Pattern Recognition and Machine Learning

Bonus-Project : Bitcoin Prediction

Viradiya Dhruvkumar - B20CS079

Problem-Statement:

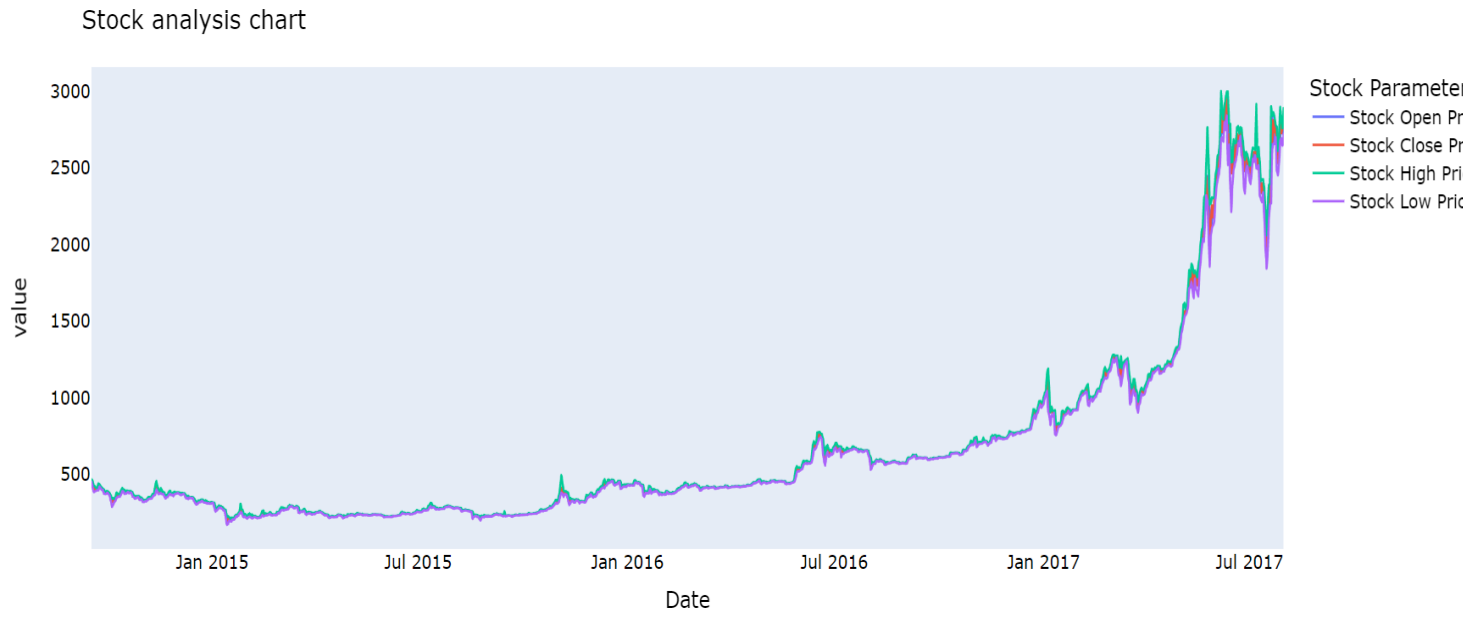
- Bitcoin dataset is provided, where the data of year 2013-2017 is given. There are columns like Open, Close, High, Low, Volume, Market Capitalization.
- We're supposed to do future prediction of all columns using the dataset.

Pre-Processing:

- First, we loaded the bitcoin dataset.
- It has 1556 data points and 6 columns. Columns are **Date, Open, Close, High, Low, Volume, Market Capitalization.**
- First, we converted all the values into float data type and we date into datetime datatype.
- We see that Volume is having some null values, so we replace the volume by the mean of all values of the year 2017.

Exploratory Data Analysis:

- We plotted all the Stock analysis.



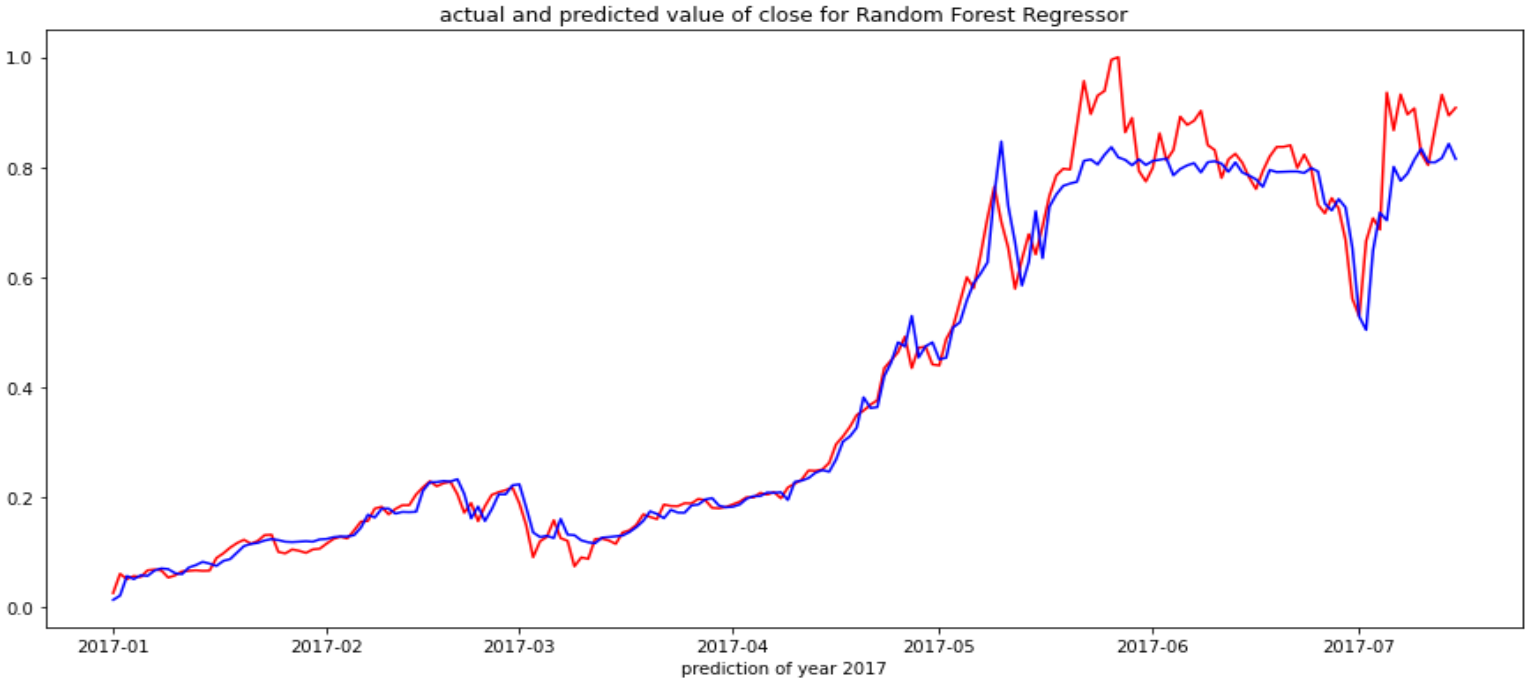
Analysis of Close Stock:

- Now, we have taken only close stock and try to analyze it.
- We, first normalize the dataset, then divide the dataset into train and test dataset, where from year 2013-2016 we've taken as training dataset and year 2017 we've taken as testing dataset
- **Creating train-test dataset:**
  - For analysis we've taken the first 14 values for training and created 14 columns using them, and taken the 15th value for prediction.
  - This way we've created the whole dataset.

Training the RandomForest Regressor:

- We trained the random forest regressor on the training dataset, and then reported the r2-score and mean_squared_error.

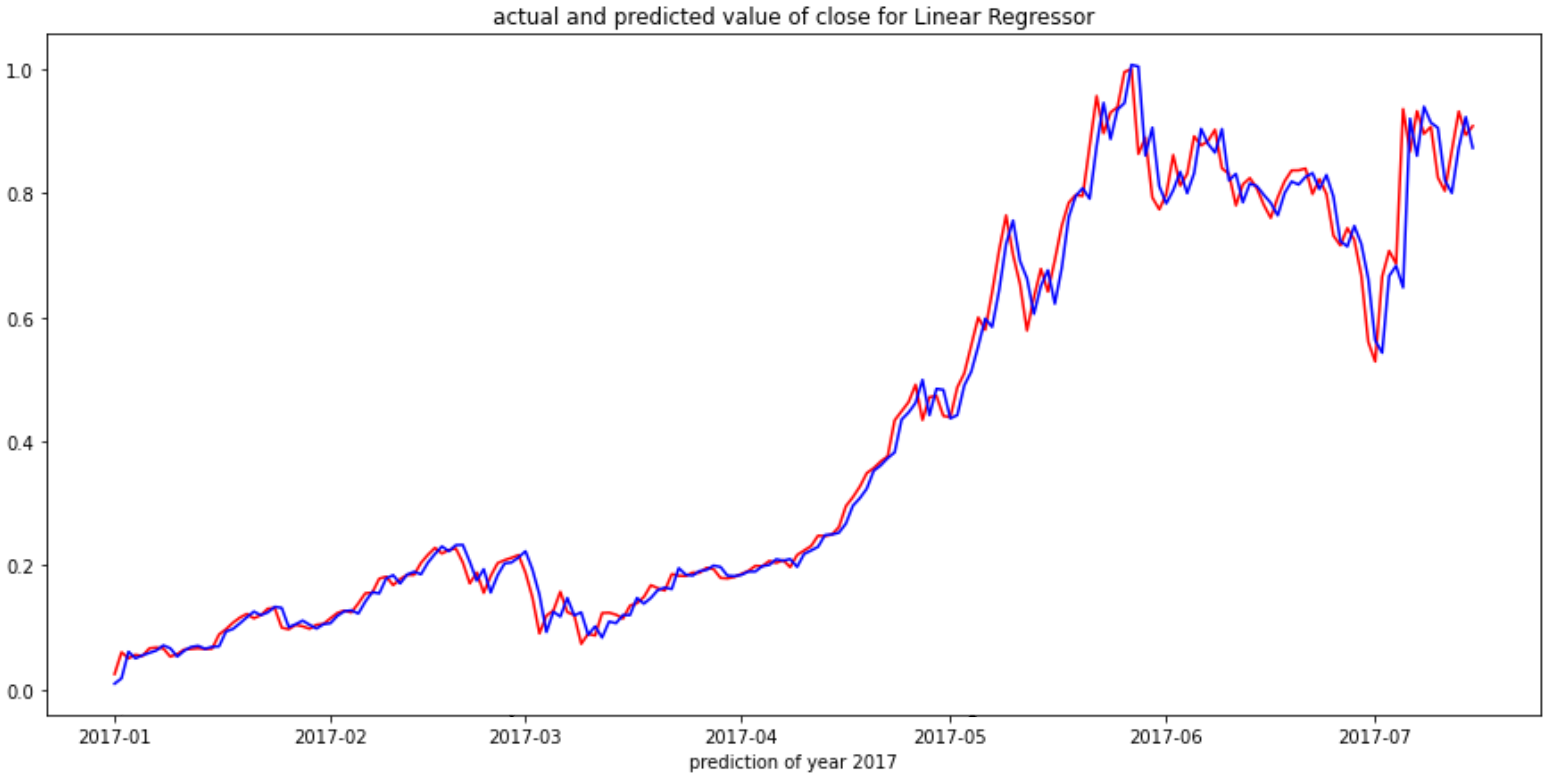
```
R2-score of close stock for Random Forest Regressor:  0.9751746613947124
Mean Squared Error of close stock for Random Forest Regressor:  0.0024228578538561937
```



Training the Linear Regressor:

- We trained the Linear regressor on the training dataset, and then reported the r2-score and mean_squared_error.

```
R2-score of close stock for Linear Regressor:  0.9838892663618023
Mean Squared Error of close stock for Linear Regressor:  0.0015723458256629275
```

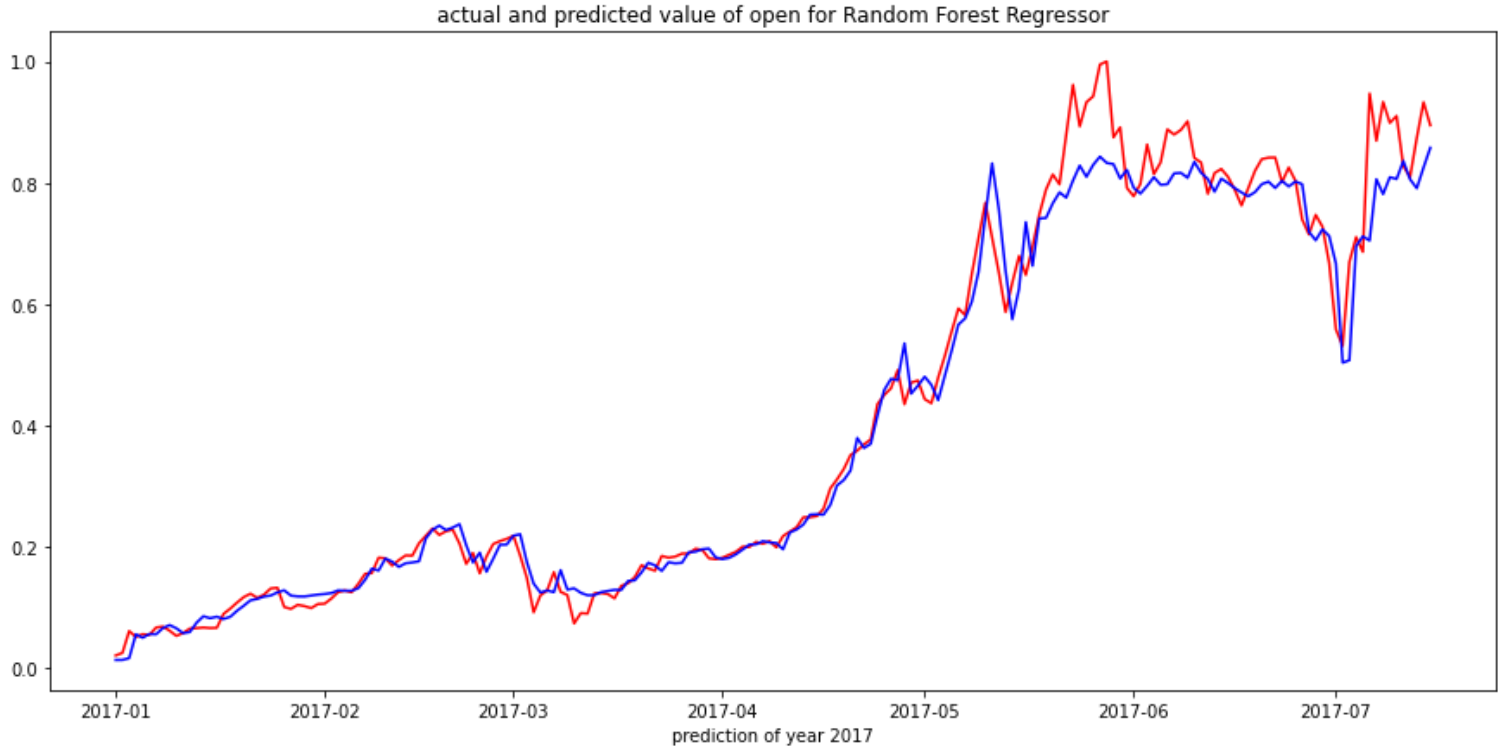


Now, we've done the similar analysis for all the other columns and plotted the results.

Analysis of Open-stock:

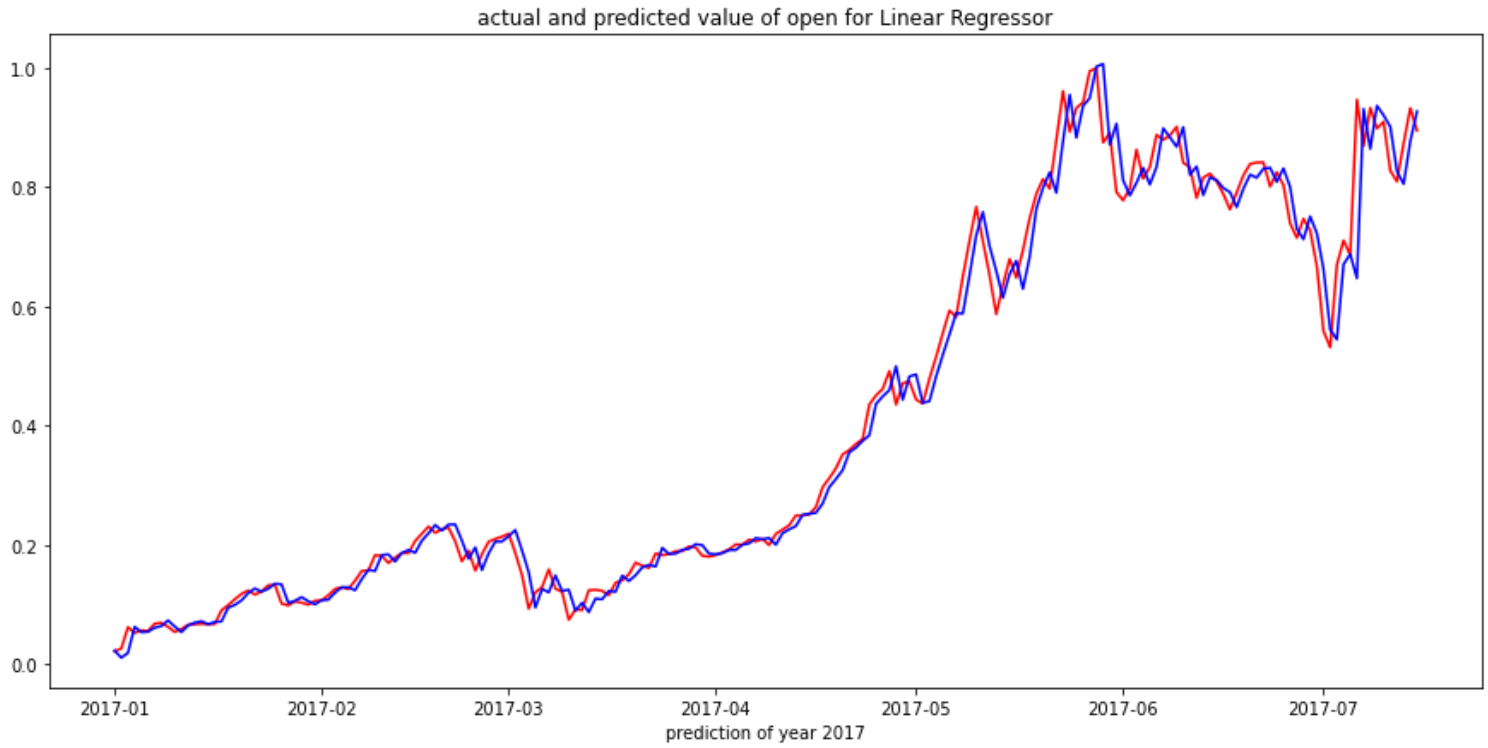
RandomForest Regressor for open stock:

```
R2-score of Open stock for Random Forest Regressor:  0.976524342876003
Mean Squared Error of Open stock for Random Forest Regressor:  0.002289122599889249
```

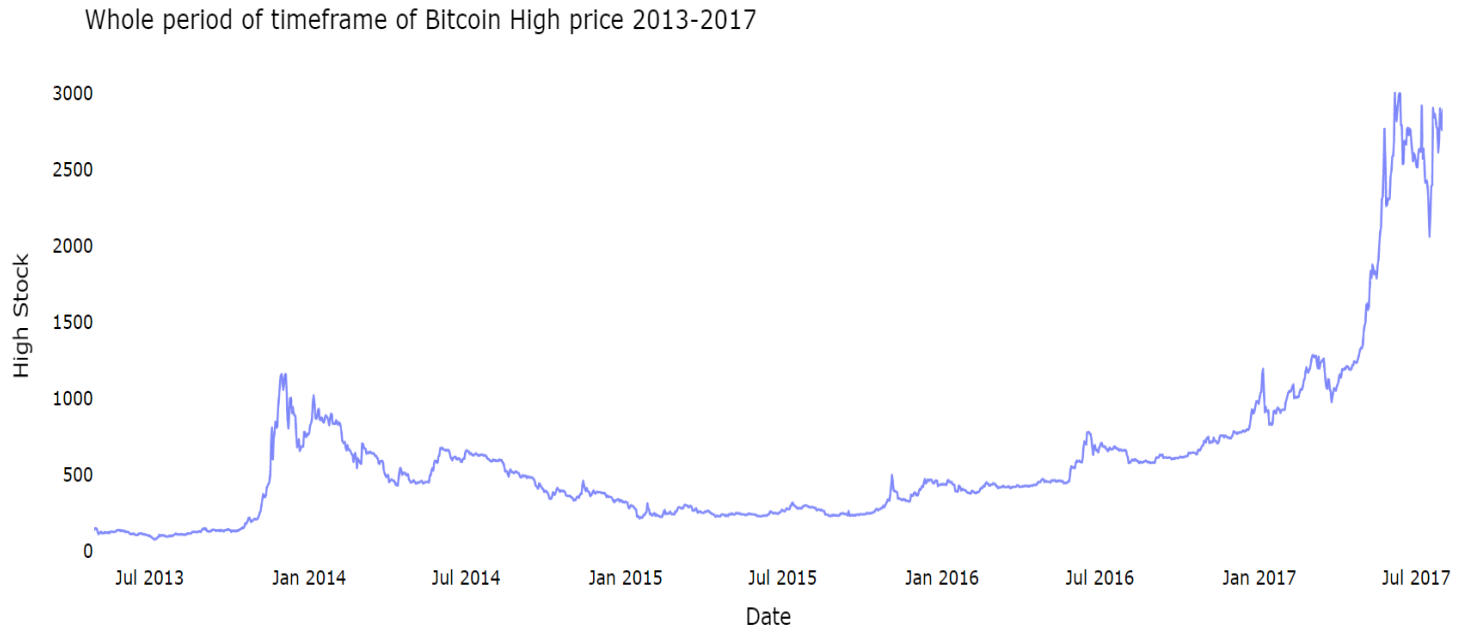


Linear regression for Open stock:

```
R2-score of Open stock for Linear Regressor:  0.983855894748969
Mean Squared Error of Open stock for Linear Regressor:  0.0015742194559209713
```

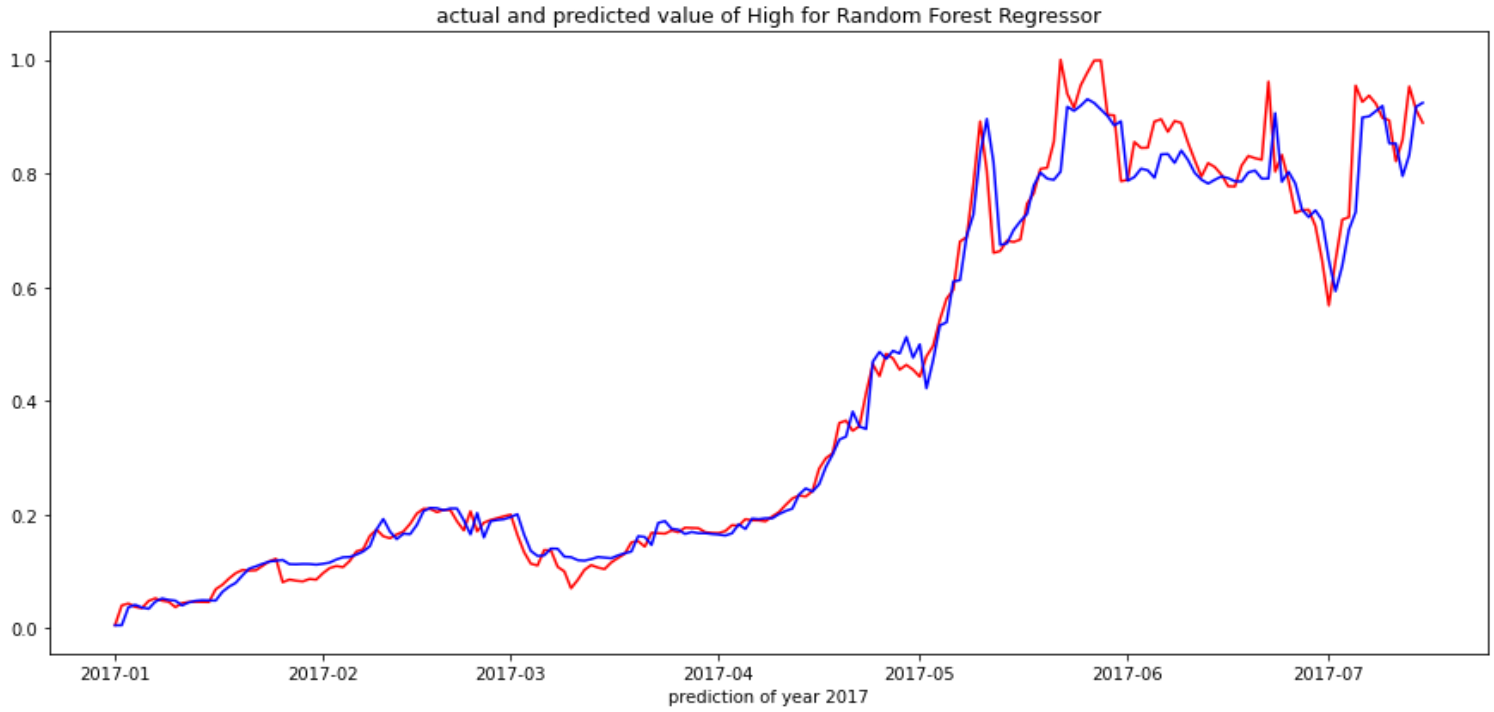


Analysis of High Stock:



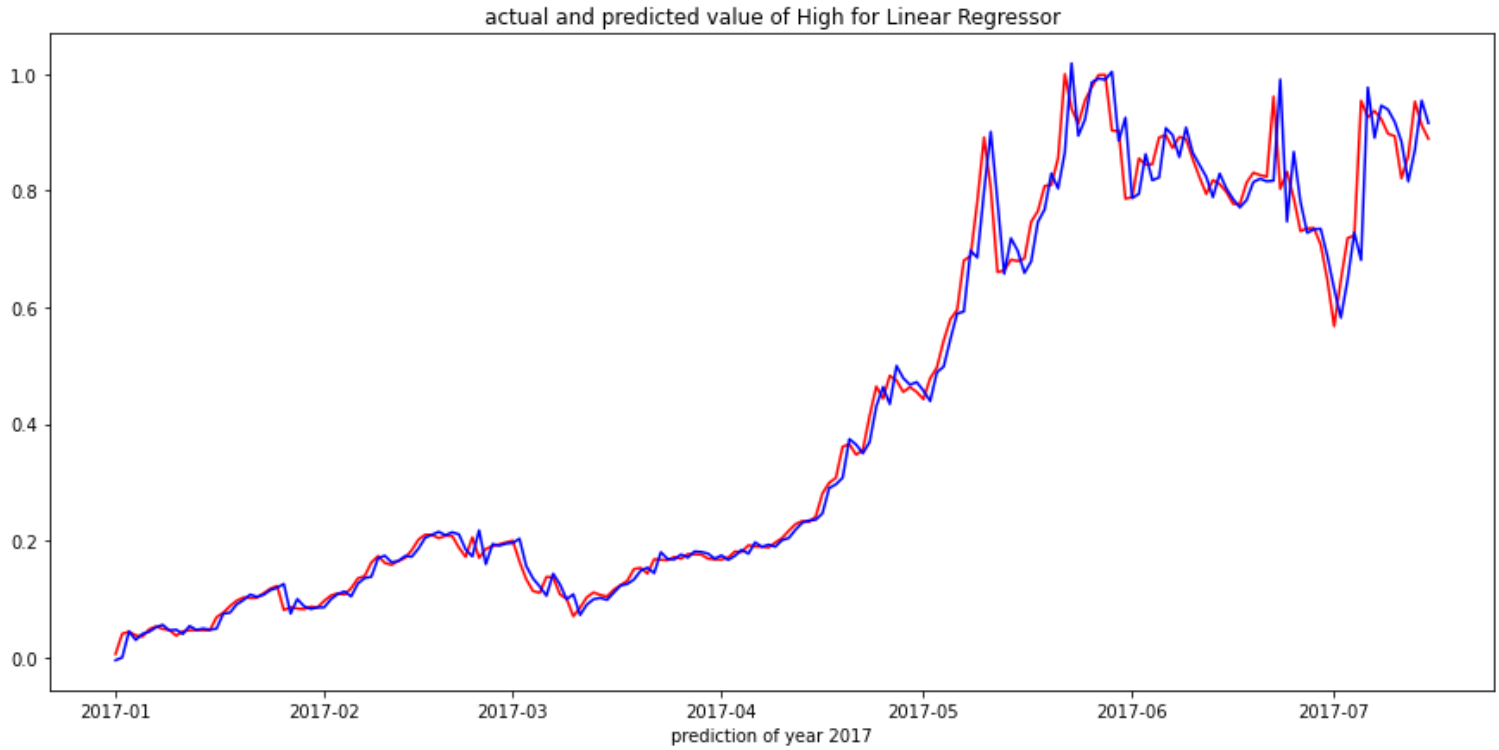
RandomForest Regressor for High stock:

```
R2-score of High stock for Random Forest Regressor:  0.983439346284942
Mean Squared Error of High stock for Random Forest Regressor:  0.0017833351827924357
```



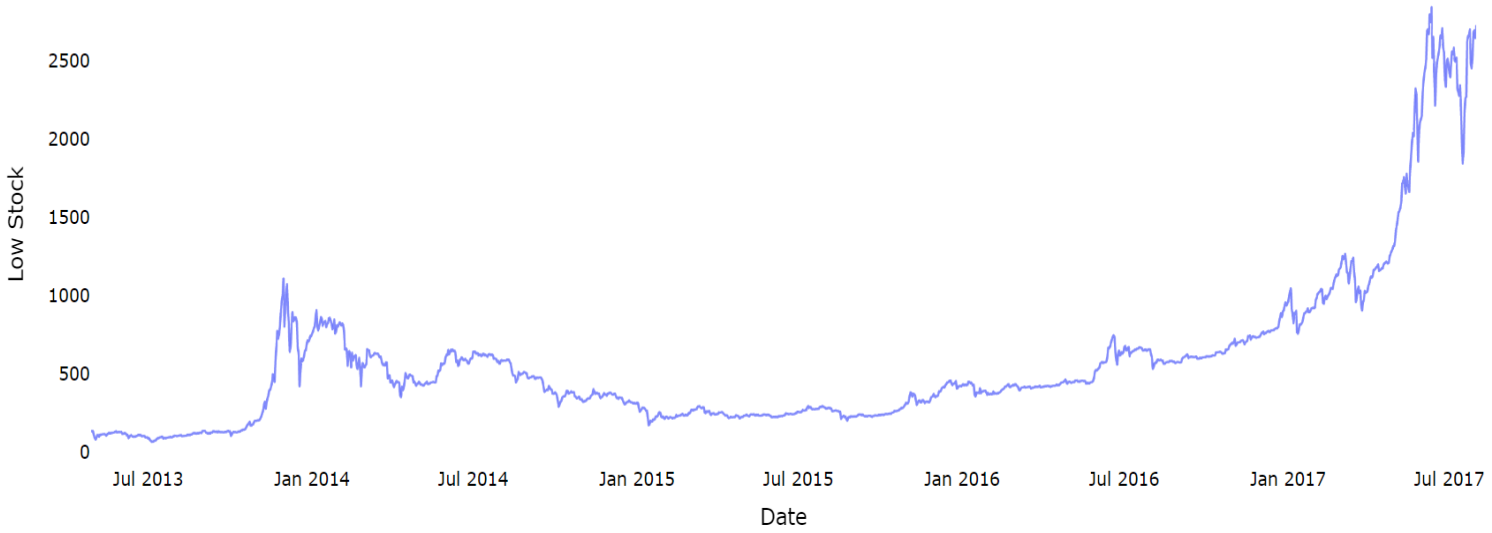
Linear Regressor for High stock:

```
R2-score of High stock for Linear Regressor:  0.9831367966443475
Mean Squared Error of High stock for Linear Regressor:  0.0018159152625342643
```



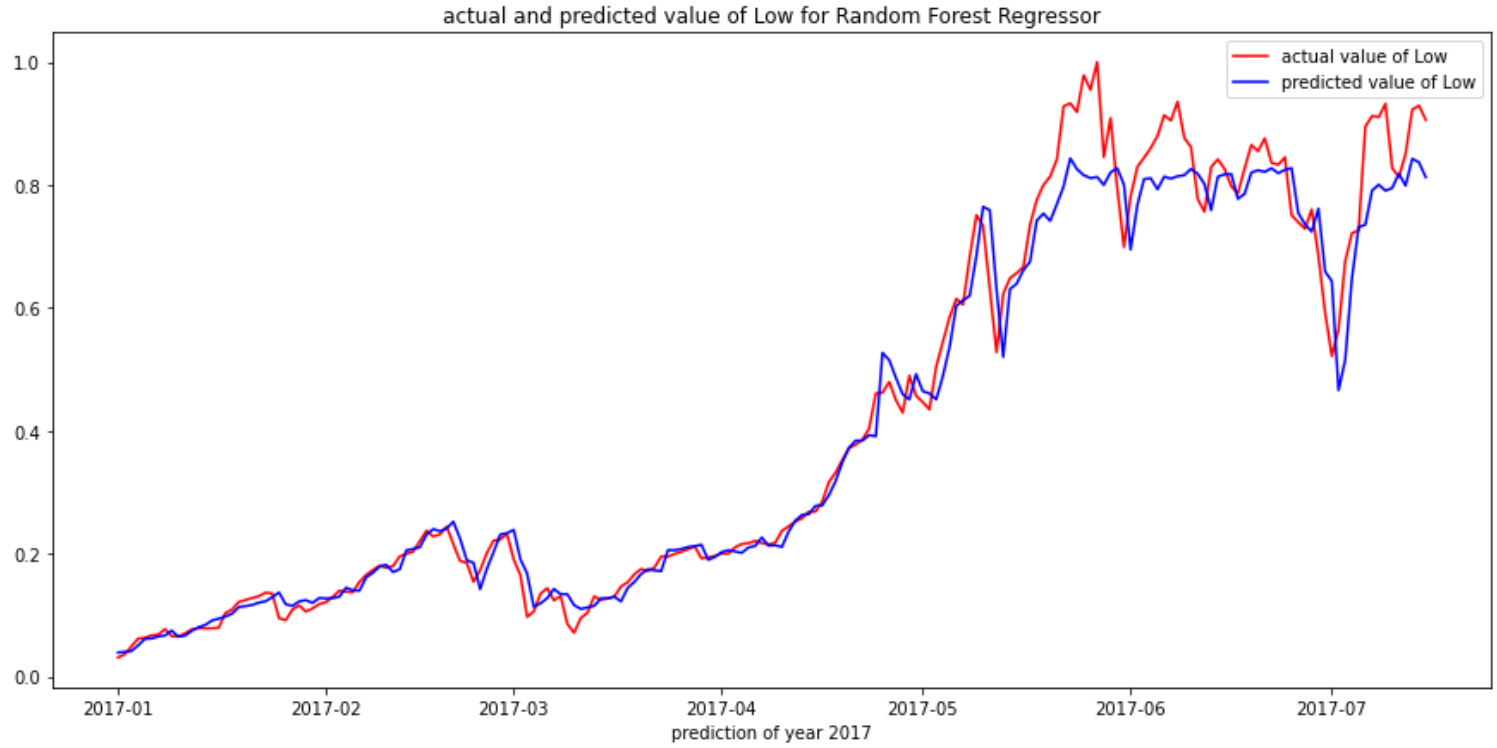
Analysis of Low-stock:

Whole period of timeframe of Bitcoin Low price 2013-2017



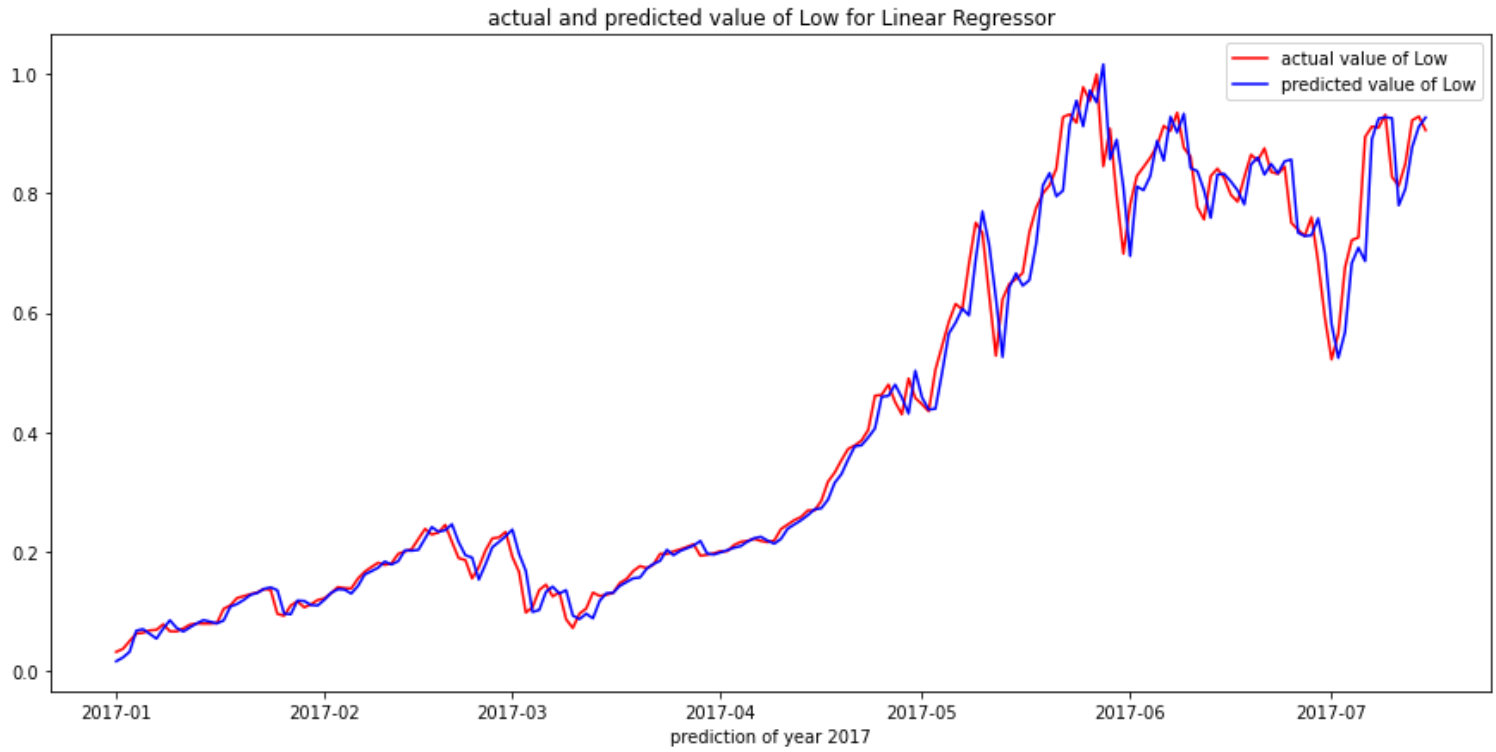
RandomForest Regressor for Low stock:

```
R2-score of Low stock for Random Forest Regressor:  0.9730614470171944
Mean Squared Error of Low stock for Random Forest Regressor:  0.0025995992381698434
```

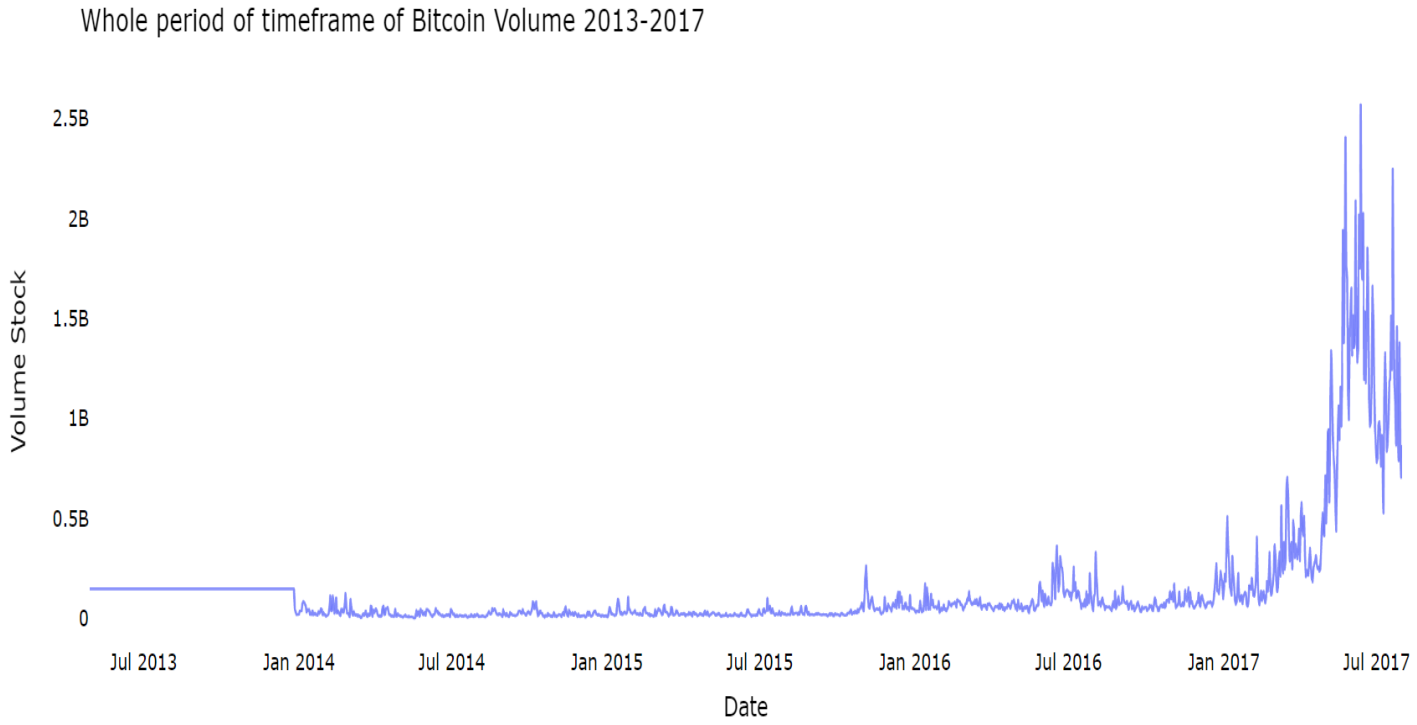


Linear Regressor for Low stock:

```
R2-score of Low stock for Linear Regressor:  0.9831074168641207
Mean Squared Error of Low stock for Linear Regressor:  0.0016301523797058477
```

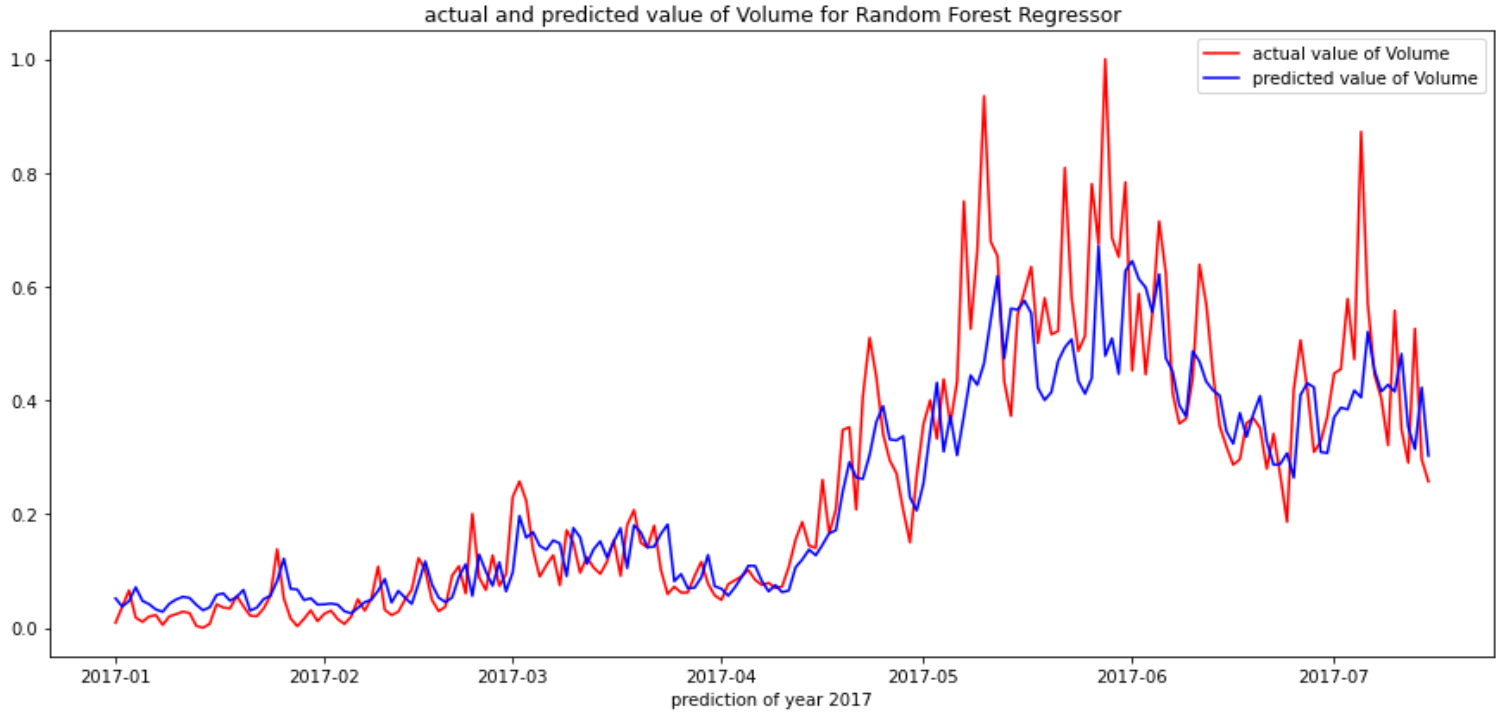


Analysis of Volume:



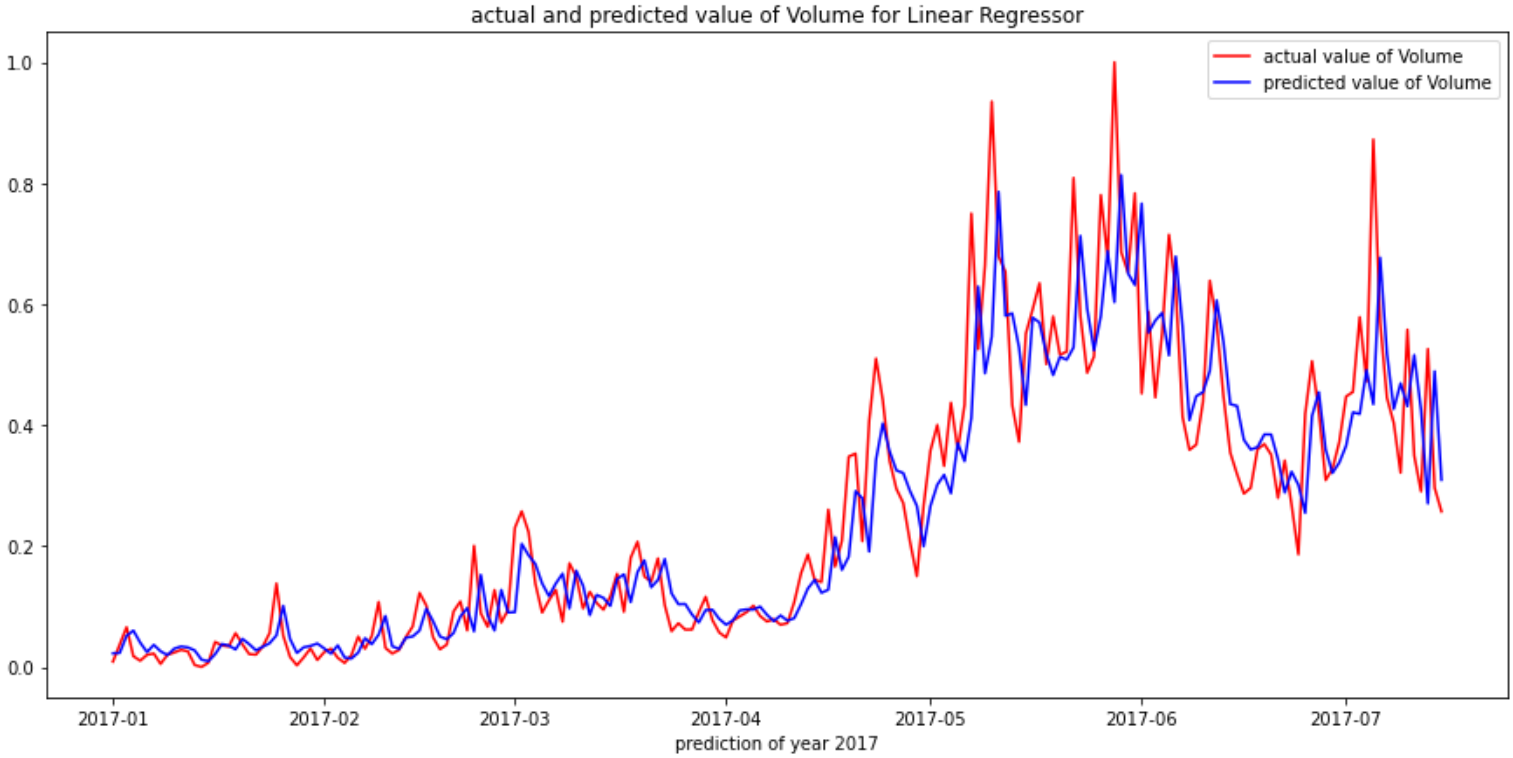
RandomForest Regressor for Volume:

```
R2-score of Volume stock for Random Forest Regressor:  0.7904472530819766
Mean Squared Error of Volume stock for Random Forest Regressor:  0.010936367654018977
```



Linear Regressor for Volume:

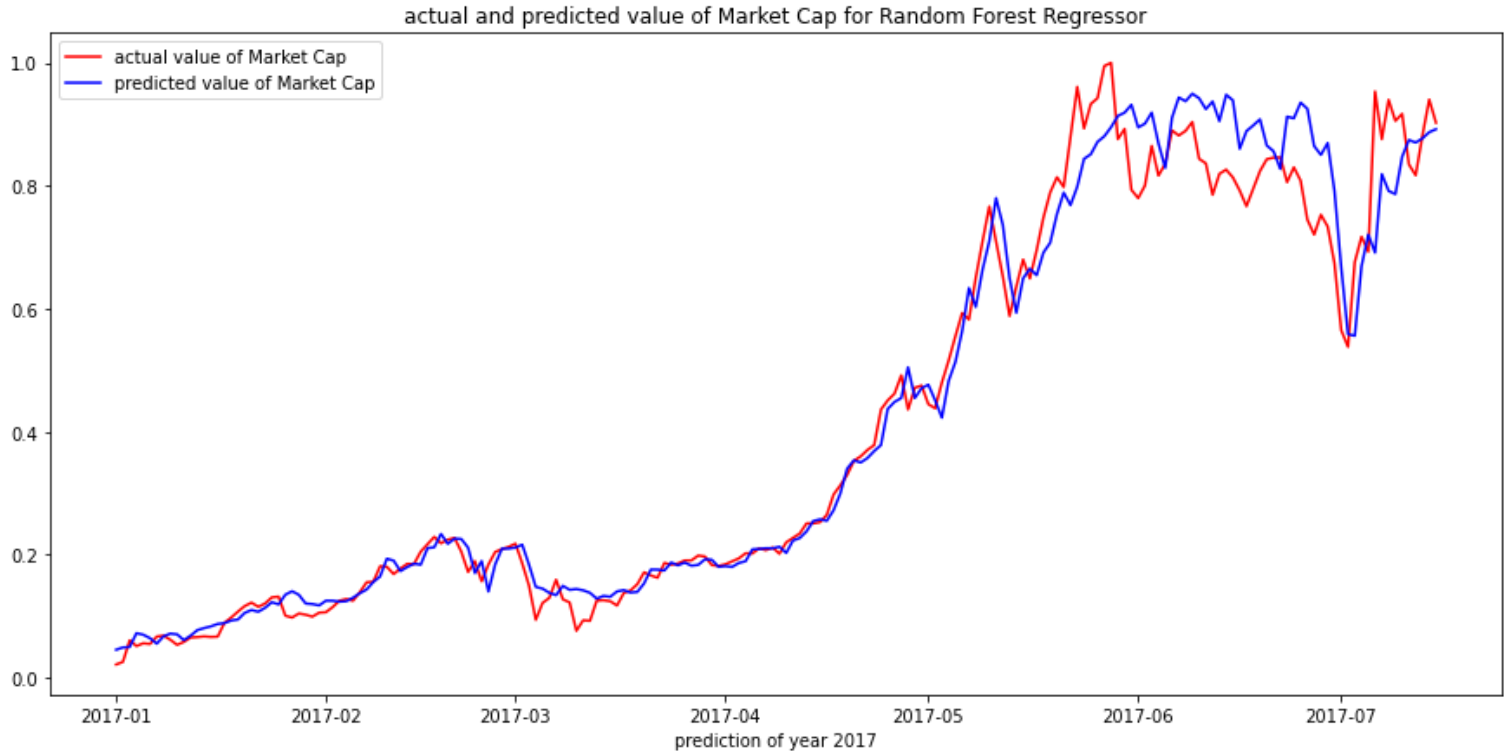
```
R2-score of close stock for Linear Regressor:  0.8186028000709457
Mean Squared Error of close stock for Linear Regressor:  0.00946695521299844
```



Analysis of Market Capitalization:

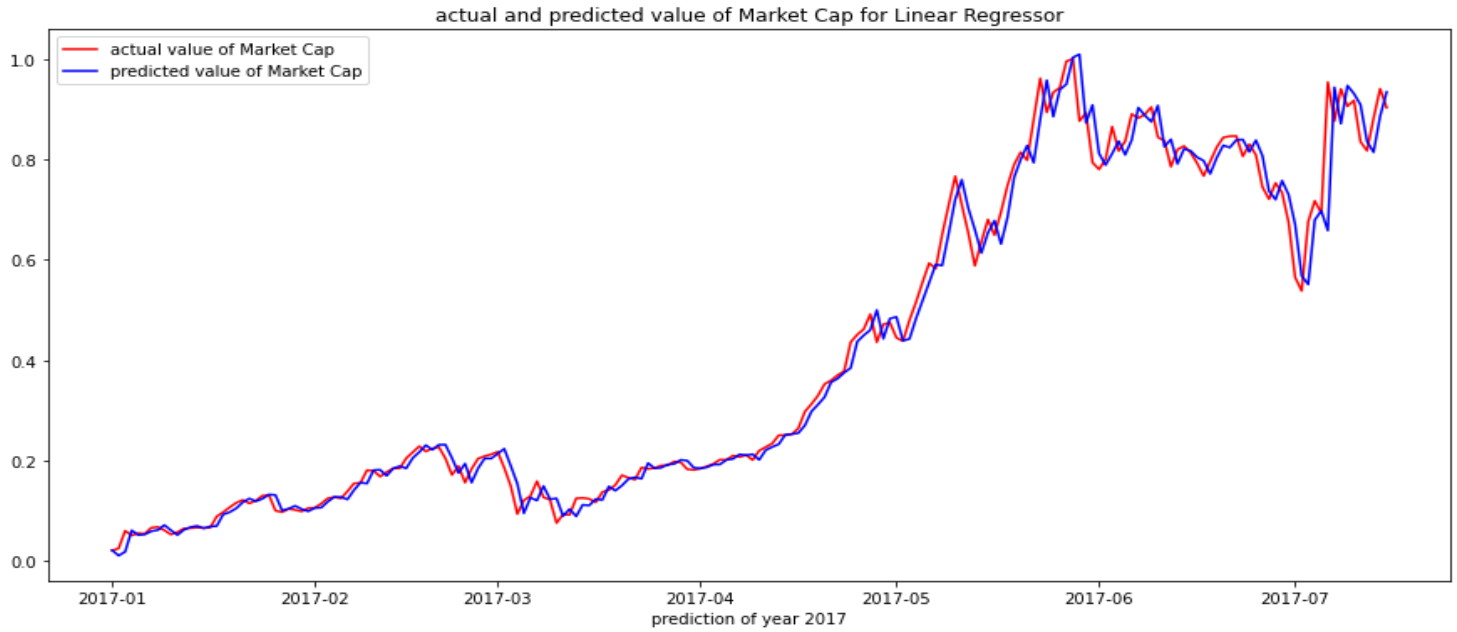
RandomForest Regressor for Market Capitalization:

```
R2-score of Market Cap for Random Forest Regressor:  0.9672530102560993
Mean Squared Error of Market Cap for Random Forest Regressor:  0.003222016710100199
```



Linear Regressor for Market Capitalization:

```
R2-score of Market Cap for Linear Regressor:  0.9842640017242529
Mean Squared Error of Market Cap for Linear Regressor:  0.0015482842786796497
```



R2-score and MSE for all attributes:

1) For Random Forest Regressor:

	Close	Open	High	Low	Volume	Market Cap
R2-score	0.975175	0.976524	0.983439	0.973061	0.790447	0.967253
Mean-squared-error	0.002423	0.002289	0.001783	0.002600	0.010936	0.003222

2) For Linear Regressor:

	Close	Open	High	Low	Volume	Market Cap
R2-score	0.983889	0.983856	0.983137	0.983107	0.818603	0.984264
Mean-squared-error	0.001572	0.001574	0.001816	0.001630	0.009467	0.001548