Web Prototype for modeling stocks
Project 1 for MA471

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 commercialization without mentor's consent. My contribution won't demand any claim in the future for further progress of Mentor's development and innovation along the direction unless there is a special continuous involvement"

There are several ways to understand time series.
Some of the models are described below.

Autoregressive model

The notation AR(p) indicates an autoregressive model of order *p*.
AR(*p*) model is defined as

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

Where φ_i are parameters of the model and c is a constant and ε_t is white noise.
Here $Cov(\varepsilon_t, \varphi_{t-j})=0$ is assumed.

Moving average model

The notation MA(q) indicates a moving average model of order q.
MA(q) model is defined as

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

Where θ_i are the parameters of the model, μ is the mean, and ε_t is the white noise of the series.

Combining these 2 series, we have,

Autoregressive moving average model (ARMA)

The notation ARMA(p, q) indicates to the model with p autoregressive terms and q moving-average terms.

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}.$$

In this project, we are using the ARMA model for modeling stocks.

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.

$|\varphi_i|<1$ for the AR coefficients and $|\theta_i|<1$ for the MA coefficients.
If the log-returns of the stationary conditions are not satisfied the model shows an error.
(These conditions are derived from the characteristic equation of the model)

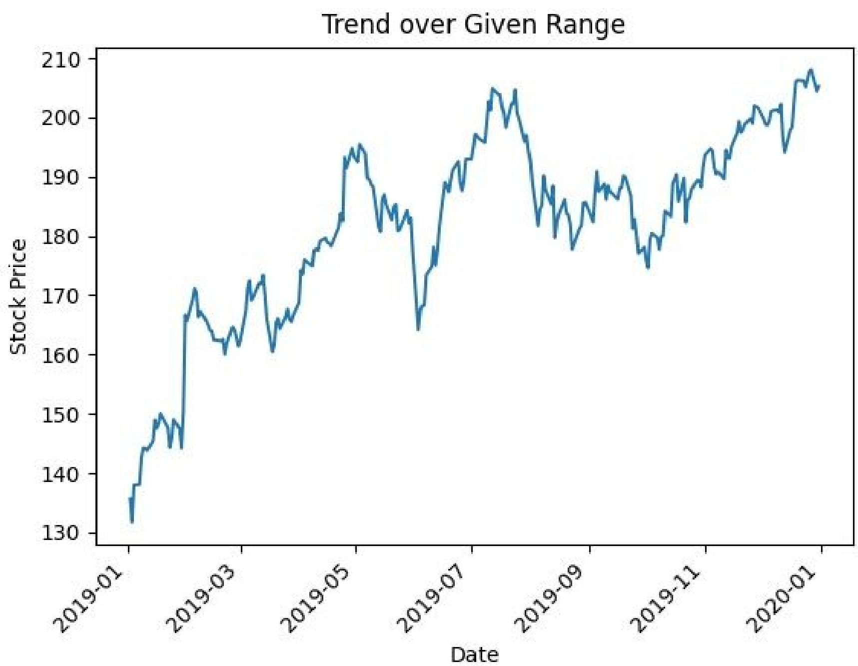
Choosing the order of ARMA(p,q)
Several methods exist to select the order for the ARMA model.

Akaike information criterion(AIC)
$AIC(l) = \ln(\tilde{\sigma}^2) + l(2/T)$,
This metric uses a penalty on the number of parameters added.
So we select an order which has the minimum AIC and fit the model.

Another commonly used criterion function is BIC
Bayesian information criterion (BIC)
$BIC(l) = \ln(\tilde{\sigma}^2) + l(\ln(T)/T)$.
BIC uses log(T) penalty whereas AIC uses 2 for each extra parameter.

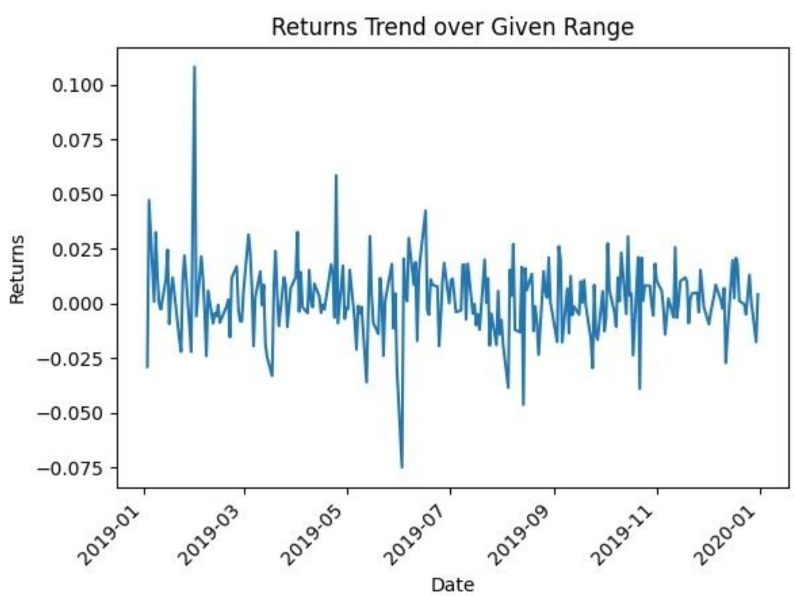
Here we use the AIC metric to select the order.
A summary is generated after we run the model which can be seen below.

Below are the insights for **Facebook** from Jan 01, 2019, to Jan 1, 2020.



Here it is important to note that the returns of the stock are stationary which is a key factor for the ARMA model.
If the returns do not follow the stationary rule, the model might show an error.

The summary of the ARMA model is as follows.



```
                                     SARIMAX Results
==============================================================================
Dep. Variable:                  Close   No. Observations:                  252
Model:                 ARIMA(5, 0, 3)   Log Likelihood                -634.162
Date:                Thu, 15 Oct 2020   AIC                           1288.323
Time:                        16:07:07   BIC                           1323.618
Sample:                             0   HQIC                          1302.525
                                - 252
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const        181.4428     39.946      4.542      0.000     103.149     259.736
ar.L1          0.4634      0.088      5.278      0.000       0.291       0.635
ar.L2          0.0419      0.050      0.836      0.403      -0.056       0.140
ar.L3         -0.3948      0.063     -6.260      0.000      -0.518      -0.271
ar.L4          0.9485      0.052     18.375      0.000       0.847       1.050
ar.L5         -0.0659      0.077     -0.853      0.394      -0.217       0.086
ma.L1          0.5349      0.071      7.510      0.000       0.395       0.675
ma.L2          0.5511      0.064      8.634      0.000       0.426       0.676
ma.L3          0.9646      0.074     13.097      0.000       0.820       1.109
sigma2         8.6960      0.647     13.449      0.000       7.429       9.963
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):               208.60
Prob(Q):                              0.97   Prob(JB):                         0.00
Heteroskedasticity (H):               0.63   Skew:                             0.02
Prob(H) (two-sided):                  0.04   Kurtosis:                         7.46
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

The order of the model is **(5,3)** so, we have,a constant bias of AR model, 5 coefficients relating to the AR model and 3 coefficients relating to MA model and variance of the white noise

Since ARIMA model with d=0 is basically ARMA, ARMA(p,q) is simulated with statsmodel library in python with ARIMA(p,0,q)

`const`

The constant term in the Autoregressive model.

`ar.L1`

The first coefficient of the Autoregressive part of the model.
`ar.Li` is the ith coefficient of the autoregressive part of the model.

`ma.L1`

The first coefficient of the Moving average part of the model.
`ma.Li` is the ith coefficient of the Moving average part of the model.

`Skew`

Skew index of the stock price plot indicates the shape of the distribution

`Kurtosis`

Kurtosis index is the measure of the degree to which value cluster in the tails or the peak of the frequency distribution.

The plot below shows the actual stock price of Facebook along with the forecast values from the ARMA(5,3) model.