Bitcoin Price Prediction

-By Akshat Jain (B20AI054)

Deployed Website: https://bitcoin-predictor-akshat.herokuapp.com/
Github Repository:https://github.com/akshatjain1004/bitcoin-price-predictor

Abstract: This paper reports my experience with building a regression pipeline to predict the **future price of bitcoin.**

1. Introduction

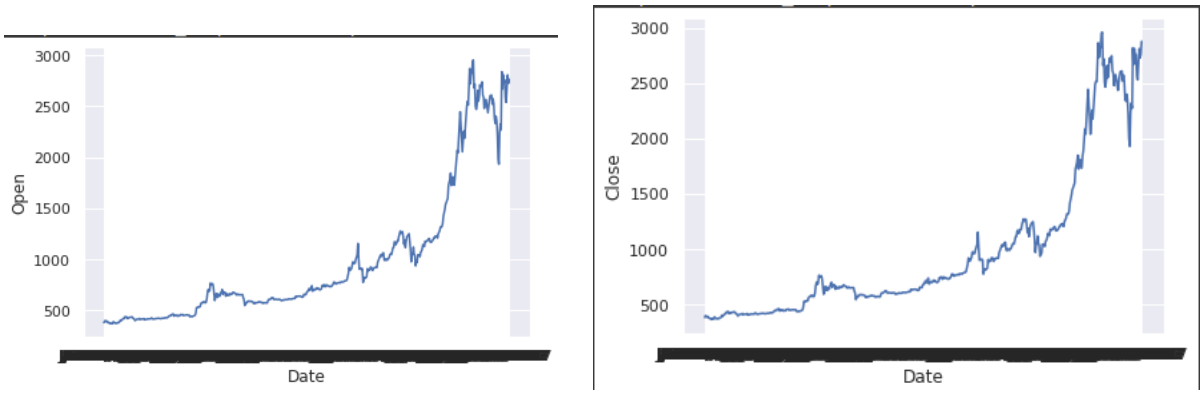
Bitcoin is a decentralized digital currency, without a central bank or single administrator, that can be sent from user to user on the peer-to-peer bitcoin network without the need for intermediaries. Bitcoin price is highly volatile. My pipeline helps in predicting the **Open, Close and Market Volume** of a bitcoin-based on the last 9 years of publicly available BTC data from **17-09-2014** to **15-04-2022.**

2. Data Analysis

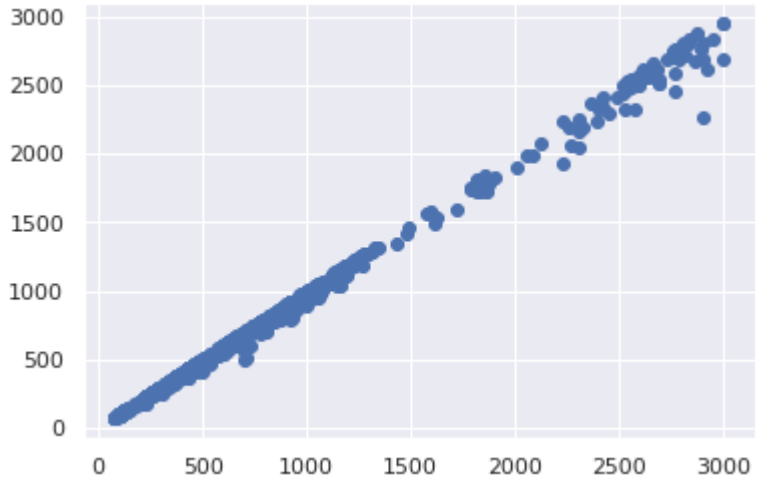
	Date	Open	High	Low	Close	Volume	Market Cap
0	Jul 31, 2017	2763.24	2889.62	2720.61	2875.34	860,575,000	45,535,800,000
1	Jul 30, 2017	2724.39	2758.53	2644.85	2757.18	705,943,000	44,890,700,000
2	Jul 29, 2017	2807.02	2808.76	2692.80	2726.45	803,746,000	46,246,700,000
3	Jul 28, 2017	2679.73	2897.45	2679.73	2809.01	1,380,100,000	44,144,400,000
4	Jul 27, 2017	2538.71	2693.32	2529.34	2671.78	789,104,000	41,816,500,000

```
#   Column      Non-Null Count  Dtype
---  ------      --------------  -----
0   Date        1556 non-null   object
1   Open        1556 non-null   float64
2   High        1556 non-null   float64
3   Low         1556 non-null   float64
4   Close       1556 non-null   float64
5   Volume      1556 non-null   object
6   Market Cap  1556 non-null   object
dtypes: float64(4), object(3)
```

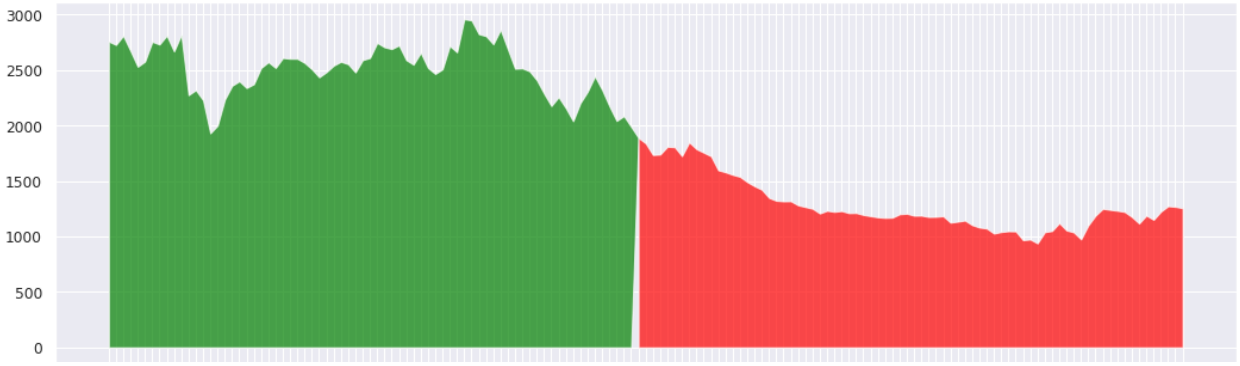
The data has a number of labels such as **High, Low, Open, Close, Volume…**etc. And we have to predict them on the basis of our feature: **Date**



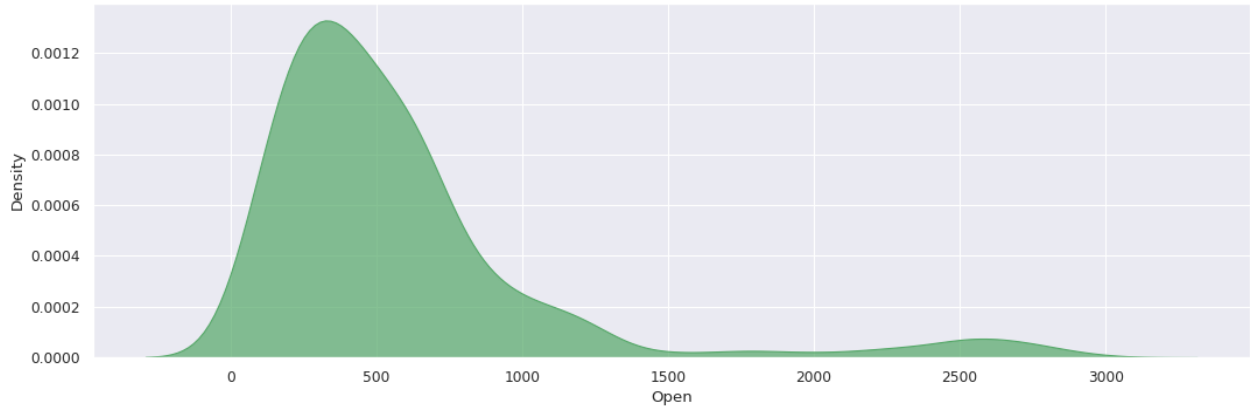
Time Series of Open and Close price: We can see a big recent spike in the BTC prices.



Linear relation between High and Open Prices.



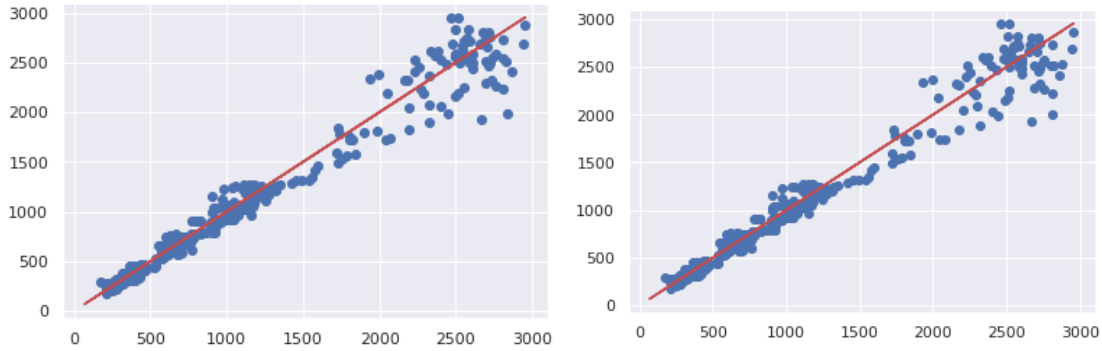
Recent Price Analysis: I took last 150 days of data and compared it with their mean values. The region in red is where the price is lower than the mean values and the region in green is where the price is higher.
Here we can see that the green and the red regions are continuous without any discrepancy. This gives us a hint that the price depends on the previous price.



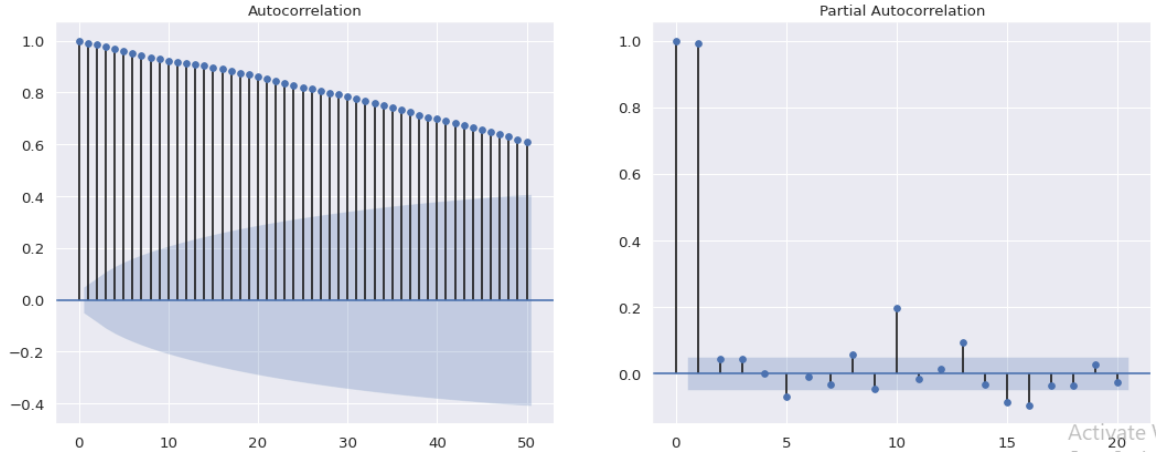
The Open Price Density Plot: Here most of the data points correspond to the lower values of Open Price, this is because of the sudden exponential increase in the bitcoin price.

The Peaks and The Troughs Analysis

3. Methodology:



Open and Close Price vs. The Open and Close Price 5 days later



Autocorrelation (ACF) and Partial Autocorrelation (PACF) Plot

Here we can clearly verify our intuition from the previous parts. Suppose if we're to predict a price 5 days later, it depends linearly on today's price. The ACF plot shows the correlation of the time series with its own lags. Each vertical line (on the autocorrelation plot) represents the correlation between the series and its lag starting from lag 0. Now, from the time series, it is pretty obvious that price vs date trend is very noisy and hence we cannot use simple date directly as a feature. Therefore I adopted following steps:

- Convert dates from Strings to Date-Time objects and Calculate the number of days between the date on which the price is to be predicted and the last date in our training dataset.
- Now, once we have the difference, I shift the dataset by that value.
- I now use the unshifted column as feature (X) and the shifted colum(Y) as the label
- Now I trained different models on this new training dataset and compared their accuracy.

Models Used:

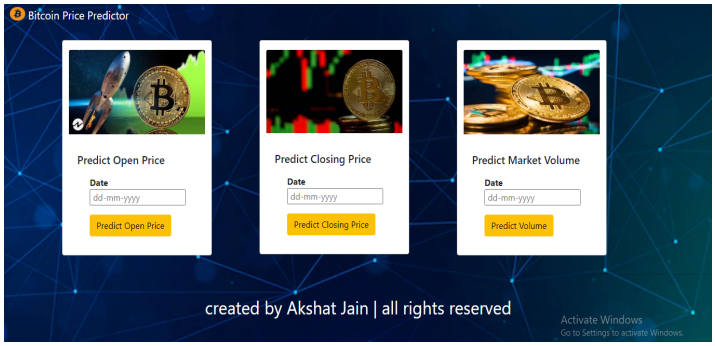
- KNN Regressor
- Random Forest Regressor
- SVR

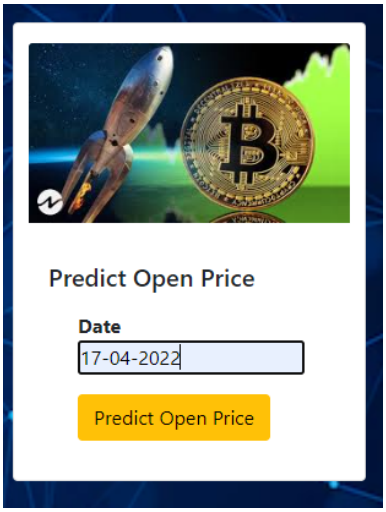
4. Comparison of Models

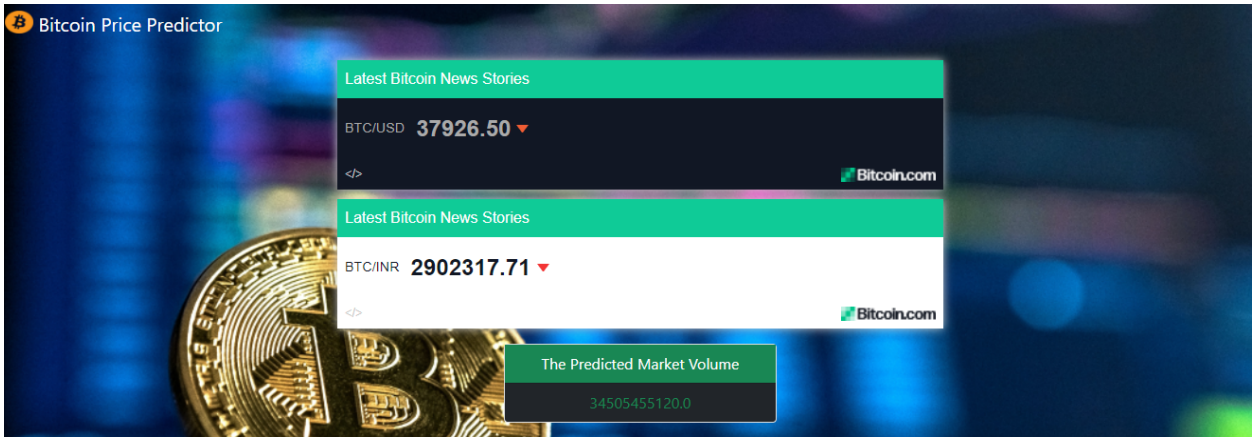
Metrics	SVR	KNN	RandomForest
MEAN ABSOLUTE PERCENTAGE ERROR	0.0437	0.0539	0.4148
MEAN ABSOLUTE ERROR	27.17	32.9	131.93
MEAN SQUARED ERROR	3551.5	4223.7	26374.2

By all the three metrics, SVR performs the best.

Functionalities of the website: Fully Responsive Real-time price display in INR and USD, BTC price plots. Prediction of Open Price, Close Price and Market Volume Through ML engine.
TechStacks used for Website: Bootstrap, Html,CSS, js and flask







References:

-https://www.machinelearningplus.com/plots/top-50-matplotlib-visualizations-the-master-plots-python/

- https://www.bitcoin.com/tools/widgets/