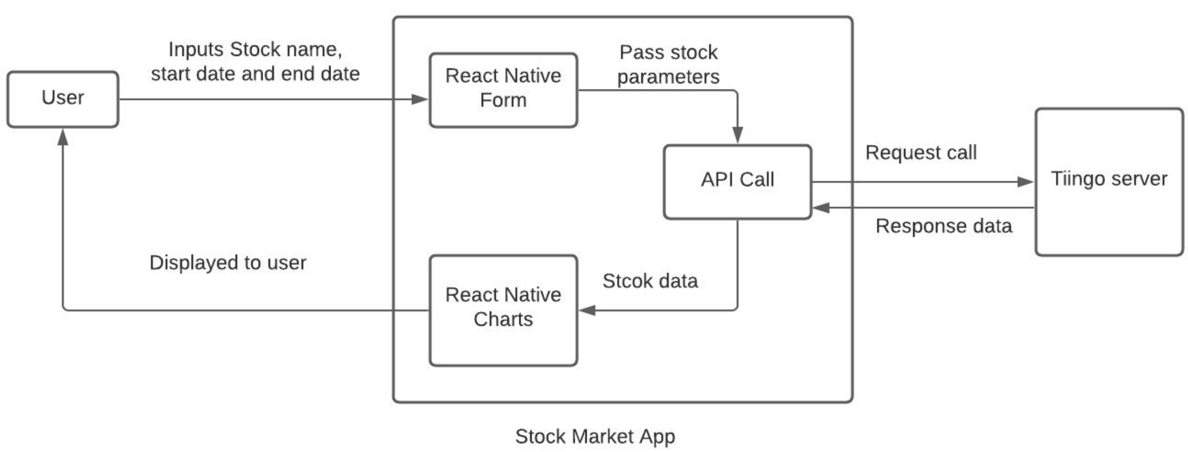
Stock Market Modelling App
REPORT

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.

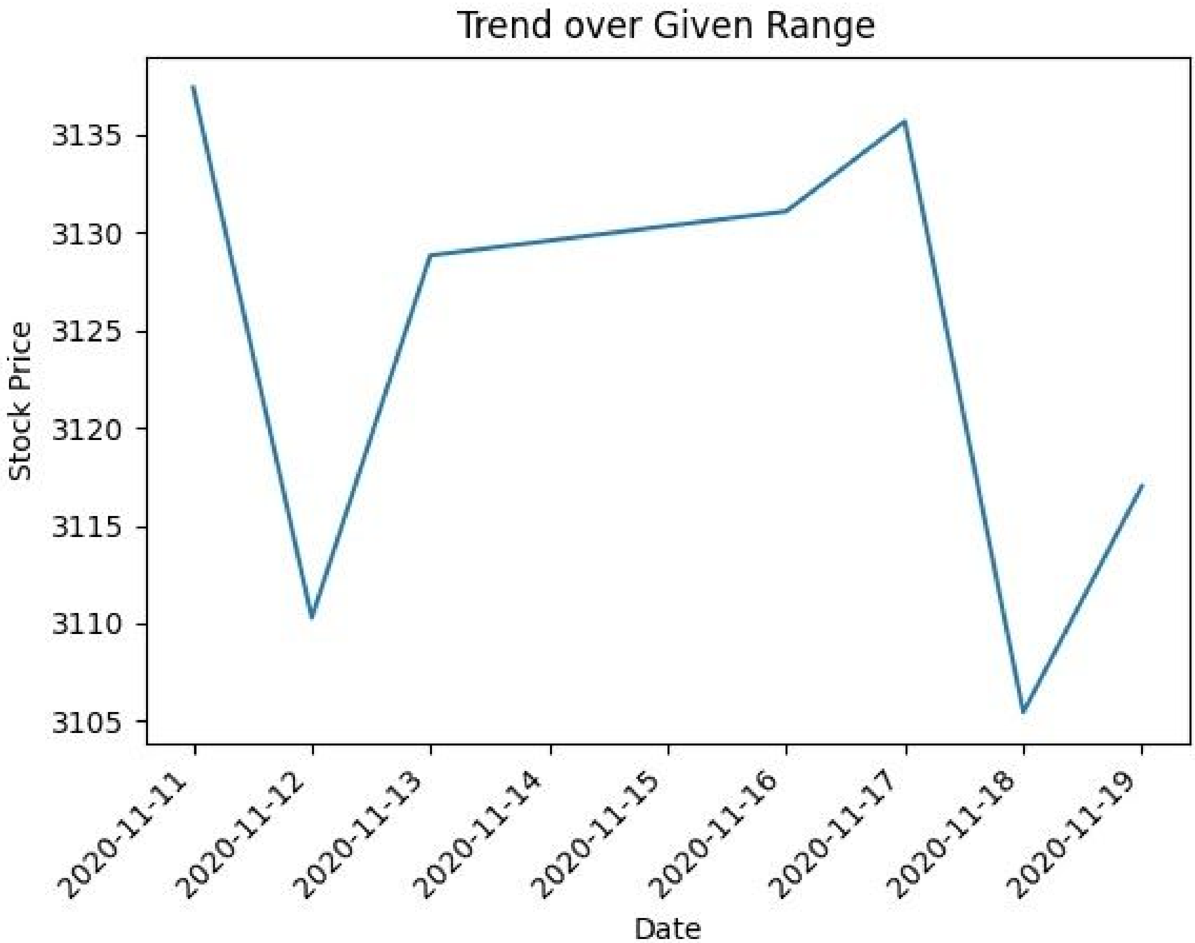
The fitted model on a sample input stock ticker and the predictions made by the model are shown below.

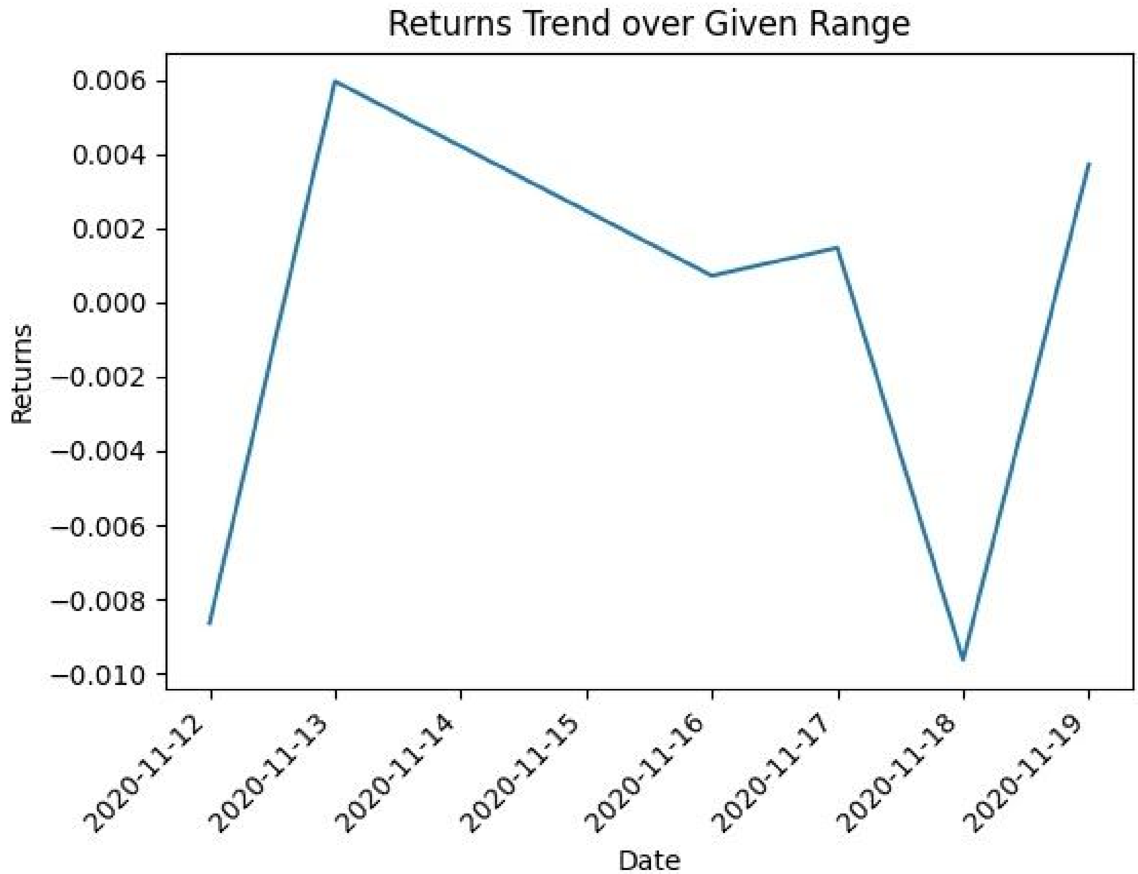
GARCH model (Generalised Autoregressive Conditional Heteroskedasticity): GARCH is used extensively within the financial industry as many asset prices are conditional heteroskedastic. Just like ARCH(p) is AR(p) applied to the variance of a time series, GARCH(p, q) is an ARMA(p,q) model applied to the 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.

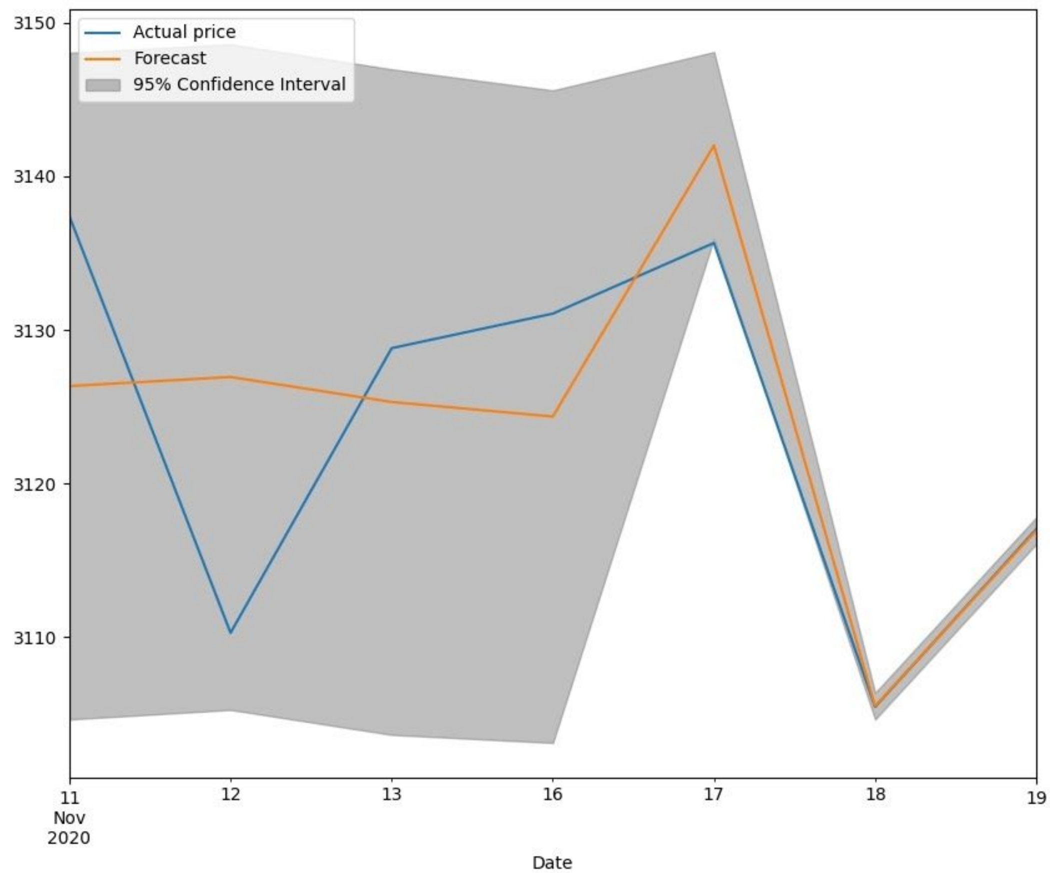
The GARCH(p,q) model is defined by the following system of equations:
$a_t = \sigma_t \epsilon_t${t} = \sigma_{t}\epsilon_{t} \\$t = \sigma_t \epsilon_t$} = \sigma_{t}\epsilon_{t} \\ = \sigma_{t}\epsilon_{t} \\$= \sigma_t \epsilon_t$\sigma_{t}\epsilon_{t} \\$\sigma_t \epsilon_t$igma_{t}\epsilon_{t} \\$igma_t \epsilon_t$ ma_{t}\epsilon_{t} \\$ma_t \epsilon_t$_{t}\epsilon_{t} \\$_t \epsilon_t$t}\epsilon_{t} \\ \epsilon_{t} \\$\epsilon_t$psilon_{t} \\$psilon_t$ilon_{t} \\$ilon_t$on_{t} \\$on_t$_{t} \\$_t$t} \\ \\$a_t = \sigma_t \epsilon_t$

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

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


Trend over Given Range
3135
3130
3125
3120
3115
3110
3105
Stock Price
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
Returns
0.006
0.004
0.002
0.000
−0.002
−0.004
−0.006
−0.008
−0.010
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:                        16:17:03   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
===================================================================================
```