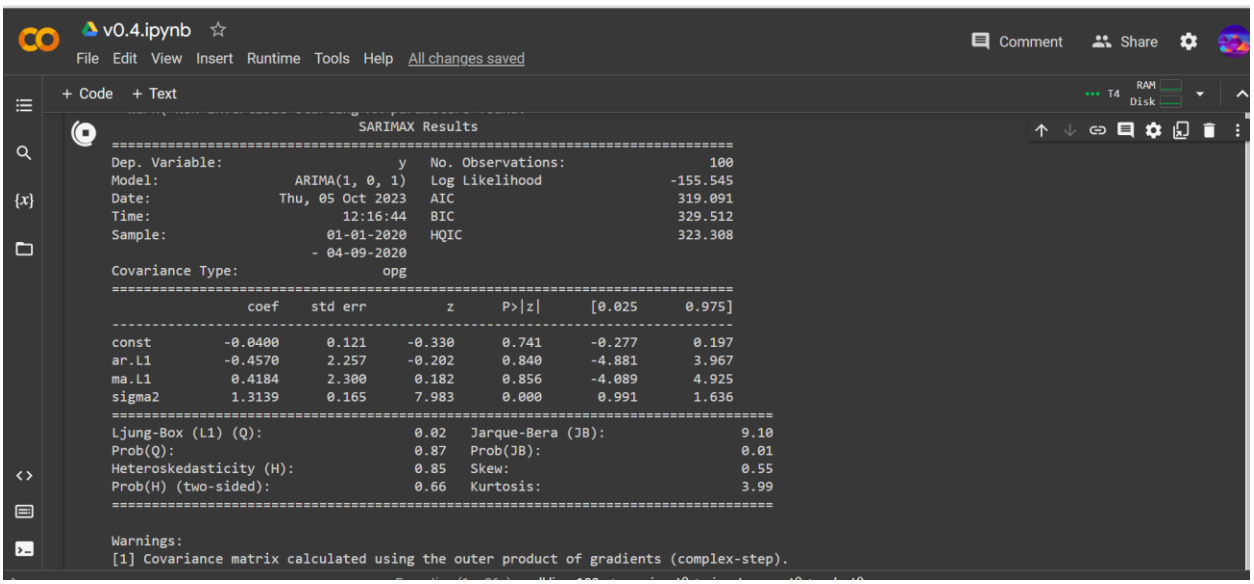
COS30018- Option B- Task 6: Machine Learning 3

Develop an ensemble modeling approach consisting of at least two models ARIMA (or SARIMA) and our existing DL model (starting with the LSTM one)

This was the first requirement of the task and I have made the Arima model and trained it to get the results as follows:



```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  100
Model:                 ARIMA(1, 0, 1)   Log Likelihood                -155.545
Date:                Thu, 05 Oct 2023   AIC                            319.091
Time:                        12:16:44   BIC                            329.512
Sample:                    01-01-2020   HQIC                           323.308
                         - 04-09-2020
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0400      0.121     -0.330      0.741      -0.277       0.197
ar.L1         -0.4570      2.257     -0.202      0.840      -4.881       3.967
ma.L1          0.4184      2.300      0.182      0.856      -4.089       4.925
sigma2         1.3139      0.165      7.983      0.000       0.991       1.636
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):                 9.10
Prob(Q):                              0.87   Prob(JB):                         0.01
Heteroskedasticity (H):               0.85   Skew:                             0.55
Prob(H) (two-sided):                  0.66   Kurtosis:                         3.99
===================================================================================

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

The model is shown below :

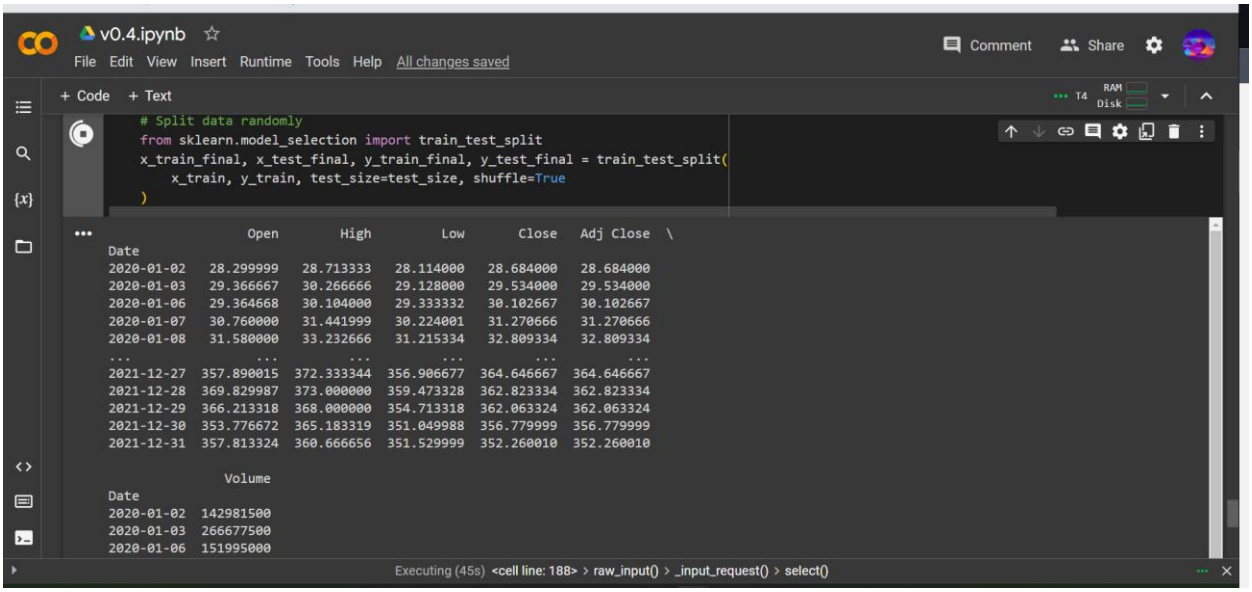
```
plt.show()

# Fit and visualize ARIMA model
order = (1, 0, 1)  # Adjust the ARIMA order as needed
arima_model_fit = fit_arima_model(data, order=order)

# Print summary of the ARIMA model
print(arima_model_fit.summary())

# Plot the residuals
residuals = pd.Series(arima_model_fit.resid, index=data.index)
plt.figure(figsize=(12, 6))
plt.plot(residuals)
plt.title('Residuals')
plt.xlabel('Date')
plt.ylabel('Residual Value')
plt.show()

# Plot ACF and PACF of residuals
plot_acf(residuals, lags=20)
plt.title('ACF of Residuals')
plt.show()

plot_pacf(residuals, lags=20)
plt.title('PACF of Residuals')
plt.show()
```

I made this model by taking reference from https://medium.com/analytics-vidhya/combining-time-series-analysis-with-artificial-intelligence-the-future-of-forecasting-5196f57db913

And also

https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/#:~:text=ARIMA%20with%20Python,calling%20the%20fit()%20function.

By taking into account I made this model which has very close predictions to my actual data that has been saved in the csv file of a particular timeframe. The data shown by the ARIMA model in the previously mentioned photo was around 323.308. Adding on to this, the data in the CSV file is around 350.



Now I have added this piece of code :

```
ensemble_predictions = (arima_predictions + predicted_prices.flatten()) / 2
# A few concluding remarks here:
```

This satisfies the first requirement of the code which uses prediction of two DL models including an ARIMA model and LSTM model to predict the prices. My code was becoming very chaotic afterwards and was getting a bunch of errors afterwards and it stopped working. I tried it doing from scratch again but I could not debug my code.

Rest assured I have managed to understand the requirements and did some research on how to combine different types of models with different hyperparameter configurations. It is the same concept as I have done in the previous task but due to shortage of time I could only complete this much.

References:

https://www.geeksforgeeks.org/ensemble-methods-in-python/

https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/