Predictively Modeling Stocks and Bitcoin with Recurrent Neural Networks

Tyler Artz

Abstract

The financial world has seemed to warm up to the idea of digital currencies in the past few years. Due to the many previous conclusions of other researchers, a recurrent neural network has been chosen for this project both due to the accuracy of previous research and the inherent benefit of using backpropagation in the proposed question as well as the ability of recurrent neural networks to consider previously calculated outputs and use them as an additional input stream. After testing each network and comparing, it is found that there is a positive effect on performance for cryptocurrency price movement predictions when using stock and cryptocurrency input data and the opposite is true for stock price movement predictions.

1. Introduction

THE world of cryptocurrencies has moved its way into mainstream financial markets as well as has been increasingly talked about in daily life. Although there seem to be some similar characteristics between the cryptocurrency market and the stock market, there is very little research comparing the two, specifically when it comes to predictively modeling cryptocurrencies like there has been with similar research regarding the stock markets of the world.

The adoption of cryptocurrency by the layman has been slowly increasing and has similar movements to the stock market. Due to the lack of research in this field, a connection between the two markets will attempt to be explored using similar means to that of those who have done previous financial market research using machine learning, specifically using Bitcoin to compare with Apple and Tesla.

2. Literature Review

2.1 US Stock and cryptocurrency prices

To understand why predictive modeling on cryptocurrencies using US stock data may provide a benefit to the accuracy of said predictions, it is necessary to be aware of the possible relationship between the two, coming from both the direct relationship in volatility rates of the two and the increasing adoption of the cryptocurrency.

a. Volatility Relationships

Using a mathematical model, previous research has shown that a volatility relationship does exist within the three financial markets of gold, US stocks, and cryptocurrencies while applying a GARCH-DCC model to gold, the S&P 500, and Bitcoin volatility "with the current day and one-day lagged prices," [1]. This means that the volatility, or rate at which the price of an asset is predicted or can change over a period of time, of each of gold, the S&P 500, and Bitcoin shifts together whether that be because one of the valuations leans heavily into the others, or the valuations of each rely on similar external factors.

b. Rising Bitcoin Adoption

The amount of people in the US who are investors that also have positions in Bitcoin has been increasing for at least the past 3-4 years. The percentage of people aged 18-49 who have over $10,000 invested in stocks or other financial instruments that have investments in Bitcoin has gone from 3% in May 2018 to 13% in June 2021, which is an increase of over 300% in investors who also hold Bitcoin [8]. This shows that there is increasing overlap in the people who have traditional investments and who hold Bitcoin.

2.2 Neural Networks

Neural networks have become a popular approach in classification or identification problems due to the nature of them attempting to replicate the learning or training process that the human brain undergoes when identifying something, such as handwritten numbers [4]. Figure 1 shows the basic construction of a neural network.

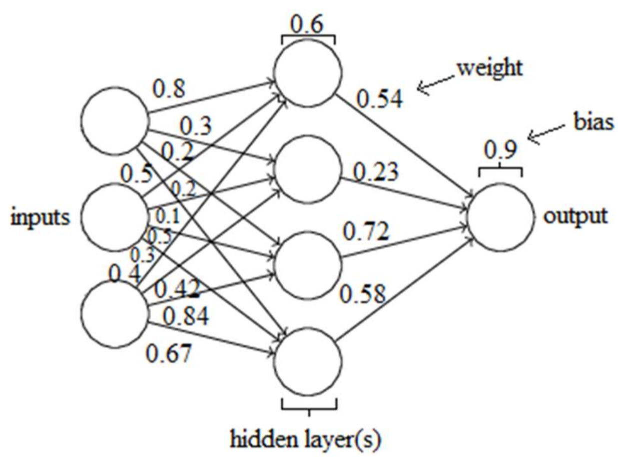

Figure 1. Basic Model for Neural Networks [4] with added labels.

Support vector machines were another reoccurring proposed solution but were not chosen for this project because of the noticeable difference in recall for the model constructed by Wang [9]. Auto-Regressive Integrated Moving Average (ARIMA) was also implemented by Pang et al. [7] as well as Shah et al. [6]. This methodology was not selected as it is more of a strictly statistical model and neural networks are more proficient overall for our proposed problem.

- #### *a. Recurrent Neural Networks*

Recurrent neural networks excel at problems in which previous outputs need to consider for future outputs. "Recurrent neural networks not just take the current input data they see (t), but also what they have perceived previously in time (t-1). It includes neurons, hidden layers, and activation functions to compute the output value using back propagation," [7]. Taking this into account, recurrent neural networks seem to be the best choice for the task ahead as using prices that change on a minute-by-minute basis will obviously be basing a portion of their current valuation on the previous valuations of anytime from the past minute to the past few months, maybe even multiple years if a situation calls for it.

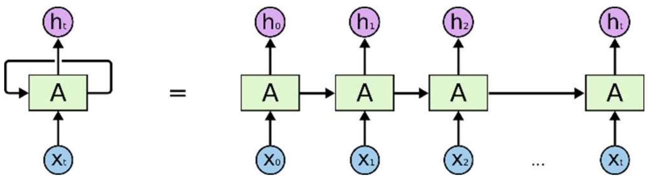

Figure 2. Simple Recurrent Nueral Network Model [5].

b. Backpropagation

Backpropagation in recurrent neural networks is essential, as it allows the neural network to learn from its input and previous input to refine decision making. "The goal of backpropagation is to optimize the weights so that the neural network can learn how to correctly map arbitrary inputs to outputs," [2]. Without backpropagation, neural networks would simply not work as it allows the weights of each neuron in each layer to change based on the how each neuron reacts to its input. Without this, neural networks would struggle to learn anything from the information as it is being processed, and it is almost impossible for humans to explain, in a way that we can understand, what a specific weight or a specific neuron is actually for. Below are the formulas used to compute the changes in weight for backpropagation in order:

- $net_{h1} = w_1 * i_1 + w_2 * i_2 + b_1 * 1$
- $out_{h1} = \frac{e^{net_{h1}-e^{-net_{h1}}}}{e^{net_{h1}+e^{-ne\ h1}}}$
- $E_{01} = \frac{1}{2}(target_{01} - out_{01})^2$
- $Etotal = E_{01} + E_{02} + \cdots + E_{ox}$
- $w_5^+ = w_5 - \eta * \frac{\partial Etotal}{\partial w_5}$

The first equation uses the weights (w), inputs (i), and biases (b) to calculate the net input to a node in the neural network. Equation 2 uses that net input with the activation function of the neural network (in this case tanh) to compute the output that the node will send as input to the next layer in the network. Equation 3 calculates the error of a single node using the target value that is expected and the actual output which is then added with all other error values to get the error total computed in equation 4. Equation 5 computes the new weight of the node by multiplying the learning rate of the network with the gradient calculated by the partial derivative of the loss function (mean squared error will be used for this project).

2.3 Predictive Modeling using neural networks

Predictive modeling of the movement of stock prices have used recurrent neural networks in their studies consistently for at least the past few years, pointing this research project towards said method. Not only have recurrent neural networks been used consistently in previous years of research, but they have also been one of the best performers in terms of accuracy of price movement predictions in the stock market, competing primarily with Long Short-Term Memory.

To be more specific regarding the performance of neural networks in this area of research, artificial neural networks were able to achieve an average accuracy of 87.275% across four different sections of the stock market, recurrent neural networks were able to achieve an average accuracy of 89.16% across those same four sections of the stock market [3]. A different study was also successful in the prediction of stock price movements with recurrent neural networks, but with the addition of social sentiment data around the stocks as an added input stream [7]. Along with much more research having been done using recurrent neural networks to solve similar problems, recurrent neural networks seem to be a both interesting and somewhat optimal technique.

3. Primary Objective

To investigate the impact of predictive performance of Recurrent Neural Networks on historical stock price data and historical cryptocurrency price data individually and when the data are combined for modeling (1.5 person-weeks over 1 semester).

4. Solution Description

The first step to complete the objective is to gather our data that the networks are learning from and to preprocess it. The dataset, which is comprised of AAPL and TSLA for the stock data along with BTC-USD and ETH-USD for the cryptocurrency data, was obtained from Yahoo! Finance and are from the date range Sept. 17^{th}, 2014 – Aug. 30^{th}, 2022.

Each of the four historical data subsets contain the following features: Open, High, Low, Close, Adj. Close, and Volume. The subsets are each then preprocessed for the network as such:

- Remove Adj. Close
- Normalize the datapoints in relation to each other
- Establish sequences so that the data is in time-series
- Classify each sequence as either a buy or sell (the price moves either up or down, 0 or 1)
- Shuffle the sequences
- Split the sequences into a train : test split

Once all the data is preprocessed, the networks are created and trained according to the experiment design and results gathered.

5. Goal Tree and Hypotheses

Due to the nature of cryptocurrency and how it is very reliant of news surrounding itself and public perception, it is hypothesized that there will be a noticeable improvement to the Recurrent Neural Network performance when using both stock and cryptocurrency input data to predict cryptocurrency prices, but there will not be a noticeable improvement to the Recurrent Neural Network performance when using both stock and cryptocurrency input data to predict stock prices.

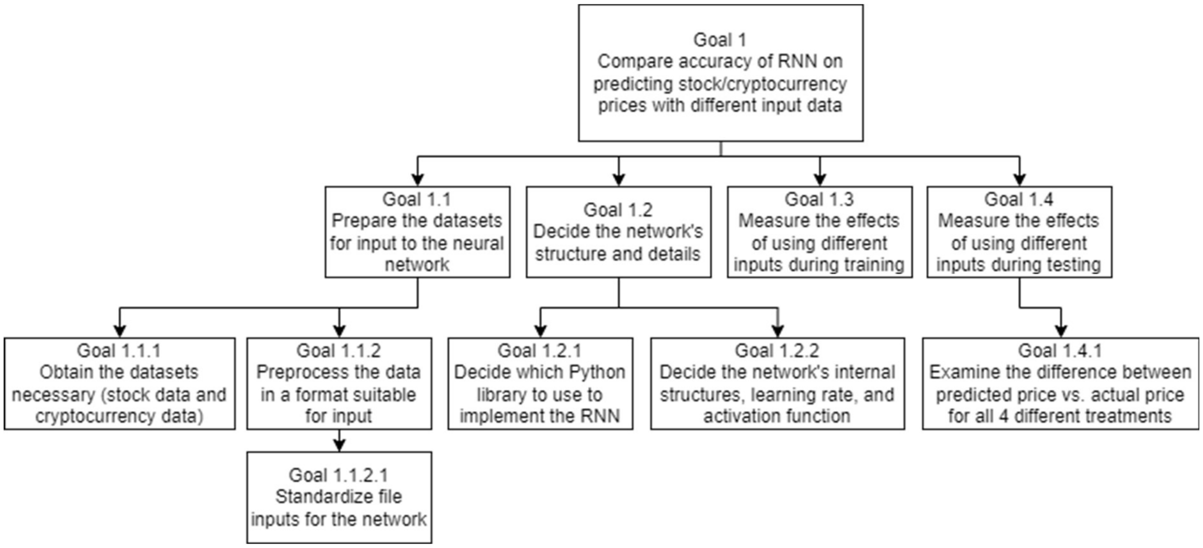

Figure 3. Goal Tree

6. Experiment Design

Factors	Values
Input Streams	{Stock data, Bitcoin Data, Both}
Model Type	{Stock Model, Bitcoin Model}
Sample Size	10 Random seeds per treatment

Block Design		Input Stream		
		Stock Data	Bitcoin Data	Both
Model Type	Stock Modeling	x		x
	Bitcoin Modeling		x	x

Figure 4. Experiment Block Design

For the experiment itself, both the input streams that the neural network will receive and the model that the neural network produces will vary. Under the block design, it is visible which combinations of what will be used and that there will be two formal experiments.

In each experiment, the two networks that have the same model type will be compared. For example, experiment 1 is reliant on the stock modeling which means that the baseline network is modeling stock price movements with only stock data as input while the experimental network is predicting stock price movements with both stock and cryptocurrency data. The same applies for experiment 2 (with the difference being it modeling cryptocurrency price movements). After the experiments are concluded, the final accuracies of the baseline and experimental networks are compared to determine if there are any difference in performance due to the input data.

7. Results

The results of the two experiments were in line with the original hypotheses of experiment 1 resulting in no improvement between input data while experiment 2 did show significant improvement between input data.

The plots in Fig. 5 and Fig 6. Represent the average and standard deviation of the accuracies between the baseline and experimental networks for experiments 1 and 2 respectively. Is it clearly visible that, for experiment 1, the result was not only that there was no improvement, but also that the experimental network proved to be worse at modeling the stock price movements. For experiment 2, the result clearly confirms the hypothesis that there is an improvement in the network's modeling of cryptocurrency price movements. Additional plots representing the distribution of the data can be found in the Appendix.

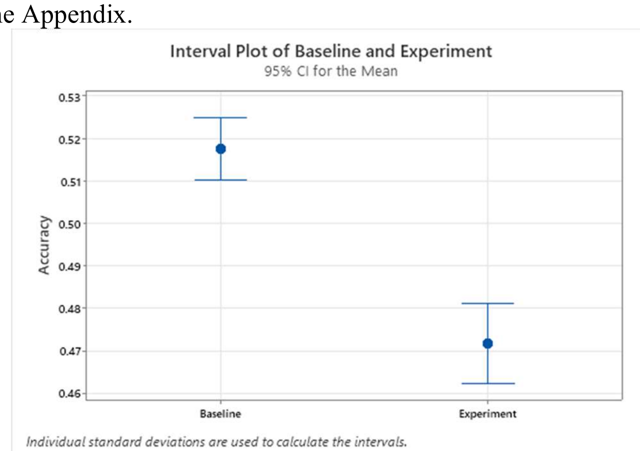

Figure 5. Average accuracy of baseline and experiment network for Experiment 1, with standard deviation.

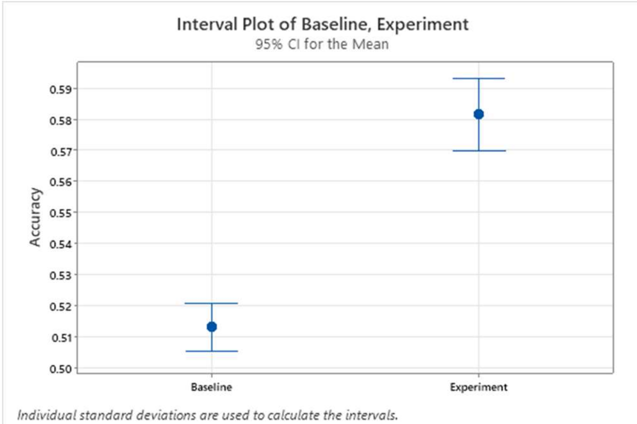

Figure 6. Average accuracy of baseline and experiment network for Experiment 2, with standard deviation.

8. Conclusion

With the results that were collected during the experiments, it can be concluded that using cryptocurrency data in addition to stock data to predictively model stock price movements is detrimental to the network's predictive performance, while using stock data in addition to cryptocurrency data to predictively model cryptocurrency price movements is beneficial to the network's predictive performance.

From these conclusions, more questions are asked such as why cryptocurrency data causes decreased performance for stock modeling or why stock data causes increased performance for cryptocurrency modeling. To speculate, these two questions go hand in hand in that there may be a possible time lag in the relationship between the stock market and cryptocurrency market. This is to say that the relationship is only able to improve performance in the direction of the time lag and worsens performance in the opposite direction.

9. Future Work

Moving forward, there is a lot left unfinished regarding this research. First and foremost is conducting true out of sample testing. This was unable to be done due to time constraints and, more importantly, the amount of data that was available due to cryptocurrency not existing for as long as the stock market has. After true out of sample testing, a forward test is another avenue to explore, allowing for a measure of how the networks perform on their own and compare them to the real-world price movements as time passes.

Apart from additional testing on these networks that model only price movements, another direction this research can be taken is the prediction of the movements that include a magnitude of change. In other words, predicting the specific price the asset would be after it moves. Based on this research though, this would only work for experimental network of experiment 2, and not the experimental network of experiment 1.

10. References

[1] J.-M. Kim, S.-T. Kim, and S. Kim, "On the Relationship of Cryptocurrency Price with US Stock and Gold Price Using Copula Models," *Mathematics*, vol. 8, no. 11, Art. no. 11, Nov. 2020, doi: 10.3390/math8111859.

[2] M. Mazur, "A Step by Step Backpropagation Example", 2015, https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/

[3] M. Nabipour, P. Nayyeri, H. Jabani, S. S., and A. Mosavi, "Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis," *IEEE Access*, vol. 8, pp. 150199–150212, 2020, doi: 10.1109/ACCESS.2020.3015966.

[4] M. A. Nielsen, "Neural Networks and Deep Learning", Determination Press, 2015, http://neuralnetworksanddeeplearning.com/.

[5] C. Olah, "Understanding LSTM networks," *Understanding LSTM Networks*, Aug. 2015, https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[6] D. Shah, H. Isah, and F. Zulkernine, "Stock Market Analysis: A Review and Taxonomy of Prediction Techniques," *International Journal of Financial Studies*, vol. 7, no. 2, Art. no. 2, Jun. 2019, doi: 10.3390/ijfs7020026.

[7] Y. Pang, G. Sundararaj, and J. Ren, "Cryptocurrency Price Prediction using Time Series and Social Sentiment Data," in *Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, New York, NY, USA, Dec. 2019, pp. 35–41. doi: 10.1145/3365109.3368785.

[8] L. Saad, "Bitcoin Making Inroads With Younger U.S. Investors," *Gallup.com*, Jul. 22, 2021. https://news.gallup.com/poll/352508/bitcoin-making-inroads-younger-investors.aspx (accessed Feb. 24, 2022).

[9] S. Wang, "The Prediction of Stock Index Movements Based on Machine Learning," in *Proceedings of the 2020 12th International Conference on Computer and Automation Engineering*, New York, NY, USA, Feb. 2020, pp. 1–6. doi: 10.1145/3384613.3384615.

[10] "Yahoo! Finance Historical Data," Yahoo! Finance, Dec. 06, 2021. https://finance.yahoo.com/quote/AAPL/history?p=AAPL. (accessed Dec. 6, 2021).

11. Appendix

The plots in Fig. 7 and Fig 8. represent the overall distribution of accuracies for experiment 1.

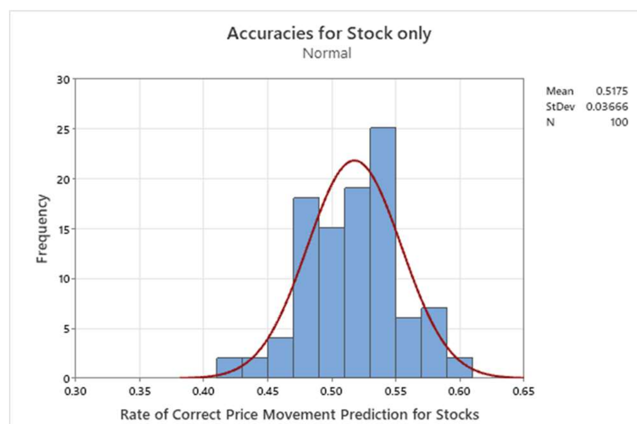

Figure 7. Accuracies for Baseline in experiment 1.

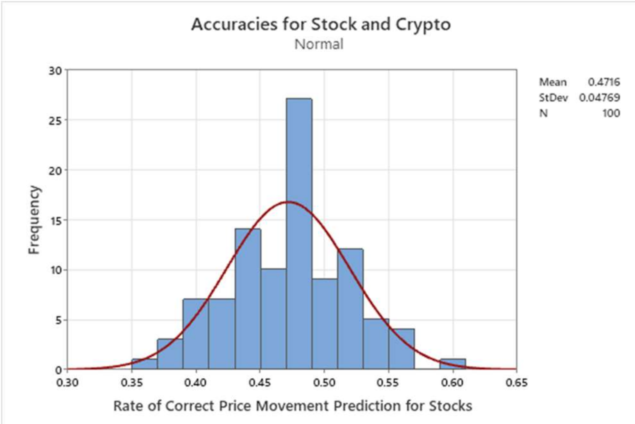

Figure 8. Accuracies for Experiment in experiment 1.

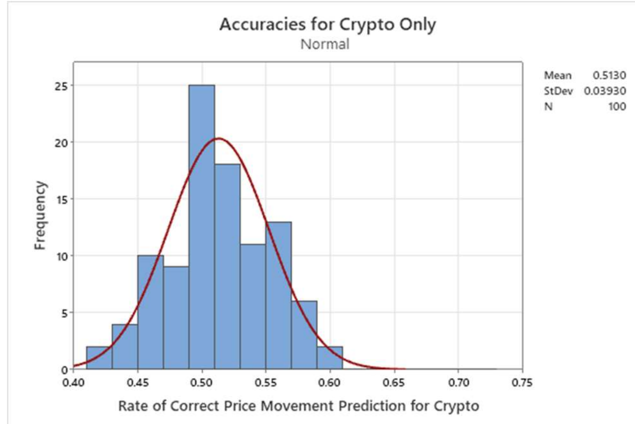

Figure 9. Accuracies for Baseline in experiment 2.

The plots in Fig. 9 and Fig. 10 represent the overall distribution of accuracies for experiment 1.

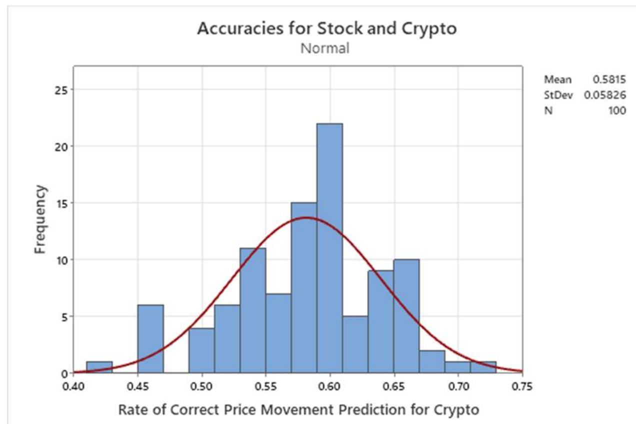

Figure 10. Accuracies for Experiment in experiment 2.