

# PREDICTING CRYPTO PRICE USING LSTM MODEL

Qiuyang Wang and Xinyu Jiang and Yan Zhan

University of Colorado Boulder

## Introduction

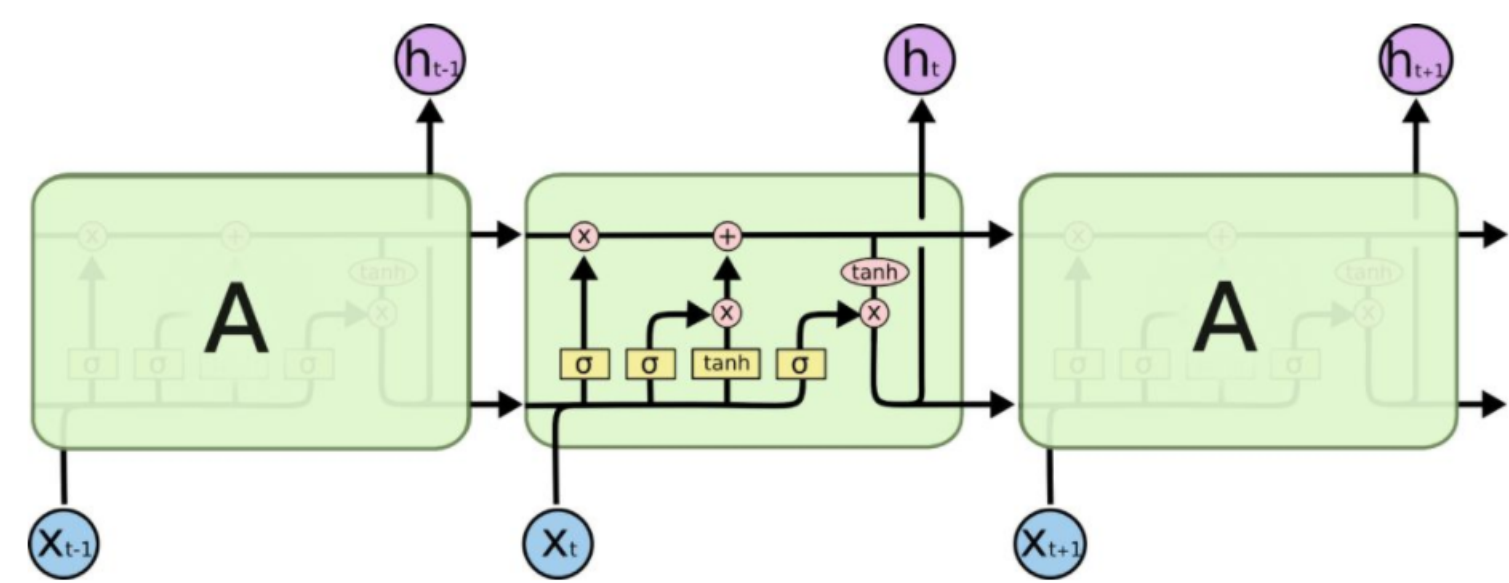
Researchers from different fields have been trying to predict the price of Bitcoin with various methods. In this project, we follow We et al(2018) and try to predict the price of Bitcoin in a short-term future with LSTM. As a time series approach, LSTM outperform others such as Autoregressive (AR), univariate Moving Average (MA) and Autoregressive Integrated Moving Average (ARIMA) in terms of the Bitcoin price prediction. Those methods are more suitable for data with seasonal trends pattern which is not observed in Bitcoin price. LSTM is favored over other time series method due to the temporal nature of the more advanced algorithms.



image from <https://www.techrepublic.com/article/9-things-to-know-about-cryptocurrency-such-as-cardano-binancecoin-and-ethereum/>

## Approach

After discussing, we decided to use LSTM as our model and also to change the scope of the project, from predicting trend to predicting crypto price in the next hour. To reach the goal, we also need to reshape and edit the raw data to make sure we can feed the correct data to our model. In terms of predicting the crypto price in the next hour, we found a way to create the label which acts like the crypto price in the next hour. Then we fit the LSTM model with the pre-processed data and feed test data to the model as well. The model generated the predicted crypto prices in the next hour just after the last close time of the last row of the data. And we validated the predicted labels with our test y, and generated MSE, loss, residual plots to evaluate the performance of our model.

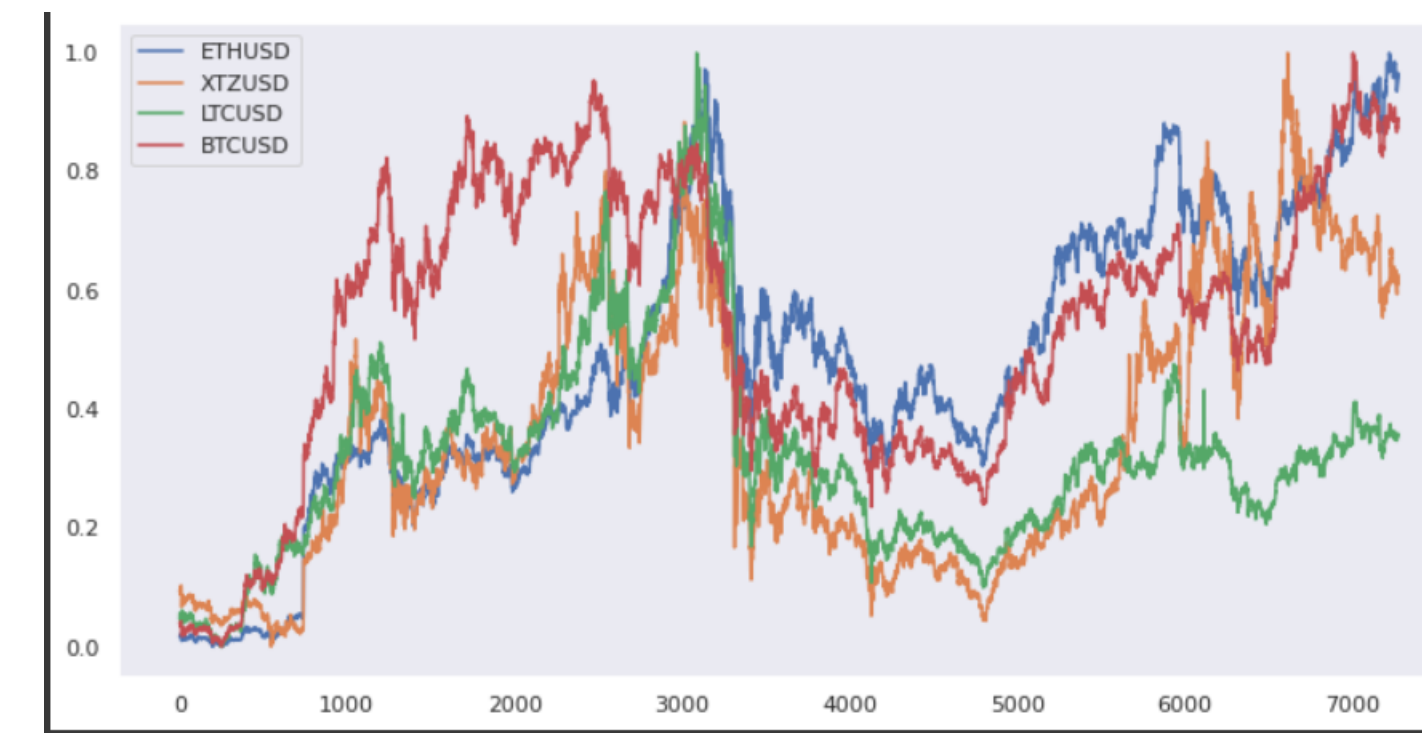


## Data

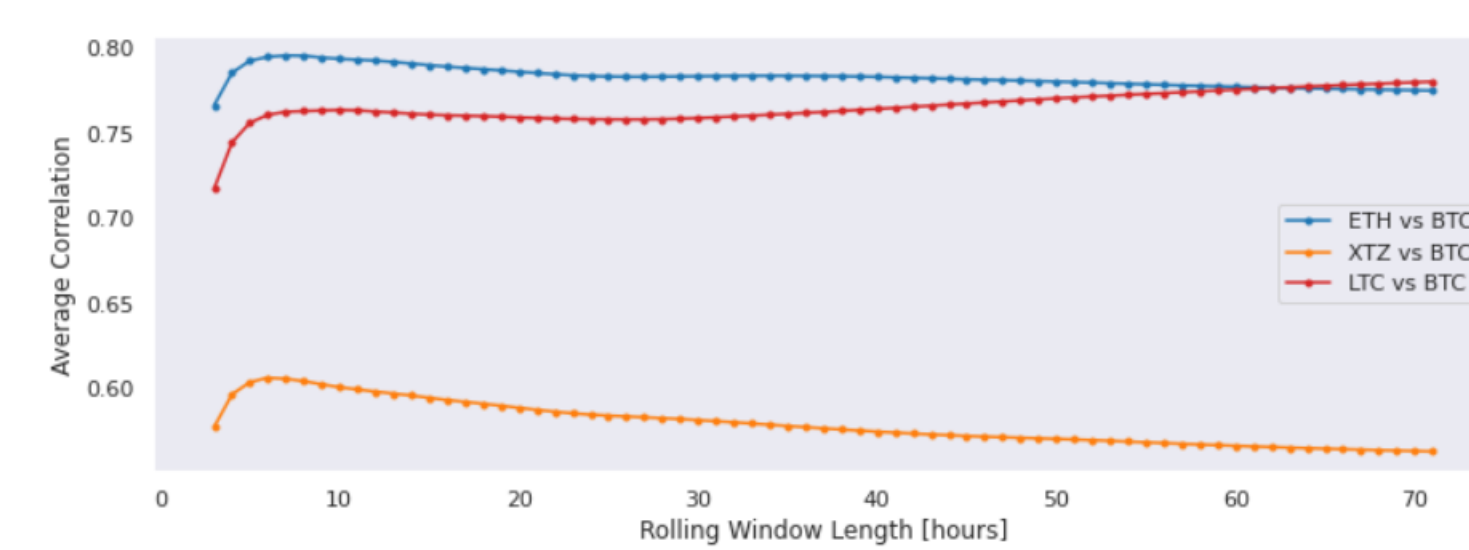
We used the data from Binance Data Collection which is a website provides free crypto data sets in daily and monthly with complete time resolutions(minutes, hours, days, months). The we used klines data sets in 1-hour resolution, the time range of the data sets are from December 2020 to November 2021.

Open time	Open	High	Low	Close	Volume	Close time
1499040000000	0.01634790	0.80000000	0.01575800	0.01577100	148976.11427815	14996447999

And we visualize the different crypto prices in plot to find the relationship between different cryptos.

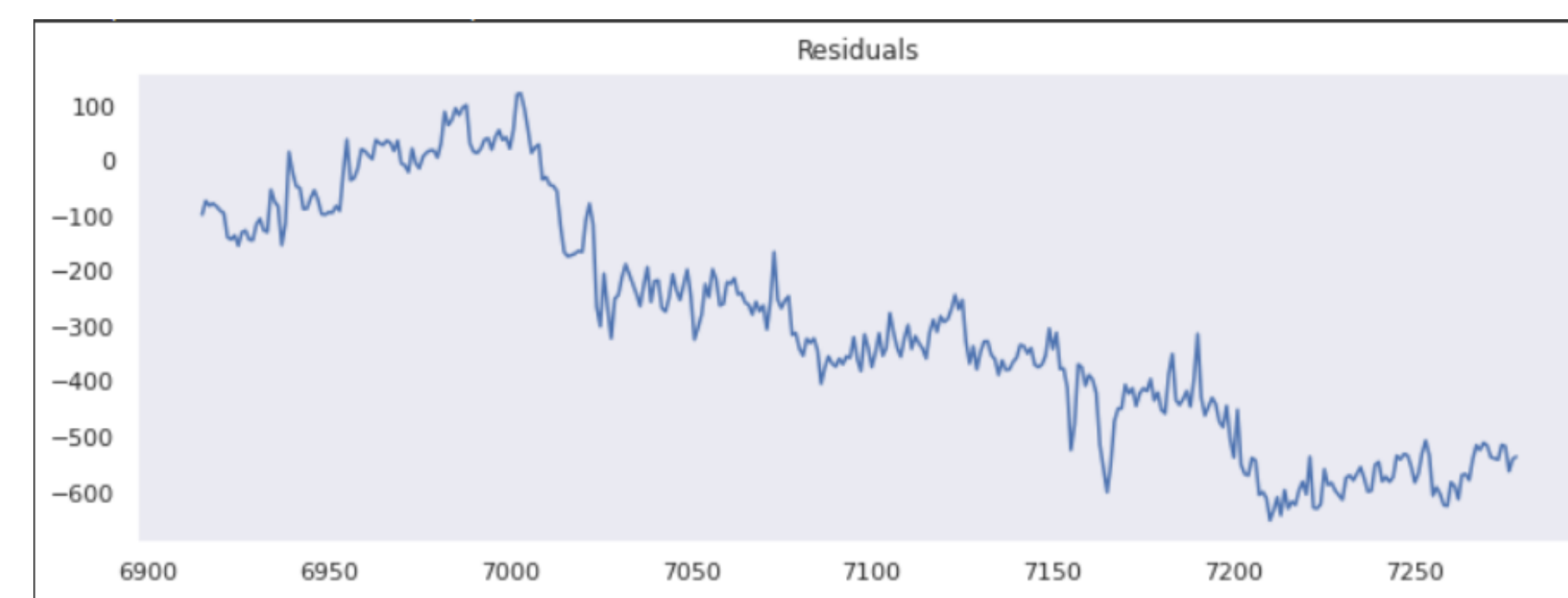


We also create co-relation plot to confirm the relationship between the crypto prices we have chosen.



## Results

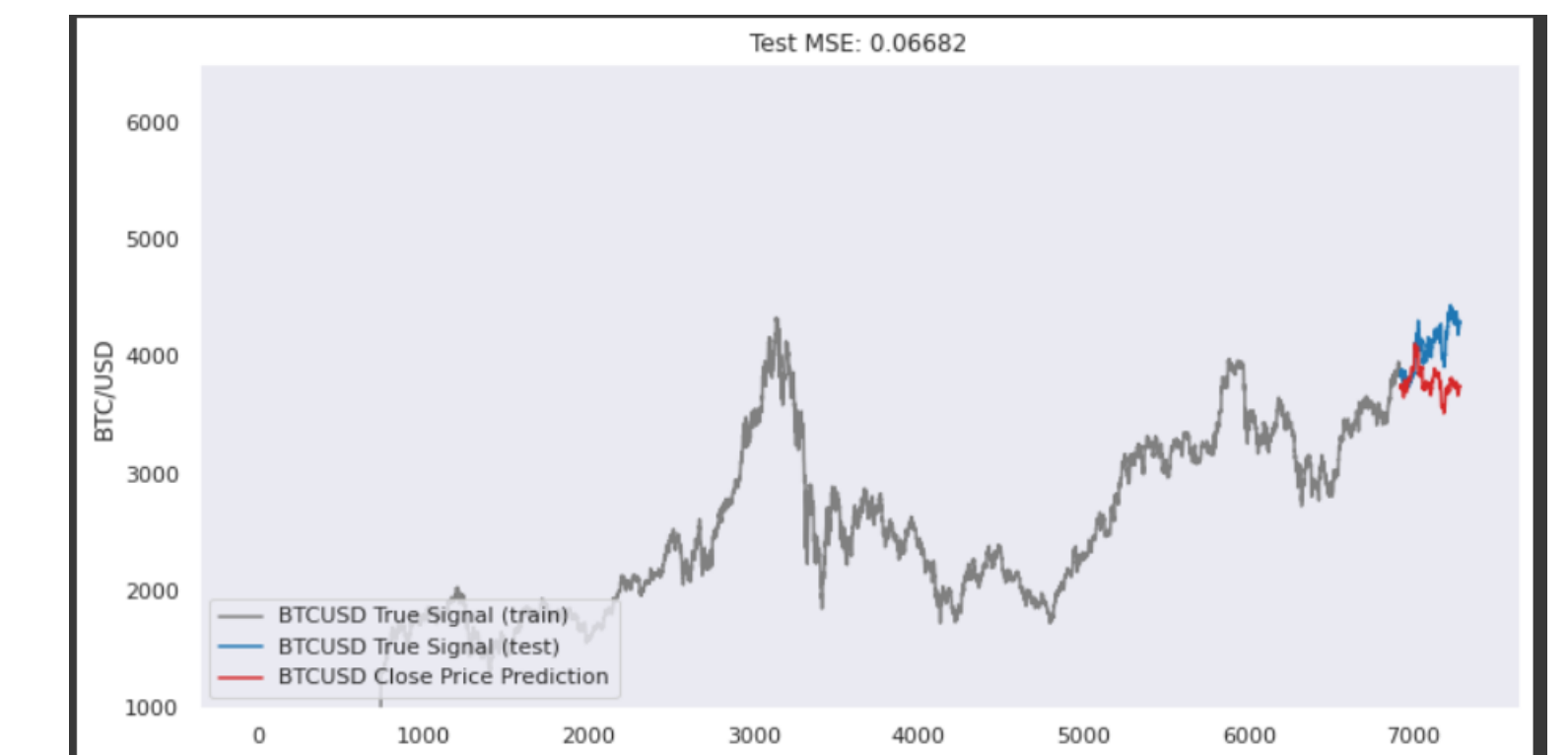
**Residuals** The graph below represents the residuals of difference between the predicted value and the actual Bitcoin closing price.



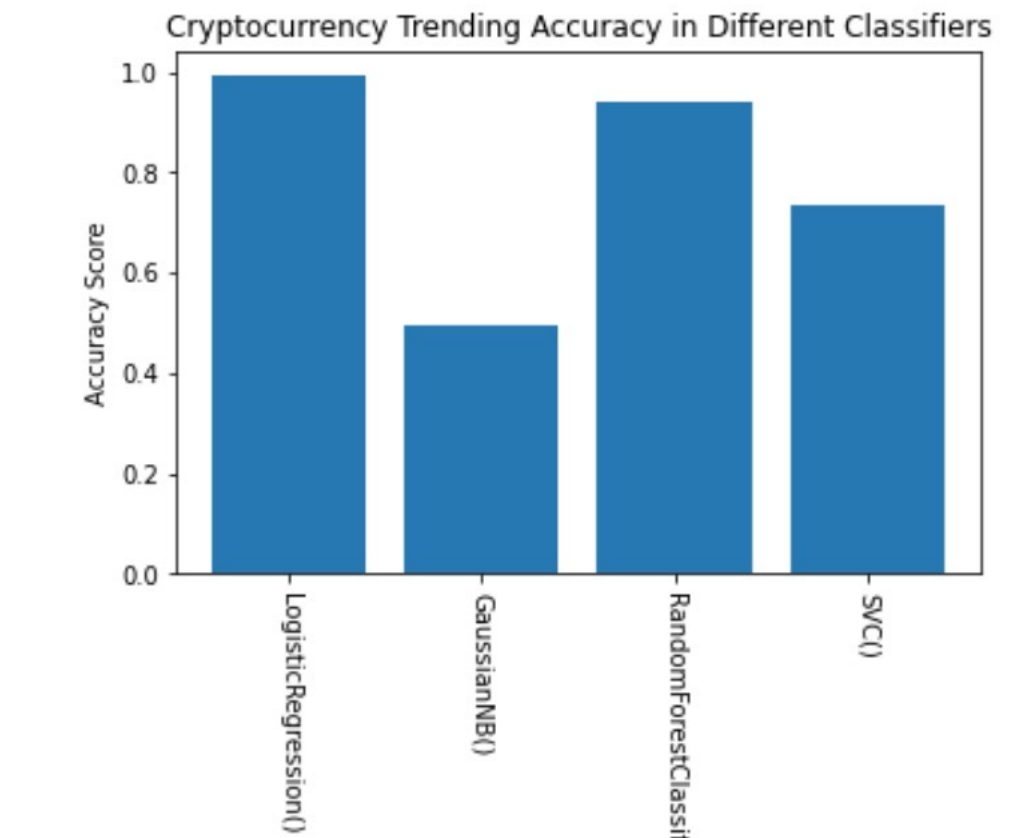
**MSE** In order to test the prediction of bitcoin price and actual value, we use the mean square error regression loss (MSE) to measure. Based on the graph below, the test MSE result we got is 0.06682 which represents the lowest error measurement.

## Discussion

1. neural network is the better model compared to regression.
2. Our model can't predict exact price when the external factors are influential, but is capable of predicting the trend.
3. The model is able to predict the Bitcoin price in the next hour with certain range when there is no large external factors.
4. Our data might have been tainted since there are just too many things happened about crypto. The market was always in shcok during the last year.
5. We have also created a new column of data called trend, which detects the current cryptocurrency and previous prices, and uses traditional machine learning methods (logistic regression, svm, random forest, naive bayes) to predict and classify



```
{'LogisticRegression()': 0.9921077065923862, 'GaussianNB()': 0
```



## Future works

1. We want to use the data before COVID and after COVID to train the model again and observe the residuals plot to see if the influence of black swan external factor(like the pandemic) has impact on crypto market.
2. We still need to adjust the variables in the model to make sure it perform its best.