Financial Analytics - Assignment

Piyush Singh

2001CS51

OVERVIEW

I have predicted the value of `NIFTY 50` index using ARIMA. The methodology adopted and the steps taken are described below. The value of the index for a period of one year has been obtained from here:
https://finance.yahoo.com/quote/%5ENSEI/history?p=%5ENSEI

The link for Google Colab Notebook created is as follows:
https://colab.research.google.com/drive/1jvttmqQAU74L-MAS2AgYgQFdNazZOhwU?usp=sharing

STEPS

- `ARIMA` model has been used for predicting the value. The prediction has been done for the closing value of the `NIFTY 50` index on 31st March 2023.
- `statsmodels` library in python has been used for the purpose which is a is a Python module that provides classes and functions for the estimation of many different statistical models.
- `NIFTY 50` over a period of 1 year from `2022-03-30` to `2023-03-29` show below.
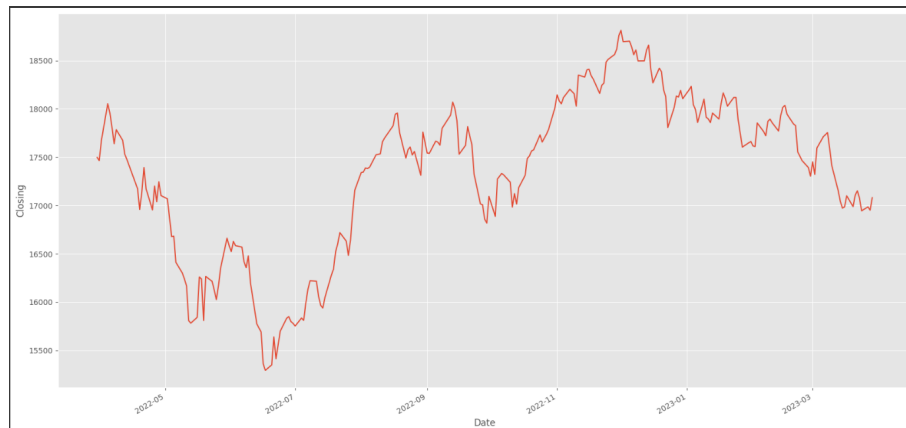
- Next we compute all possible combination of values for `p, q, d`

```
# Creating tuples of all possible combinations of p, d, q
# in range 0 to 4 (inclusive)
p = d = q = range(0,5)
pdq = list(itertools.product(p, d, q))
print(pdq)
```

- Next, we check the combination of the tuple (`p, q, d`) for which the **AIC(Akaike information criterion)** value is minimum.

```
AIC = []
ARIMA_model = []
for param in pdq:
    mod = statsmodels.tsa.arima.model.ARIMA(train_data,
                                            order=param,
                                            enforce_stationarity=False,

                                       ↪   enforce_invertibility=False)
    res = mod.fit()
    print('ARIMA{} - AIC:{}'.format(param, res.aic), end='\r')
    AIC.append(res.aic)
    ARIMA_model.append([param])
print('The smallest AIC is {} for model
 ↪   ARIMA{}'.format(min(AIC),ARIMA_model[AIC.index(min(AIC))][0]))
```

- The smallest AIC value comes out to be 3138.8660 for the model `ARIMA(2,2,4)`
- We fit the model and the summary is as below(next page).
- Checking for stationarity for second order lag using Augmented Dickey–Fuller test. For no lag, the result comes out to be.

```
ADF Test Statistic : -1.4904959178388795
p-value : 0.5382075239865683
weak evidence against null hypothesis, time series has a unit root,
 ↪   indicating it is non-stationary
```

With no differencing, the series is not stationary.

- With second order differencing, the result of Augmented Dickey–Fuller test, comes out to be.

	coef	std err	z	P>|z|	[0.025	0.975]
Dep. Variable:	Close		No. Observations:	250		
Model:	ARIMA(2, 2, 4)		Log Likelihood	-1562.433		
Date:	Thu, 30 Mar 2023		AIC	3138.866		
Time:	15:44:02		BIC	3163.317		
Sample:	0		HQIC	3148.715		
	- 250					
Covariance Type:	opg					

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.4573	0.021	-22.071	0.000	-0.498	-0.417
ar.L2	-0.9271	0.017	-56.021	0.000	-0.960	-0.895
ma.L1	-0.4856	23.661	-0.021	0.984	-46.861	45.890
ma.L2	0.4737	23.135	0.020	0.984	-44.871	45.818
ma.L3	-1.0103	0.539	-1.876	0.061	-2.066	0.045
ma.L4	0.0223	0.062	0.359	0.719	-0.099	0.144
sigma2	2.149e+04	0.001	2.55e+07	0.000	2.15e+04	2.15e+04

| Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	0.26
Prob(Q):	0.95	Prob(JB):	0.88
Heteroskedasticity (H):	0.53	Skew:	0.05
Prob(H) (two-sided):	0.01	Kurtosis:	3.12

```
ADF Test Statistic : -4.976680051122677
p-value : 2.4676521592553575e-05
strong evidence against the null hypothesis(Ho), reject the null
↪  hypothesis. Data has no unit root and is stationary
```

In this case, the `p-value` is much lower than 0.05 and the null hypothesis can be rejected, that is the series is stationary.

- The plot for second differenced closing value v/s time is plotted as follows.
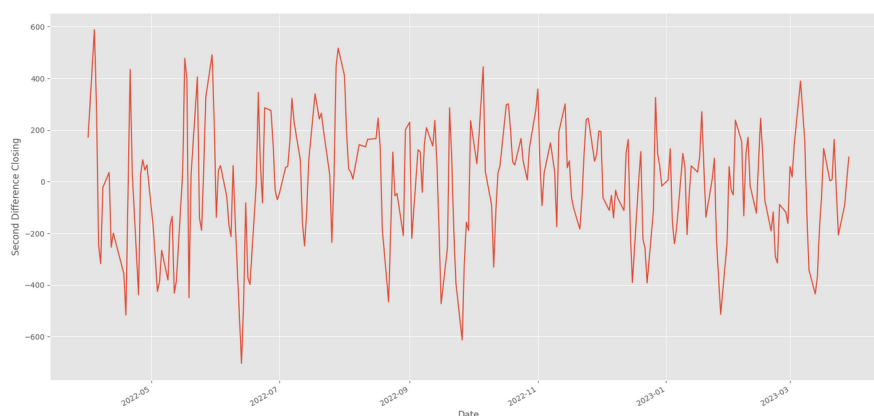
- The correlogram for `ACF` and `PACF` are as follows.
- Predicting the value for 31st March 2023 using `statsmodels.tsa.arima.model.ARIMAResults.predict`

Autocorrelation

Partial Autocorrelation
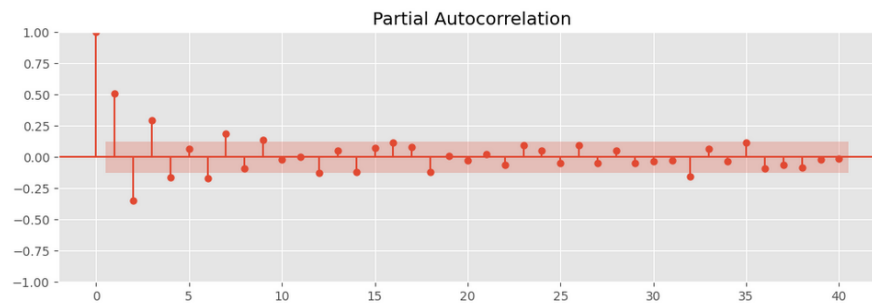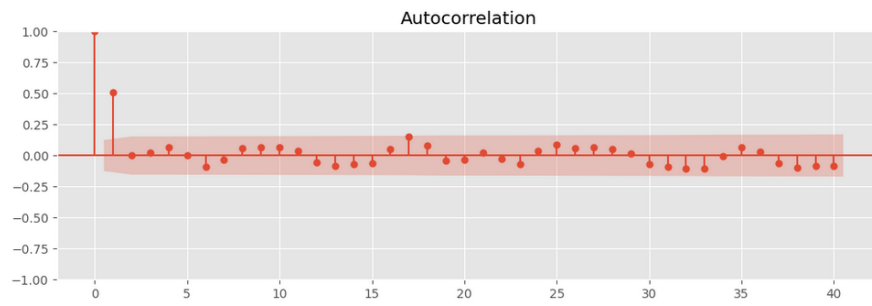
RESULTS

The predicted value for the `NIFTY 50` index on 31st March 2023 is: `17066.52`