**Forecasting High-Frequency Bitcoin Prices Using Bahdanau Attention Mechanism within LSTM Model.**

**Introduction**

High frequency trading was incredible popular in the last 20 years. Unfortunately, it was impossible to create model to predict all possible price swings on stock market. However, some of high frequencies algorithms are used as supplement of investment strategy for most investment decision[[**1**](https://www.mdpi.com/2504-3110/7/2/203#B13-fractalfract-07-00203)].

This paper will consider possibility of Bitcoin pricing prediction with usage Bahdanau (additive) attention mechanism and Modified Bahdanau (additive) attention. Modified Bahdanau attention mechanism was introduced and tested. I compared 3 algorithms without tanh function and sigmoid for bitcoin pricing. Paper has done research on Bitcoin prices on minutes interval.

**2. Literature Review**

Financial data forecasting can be done with machine learning. In fact, numerous publication showed advantage to use machine learning methods in comparison of traditional econometric methods.

From this aspect, by training a machine learning model on historical cryptocurrency price data, it may be possible to predict future price movements with some degree of accuracy. Prior research has shown that machine learning based techniques have a number of advantages over traditional forecasting models, including the ability to give results that is nearly or exactly the same as the actual result while also improving the accuracy of the results [[**3**](https://www.mdpi.com/2504-3110/7/2/203#B12-fractalfract-07-00203)]. There are several different machine learning techniques that can be used for this purpose, including decision trees, support vector machines (SVM), and neural networks (NN). The authors of [[**4**](https://www.mdpi.com/2504-3110/7/2/203#B13-fractalfract-07-00203)] reveal that inclusion of cryptocurrencies in multi-asset portfolios improves the effectiveness of the portfolio in different ways. First, it enhances the minimum variance of the portfolio and also moves the efficient frontier into a better position. Furthermore, the standard deviation of the portfolio decreases, and the Sharpe ratio increases by including cryptocurrency assets into the portfolios.

## 3. Materials and Methods

In this section, I outline the procedures employed in the pre-processing and modeling phase of the study. Subsequently, a demonstration of the prediction plot results for a selection of cryptocurrencies is presented. Finally, I provide a comprehensive evaluation of the study’s performance and analysis.

.

#### **3.1. Dataset**

In this study, we proposed a simple three-layer network architecture for each deep learning model, consisting of 200-neuron deep learning layers (LSTM). I used the dataset of 15 minutes[1].

Dataset has open price, high price, close price, low price, volume and volume weighting adjusted(VWAP)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Open** | **High** | **Low** | **Close** | **Volume** | **VWAP** |
| **id** |  |  |  |  |  |  |
| **0** | 0.2017 | 0.2022 | 0.2015 | 0.2017 | 447 | 0.202 |
| **1** | 0.2025 | 0.2028 | 0.2023 | 0.2024 | 592 | 0.2024 |
| **2** | 0.2019 | 0.2021 | 0.2017 | 0.2021 | 470 | 0.202 |
| **3** | 0.2016 | 0.202 | 0.2012 | 0.2019 | 515 | 0.2015 |
| **4** | 0.2016 | 0.2019 | 0.2007 | 0.201 | 682 | 0.2012 |

I have implemented percentage change, and we can observe that the time series is negatively autocorrelated according to the Durbin-Watson test (resulting in a statistic of 2.8). Additionally, it exhibits stationarity based on the Kwiatkowski-Phillips-Schmidt-Shin test for a 50-period forecast.

I was using AIC test to define distribution: prices of bitcoin and volume are normally distributed.

A group of graphs showing different sizes of data

Description automatically generated with medium confidence

**Fig.1 Normal distribution volume prices.**

Dataset has Open, High, Low, Close, Volume, WVAp data. Prices :Open, Close, High, Low has the highest correlations (0.99). PCA analysis reveal that there are 3 components to describe dataset.

A graph of a number of principal components

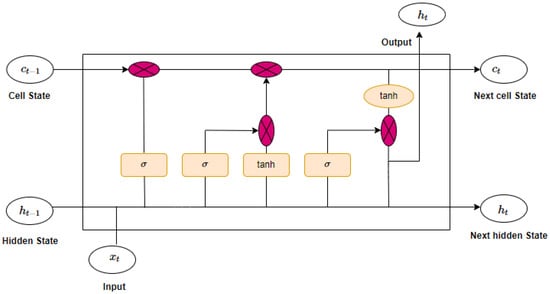
Description automatically generated

**Fig.2 PCA prices analysis**.

#### **3.2. Deep Leaning Algorithms**

#### **3.2.1. Long Short-Term Memory—LSTM**

LSTM is an updated version of RNN. They are specifically designed to avoid long-term dependence problems, whilst solving the vanishing gradient problem with an added mechanism, for regulating information, allowing it to be retained for long periods of time [[**5**](https://www.mdpi.com/2504-3110/7/2/203#B24-fractalfract-07-00203)]. In short, the LSTM architecture is made up of a number of memory blocks that are recurrently connected sub networks



**Figure 3.** The structure of a long short-term memory (LSTM) algorithm.

**.**

**4. 1 Results and discussions.**

**4.1.1 Bitcoin prices are normally distributed.**

It is possible to issue options on bitcoin. This lead to multiple crypto derivatives exchange. Bitcoin prices are stationary. It leads to approach Different models including ARIMA have been tested.

#### **4.1.2 Hypothesis**

I am going to test several hypothesis.

Bitcoin Time series are not stationary for period >8000 hours. Augmented Dickey-Fuller test gave us p=0.48. Therefore, based on the Augmented Dickey-Fuller test, the Bitcoin time series appears to be non-stationary around a deterministic trend for 8000 hours.

However, it's essential to note that different statistical tests may yield different results, and the choice of which test to use can depend on various factors stationary around a deterministic trend.

LSTM model was testing against random prices. I use Monte Carlo methods for random walk. This method is used for option pricing. I used Monte-Carlo with standard deviation (volatility) of sample to model future prices

Hypothesis for testing was following:

LTSM(no attention mechanism) follows Random Walk p-value=0.9x10-22<0.05

Null hypothesis rejected.

LTSM (no attention mechanism) outperforms Arima model prices p-value=0.5x10-6

Null hypothesis rejected

|  |  |  |
| --- | --- | --- |
| **Hypothesis** | **p-value** | **H** |
| *LTSM (no attention mechanism) prices Random Walk prices* | 0.9x10-22 | **Rejected** |
| Arima model predicted prices better than LTSM (no attention mechanism) prices | 0.5x10-6 | **Rejected** |

4.1.2. Model without attention mechanism with epochs=10 leads to the best results.

|  |  |  |
| --- | --- | --- |
| Models | MAPE | MSE |
| Model (Bahdanau attention) | 0.9912 | 0.035 |
| Model(Bahdanau attention sigmoid) | infinity | infinity |
| Model(without attention) | 0.3867 | 0.007 |

A graph with blue and orange lines

Description automatically generated

**Figure 4. Prices of bitcoin with future price prediction**

**4.1.3. Optimal period for forecast.**

We used a 50-period horizon in our calculations for predicting Bitcoin prices. This choice is highly constrained.

I employed Particle Swarm Optimization to determine the optimal period for enhancing the LSTM model's performance.

The optimal value obtained was 51, which is close to my current value. However, for periods ranging from 55 to 100, the optimal period is 72. This suggests that we can utilize different periods and combine the predictions for better results.

**Conclusion:**

Bitcoin prices are normally distributed, allowing us to use them as financial assets. Exchanges can effectively price options on bitcoins and Bitcoin ETFs. Bitcoin prices exhibit stationarity for several hours, and they are autocorrelated. This explains why an LSTM model is possible, or any directional model for that matter.

The shapes of prices predicted by LSTM, LSTM with attention, and ARIMA are approximately similar. This leads us to believe that Bitcoin can be hold for the medium term (1-3 months) with positive return. Prices can be predicted by the LSTM model with a mean squared error (MSE) around 0.007.

Despite the recent popularity of attention mechanisms, an LSTM without attention mechanism performed better in my environment.

The forecasted period for Bitcoin has been optimized using the Particle Swarm Optimization algorithm. The optimal value for the forecasted period is 51, which is close to my current value.

My general observation is that Bitcoin is a new financial instrument, made possible by sophisticated technology. Therefore, we need sophisticated technology to price, analyze, and manage risk associated with cryptocurrencies.

**References.**

1. https://www.kaggle.com/competitions/g-research-crypto-forecasting?rvi=1
2. Bulíř, A. Income inequality: Does inflation matter? *IMF Staff. Pap.* **2001**, *48*, 139–159. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Income+inequality:+Does+inflation+matter?&author=Bul%C3%AD%C5%99,+A.&publication_year=2001&journal=IMF+Staff.+Pap.&volume=48&pages=139%E2%80%93159)]
3. Hitam, N.A.; Ismail, A.R.; Samsudin, R.; Alkhammash, E.H. The Effect of Kernel Functions on Cryptocurrency Prediction Using Support Vector Machines. In Proceedings of the International Conference of Reliable Information and Communication Technology; Springer: Cham, Switzerland, 2022; pp. 319–332. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=The+Effect+of+Kernel+Functions+on+Cryptocurrency+Prediction+Using+Support+Vector+Machines&conference=Proceedings+of+the+International+Conference+of+Reliable+Information+and+Communication+Technology&author=Hitam,+N.A.&author=Ismail,+A.R.&author=Samsudin,+R.&author=Alkhammash,+E.H.&publication_year=2022&pages=319%E2%80%93332)]
4. Andrianto, Y.; Diputra, Y. The effect of cryptocurrency on investment portfolio effectiveness. *J. Financ. Account.* **2017**, *5*, 229–238. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=The+effect+of+cryptocurrency+on+investment+portfolio+effectiveness&author=Andrianto,+Y.&author=Diputra,+Y.&publication_year=2017&journal=J.+Financ.+Account.&volume=5&pages=229%E2%80%93238&doi=10.11648/j.jfa.20170506.14)] [[**CrossRef**](https://doi.org/10.11648/j.jfa.20170506.14)][[**Green Version**](http://article.sciencepublishinggroup.com/pdf/10.11648.j.jfa.20170506.14.pdf)]
5. Wu, C.H.; Lu, C.C.; Ma, Y.F.; Lu, R.S. A new forecasting framework for bitcoin price with LSTM. In Proceedings of the 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018; pp. 168–175. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=A+new+forecasting+framework+for+bitcoin+price+with+LSTM&conference=Proceedings+of+the+2018+IEEE+International+Conference+on+Data+Mining+Workshops+(ICDMW)&author=Wu,+C.H.&author=Lu,+C.C.&author=Ma,+Y.F.&author=Lu,+R.S.&publication_year=2018&pages=168%E2%80%93175)]