Title, author, abstract, keywords, introduction, methods, results, discussion, and references.

Title

Bitcoin Price Analysis

Author

Rohit Gampa(rgampa@iu.edu), Mitali Tavildar(mtavilda@iu.edu)

Abstract

The idea behind bitcoin price analysis is to ask very interesting questions regarding the price of bitcoin as well as to look at the analysis from an investor perspective to see if there are opportunities that can be observed from the historical bitcoin price data.

Keywords

Bitcoin, Binance, time-series, analysis, Cryptocurrency, trading, investment, backtest, prophet, python, tensorflow, LSTM, dropout, google trends, pytrends.

Introduction

We want to be able to analyze price data to gain understanding on how the price changes over a period as well as understanding the effect of google searches on the price via google trends data.

For this we had looked at traditional stocks as well as cryptocurrencies, to choose which will be better for analysis.

We have chosen to analyze bitcoin price due to the following reasons:

* Traditional Stocks trade only for 250 days in a year, for 8 hours a day Vs Cryptocurrencies like bitcoin trade 24 hours a day, 7 days a week.
* Good Quality historical data for the traditional markets is hard to get and is expensive as well, Eg: Yahoo finance has data for stocks daily candle data, but it has been known to have issues sometimes in Data Quality.
* Also getting data for traditional stocks for a hourly candles seems to be a very costly, as we need to pay a monthly subscription to get this kind of data.
* After filtering out the traditional stock market, we looked at various cryptocurrencies and choose bitcoin due to its highest volume of trades and it being a popular cryptocurrency.
* We choose bitcoin as it is known to almost everyone , and it makes it easier to get google trends data.

1.Methods

1.a Getting the hourly and Daily Bitcoin Data

Using Binance API 'https://api.binance.com/api/v3/klines' we created a custom code ‘download\_candle\_stick\_data.py’to download the download the bitcoin data, as like all companies binance does not allow to download a large amount of data, as it can cause server load issues. To circumvent this we are not using any tokens generated by binance, as we do not need any proprietary information from the account and created a loop for all the historical data we want, since we want the data from the beginning of the exchanges, we loop though this api with small chunks of time to get the data and we append this data to the main dataframe till we get the latest information, then the dataframe is stored as a csv.

This file can then be used to load the data into python for analysis.We used the above code to download the historical data for the traing pair “BTC-USDT”, as this has the highest volume of trades.

1.b. Exploratory data analysis

To EDA, we considered the records that described the highest price, lowest price, open time, close time, open price, and close price for each day.

The preprocessing steps:

* Converting the columns consisting of dates from Unix timestamp to date time format.
* Creating customized columns based on data: -
* To analyze the close price and open price for each day.
* To analyze the open price and the highest price for each day.

Using the customized columns, we use the below mentioned formula to calculate gain each day:

**Gain(at each day) = Close price(at each day) – Open price(at each day)**

**If Gain > 0, Gain\_Achieved = TRUE else,**

**Gain\_Achieved = FALSE**

Using the customized columns, the correlation between Gain and difference between (open and highest prices) was inferred to have a value of 0.60.

This provided an insight that the Gain and increase in highest rate for the day greater than open rate in positively correlated, increase in one could possibly increase the probability of increase the other.

1.c. Predicting gain using classic algorithms:

Feature columns: open, low, high of the day

Label column: Gain

i. Logistic regression

Using the logistic regression as a baseline classification algorithm, we fitted our data to the model to predict if there is a probability of achieving gain at the end of the day.

ii. KN neighbors classifier

Fill with some information abount KNN, why it was chosen, etc

2.c. Trend Analysis using Prophet

Prophet is an open-source algorithm for generating time-series models that uses a few old ideas with some new twists. It is particularly good at modeling time series that have multiple seasonalities and doesn’t face some of the above drawbacks of other algorithms.[[1]](#footnote-1)

For trend analysis and prediction, records that described the highest price, lowest price, open time, close time, open price and close price for every hour were utilized.

The Fig3. Shows the general trend followed by our data.

Following the steps mentioned in the Prophet documentation[[2]](#footnote-2), we fit our data into the Prophet model. The resultant trend prediction is displayed in Fig5, the Prophet was able to catch the upward trend and predict the same.

Prophet.plot\_components method was used to see the trend, yearly seasonality, and weekly seasonality of the time series. Fig4 displays the output from Prophet.

1.d Getting google trends Data.

For getting google trends data, I have use the pytrends package. I have downloaded two types of Google trends data using this method.

1. Hourly breakup of google search data
2. Daily breakup of google search data.

In pytrends methods, we need to input the duration for which we need the data, for what type of google search , as well as the geographic information. Since we need the google trends data for the time we have historical data, I have downloaded data from 2017 August to present, setting the search to only google searches, not including image searches, and setting the geographic to “”, which stands for global search volumes.

The code below should be able to explain how the data is downloaded.

Text

Description automatically generated

A picture containing graphical user interface

Description automatically generated

For downloading the daily google trends data, I cut the total time we needed the data into years and downloaded them individually, as the google api’s were giving timeout errors.

**Results:**

**Results for EDA by Comparison using visualization:**

Chart, line chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

Fig1. Prices at open time vs prices at close time at each day

Fig2. Prices at open time vs highest price each day

We observe the overlap in the graphs for both the open price and close graph each day Fig1 with a lag of 1. The prices are continuous, which means the closing rate from previous record is same as the opening rate for the next record, this helps in feature selection. As compared to traditional stock market the open price of the stock next day changes from the previous day’s close.

There is visible difference in the open price and highest price of the day Fig2. We see that the highest price (orange line graph) has risen above the price. This shows the volatility of the price of bitcoin.

We observed that if a person invested at the start of the day and sell by the end of the day, he would have made 53% profitable trades.

Accuracy: 87%

Accuracy: 71%

Rather than using the above strategy of investing all day, if we created a model, which helps the person investing know if the following day will be profitable of not, for this we created a logistic regression model, and a KNN model as specified previously in methods.

Here are the results:

1. Accuracy of 87% achieved from logistic regression
2. Accuracy of 71% achieved from KNN method.

From the above results we conclude that Logistic regression method performed the best.

**Results for the Trend Analysis using Prophet.**

Using the outputs, analysis on general trend was done.

* The Fig4 clearly displays a drop in trend during Fridays, which suggests approximately the best time to buy.
* Similarly, the trend suggests the best time to sell is Monday.
* By looking at the trend for the day, we see that 0100 is the best time to buy bitcoin and 2000 is the best time of the day to seel bitcoin.

Looking into the trend of bitcoin, we see that the trend is increasing which means currently we are in a bitcoin bull market.

Chart, histogram

Description automatically generated Chart

Description automatically generated

Fig3. Prices at close time at each hour of the day

Chart, histogram

Description automatically generated

Fig4. Trend analysis and prediction for varied time components

Fig5. Trend prediction by Prophet

1. [↑](#footnote-ref-1)
2. [↑](#footnote-ref-2)