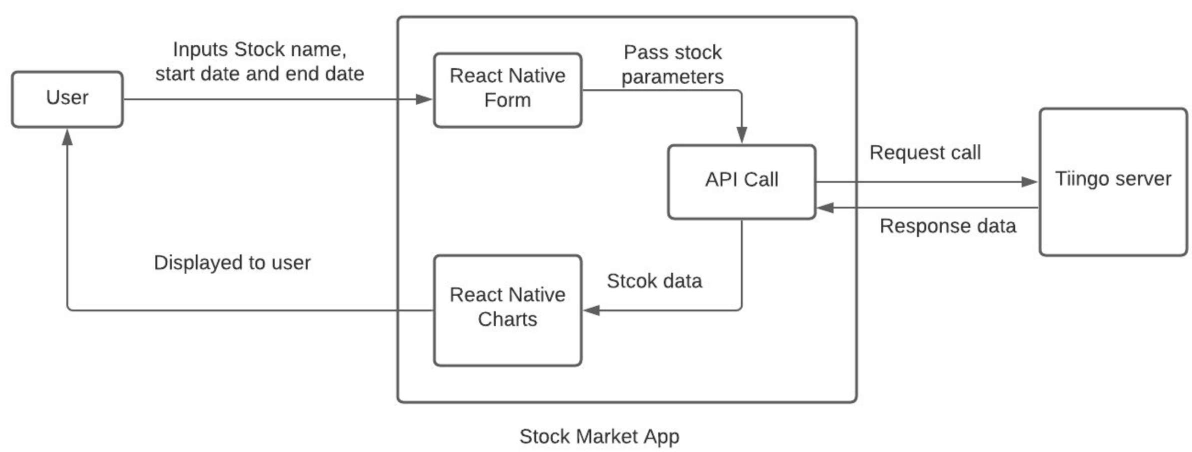
Stock Market Modelling App
REPORT

Contributors: Ankur Ingale, Arun Kumar, Garvit Mehta, Kommineni Nikhil, Tanya Chauhan

Disclaimer: The content of this report and the product made is only meant for the learning process as part of the course. This is not for use of making publication or making commercialisation without the mentor's consent. My contribution won't demand any claim in future for further progress of Mentor's development and innovation along the direction unless there is a special continuous involvement.

Problem Statement: To build a React Native mobile app with Django backend which models stock data price.

High Level Diagram:



This project is based on Stock Market modelling. It is a mobile application developed using React Native which is a popular choice for cross platform development. The app currently fetches stocks data from an API and uses the fetched data to display a histogram of stock prices. The tools and technologies used in making the app were explained in the project 1 report.

In project 2, we have added a Django backend. We have fitted ARMA and GARCH models to the data fetched from the API. We have selected the model with least AIC value. Here, AIC stands for the Akaike information criterion. It is an estimator of out-of-sample prediction error and thereby relative quality of statistical models for a given set of data. We have displayed the plots for the predictions made by the models.

ARMA model (Autoregressive moving average model): ARMA(p, q) indicates that there are p autoregressive terms and q moving-average terms in the model. The AR part involves regressing the variable on its own lagged (i.e., past) values. The MA part involves modeling the error term as a linear combination of error terms occurring contemporaneously and at various times in the past. The equation is as follows:

$$X_t = c + \epsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

where c is a constant, ϕ_i and θ_i are the parameters of the model and the random variable ϵ_t is white noise.
These models assume that the log-returns of the stock price is stationary. Therefore the parameters obtained while fitting the model should satisfy the stationary conditions.

The fitted model on a sample input stock ticker and the predictions made by the model are shown below.

GARCH model (Generalised Autoregressive Conditional Heteroskedasticity): GARCH is used extensively within the financial industry as many asset prices are conditional heteroskedastic. Just like ARCH(p) is AR(p) applied to the variance of a time series, GARCH(p, q) is an ARMA(p,q) model applied to the variance of a time series. The AR(p) models the variance of the residuals (squared errors) or simply our time series squared. The MA(q) portion models the variance of the process.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{i-1}^{s} \beta_j \sigma_{t-j}^2$$

Generalised Autoregressive Conditional Heteroskedastic Model of Order p, q

A time series $\{\epsilon_t\}$ is given at each instance by:

$$\epsilon_t = \sigma_t w_t$$

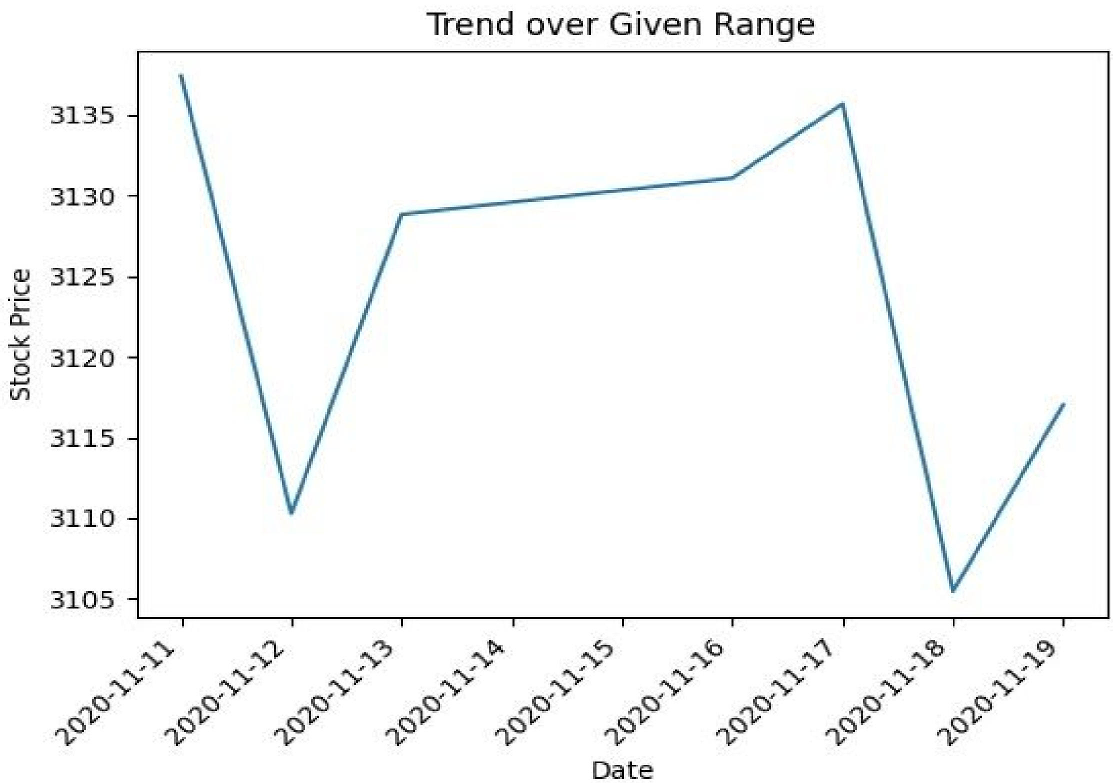
Where $\{w_t\}$ is discrete white noise, with zero mean and unit variance, and σ_t^2 is given by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

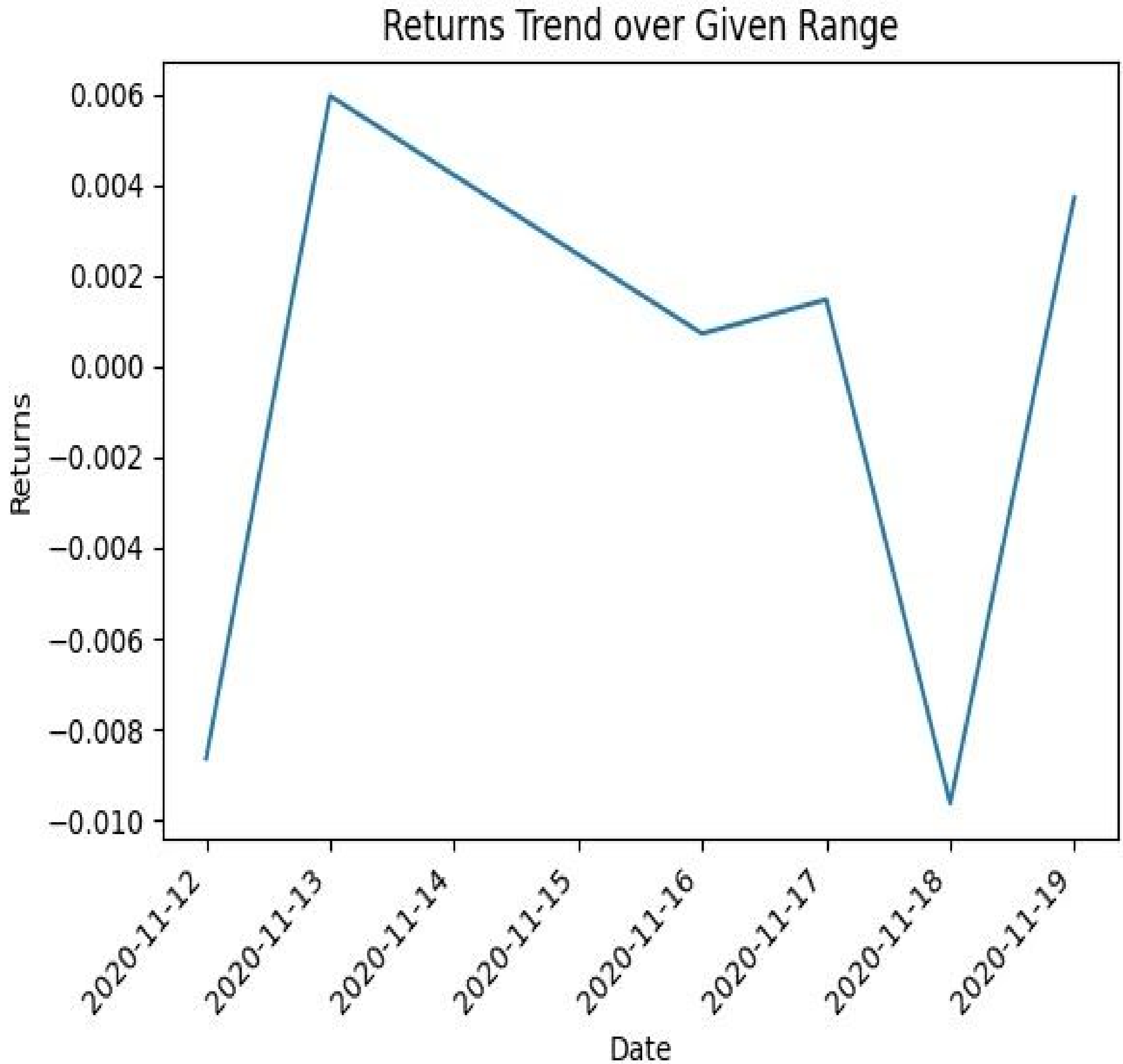
Where α_i and β_j are parameters of the model.

We say that $\{\epsilon_t\}$ is a *generalised autoregressive conditional heteroskedastic model of order p,q*, denoted by GARCH(p,q).

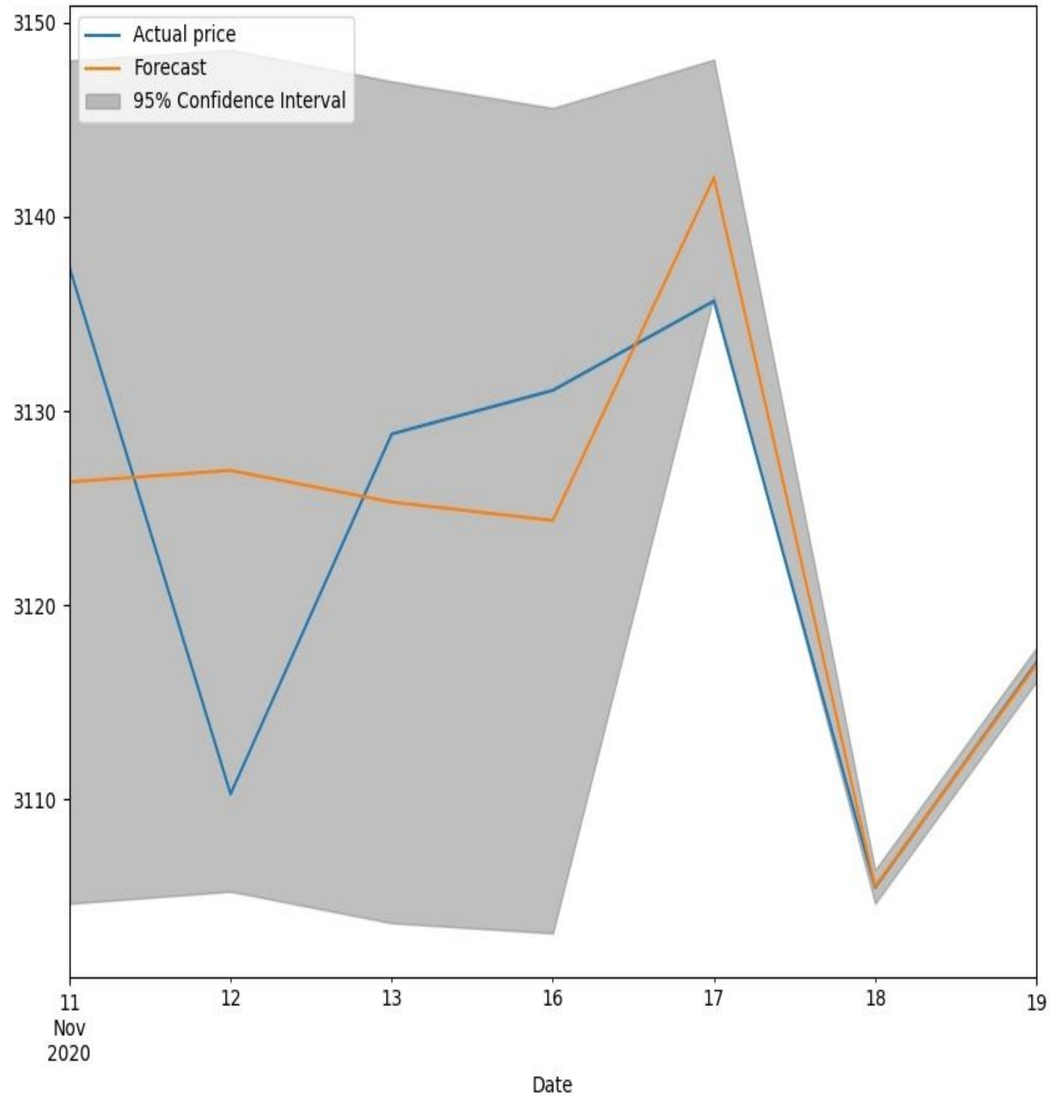
The fitted model on a sample input stock ticker = “AMZN” from Nov 11,2020 and Nov 19,2020 and the predictions made by the model between the given dates are shown below.



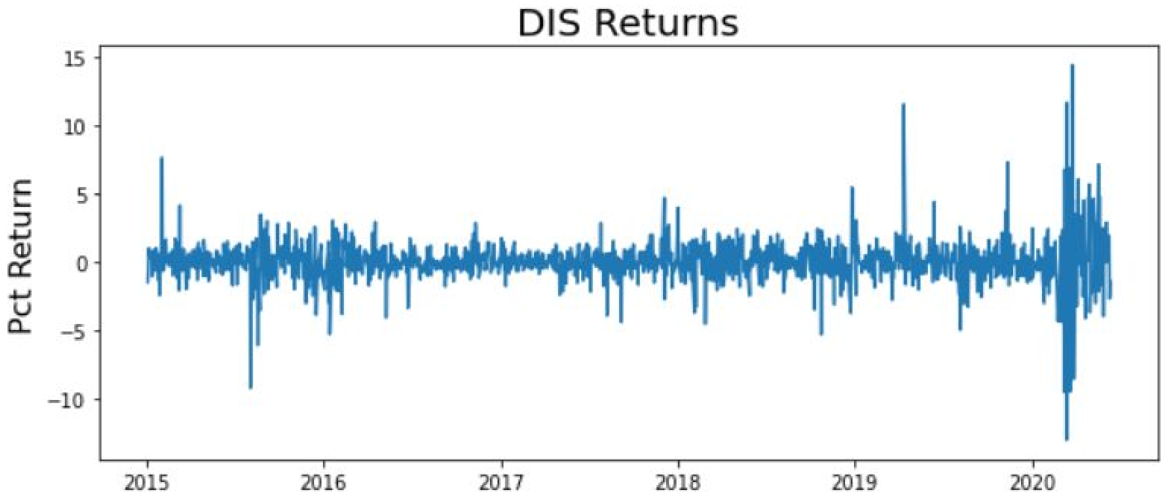
Trend over Given Range
Stock Price
3135
3130
3125
3120
3115
3110
3105
2020-11-11
2020-11-12
2020-11-13
2020-11-14
2020-11-15
2020-11-16
2020-11-17
2020-11-18
2020-11-19
Date

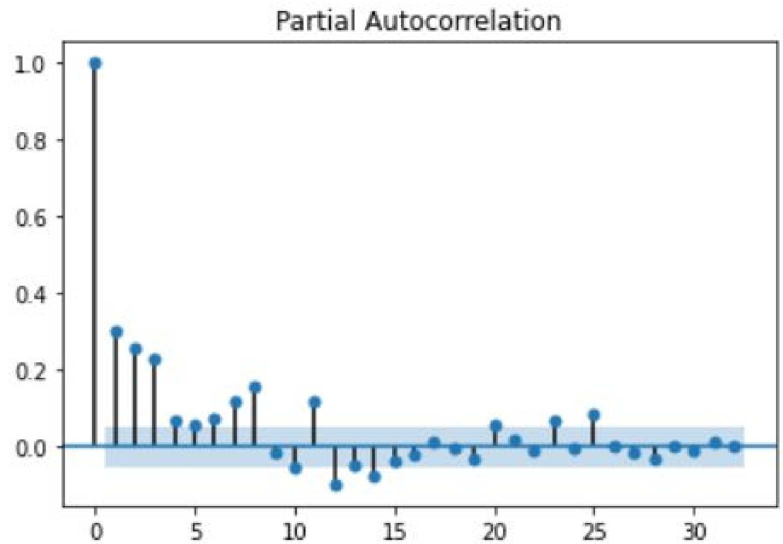


Returns Trend over Given Range
0.006
0.004
0.002
0.000
−0.002
−0.004
−0.006
−0.008
−0.010
Returns
2020-11-12
2020-11-13
2020-11-14
2020-11-15
2020-11-16
2020-11-17
2020-11-18
2020-11-19
Date

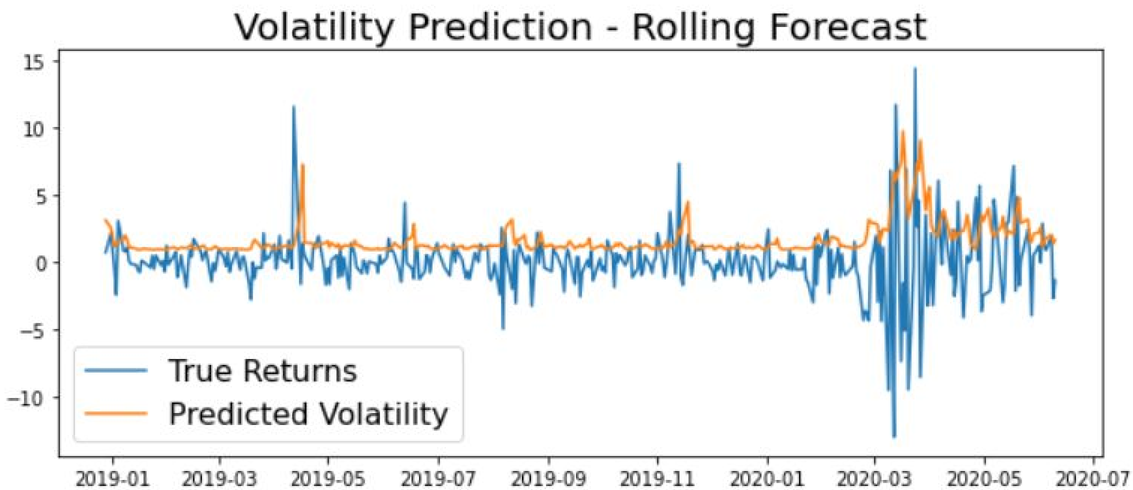

Actual price
Forecast
95% Confidence Interval
3150
3140
3130
3120
3110
11
Nov
2020
12
13
16
17
18
19
Date

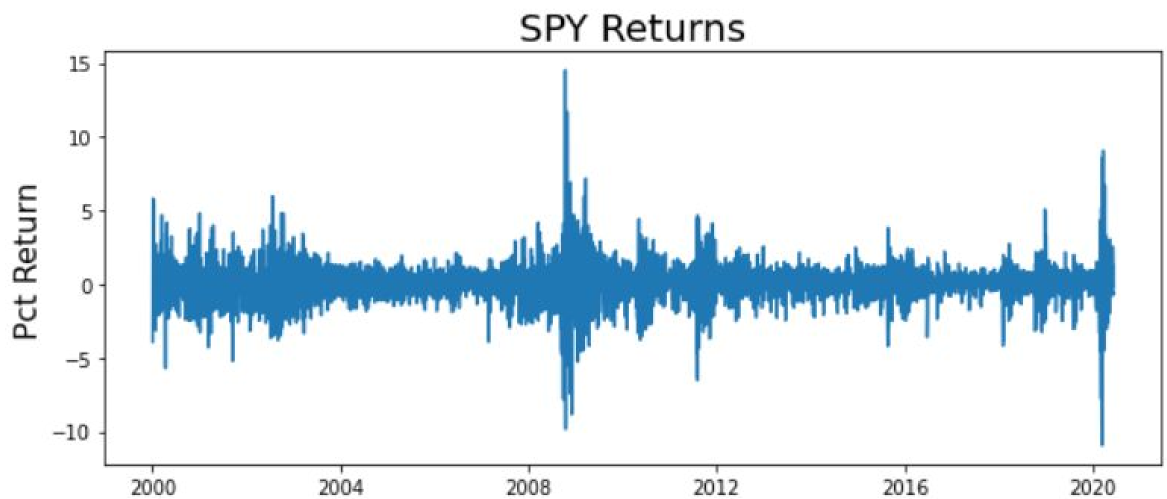
```
                               SARIMAX Results
==============================================================================
Dep. Variable:                  Close   No. Observations:                    7
Model:                 ARIMA(5, 0, 0)   Log Likelihood                 -19.536
Date:                Thu, 26 Nov 2020   AIC                             53.072
Time:                        17:38:51   BIC                             52.694
Sample:                    11-11-2020   HQIC                            48.392
                         - 11-19-2020
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       3126.3438     11.121    281.125      0.000    3104.547    3148.140
ar.L1          1.1886      3.407      0.349      0.727      -5.489       7.866
ar.L2         -0.2440      0.329     -0.742      0.458      -0.889       0.400
ar.L3         -0.2466      0.346     -0.712      0.477      -0.925       0.432
ar.L4          1.1960      3.463      0.345      0.730      -5.592       7.984
ar.L5         -0.9897      0.054    -18.328      0.000      -1.096      -0.884
sigma2         0.1969     12.712      0.015      0.988     -24.718      25.111
===================================================================================
Ljung-Box (L1) (Q):                   2.17   Jarque-Bera (JB):                 0.87
Prob(Q):                              0.14   Prob(JB):                         0.65
Heteroskedasticity (H):               0.02   Skew:                            -0.70
Prob(H) (two-sided):                  0.03   Kurtosis:                         2.00
===================================================================================
```

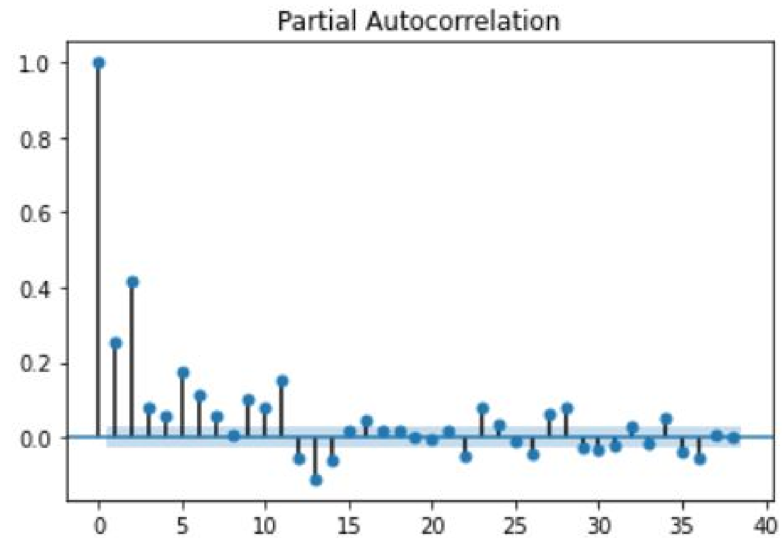

DIS Returns
Pct Return
15
10
5
0
−5
−10
2015
2016
2017
2018
2019
2020



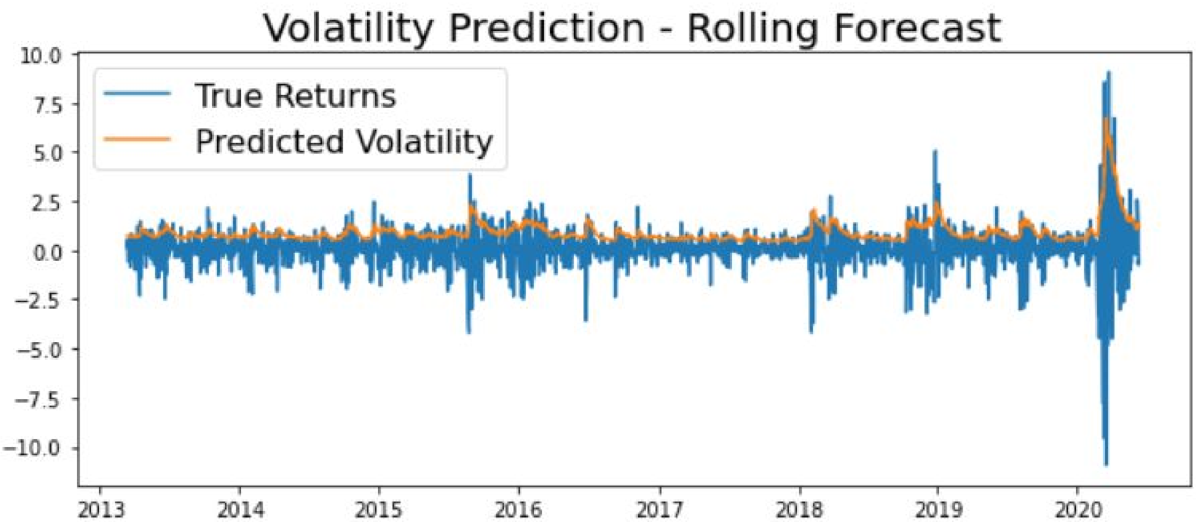

```
Iteration:      1,   Func. Count:      7,   Neg. LLF: 2373.535733608614
Iteration:      2,   Func. Count:     17,   Neg. LLF: 2371.314534935812
Iteration:      3,   Func. Count:     26,   Neg. LLF: 2368.4827476952723
Iteration:      4,   Func. Count:     34,   Neg. LLF: 2348.761123360837
Iteration:      5,   Func. Count:     42,   Neg. LLF: 2333.340283627342
Iteration:      6,   Func. Count:     50,   Neg. LLF: 2329.706255784284
Iteration:      7,   Func. Count:     58,   Neg. LLF: 2321.491751235455
Iteration:      8,   Func. Count:     67,   Neg. LLF: 2319.814027518538
Iteration:      9,   Func. Count:     74,   Neg. LLF: 2314.8798164018094
Iteration:     10,   Func. Count:     81,   Neg. LLF: 2314.2849034934216
Iteration:     11,   Func. Count:     88,   Neg. LLF: 2314.143999368298
Iteration:     12,   Func. Count:     95,   Neg. LLF: 2314.0779726786495
Iteration:     13,   Func. Count:    102,   Neg. LLF: 2314.0771981485614
Iteration:     14,   Func. Count:    109,   Neg. LLF: 2314.077192862747
Optimization terminated successfully.    (Exit mode 0)
            Current function value: 2314.0771928627173
            Iterations: 14
            Function evaluations: 109
            Gradient evaluations: 14
```

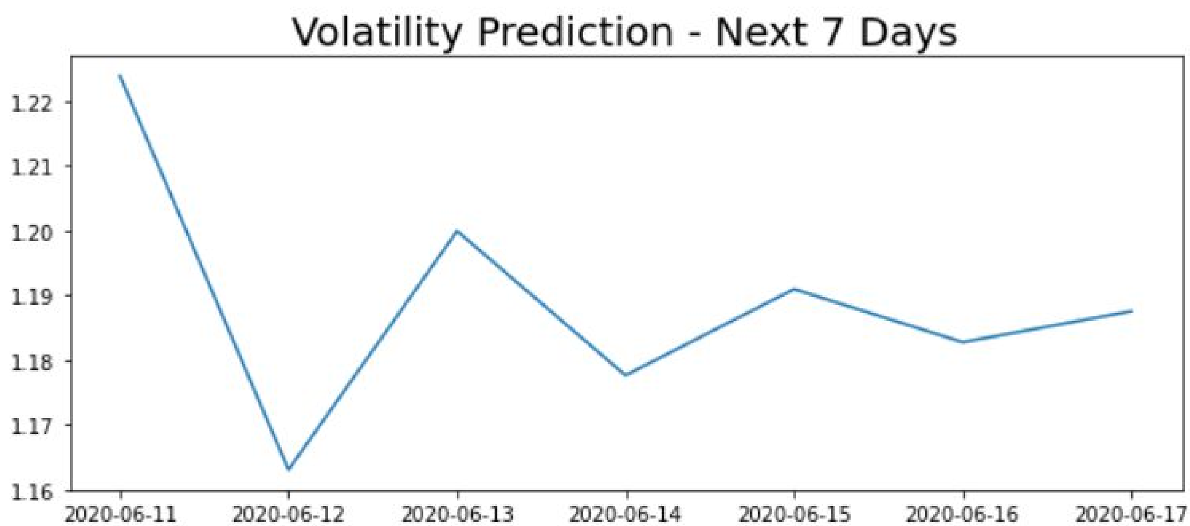

Volatility Prediction - Rolling Forecast
True Returns
Predicted Volatility
15
10
5
0
−5
−10
2019-01
2019-03
2019-05
2019-07
2019-09
2019-11
2020-01
2020-03
2020-05
2020-07


SPY Returns
Pct Return
15
10
5
0
−5
−10
2000
2004
2008
2012
2016
2020




```
Iteration:      1,   Func. Count:      8,   Neg. LLF: 7063.849269096254
Iteration:      2,   Func. Count:     21,   Neg. LLF: 7059.949362433973
Iteration:      3,   Func. Count:     33,   Neg. LLF: 7059.912631632853
Iteration:      4,   Func. Count:     43,   Neg. LLF: 7059.364150027669
Iteration:      5,   Func. Count:     53,   Neg. LLF: 7054.846718303968
Iteration:      6,   Func. Count:     62,   Neg. LLF: 7054.357146455823
Iteration:      7,   Func. Count:     71,   Neg. LLF: 7053.885873265253
Iteration:      8,   Func. Count:     80,   Neg. LLF: 7053.660000757351
Iteration:      9,   Func. Count:     90,   Neg. LLF: 7053.649800194631
Iteration:     10,   Func. Count:     99,   Neg. LLF: 7053.570243379768
Iteration:     11,   Func. Count:    107,   Neg. LLF: 7053.566838787082
Iteration:     12,   Func. Count:    115,   Neg. LLF: 7053.566735130049
Optimization terminated successfully.    (Exit mode 0)
            Current function value: 7053.566735132504
            Iterations: 12
            Function evaluations: 115
            Gradient evaluations: 12
Text(0.5, 1.0, 'Volatility Prediction - Next 7 Days')
```


Volatility Prediction - Rolling Forecast
True Returns
Predicted Volatility
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
2013
2014
2015
2016
2017
2018
2019
2020


Volatility Prediction - Next 7 Days
1.22
1.21
1.20
1.19
1.18
1.17
1.16
2020-06-11
2020-06-12
2020-06-13
2020-06-14
2020-06-15
2020-06-16
2020-06-17