max transformation.

Additionally, we prepared the data for 10 time steps(lookback).

We perform all these steps on the closing price of the dataset (only 1 column is being used as this is univariate time series forecasting).

C. LSTMs

Long Short-Term Memory (LSTM) is a recurrent neural network (RNN) architecture that was created to more precisely simulate temporal sequences and their long-range relationships than ordinary RNNs. They were introduced mainly to counter the problem of vanishing gradients in RNNs.



A typical LSTM network is made up of various memory blocks known as cells (the rectangle that we see in the image). The cell state and the concealed state are the two states that are passed to the following cell. The memory blocks are in charge of remembering things, and they are manipulated through three basic mechanisms known as gates.

D. Time2Vec

Time2Vec[6] is a model-agnostic, general-purpose time representation that can be used in any architecture. It gives a vector embedding representation of time in order to streamline the feature engineering process and improve model time. Mathematically, Time2Vec looks like

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

Here The Time2Vec dimension is k, the raw time series is tau, the periodic activation function is F, and the learnable parameters are omega and phi. For our research, we have implemented our Time2Vec layer and stacked it with a LSTM cell.

IV. ARCHITECTURE

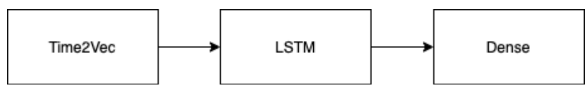


Figure 2.

The architecture includes a 3 layered network along with Dropout and recurrent dropouts incorporated in the LSTM layer to handle overfitting.

A. Loss Function

The objective loss function used for this regression task is "mean absolute error".

$$\mathbf{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}.$$

Where y_i is the prediction and x_i is the true value and e_i is the absolute error and n is the number of samples.
The mean absolute error is calculated on the same scale as the data. In time series analysis, the mean absolute error is a popular metric of forecast error.

B. Model
The model architecture for this specific task consists of three layers.

The general architecture of the first model is provided in Figure 2.

The Time2Vec layer's weights were initialised using standard Initialization, while the LSTM layer's weights were initialised using the default kernel initializer, "glorot uniform".

To manage overfitting in the other variant of the model that we used in the second dataset (2019 to 21), we employed dropout and recurrent dropout in the LSTM layer as a regularizer.

The optimizer used in this task is Adam Optimizer. We have used a dense layer as the final layer for the output.

V. RESULTS AND DISCUSSION

We have used four performance metrics: Root mean squared error(RMSE), MAE(Mean Absolute error), R^2 and Mean Absolute Percentage Error(MAPE).

A. Model Trained on All Data(2015-21)

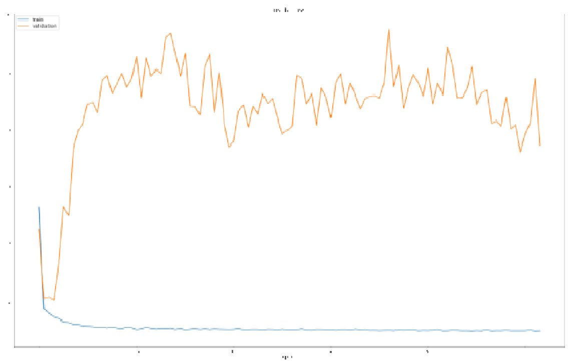


Figure 3. Loss convergence Curve while Training (Blue for train and Yellow for validation)

The model based on the data from 2015 to 2021 was trained on 500 epochs with early stopping at 104. We achieved the following results:

- MAE: 63.58
- R^2: 0.76
- RMSE: 65.85
- MAPE: 14.42

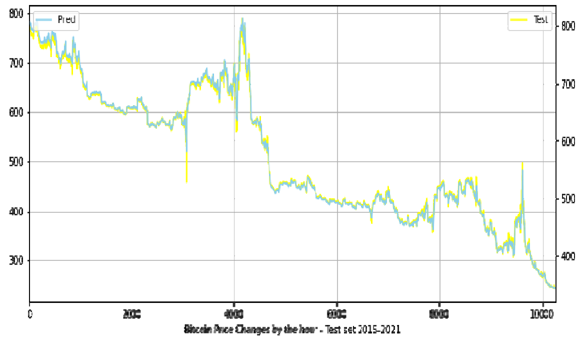


Figure 5. The forecast on predicted values using the test set. (Skyblue for predicted and Yellow for actual values)

B. Model Trained on Selected Data(2019-21)

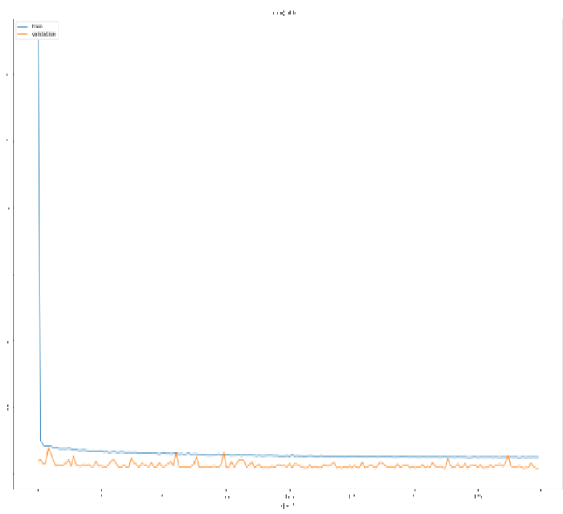


Figure 4. Loss convergence curve while training (Blue for train and Yellow for validation)

Using data from 2019 to 2021 we trained another model for 200 epochs and got the following results:

- MAE: 33.02
- R^2: 0.99
- RMSE: 67.45
- MAPE: 0.48

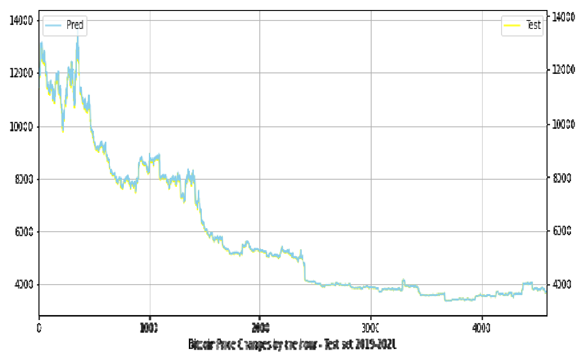


Figure 6. The forecast on predicted values using the test set.(Skyblue for predicted and Yellow for actual values)

C. Ablation Analysis on Selected Data(2019-21)

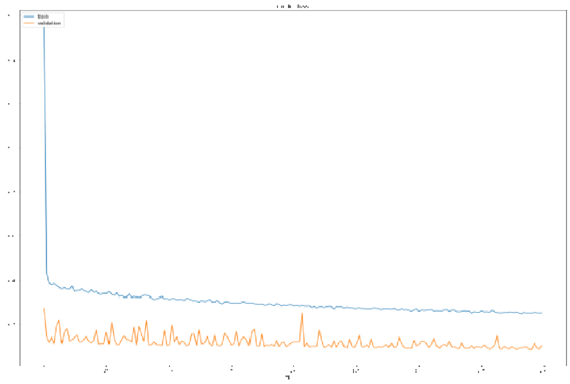


Figure 7. Loss convergence curve while training (Blue for train and Yellow for validation)

On training without the Time2vec layer in the architecture, we achieved the following results:

- MAE: 56.25
- R^2: 0.99
- RMSE: 81.67
- MAPE: 0.99

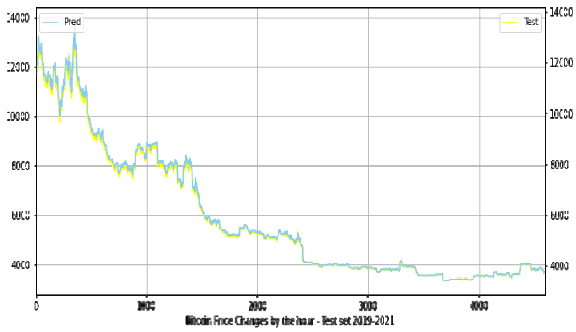


Figure 8. The forecast on predicted values using the test set.(Skyblue for predicted and Yellow for actual values)

VI. CONCLUSION AND FUTURE WORK

In this paper, we used a deep learning model to estimate short-term bitcoin prices while adding localised periodic patterns and vector representations. It is the first study to produce accurate hourly estimates using data up to August 2021.

We have conducted our research on a data with two distinct time periods, for the model trained on the dataset that includes data from 2015 to 2021, we have achieved an MAE of 63.58, and for the model trained on the dataset that includes data from 2019 to 2021, we have achieved an MAE of 33.02. Performance evaluation results reveal an improvement over the recent study in short term price forecasts. In doing ablation analysis with data from 2019 to 2021 while removing the time2vec layer, we have found MAE of 56.25, and conclude vector representations(tim2vec) produce a better fit than bare LSTM cells. In comparison with [10] we have improved the mean absolute percentage error from 3.52% to 0.48%.

After carefully examining the forecast graph of future predicted vs observed values (Figures 5 and 6), we conclude that, while it is difficult to predict short term prices using a much longer range (Figure 5) for 2015 to 2021, it is possible to predict short term prices using a much shorter range (Figure 6) for 2019 to 2021.
In the future, we can try to incorporate multi-headed attention in our similar to deep-transformer models[9], which will allow us to keep much longer periodic patterns as the timesteps get longer.

We might also consider implementing a multivariate data architecture that includes more indications, such as volume, as well as incorporating global happenings from social media data, which may influence prices to some extent.

REFERENCES

[1] Abraham J, Higdon D, Nelson J, Ibarra J (2018) Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Sci Rev 1(3):1
[2] Huang JZ, Huang W, Ni J (2019) Predicting Bitcoin returns using high-dimensional technical indicators. J Finance Data Sci 5(3):140–155
[3] I. M. Wirawan, T. Widiyaningtyas and M. M. Hasan, "Short Term Prediction on Bitcoin Price Using ARIMA Method," 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), 2019, pp. 260-265, doi: 10.1109/ISEMANTIC.2019.8884257.
[4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in *Proc. ECOC’00*, 2000, paper 11.3.4, p. 109.
[5] McNally S, Roche J, Caton S (2018) Cryptocurrency forecasting with deep learning chaotic neural networks. IEEE, pp 339–343
[6] Seyed Mehran Kazemi and Rishab Goel and Sepehr Eghbali and Janahan Ramanan and Jaspreet Sahota and Sanjay Thakur and Stella Wu and Cathal Smyth and Pascal Poupart and Marcus Brubaker, Time2Vec: Learning a Vector Representation of Time, 2019, 1907.05321, arXiv, cs.LG
[7] Lahmiri S, Bekiros S (2019) Cryptocurrency forecasting with deep learning chaotic neural networks. Elsevier, 35–40
[8] Jelinek, Srdjan & Poledica, Ana & Petrovic, Bratislav & Milosevic, Pavle. (2019). Forecasting Cryptocurrency Time Series Using Fuzzy Transform, Fourier Transform and Fuzzy Inference System. 10.2991/eusflat-19.2019.88.
[9] Attention Is All You Need, Ashish Vaswani and Noam Shazeer and Niki Parmar and Jakob Uszkoreit and Llion Jones and Aidan N. Gomez and Lukasz Kaiser and Illia Polosukhin},
[10] 2017, 1706.03762, arXiv, cs.CL
[11] Schulte, Maximilian and Eggert, Mathias, "Predicting Hourly Bitcoin Prices Based on Long Short-term Memory Neural Networks" (2021). Wirtschaftsinformatik 2021 Proceedings. 3. https://aisel.aisnet.org/wi2021/XStudent/Track03/3

★★★