Analysis on Correlation Between Bitcoin and Stock Market

Tejaswini Yella
Department of Information Systems
University of Maryland , Baltimore County
Baltimore, Maryland 21250
Email: yella1@umbc.edu

Abstract - **Bitcoin had a serious rise and fall in its prices over past couple of years. Some believe that they had lifted other stocks who are dependent on it for improving their businesses. Some also believe that the dependent companies stocks also might crash when bitcoin crashes. In this paper, I will work on analyzing the hypothesis that "Bitcoin and its dependent business stocks like Nvidia, Overstock, AMD prices are highly correlated with each other".**

Keywords - Nvidia, Overstock, AMD, Bitcoin, Time series analysis, stocks, model

I. Introduction and Motivation

Bitcoin has been on a breathtaking ride from 2014 till now. Like all rising tides, bitcoin's success has lifted other boats[6]. But for investors who aren't so sure about stock market, bitcoin also acted like a down-ballot that reach into the stock market. Some of the publicly traded companies which offer, accept, and help improve access to bitcoins have seen their stocks triple or more in value in the past couple of years. some believe that the value of bitcoin is headed for a crash. And if bitcoin bursts, it's not just crypto-believers that will suffer. It is also said that overstock, Nvidia are soaring because of bitcoin[6] .The Nvidia and AMD graphic cards, which people use to mine cryptocurrencies were of more demand once from bitcoin seekers. The dramatic jump in digital currency prices is driving demand for its graphics cards. These stocks were the top performers in the S&P 500 during c course of time[8]. In 2014, Overstock became the first major retailer to accept bitcoin for transactions and Overstock shares have risen 260% rise since August[6]. Through this project, I am focusing on getting some insights and prove hypothesis like

- The GPU's developed by Nvidia were once used for bitcoin mining, so the fall of crypto currencies affects the stock price of Nvidia.
- Similarly Overstock has started accepting bitcoins, so the stock rates would have been affected by bitcoin prices.
- As AMD graphic cards are used by crypto currencies the fall of Bitcoin might affect the stock price of amd.

Through the visualization I am also explaining about certain period of time where these stocks and bitcoin prices followed very similar trend, like rising together and falling at the same time. In addition to this, as I have data collected over a period of time for all these stocks, I performed time series analysis and built an forecasting model to predict future values of prices based on historical data.

II. Dataset

A. Bitcoin Historical Prices Dataset :

Bitcoin is the longest running and most well-known cryptocurrency, first released as open source in 2009 by the anonymous Satoshi Nakamoto. Bitcoin serves as a decentralized medium of digital exchange, with transactions verified and recorded in a public distributed ledger (the blockchain) without the need for a trusted record keeping authority or central intermediary[9].
I got this dataset from the website Coin market cap[9].This dataset was imported into excel and has the bitcoin historical prices from April 2013 to Oct 2018 with the attributes as Date Open High Low Close Volume Market. Cap Coin Delta

B. Nvidia dataset:

This company's some of the main products are designs graphics processing units (GPUs) for the gaming and professional markets, as well as system on a chip units (SoCs) for the mobile computing and automotive market[10].
The sources of stock prices dataset are Nasdaq[2] and Kaggle [1].This dataset was imported into excel and the Nvidia stock prices are from 1999 to Oct 2018 with the attributes as Date Open High Low Close Volume Market. Cap

C. Overstock dataset:

Overstock.com initially sold exclusively surplus and returned merchandise on an online e-commerce marketplace, liquidating the inventories of at least 18 failed dot-com companies at below-wholesale prices[12].
The source of stock prices dataset are from Nasdaq[3] and Kaggle [1].This dataset was imported into excel and the Nvidia stock prices are from 2005 to Oct 2018 with attributes as
Date Open High Low Close Volume Market. Cap.

D. AMD dataset:

AMD's is a semiconductor company and its main products include microprocessors, motherboard chipsets, embedded processors and graphics processors for servers, workstations and personal computers, and embedded systems applications[11].The sources of stock prices dataset are Nasdaq [4] and Kaggle [1].
This dataset was imported into excel and the Nvidia stock prices are from 1983 to Oct 2018 with attributes as
Date Open High Low Close Volume Market. Cap.

III. RELATED WORK

While exploring about the previous work on the relationship between stocks and cryptocurrencies, I came across few articles which got published on various websites with headlines as Bitcoin-Related Stocks Aim to Capitalize on Cryptocurrency Craze[7] and 3 Stocks That Are Soaring Because of Bitcoin[6]. These articles gave me a positive insight into the hypothesis proposed in this project. Some works mainly mentioned that besides Nvidia making high-performance graphics processors for video games, this company really saw that their stocks have doubled in 2017 when they have been co-opted by mining operations that generate bitcoin and other digital currencies. As a result, bitcoin mining outfits have taken off around the world, Nvidia's revenue has climbed 40% for fiscal 2017. And its graphics cards have sold out around the world because the demand among miners is so high[7]. similarly, AMD has quickly earned a spot in the cryptocurrency hall of fame because of its graphics processors. It was also mentioned that the dramatic rise in the price of bitcoin and other cryptocurrencies is a big reason Overstock is pushing forward with blockchain deals and de-emphasizing the retail business. But that's not without risk. Shares of Overstock have been riding the bitcoin wave in 2017. The stock has soared more than 130% over, but it has plunged nearly 30% in 2018[13]. So these are some takeaways from these articles that showed positivity towards the further analysis

IV. ASSUMPTION AND LIMITATIONS

A. Assumptions

- Out of multiple columns like close price and all considered Open stock price and Date as attributes for time series analysis.
- During the time series analysis, even though Nvidia, AMD and Overstock were in existence before Bitcoin, during the trend analysis mainly the comparison in the trends were observed after the bitcoin existence (i.e. from 2013).

B. Limitations

- This study is limited to Bitcoin crypto currency
- The forecasting models which are developed in this project are limited to the stocks Nvidia, AMD and overstock. As the time series analysis is performed for only these three stocks.
- The analysis and conclusions made in this project are limited to only these three stocks and bitcoin. The same conclusions may not be valid for other coins or stocks.
- This project only explains the correlations of stocks and bitcoin, but not the correlation between various stocks.
- As the data on bitcoin prices is limited to 6 years (i.e 2013 to 2018) forecasting model was not developed for bitcoin but it was built for other stocks.
- Correlation factor is calculated from 2013 o 2018 as the bitcoin data is availble from 2013.
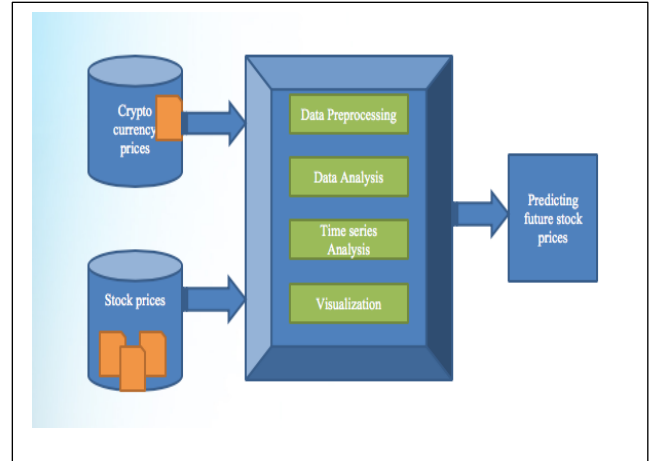
V. SYSTEM ARCHITECTURE

Fig.1 System Architecture

The overall process and steps followed in this project are as depicted above.

A. Data Collection :

Initially historical data of stocks and crypto currency are collected from various sources like Nasdaq and Kaggle and merged into a common csv file using Microsoft Excel. So the inputs to the data preprocessing stage are structures csv files of stocks(Nvidia, AMD, Overstock) and crypto-currency(Bitcoin).

B. Data Preprocessing:

To achieve better results initially performed few data cleaning steps.

- Observed the missing values using python and as the data is well structured, only found few missing values for earlier data(i.e. during 1987 for AMD stocks) So, the dataset has less than 2% of missing values. Hence removed those rows.

- Secondly, observed the inconsistency in date format, hence using excel converted the dates which are in string format to MM/DD/YY and ensured consistency throughout the data.

- Next, as the price range for bitcoin and various stocks are varied (i.e. 2 digits for AMD to 5 digits for bitcoins) and also the prices varied over course of years individually for each stock are at different ranges(one digit for Nvidia in starting to 3 digits now). So performed normalization on open prices and normalized data to [0,1] range while giving inputs to the forecasting model during training.

- Ensured that all the column types are nominal instead of string except date column (i.e. close, high, low)

So the Output of Data preprocessing phase and input for analysis phase is a cleaned csv file without NA values, with

consistent date format with date datatype and nominal values for other columns. And the input for forecasting model is a consistent date format and normalized Open attribute values.

VI. ANALYSIS

As part of data analysis, Initially visualized the open prices attribute of Bitcoin and three stocks on a single plot with date as reference and open prices as dependent variable. Found some commonality in trends between Nvidia, Overstock and Bitcoin.
Now in order to find the correlation between three stocks prices and bitcoin used Pearson correlation coefficient.

The monthly, quarterly and annual average open stock prices are calculated using pivot table in excel. Inputs to pivot table are daily open stock prices and Date.

An Input to weka tool is a csv file with the average monthly open prices of three stocks and bitcoin for finding correlation factor on monthly data.
Using weka tool I found the correlation between stocks and bitcoin for average monthly open prices.
From Table I the average monthly stock prices correlation, it is clearly evident that Nvidia and overstock stocks are highly correlated with bitcoin.

To find the correlation factor on quarterly data, used average quarterly prices of all the three stocks and bitcoin.
From Table II the average quarterly stock prices correlation, it is clearly evident that Nvidia and overstock stocks are highly correlated with bitcoin again with 0.916 and 0.815 respectively and poor correlation with amd when compared to other two stocks.

Table I. Correlation between average monthly prices

| | correlation factor on average monthly prices | | | |
	Bitcoin	*Nvidia*	*Overstock*	*AMD*
Bitcoin	-	0.841	0.803	0.654
Nvidia	0.841	-	0.661	0.912
Overstock	0.803	0.661	-	0.504
AMD	0.654	0.912	0.504	-

Table II. Correlation between average quarterly prices

| | correlation factor on average quarterly prices | | | |
	Bitcoin	*Nvidia*	*Overstock*	*AMD*
Bitcoin	-	0.916	0.815	0.749
Nvidia	0.916	-	0.708	0.928
Overstock	0.815	0.708	-	0.562
AMD	0.749	0.928	0.562	-

VII. TIME SERIES ANALYSIS

Time-series analysis is a basic concept within the field of statistical learning that allows the user to find meaningful information in data collected over time. Time series analysis is performed for three stocks as I have large amount of data collected for these three stocks and limited data on bitcoin prices due to its existence after 2013.
The inputs to the time series analysis is the daily open prices of all three stocks.
Following step by step procedure is used to perform time series analysis.

- Jupyter notebook IDE is used to perform time series analysis and entire code is written in Python.
- Loaded the csv files of stocks and used the normalized open prices to feed as input to machine learning model.
- Divided the dataset into test and train in ratio based on length , so that data after 2014 comes under test set.
- Used sliding window concept to get the train and test data with look back of nine values in order to predict the tenth value.
- Used deep learning architecture LSTM(deep learning model for sequential time series analysis) of stateful mode to train the model.
- Mean squared error is used as loss function and optimizer used in Adam.
- Predicted the results on test data and train data using the trained model.
- Calculated the mean_squared_error and test and train scores.
- Once we obtain the predicted values, these are exported into csv file along with date for visualization in tableau.

This machine learning model in order to predict the stock prices of today's value which is the output, will take the input of previous nine stock prices.
we have the actual values(i.e. the values from dataset) after 2013 for all the three stocks and from this model we will get the predicted values of all the three stocks after 2013.

A. Results

The Table III shows the RMSE of the machine learning model on train and test dataset. The actual and predicted prices which are varied over time are visualized in the next part visualization.
The forecasting time series analysis model is predicting with more accuracy for Nvidia, AMD and little less in Overstock. We can understand more about these values when we visualize these results in the next section

Table III. Time series analysis scoring model.

| Stocks | Scores of model | |
	Train score	*Test score*
Nvidia	0.54 RMSE	3.81RMSE
AMD	0.74 RMSE	3.36 RMSE
Overstock	0.94 RMSE	9.96RMSE

VIII. VISUALIZATION

As part of visualization, Initially I visualized the analysis results and then the results of time series analysis.

A. Trends in three stocks and bitcoin over a period of time

Fig.2 is for yearly trend analysis of average stock prices and bitcoin. From this plot, it is clearly evident that Nvidia, Overstock and bitcoin has similar trends from 2015.It is also evident that even AMD followed similar trend from 2015 to 2016 but couldn't catch the peak value of bitcoins during late 2017.

Fig.3 is for monthly and quarterly trend analysis of average stock prices and bitcoin. From this plot it is clearly evident that from 2017 Q3 to 2018 Q2 both Bitcoin and Overstock has a similar trend and they rose and fall together. Whereas for bitcoin and Nvidia, Nvidia rose along with bitcoin but did not share it's fall. We can also observe that bitcoin and AMD did not follow similar trend in this time frame.

In Fig.4 even though Y-axis has Year, the green points in the graph represents the monthly prices. We can observe that 2017 is rising tide year for bitcoin with starting from 814USD to 14,748USD by the end of year. From then it has seen major downfall for next 3 months and from then maintained some consistency with only small increments and small downs.

Fig.4 Monthly price variation of bitcoin.

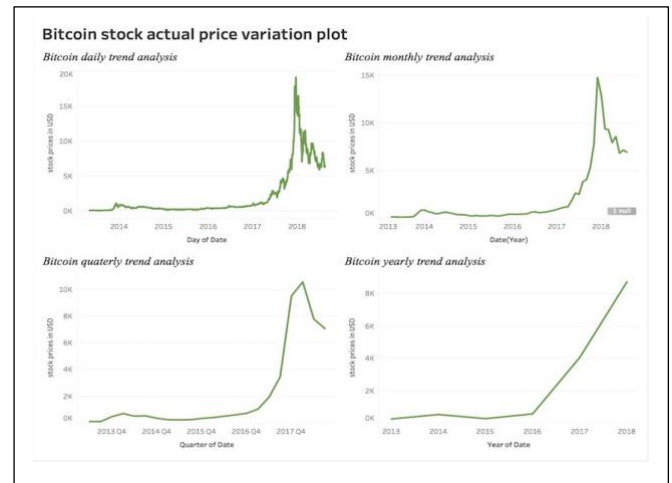Fig.5 Trend analysis of bitcoin.

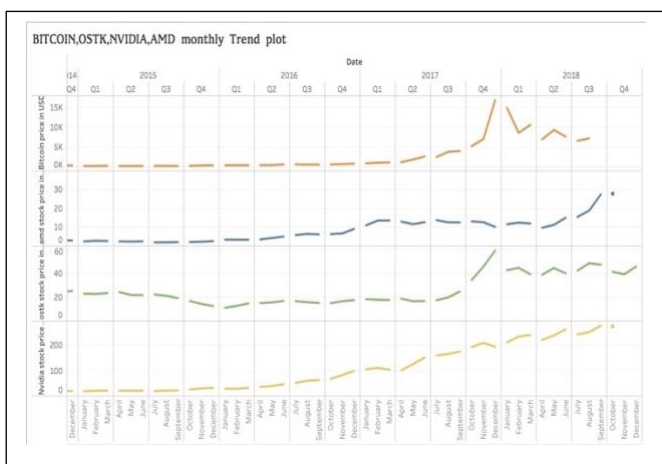Fig.2 Trend plot of average stock prices and bitcoin yearly

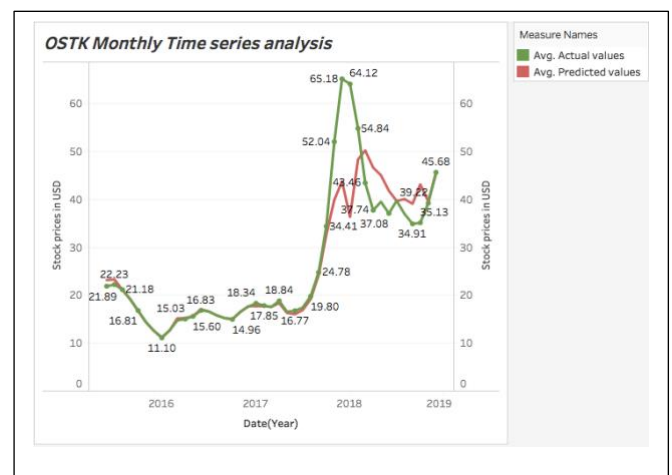Fig.3 is for average monthly prices of bitcoin and stocks.

Fig.6 Monthly price variation of overstock

B. Time series analysis of three stocks

Fig.6 is for monthly average stock prices of overstock. Here I want to make a comparison between overstock and bitcoin by considering actual values. We can observe that 2017 is rising tide year for overstock also with starting from 18USD to 65USD by the end of year. From then even overstock has seen major downfall for next 3 months and later maintained some consistency with only small increments and small downs. Thus it is clearly visible that bitcoin and overstock followed same trend from 2017 to start of 2018 year.

Fig.7 is dashboard of time series analysis of overstock stock prices over different time frames(i.e. daily, monthly, quarterly, yearly). We can see from the plot clearly that predicted values are trying to catch actual values but it could not catch sometimes due to unexpected rising tides. Hence we can say this model gained almost good insights from the historical data but not exact especially during extreme situations.

Fig.8 is a dashboard of time series analysis of Nvidia stock prices over different time frames(i.e. daily, monthly, quarterly, yearly). We can see from the plot clearly that predicted values are almost equal to actual values. Hence we can say that this model gained good insights from the historical data and helps in predicting the future prices too.

Fig.9 is a dashboard of time series analysis of AMD stock prices over different time frames(i.e. daily, monthly, quarterly, yearly). We can see from the plot clearly that predicted values are almost equal to actual values till starting of 2017 year. Hence we can say that this model gained almost good insights from the historical data but not exact especially during extreme situations.
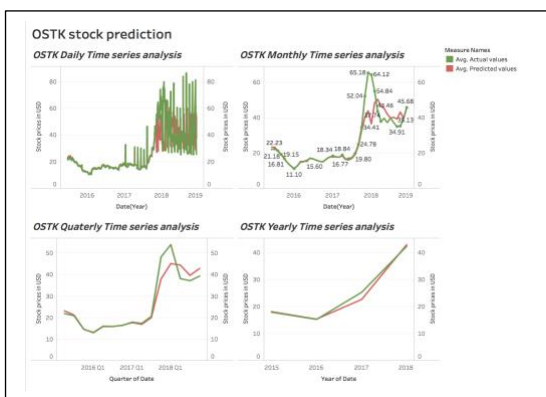

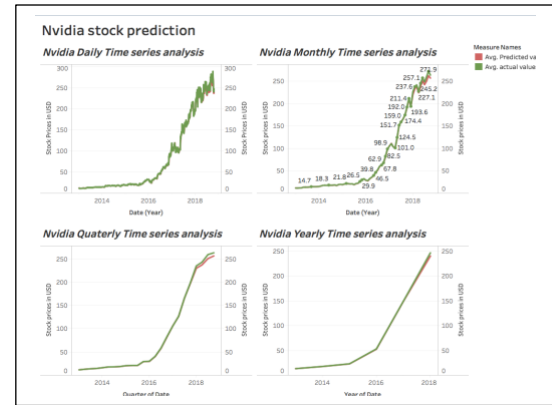Fig.7 Time series analysis of overstock stock prices

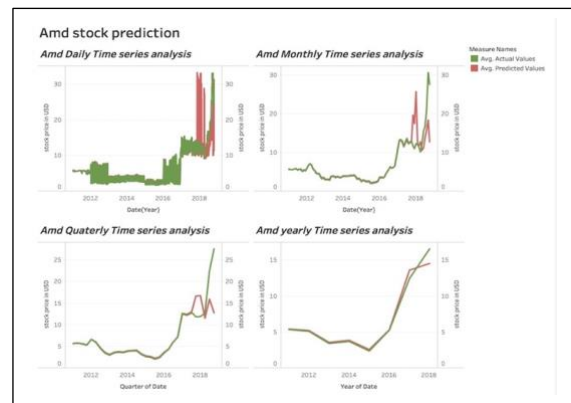

Fig.8 Time series analysis of Nvidia stock price

Fig.9 Time series analysis of AMD stock prices

From the first part of visualization it is clearly evident that bitcoin, overstock and Nvidia followed similar trend especially in 2017 year. And from the second part of visualization it is evident that how the machine learning model got insights from time series analysis.

IX. FUTURE WORK

- In the analysis phase, while finding the correlation factors, I observed that Nvidia and AMD stock has really high correlation coeffficient of 0.91. So further analysis can be done on this.
- This project scope is limited to bitcoin, but similar analysis can be done on various other crypto currencies and stocks.

X. CONCLUSION

From the analysis and visualization, due to the high correlation factor between Bitcoin, Nvidia and Bitcoin, Overstock. We may conclude that stock prices of Nvidia and overstock are impacted by bitcoin hence the hypothesis framed in case of Nvidia and overstock are proved.

But even though AMD graphic cards are used for mining in cryptocurrencies due to the poor correlation between Bitcoin and AMD, I am unable to prove that AMD stock prices are impacted by Bitcoin prices hence the hypothesis in case of bitcoin and AMD is rejected, but this may be proved in case of AMD and other crypto coins.

REFERENCES

[1] Boris, Marjanovic. (20, October 2017). Huge Stock Market Dataset. Retrieved from

https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs/home

[2] NVIDIA Corporation Common Stock Historical Stock Prices. (20, October 2018). Retrieved from

https://www.nasdaq.com/symbol/nvda/historical

[3] Overstock.com, Inc. Common Stock Historical Stock Prices. (20, October 2018). Retrieved from

[4] https://www.nasdaq.com/symbol/ostk/historical

[5] Advanced Micro Devices, Inc. Common Stock Historical Stock Prices (20, October 2018). Retrieved from

https://www.nasdaq.com/symbol/amd/historical

[6] Historical data for Bitcoin. (20, October 2018). Retrieved from https://coinmarketcap.com/currencies/bitcoin/historical-data/

[7] Ryan Derousseau. (30, November 2017).3 Stocks That Are Soaring Because of Bitcoin. Retrieved from

http://time.com/money/5042272/stocks-investing-bitcoin/

[7] Nathan, Reiff. (08, November 2017). Bitcoin-Related Stocks Aim to Capitalize on Cryptocurrency Craze. Retrieved from https://www.investopedia.com/news/bitcoinrelated-stocks-aim-capitalize-cryptocurrency-craze/

[8] Tae, Kim. (06, JUNE 2017). AMD shares are surging with bitcoin because digital currency 'miners' need its graphics cards.Retrieved from https://www.cnbc.com/2017/06/06/amd-surges-because-bitcoin-miners-need-its-graphics-cards.html

[9] Bitcoin. Retrieved from https://en.wikipedia.org/wiki/Bitcoin

[10] Nvidia. Retrieved from https://en.wikipedia.org/wiki/Nvidia

[11] Advanced Micro Devices. Retrieved from

https://en.wikipedia.org/wiki/Advanced_Micro_Devices

[12] Overstock.com. Retrieved from

https://en.wikipedia.org/wiki/Overstock.com

[13] Paul R. La MOnica.(10 August, 2018) Remember Overstock? It's basically a crypto company now. Retrieved from

https://money.cnn.com/2018/08/10/technology/overstock-blockchain-bitcoin-investment-earnings/index.html