Project 2: Draft Milestone 3 – Cryptocurrency Price Speculation & Prediction

Thip Rattanavilay
DSC-680
10/17/2021

Abstract

I have spent the entire year of 2020 understanding Blockchain and Cryptocurrency, especially Bitcoin and Ethereum. These two crypto are the most volatile markets today and has gained a lot of attention from investors across the globe. Cryptocurrency, being a novel technique for transaction system, has led to a lot of confusion among the investors and any rumors or news on social media has been claimed to significantly affect the prices of cryptocurrencies. The goal of this study is to predict prices for crypto coins using Machine Learning Techniques for the few I weeks for project 2 I will prepare a strategy to maximize gains for investors, friends and of course myself. I also aim to find out if there is a co-relation between fluctuating crypto prices and related news.

Introductions

Stock market is one of the most volatile data available in terms of Machine learning datasets. Researchers have been long trying to predict the stock market and any breakthrough in this field would result in, literally, the people being able to mint money. Cryptocurrencies, to be specific, has gained a lot of traction in the recent years from investors across the globe. Cryptocurrency being a novel technique for transaction system has led to a lot of confusion among the investors and any rumors or news on social media has been claimed to significantly affect the prices of cryptocurrencies. Bitcoin is one of the oldest and biggest cryptocurrencies being traded as of now, in terms of the volume being traded. It is so big that even now, with the advent of thousands of new cryptocurrencies, Bitcoin has a market share of more than 55% as compared to other cryptocurrencies, being followed by Ethereum at 8.57%. This says a lot about why Bitcoin might be a really interesting and important stock to predict. Also, Bitcoin prices fluctuate heavily. Over the past 2 years, Bitcoin has seen its highest price around $20000 and its lowest price around $900. It is very sporadic, and this is one of the most important reasons which attracted us to analyze and predict its price. Our rest of the paper would discuss our various attempts to predict its price and us trying to reason out how external factors like news affect its price.

Business Problem

Cryptocurrency, especially Bitcoin, is one of the most volatile markets today and has gained a lot of attention from investors across the globe. Cryptocurrency, being a novel technique for transaction system, has led to a lot of confusion among the

investors and any rumors or news on social media has been claimed to significantly affect the prices of cryptocurrencies. The goal of this study is to predict prices for Bitcoin using Machine Learning Techniques for the next day and prepare a strategy to maximize gains for investors. We also aim to find out if there is a co-relation between fluctuating Bitcoin prices and related news.

Dataset

For the purpose of this research, I will utilize data from Jan-2018 to Aug-2019, concerning the hourly prices in USD and were divided into training set consisting of data from Jan-2018 to Feb-2019 (10176 values) and testing set from Mar-2019 to Aug-2019 (4416 values). This data was taken from Kaggle.com, kraken.com website and others like Coinbase etc., which is a trading platform and data repository for cryptocurrency exchanges. Also, I will utilize those data for forecasting and predictions (number of past prices taken into consideration).

Variables:
Open
Close
Volume
Volume Currency
Weighted Price Date

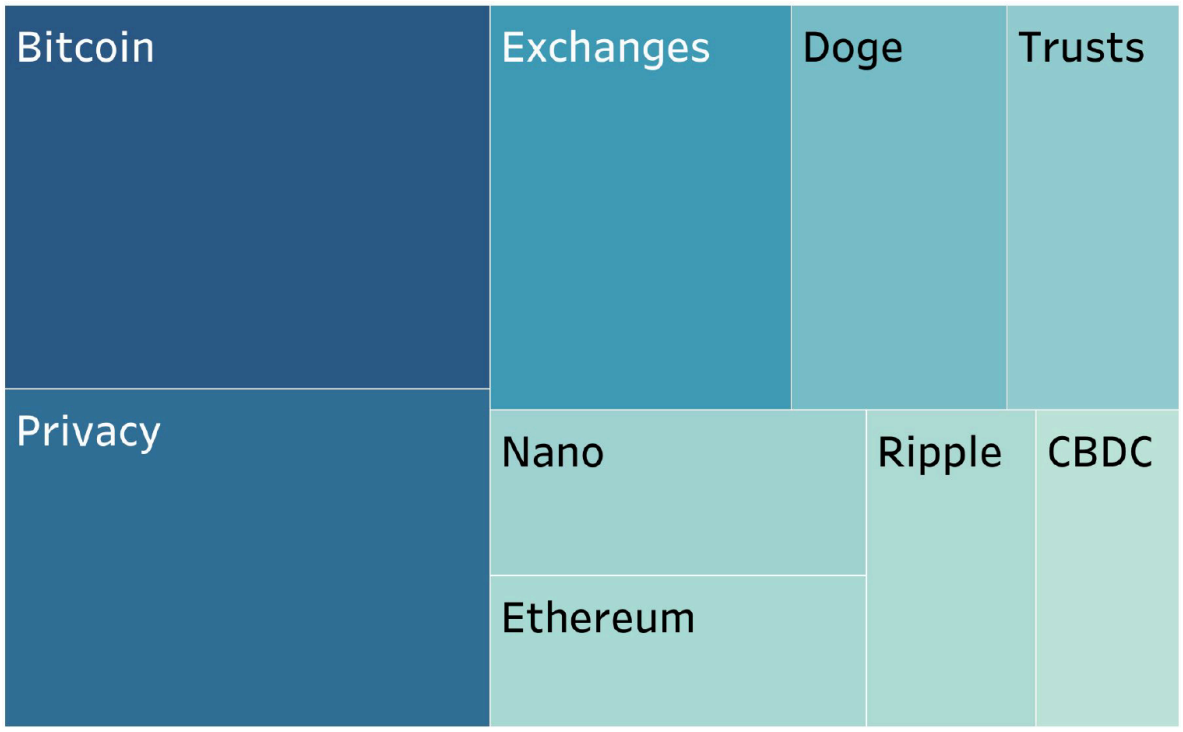
Crypto Variables:
BTC
ETH
LTC
XRP
ETC
STR
DASH

Methods

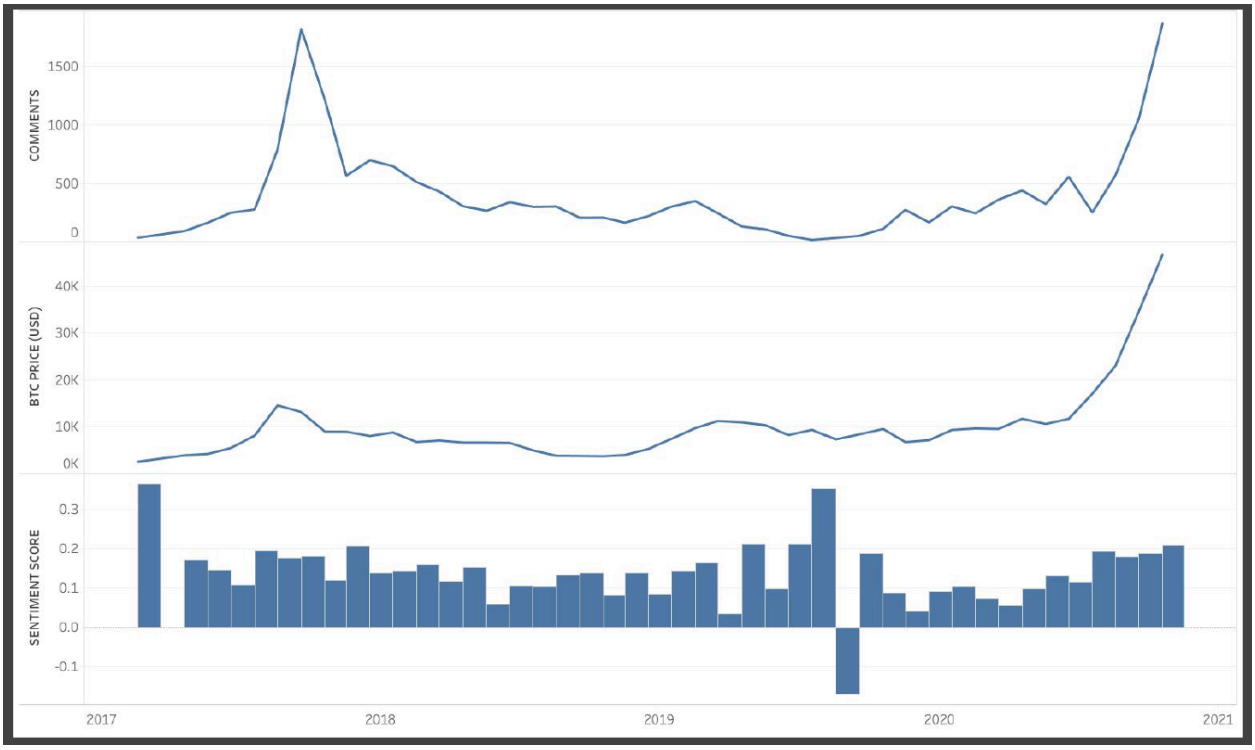
To solve the problem using Machine learning, we first tried to categorize the problem and tried to find previous solutions on how they solved it. We quickly learned that, since the problem involves prices which are changing with time, this could be modelled as a Series prediction problem. In parallel, we also tried to solve the problem as a normal machine learning problem with the features being the previous prices and the output being the price predicted for that day. We explain what models we have used and how we configure them to predict the Bitcoin prices.

Illustrations

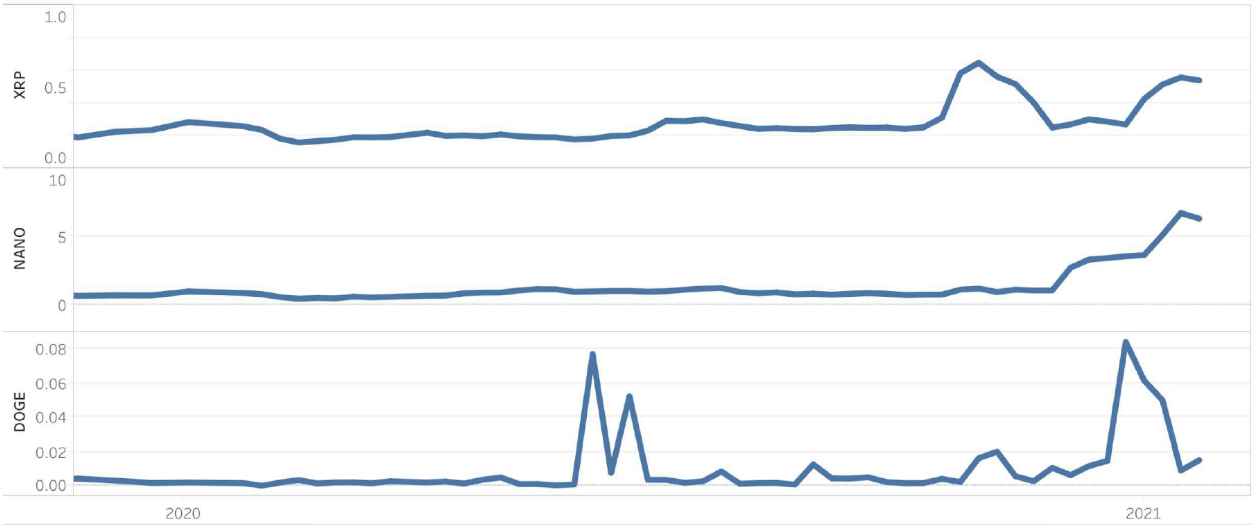
Top crypto coins on the market



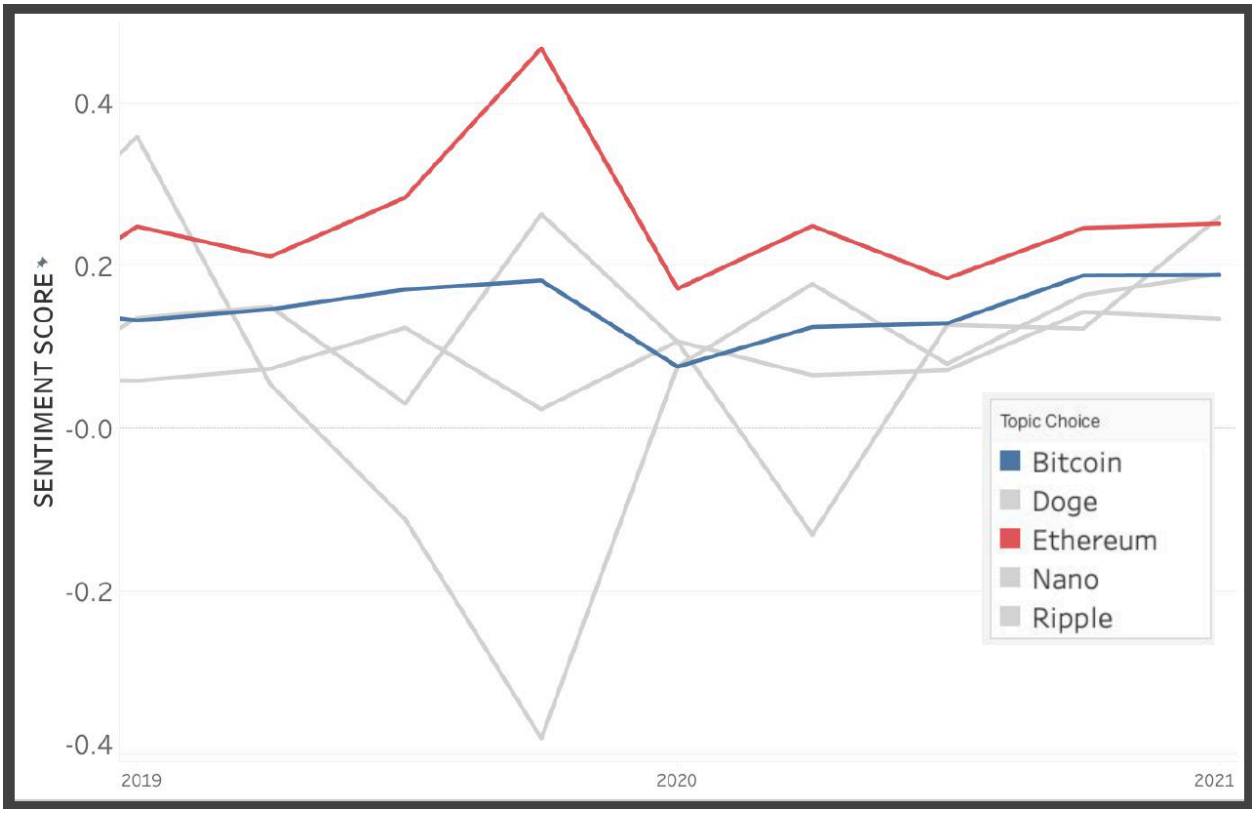
Sentiment Twitter posts



Risky Investments



Bitcoin vs Ethereum vs Doge vs Nano vs Ripple



Appendix

The input consists of a list of past Bitcoin data with step size of 256. The output is the predicted value of the future data with step size of 16. Note that since the data is ticked every five minutes, the input data spans over the past 1280 minutes, while the output cover the future 80 minutes. The datas are scaled with Min MaxScaler provided by sklearn over the entire dataset. The loss is defined as Mean Square Error (MSE).

Model	#Layers	Activation	Validation Loss	Test Loss (Scale Inverted)
CNN	2	ReLU	0.00029	114308
CNN	2	Leaky ReLU	0.00029	115525
CNN	3	ReLU	0.00029	201718
CNN	3	Leaky ReLU	0.00028	108700
CNN	4	ReLU	0.00030	117947
CNN	4	Leaky ReLU	0.03217	12356304
LSTM	1	tanh + ReLU	0.00007	26649
LSTM	1	tanh + Leaky ReLU	0.00004	15364
GRU	1	tanh + ReLU	0.00004	17667
GRU	1	tanh + Leaky ReLU	0.00004	15474
Baseline (Lag)	-	-	-	19122
Linear Regression	-	-	-	19789

1. Python
2. Tensorflow
3. Keras
4. Pandas
5. Numpy
6. h5py
7. sklearn

BTC_Close	BTC_Volume	ETH_Close	ETH_Volume
434.33	36278900.0	0.948024	206062
433.44	30096600.0	0.937124	255504
430.01	39633800.0	0.971905	407632
433.09	38477500.0	0.954480	346245
431.96	34522600.0	0.950176	219833

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 7, 512)            1058816
_________________________________________________________________
dropout_1 (Dropout)          (None, 7, 512)            0
_________________________________________________________________
lstm_2 (LSTM)                (None, 7, 512)            2099200
_________________________________________________________________
dropout_2 (Dropout)          (None, 7, 512)            0
_________________________________________________________________
lstm_3 (LSTM)                (None, 512)               2099200
_________________________________________________________________
dropout_3 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 513
_________________________________________________________________
activation_1 (Activation)    (None, 1)                 0
=================================================================
Total params: 5,257,729
Trainable params: 5,257,729
Non-trainable params: 0
```

10 Questions

What are the causes dips and spikes in crypto values?

How should I invest?

If I invest into crypto what are the risk?

What is cryptocurrency?

What is the purpose of cryptocurrency?

What is blockchain?

Is cryptocurrency taxable?

How do cryptocurrency markets look like int he coming days, weeks, or years?

Are the markets for different altcoins or largely independent affect the market cap for crypto?

How can I predict what will happen next?

Reference:

cryptocurrency prices, charts and market capitalizations. CoinMarketCap. (n.d.).Retrieved October 3, 2021, from https://coinmarketcap.com/.2. Polygon.io. (n.d.). Retrieved October 3, 2021,

Sigalos, M. K. (2021, October 2). Ethereum had a rough September. here's why and howit's being fixed. CNBC. Retrieved October 3, 2021,romhttps://www.cnbc.com/2021/10/02/ethereum-had-a-rough-september-heres-why-and-how-it-gets-fixed.html.

Yahoo! (n.d.). The Crypto Daily – Movers and Shakers – october 2nd, 2021. Yahoo!Finance. Retrieved October 3, 2021, from https://finance.yahoo.com/news/crypto-daily-movers-shakers-october-002300815.html.6. Srk. (2021, July 7).

Cryptocurrency historical prices. Kaggle. Retrieved October 3, 2021,from https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistory.
Taniaj. (2018, June 3). Cryptocurrency price forecasting. Kaggle. Retrieved October 3,2021, from https://www.kaggle.com/taniaj/cryptocurrency-price-forecasting.

Kash. (2021, September 11). Ethereum Cryptocurrency Historical Dataset. Kaggle.Retrieved October 3, 2021, from https://www.kaggle.com/kaushiksuresh147/ethereum-cryptocurrency-historical-dataset.

Sudalairajkumar. (2017, November 8). Cryptocurrency data pull. Kaggle. RetrievedOctober 3, 2021, from https://www.kaggle.com/sudalairajkumar/cryptocurrency-data-pull.

Coindesk: Bitcoin, Ethereum, crypto news and Price Data. CoinDesk Latest HeadlinesRSS. (n.d.). Retrieved October 3, 2021, from https://www.coindesk.com/learn/.