Tweets and Bitcoin Predictions

André de Oliveira Gomes

Fullstack Bootcamp JEDHA
30 November 2020

BITCOIN FACTS:

2008 - Satoshi Nakamoto's original Bitcoin paper is published

2009 - The first Bitcoin transaction is performed in block 170, from Satoshi Nakamoto to Hal Finney

2011 - 1 BTC = 1 USD

2013 - The total Bitcoin market capitalization exceeds 1 billion USD
Bitcoin mining difficulty passes 1 billion in December 2013

2017 - Bitcoin broke the $1000 mark in early January

Questions:

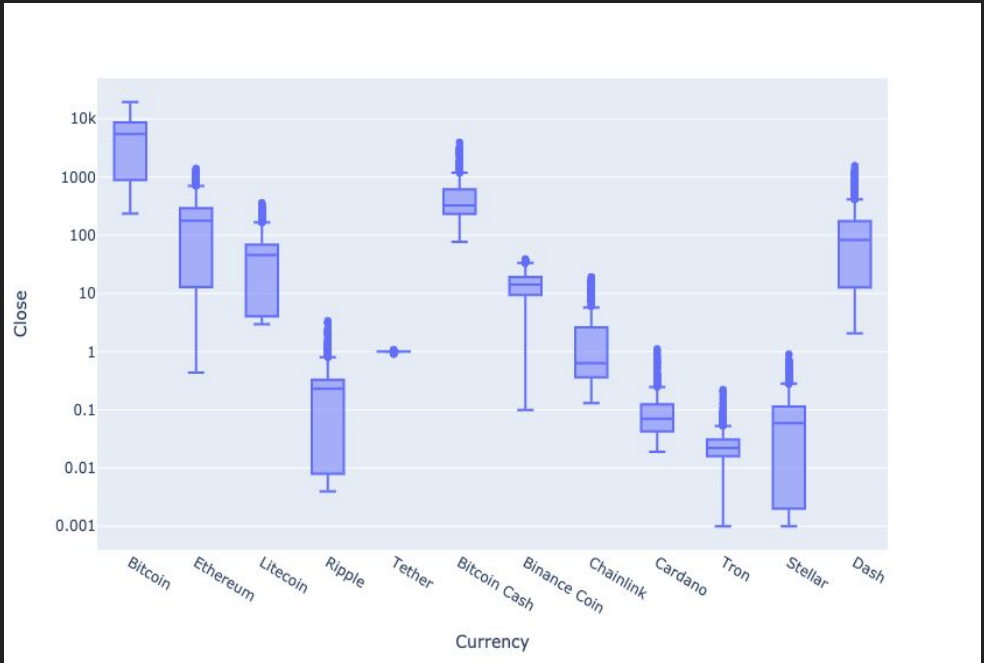
- ➔ HOW DID SOCIAL MEDIA OPINION ON BITCOIN EVOLVED DURING THIS TIME?

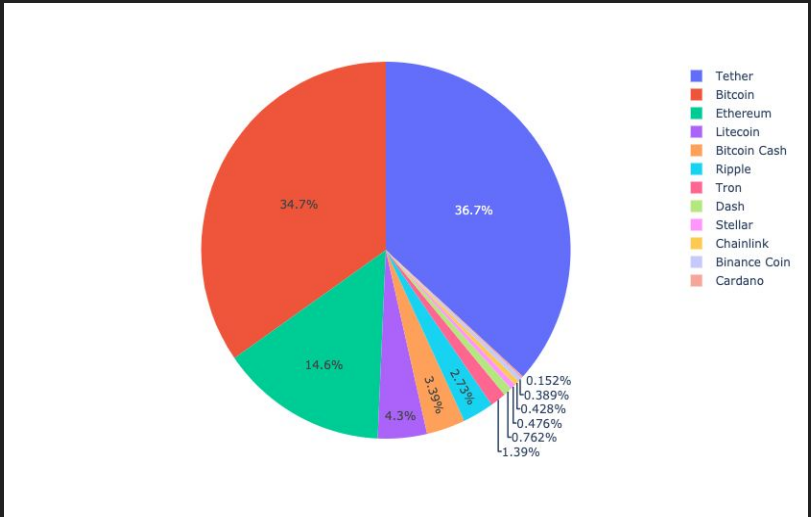
- ➔ CAN BITCOIN PRICES BE EXPLAINED BY TWEETS OF CRYPTO INFLUENCERS ?

PLAN:

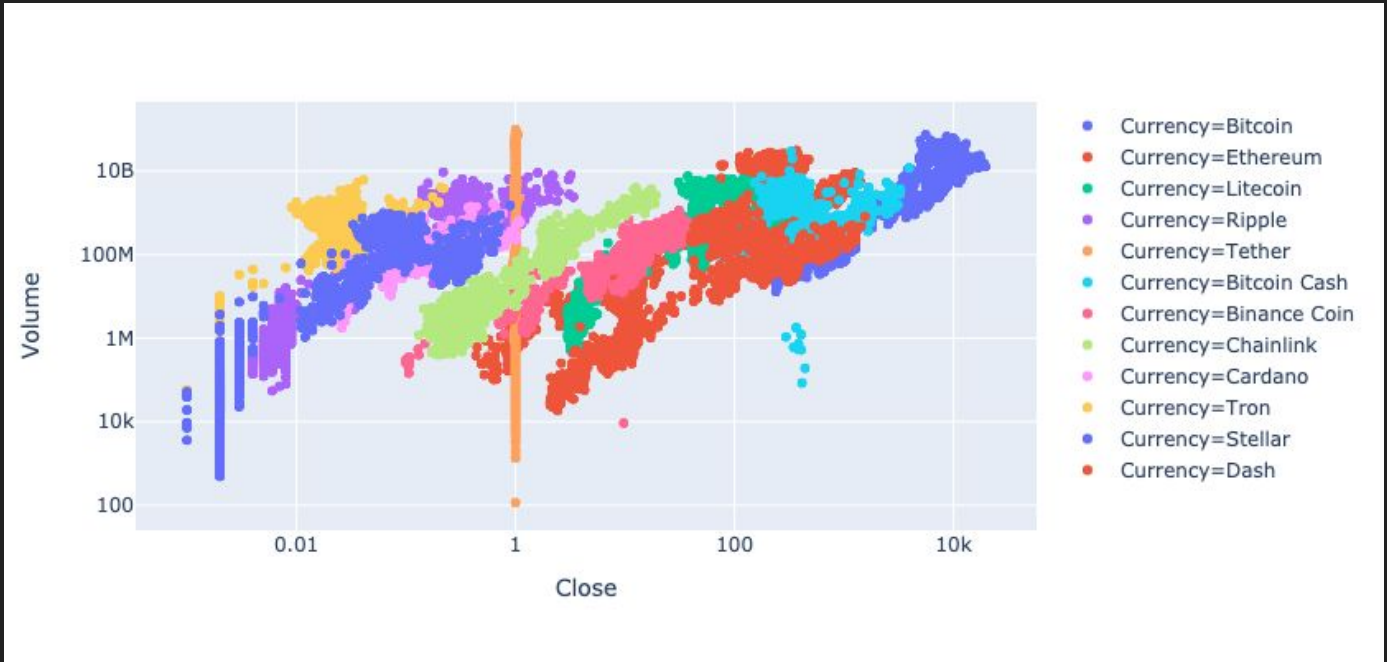
1. Exploration of crypto data.
2. Sentiment analysis on tweets.
3. Predicting prices of Bitcoin with a model is feeded by the tweets.
4. How good or bad are those predictions?
5. Brainstorm to start to think about it!

The crypto space

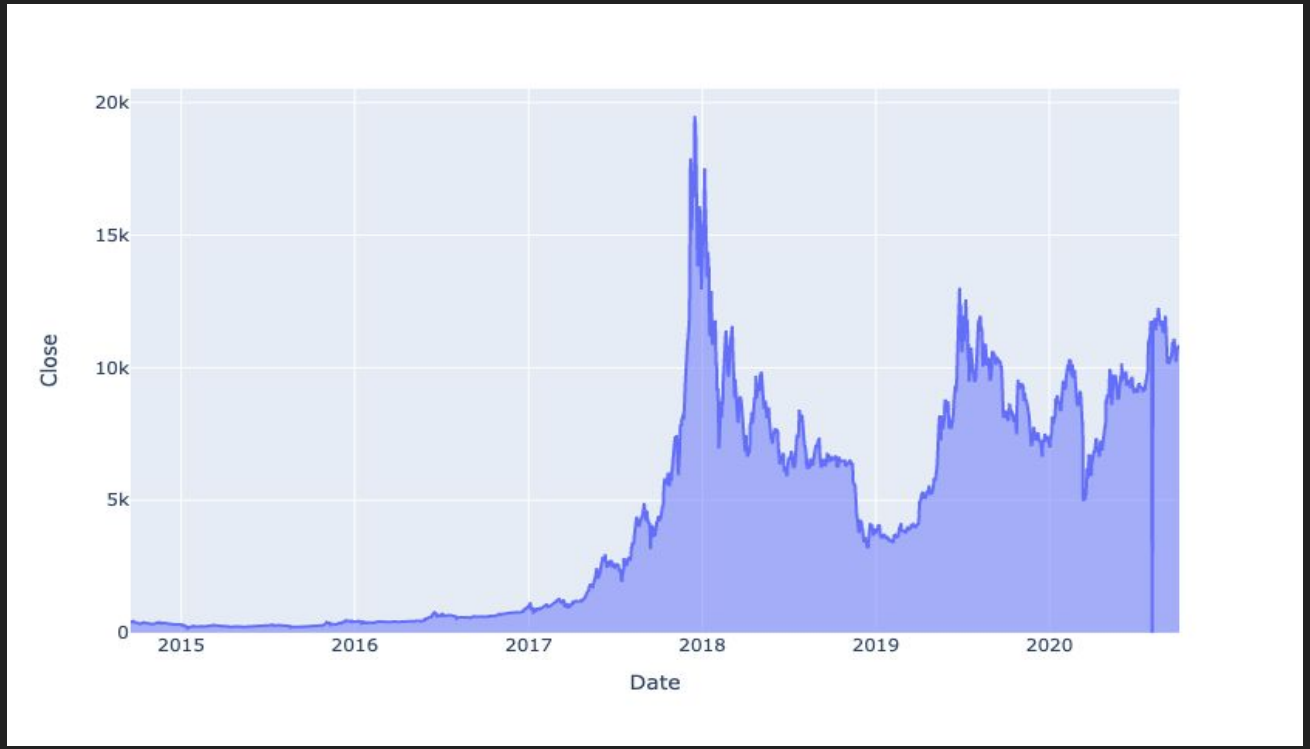




BITCOIN PERFORMANCE IN THE CRYPTO SPACE

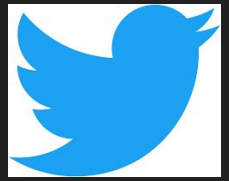


BEVOLUTION OF THE BITCOIN PRICES



Horizon of our analysis:

Start date: 16.04.2018
End date : 14.09.2020



AND ABOUT THE OPINION IN SOCIAL MEDIA?

Strategy:

1. We searched a list of crypto Twitter influencers:

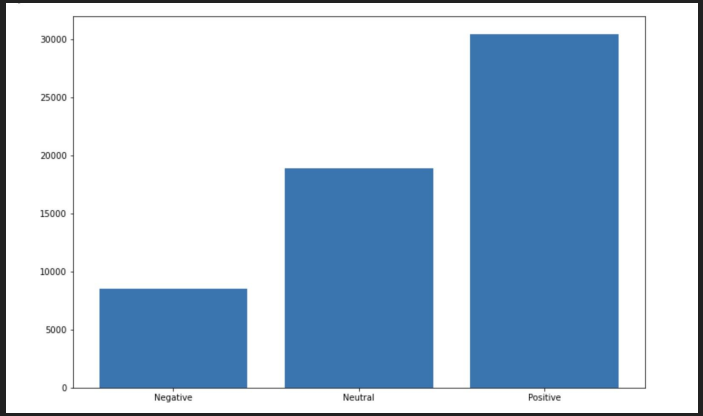
- Coinmarket
- Coindesk
- Nexo
- Vitalik Buterin
- Erik Voorhees
- Barry Silbert
- Tim Draper.
- Charlie Lee
- Brian Armstrong
- Fred Ehrsam
- Andreas Antonopoulos

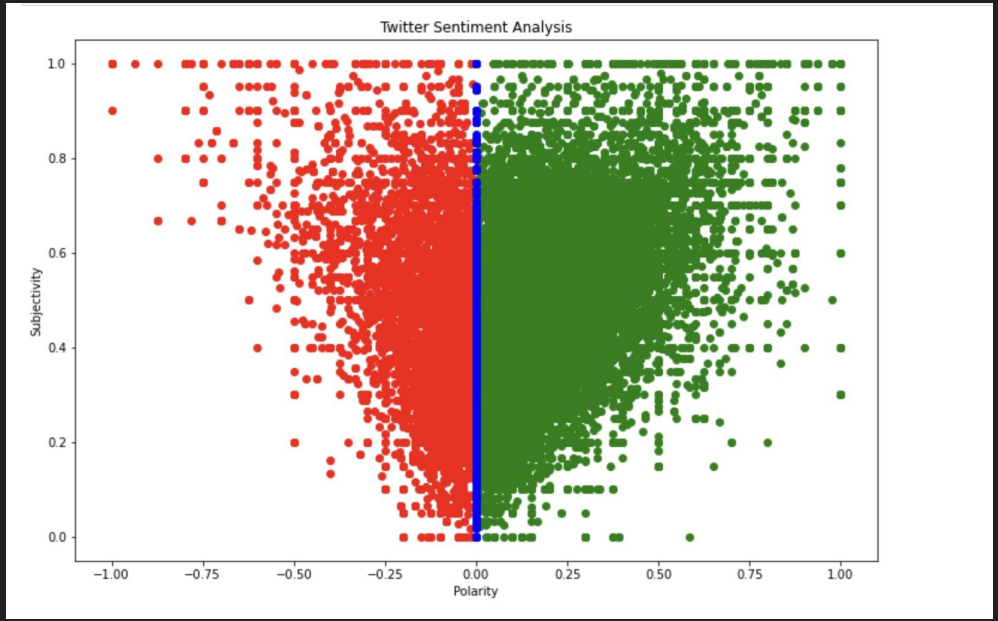
2. Scrapping with Twitter API and Tweepy during the time period from 16.04.2018 to 14.09.2020

3. Sentiment analysis on the 64129 tweets.

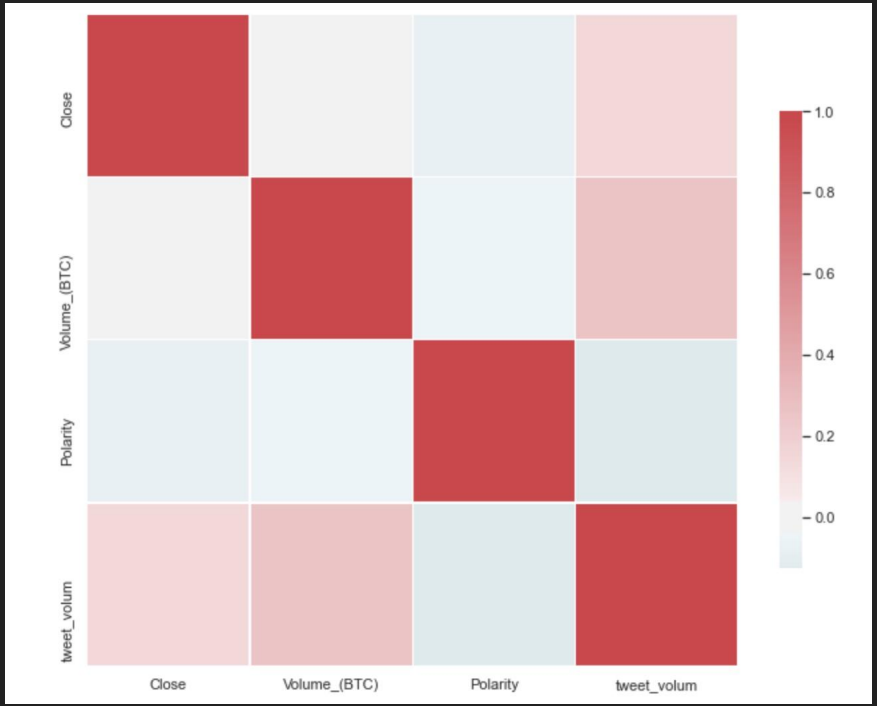
4. Daily sentiment and number tweets per day in the model of prediction of the daily close value of Bitcoin in that period.

Sentiment analysis on the tweets





Correlations



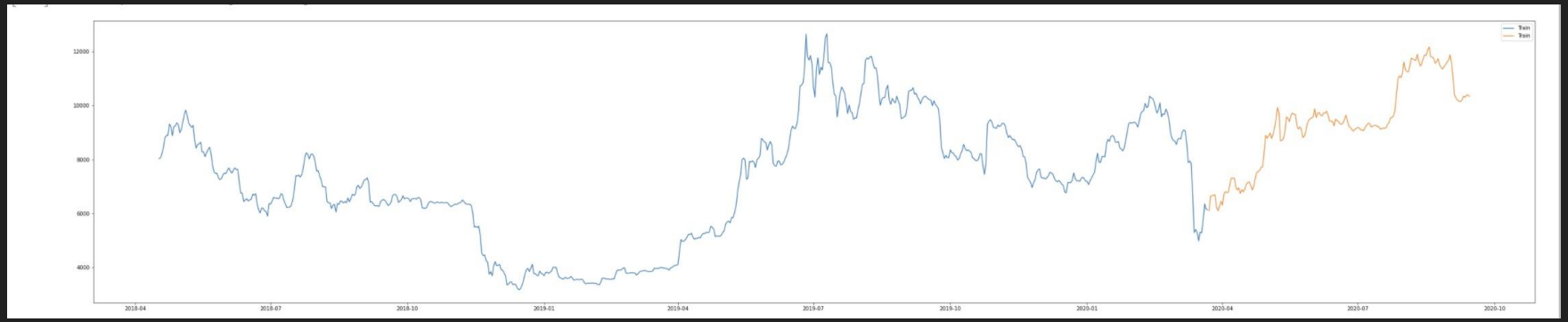
Model:

1. There is a non negligible correlation between tweet volume, volume and close price of the currency.

2. We build a neural network for the regression problem of describing the close price of BTC wrt tweets, polarity tweets, volume currency

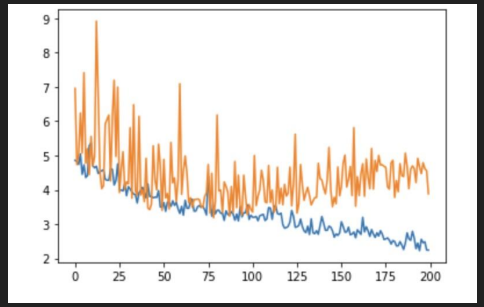
Model:

1. We separate the dataset with a 20% validation set
2. Model with several LSTM layers with a stop on the bias of the parameter of the last dense layer of the model.
3. We use the predictions on the dataset to feed the Prophet algorithm and retrieve predictions from the period 14.09.2020-28.11.2020



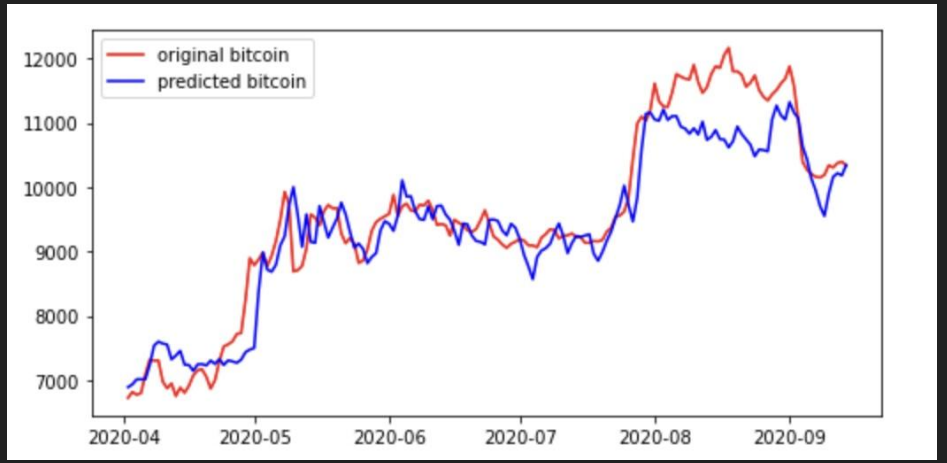
How the model fits the data:

Loss function:
MSE TRAIN SET: 2.25%
MSE TEST SET: 3.89%

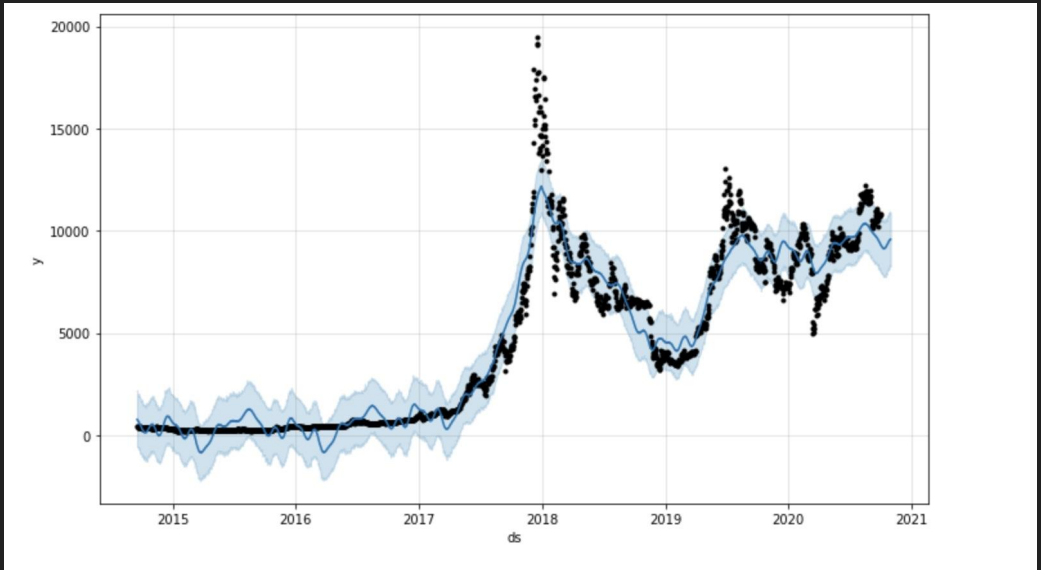


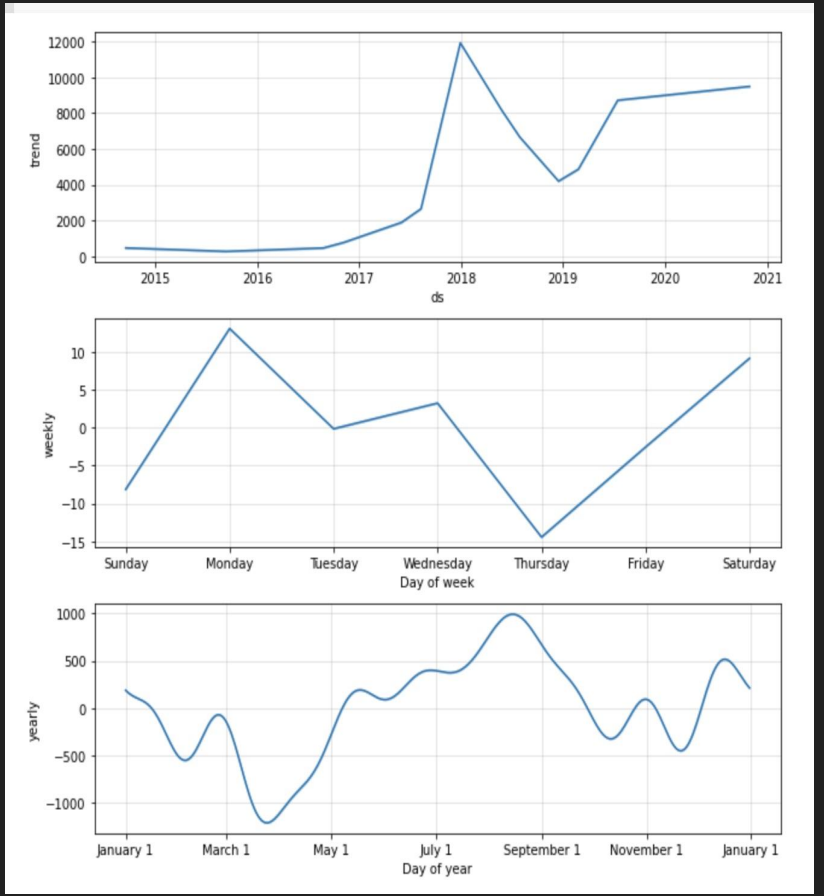
In the training set

In the test set

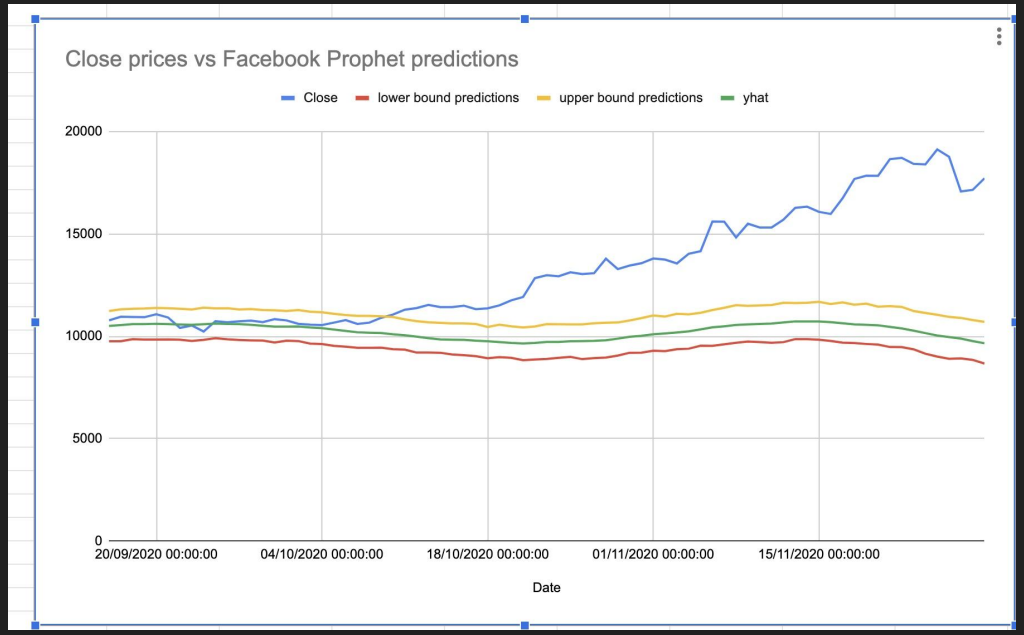


Prophetizing the future: until 28.11.2020.





Comparison with real close prices



To do:

1. Feed the LSTM model with new data until recent time and contrast the predictions with the actual close prices to see if the model continues to perform well or not so well.
2. Tune the hyper parameters.
3. Include as regressors external variables such as indexes of stock markets.
4. Include crypto internal variables such as crypto indexes
5. Perform the study for rising coins.

Conclusions:

- We use the predictions of the LSTM model from mid April until mid September to feed the Prophet model and get nes predictions until the date of yesterday. Different logic than the LSTM model before.

MERCI!
MUITO OBRIGADO

IF ANY INTEREST I AM CHARGING LOW COMMISSIONS FOR FUTURE INVESTORS. I ALSO ACCEPT BLACK FRIDAY'S COUPONS.