Nathaniel Loken
Andres Hernandez
Ali Ahsan Muhummad Muzaheed
Afzal Hossain
Nathan Gonzalez
CS487/519
Final Report
5/1/2021

Bitcoin Price Prediction

Motivation:

Bitcoin is a currency designed to exist outside of governmental entities and therefore has a lot of variability. Unlike regular currency the price is based on what people believe it to be rather than based on a counties GDP or other factors. Due to the boom we have seen in the past few years this has caused the Bitcoin market to become extremely unpredictable and volatile. Moreover, there are many alternative cryptocurrencies arising such as Ethereum, Tether, Litecoin, etc. however, this group will only be focusing on Bitcoin as it is the most prominent, but this research could be applied to other crypto currency in the future. The motivations we have as a group are based on the want to implement the code and concepts for learning, and the possibility of financial gain.

Problem Definition:

The main issue with this currency is that due to the volatility of the Bitcoin market, because of its decentralized nature, it has made the currency extremely unpredictable. For instance the price of this currency fluctuates so much that if a payment of $20,000 was made to some entity by the time the payment is processed the Bitcoin could be worth only $15,000 which is a massive difference. This creates a problem when attempting to use this currency and creating accurate pricing for companies, meaning that bitcoin cannot be used as it was intended in its current state. Therefore, from the perspective of social financial security it would be beneficial to solve this problem.

There are various calculations utilized on financial exchange information for value forecasts. The parameters influencing Bitcoin are extraordinary, hence it is important to anticipate the estimation of Bitcoin so correct venture choices can be made. Thus, from the perspective of implementation, it would be beneficial to solve this problem. Bitcoin is only the beginning and this similar problem will be the risk one is going to have to take when venturing into this new frontier of digital currency.

Solution:

The Bitcoin's worth fluctuates simply like a stock though in an unexpected way as it is highly volatile. However, the cost of Bitcoin doesn't rely upon the business occasions nor mediating government like securities exchange. Hence, predicting the future of Bitcoin is no easy task. Many people may regret not buying Bitcoin in 2012 but how were they supposed to know then? This is the dilemma we now face in regards to Bitcoin. How can we potentially solve this dilemma? Maybe ML can tell us the answer. ML models can likely give us the insight we need to learn about the future of Bitcoin. There are various calculations utilized on financial exchange information for value forecasts. The parameters influencing Bitcoin are extraordinary. Hence, it is important to anticipate the estimation of Bitcoin so the right venture choices can be made. To anticipate the worth we can use ML models to foresee the cost of Bitcoin. In this project we will try to find the feature price accuracy. The ML algorithms will improve the feature idea of crypto currencies. This will help to have an estimation about Bitcoin.

Explanation:

The purpose of the machine learning algorithms used to predict the prices are that algorithms like linear regression and learning models can take in many data points and create an accurate prediction based on trends that would otherwise be too difficult or time consuming for a human to determine. For this specific project the group decided to use Facebook's Prophet algorithm, which is similar to linear regression, in their own words "We propose a modular regression model with interpretable parameters that can be intuitively adjusted by analysts with domain knowledge about the time series". This model should help us to predict non-linear trends and seasonality.

Data Description:

This is the data as is and the attributes of the dataframe are as follows: Date, Opening price of the day, Highest price of the day, Lowest Price of the day, Close Price of the day, Adjusted Close Price of the day, Volume Sold in the day, and Currency type.

Figure 1.1

	Date	Open	High	Low	Close	Adj Close	Volume
0	2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	2.105680e+07
1	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	3.448320e+07
2	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	3.791970e+07
3	2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	3.686360e+07
4	2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	2.658010e+07
...	...	...	...	...	...	...	...
2201	2020-09-26	10702.237305	10778.500000	10682.082031	10754.437500	10754.437500	1.810501e+10
2202	2020-09-27	10752.939453	10804.732422	10643.458008	10774.426758	10774.426758	1.801688e+10
2203	2020-09-28	10771.641602	10949.123047	10716.676758	10721.327148	10721.327148	2.272037e+10
2204	2020-09-29	10712.462891	10858.939453	10665.344727	10848.830078	10848.830078	2.045987e+10
2205	2020-09-30	10845.411133	10856.528320	10689.670898	10787.618164	10787.618164	2.075962e+10

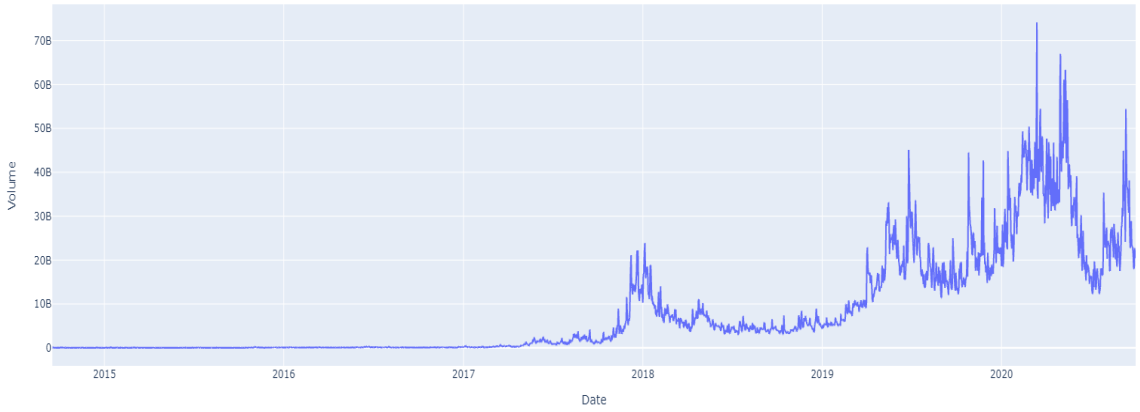
2206 rows × 7 columns

Next a graph of the closing prices was plotted, and it shows the leap in price around 2018, when there was a large bubble in which Bitcoin went from about $1,000 per coin to $20,000 per coin at its peak Moreover, you can see the trend in which after this large bubble popped you can see an almost immediate decrease to about $8,000 and goes even lower as the year goes on and finally in 2019 it stables out at about $10,000.

Figure 1.2

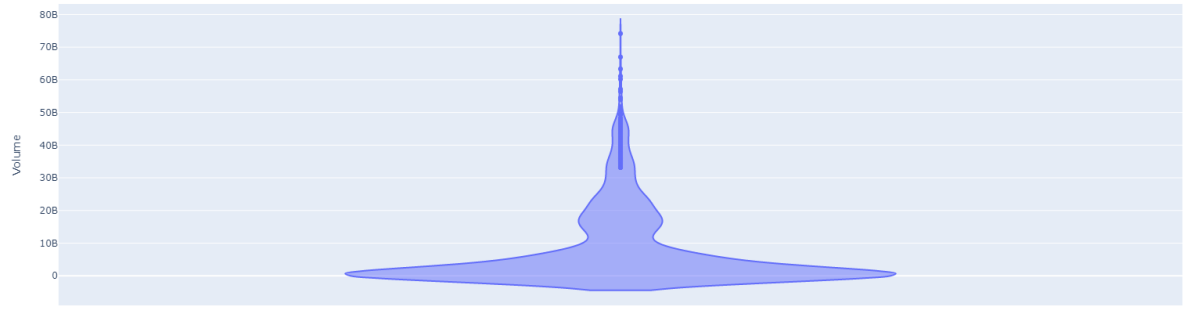
Next, this graph shows the plot of the volume of Bitcoin sold daily, and it shows that there is a large rise of Bitcoin being sold around the time of the bubble, but interestingly there is a large drop of Bitcoin being sold after the crash. This could be due to unwillingness of people to sell the coins they already owned and are waiting for another bubble. Lastly this graph shows that as the price returned to normal the volume of Bitcoin being sold skyrocketed likely due to have more popularity than ever before.

Figure 1.3



The next graph is a violin plot of the volume of the Bitcoin sold, from the data the maximum volume sold is 74.16 Billion and the minimum sold is 5.91 Million, and the median is 2.02 Billion. Moreover, it is shown that some regions are very narrow, namely the upper region, indicating that the price of Bitcoin has been the same for a small number of days, and vice versa for the larger region.

Figure 1.4



We get the following type of output using the Date and Price columns. Facebook Prophet extracts different data from input data and predicts different kinds of attributes and output those attributes through different columns . Below, We have tried to show all these columns that Facebook Prophet tried to predict.

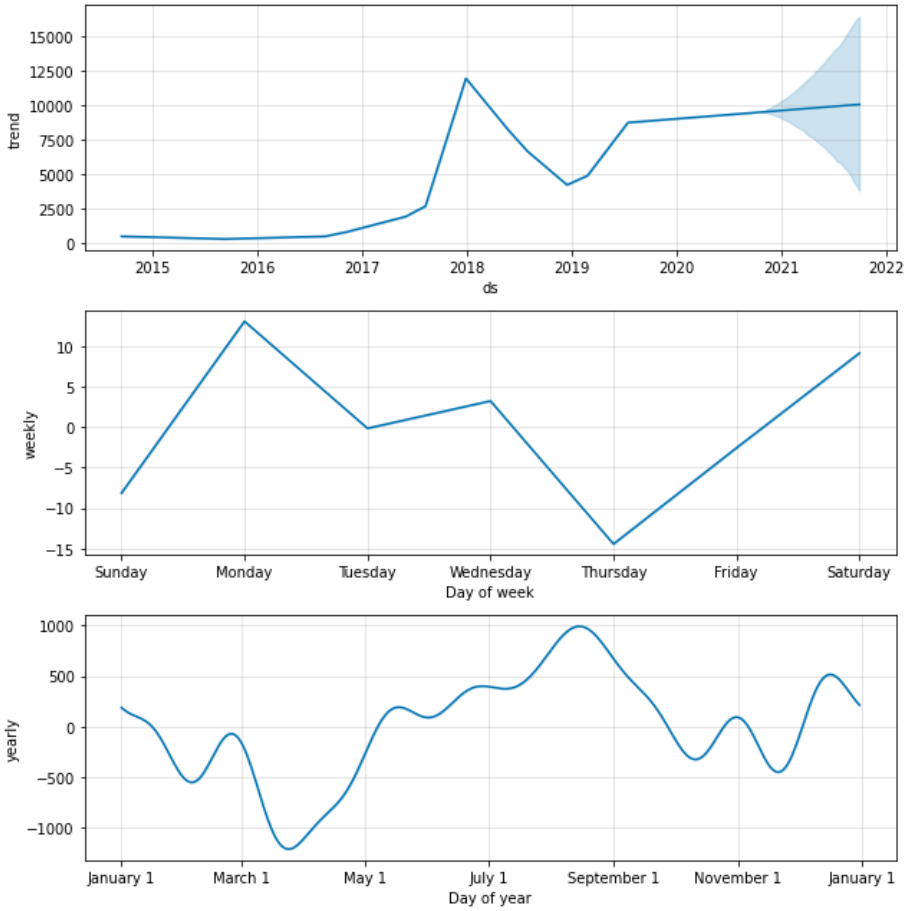
	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper	additive_terms	additive_terms_lower	additive_terms_upper
0	2014-09-17	468.142612	-550.088458	2274.399071	468.142612	468.142612	310.845504	310.845504	310.845504
1	2014-09-18	467.585449	-618.903134	2085.722521	467.585449	467.585449	270.728957	270.728957	270.728957
2	2014-09-19	467.028285	-606.271785	2068.146804	467.028285	467.028285	259.020082	259.020082	259.020082
3	2014-09-20	466.471121	-456.195706	2057.160331	466.471121	466.471121	245.864729	245.864729	245.864729
4	2014-09-21	465.913958	-758.497477	1979.472584	465.913958	465.913958	202.342195	202.342195	202.342195
...	...	...	...	...	...	...	...	...	...
2231	2020-10-26	9486.747924	8094.061303	10953.558546	9476.194968	9507.851238	44.365846	44.365846	44.365846
2232	2020-10-27	9488.381608	8166.926409	10901.166554	9474.906283	9514.952647	53.285627	53.285627	53.285627
2233	2020-10-28	9490.015292	8306.509070	10973.927106	9469.901623	9523.426679	74.675862	74.675862	74.675862
2234	2020-10-29	9491.648976	8168.829929	10882.972319	9467.400050	9530.907913	70.367889	70.367889	70.367889
2235	2020-10-30	9493.282660	8246.961152	10997.023755	9466.642664	9535.749923	90.508490	90.508490	90.508490

2236 rows × 19 columns

weekly	weekly_lower	weekly_upper	yearly	yearly_lower	yearly_upper	multiplicative_terms	multiplicative_terms_lower	multiplicative_terms_upper	yhat
3.208378	3.208378	3.208378	307.637126	307.637126	307.637126	0.0	0.0	0.0	778.988116
-14.454500	-14.454500	-14.454500	285.183458	285.183458	285.183458	0.0	0.0	0.0	738.314406
-2.583910	-2.583910	-2.583910	261.603993	261.603993	261.603993	0.0	0.0	0.0	726.048367
9.118442	9.118442	9.118442	236.746287	236.746287	236.746287	0.0	0.0	0.0	712.335851
-8.156843	-8.156843	-8.156843	210.499038	210.499038	210.499038	0.0	0.0	0.0	668.256152
...	...	...	...	...	...	...	...	...	...
13.041043	13.041043	13.041043	31.324803	31.324803	31.324803	0.0	0.0	0.0	9531.113769
-0.172609	-0.172609	-0.172609	53.458236	53.458236	53.458236	0.0	0.0	0.0	9541.667235
3.208378	3.208378	3.208378	71.467484	71.467484	71.467484	0.0	0.0	0.0	9564.691154
-14.454500	-14.454500	-14.454500	84.822389	84.822389	84.822389	0.0	0.0	0.0	9562.016865
-2.583910	-2.583910	-2.583910	93.092400	93.092400	93.092400	0.0	0.0	0.0	9583.791150

Using the data we can get different graphical representations. For example, the pick times within a week or within a year. The output using these data has been shown here one by one.

Figure 2.1 (for 'Close' price data)



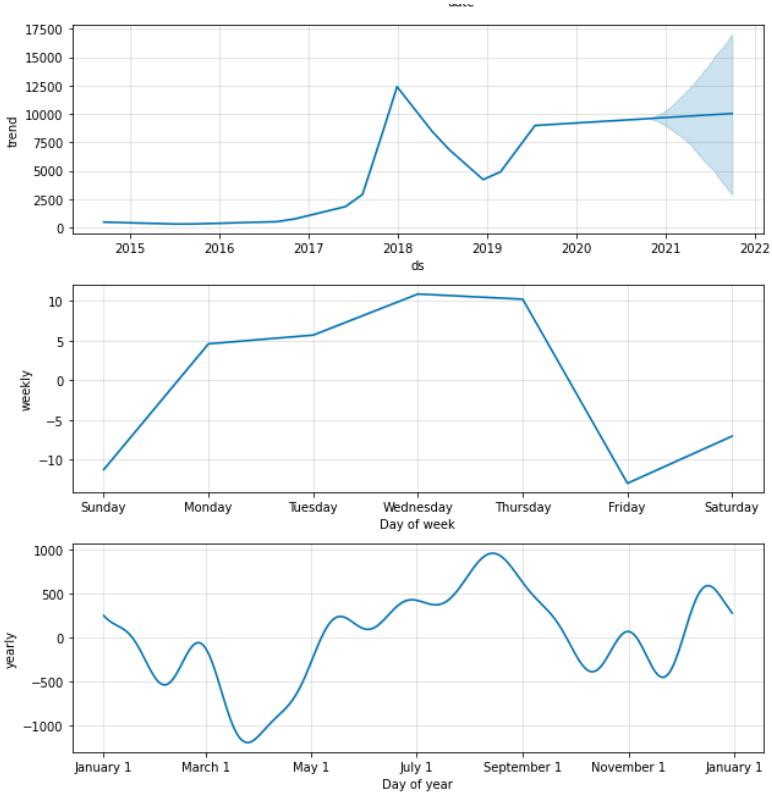
The accuracy, runtime, and forecasted values for the **'Close'** price are as follows:

```
accuracy :  0.9109592769948772
mean_squared_error :  1350317.42292312
mean_absolute_error:  700.9789675735923
```

```
running time for training : 0.0007468659998721705

Forecasted values of coming days for : Close
            ds         trend
2560 2021-09-20  10024.230029
2561 2021-09-21  10025.863713
2562 2021-09-22  10027.497397
2563 2021-09-23  10029.131082
2564 2021-09-24  10030.764766
2565 2021-09-25  10032.398450
2566 2021-09-26  10034.032134
2567 2021-09-27  10035.665818
2568 2021-09-28  10037.299503
2569 2021-09-29  10038.933187
```

Figure 2.2 (for 'High' price data)



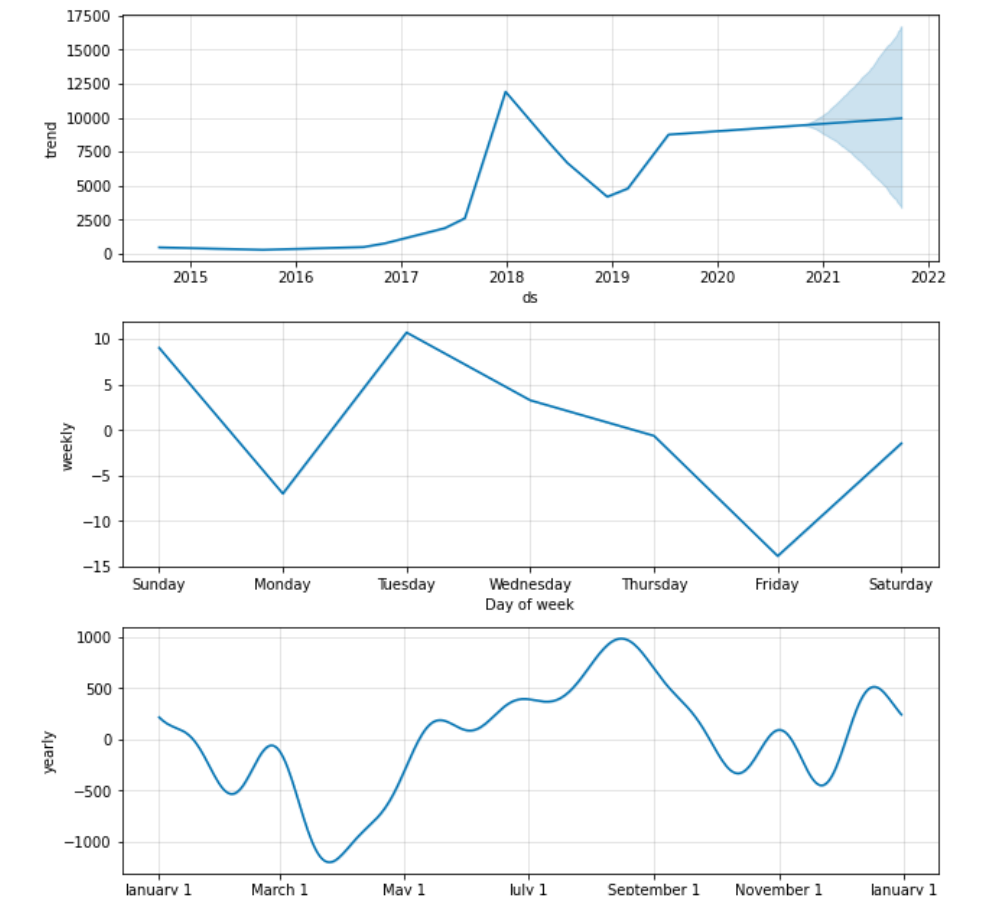
The accuracy, runtime, and forecasted values for the **'High'** price are as follows:

```
accuracy :  0.9103816027104612
mean_squared_error :  1433768.5893459765
mean_absolute_error:  706.6254881742984


running time for training : 0.0017698040001050686

Forecasted values of coming days for : High
            ds         trend
2560 2021-09-20  10018.031761
2561 2021-09-21  10019.330598
2562 2021-09-22  10020.629435
2563 2021-09-23  10021.928271
2564 2021-09-24  10023.227108
2565 2021-09-25  10024.525945
2566 2021-09-26  10025.824782
2567 2021-09-27  10027.123619
2568 2021-09-28  10028.422455
2569 2021-09-29  10029.721292
```

Figure 2.3 (for 'Open' price data)



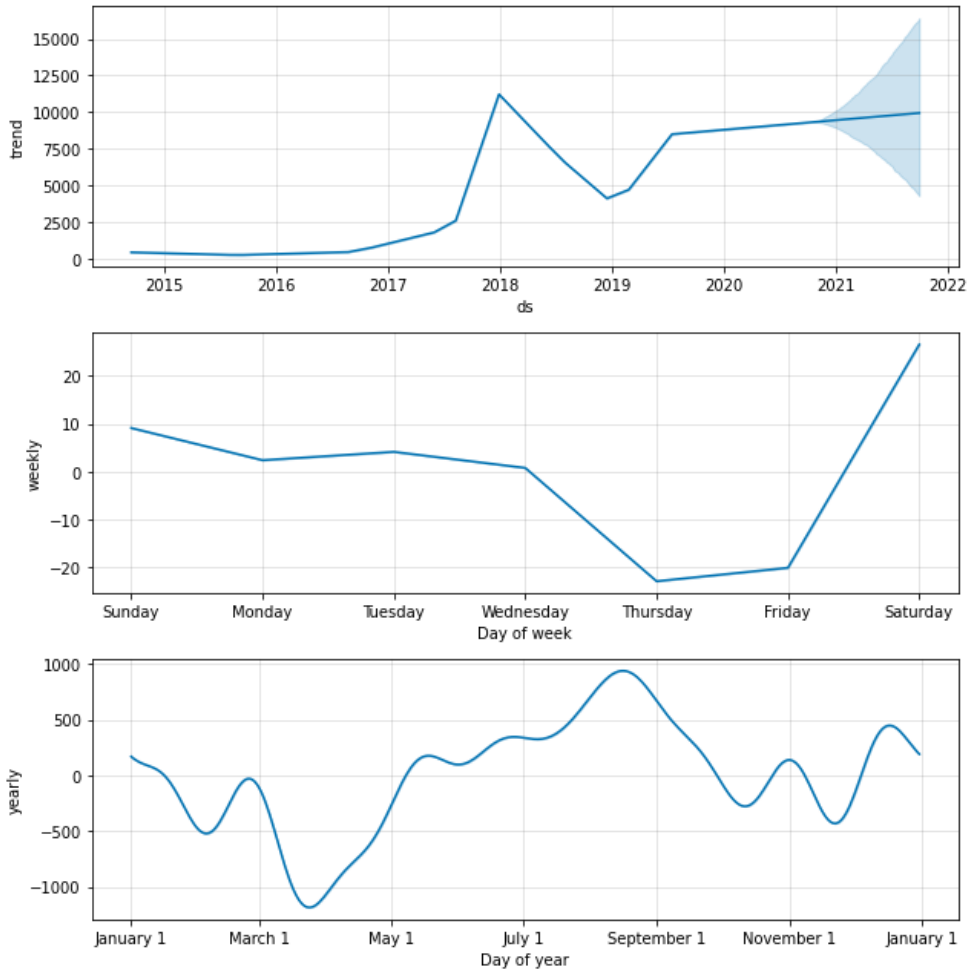
The accuracy, runtime, and forecasted values for the **'Open'** price are as follows:

```
accuracy :  0.9117448032819339
mean_squared_error :  1340289.3099724397
mean_absolute_error:  694.8852296559054


running time for training : 0.004264353000053234

Forecasted values of coming days for : Open
            ds        trend
2560 2021-09-20  9939.176462
2561 2021-09-21  9940.660043
2562 2021-09-22  9942.143623
2563 2021-09-23  9943.627204
2564 2021-09-24  9945.110784
2565 2021-09-25  9946.594365
2566 2021-09-26  9948.077945
2567 2021-09-27  9949.561526
2568 2021-09-28  9951.045106
2569 2021-09-29  9952.528687
```

Figure 2.4 (for 'Low' price data)



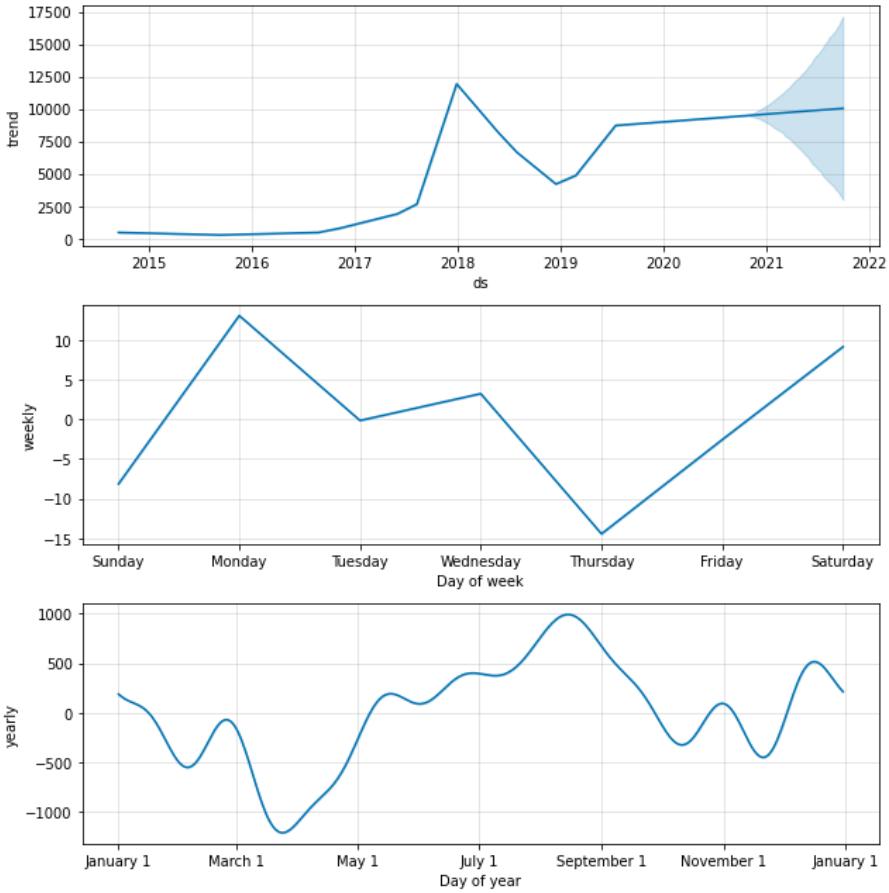
The accuracy, runtime, and forecasted values for the **'Low'** price are as follows:

```
accuracy :  0.9143219647099761
mean_squared_error :  1220724.3926872085
mean_absolute_error:  666.8820690553923


running time for training : 0.0035242100000232313

Forecasted values of coming days for : Low
             ds        trend
2560 2021-09-20  9926.446924
2561 2021-09-21  9928.234545
2562 2021-09-22  9930.022165
2563 2021-09-23  9931.809785
2564 2021-09-24  9933.597406
2565 2021-09-25  9935.385026
2566 2021-09-26  9937.172647
2567 2021-09-27  9938.960267
2568 2021-09-28  9940.747888
2569 2021-09-29  9942.535508
```

Figure 2.5 (for 'Adj Close' price data)



The accuracy, runtime, and forecasted values for the **'Adj Close'** price are as follows:

```
accuracy :  0.9109592769948772
mean_squared_error :  1350317.42292312
mean_absolute_error:  700.9789675735923
```

```
running time for training : 0.0020131829999172624

Forecasted values of coming days for : Adj Close
            ds         trend
2560 2021-09-20  10024.230029
2561 2021-09-21  10025.863713
2562 2021-09-22  10027.497397
2563 2021-09-23  10029.131082
2564 2021-09-24  10030.764766
2565 2021-09-25  10032.398450
2566 2021-09-26  10034.032134
2567 2021-09-27  10035.665818
2568 2021-09-28  10037.299503
2569 2021-09-29  10038.933187
```

It can be seen from the results that we are getting the for the ‘Low’ price data we are getting the minimum errors and for the ‘High’ price data maximum errors. we are getting maximum accuracy 91.43% when we are using the ‘Low’ price data. And for the ‘High’ price data minimum accuracy 91.03%. Consequently,

Result Analysis:

With having gone over the bitcoin dataset we were able to determine that the price is highly volatile at certain times over others. Such as the price being much cheaper towards the beginning of the year in some instances where the buyers are just selling off and not buying. In terms of trying to make a profit one would want to buy during these times versus buying at a high price and possibly missing out at buying much cheaper prices. Moreover, from our analysis with the Facebook Prophet algorithm, we have an accuracy of 91.096%, however, the mean squared error is quite high as 1,350317.423 and a mean absolute error of 700.979. This means that despite having a good accuracy of predictions, the error for incorrect predictions is extremely high with the mean error between predictions of about 700. However, because we are dealing in thousands of dollars and more often in tens of thousands of dollars, this score is actually quite accurate as $700 one way or another for something that costs $10,000 would be close enough to determine a good estimate. An example of this would be the data that was predicted shown:

Figure 3.1

Date	**Predicted Value**	**Actual Value**
2014-09-17	468.1426	457.3340
2014-09-18	467.5854	424.4400
2014-09-19	467.0282	394.7960
2014-09-20	466.4711	408.9040
2014-09-21	465.9140	398.8210

So as shown the predicted results are only slightly above the actual results ranging from between ~ $10-$70 which is quite close to the actual data. This means that this algorithm could be a viable way of predicting future Bitcoin prices just as we’ve done below. However, the issue remains that Bitcoin’s current price far exceeds the price predicted by the algorithm, this is due to outside factors that the regression algorithm could not account for which means that because this algorithms shouldn’t be used for long term predictions as its clear that this regression can’t predict the factors that caused the most recent bubble.

```
running time for training : 0.0017698040001050686

Forecasted values of coming days for : High
             ds         trend
2560 2021-09-20  10018.031761
2561 2021-09-21  10019.330598
2562 2021-09-22  10020.629435
2563 2021-09-23  10021.928271
2564 2021-09-24  10023.227108
2565 2021-09-25  10024.525945
2566 2021-09-26  10025.824782
2567 2021-09-27  10027.123619
2568 2021-09-28  10028.422455
2569 2021-09-29  10029.721292
```

```
running time for training : 0.0007468659998721705

Forecasted values of coming days for : Close
             ds         trend
2560 2021-09-20  10024.230029
2561 2021-09-21  10025.863713
2562 2021-09-22  10027.497397
2563 2021-09-23  10029.131082
2564 2021-09-24  10030.764766
2565 2021-09-25  10032.398450
2566 2021-09-26  10034.032134
2567 2021-09-27  10035.665818
2568 2021-09-28  10037.299503
2569 2021-09-29  10038.933187
```

```
running time for training : 0.0017698040001050686

Forecasted values of coming days for : High
             ds         trend
2560 2021-09-20  10018.031761
2561 2021-09-21  10019.330598
2562 2021-09-22  10020.629435
2563 2021-09-23  10021.928271
2564 2021-09-24  10023.227108
2565 2021-09-25  10024.525945
2566 2021-09-26  10025.824782
2567 2021-09-27  10027.123619
2568 2021-09-28  10028.422455
2569 2021-09-29  10029.721292
```

```
running time for training : 0.004264353000053234

Forecasted values of coming days for : Open
             ds         trend
2560 2021-09-20  9939.176462
2561 2021-09-21  9940.660043
2562 2021-09-22  9942.143623
2563 2021-09-23  9943.627204
2564 2021-09-24  9945.110784
2565 2021-09-25  9946.594365
2566 2021-09-26  9948.077945
2567 2021-09-27  9949.561526
2568 2021-09-28  9951.045106
2569 2021-09-29  9952.528687
```

```
running time for training : 0.0035242100000232313

Forecasted values of coming days for : Low
             ds         trend
2560 2021-09-20  9926.446924
2561 2021-09-21  9928.234545
2562 2021-09-22  9930.022165
2563 2021-09-23  9931.809785
2564 2021-09-24  9933.597406
2565 2021-09-25  9935.385026
2566 2021-09-26  9937.172647
2567 2021-09-27  9938.960267
2568 2021-09-28  9940.747888
2569 2021-09-29  9942.535508
```

```
running time for training : 0.0020131829999172624

Forecasted values of coming days for : Adj Close
             ds          trend
2560 2021-09-20  10024.230029
2561 2021-09-21  10025.863713
2562 2021-09-22  10027.497397
2563 2021-09-23  10029.131082
2564 2021-09-24  10030.764766
2565 2021-09-25  10032.398450
2566 2021-09-26  10034.032134
2567 2021-09-27  10035.665818
2568 2021-09-28  10037.299503
2569 2021-09-29  10038.933187
```