Week:3 ADML Assignment LSTM & RNN

LSTM

Data Preprocessing:

For this assignment, I selected the Yahoo Stock Market performance dataset. During preprocessing, I identified and handled null values, and I excluded the 'Adj. Closing Price' and 'Date Time' columns from the dataset.

Modelling Strategies:

To create the model, I initially examined the correlation plot among all variables. It became apparent that the 'Volume' variable did not exhibit significant importance, which is understandable since the volume of a stock is generally independent of its prices. Consequently, I chose to exclude the 'Volume' column, leaving us with three input columns (Open, High, Low) and one response column (Close).

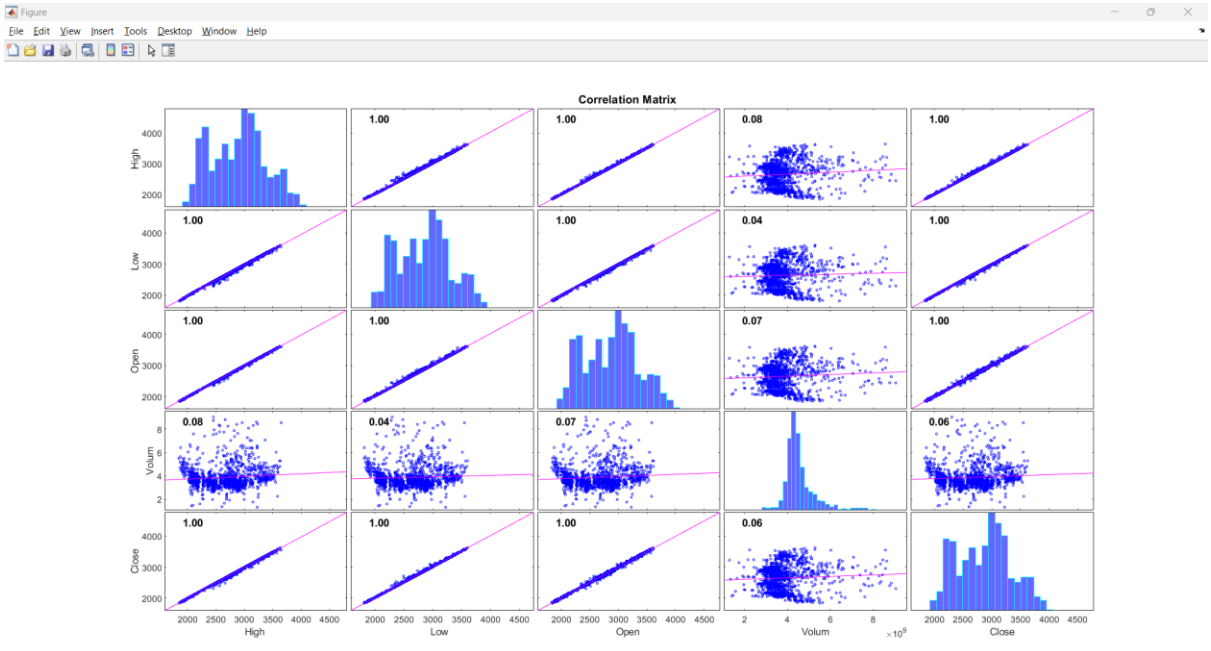

Fig.1

Additionally, I observed trends in the dataset, prompting me to focus on sequences rather than treating it as a collective time series. To address this, I implemented a code block to divide the dataset into a series of sequences. This approach allowed the model to better capture timeframes

instead of processing the entire dataset as a whole. Subsequently, I partitioned this data into training and testing sets.

For the LSTM model I set the maximum epochs at 800 and learning rate at 0.001. NumChannels was 3 as there were only 3 columns in the input vector.

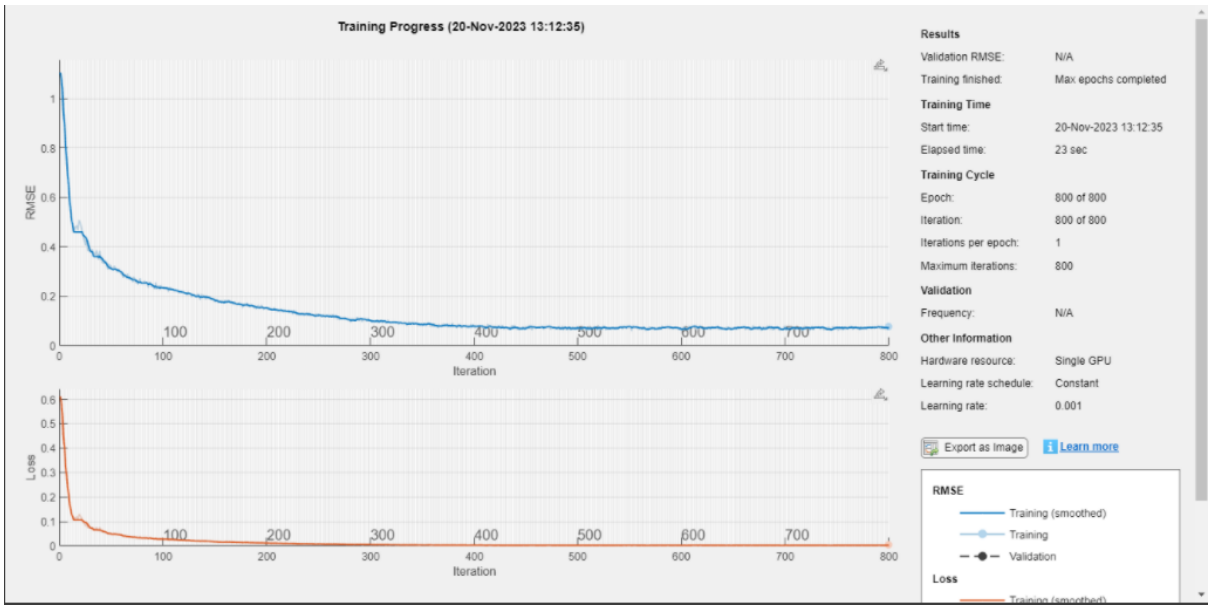

Fig.2

In Fig 2 we see that the RMSE And Loss become significantly lower after around 200 epochs for one ahead prediction.

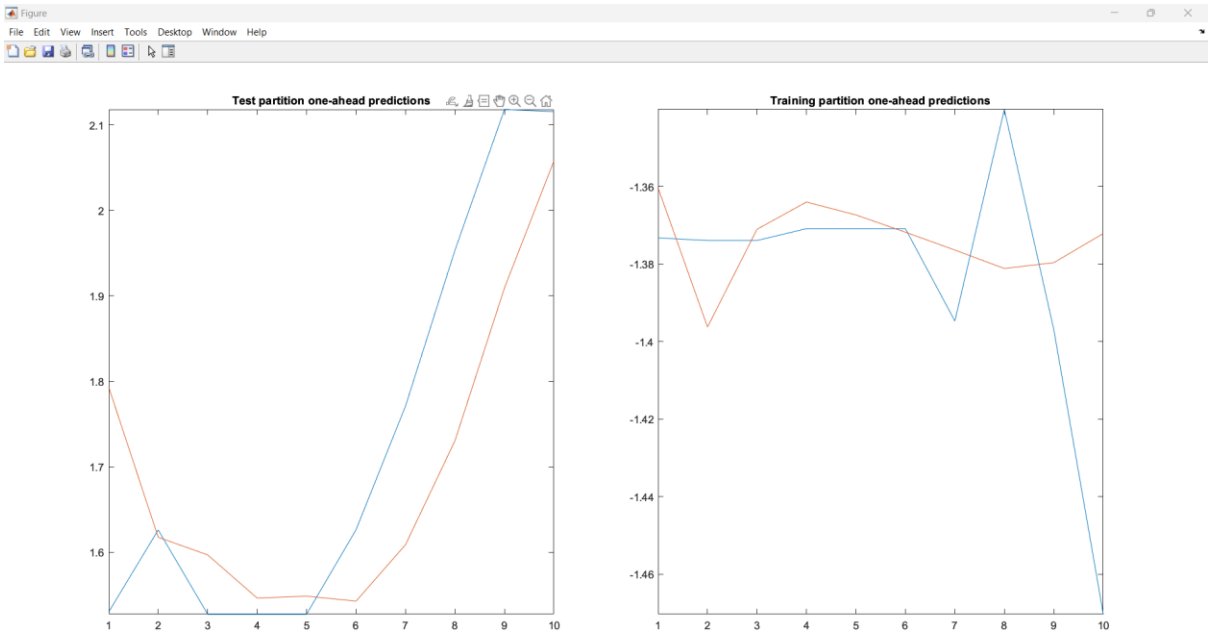

Fig.3

In Fig.3 we can see the actual performance of the one ahead prediction of our LSTM model. In this instance it doesn’t look too bad.

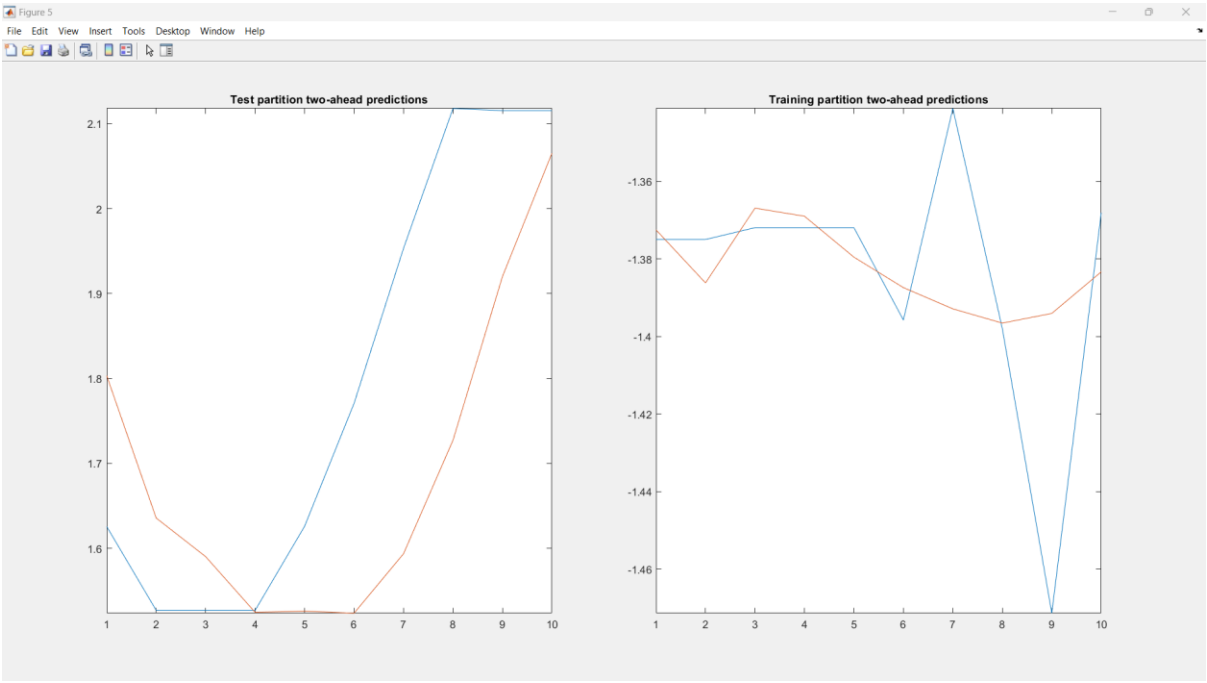

Fig.4

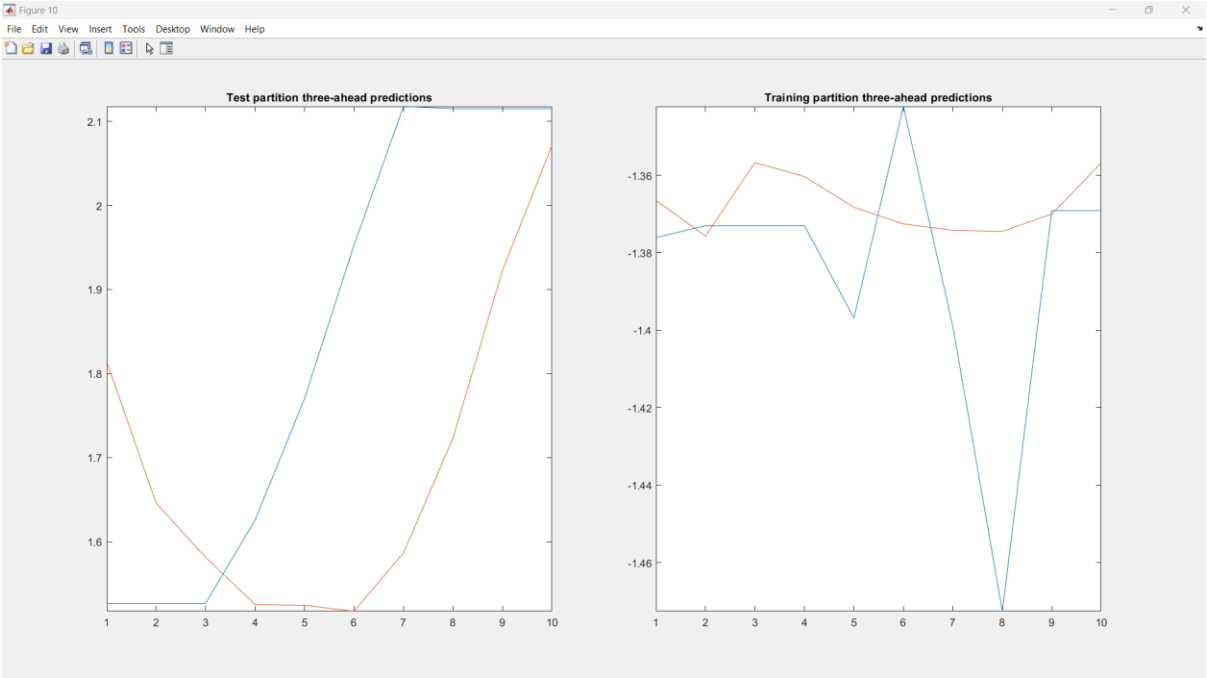

Fig.5

Figures 4 and 5 depict the model's performance in 2-ahead and 3-ahead predictions. The model demonstrates relatively good accuracy for 2-ahead predictions, but for 3-ahead predictions, the values deviate significantly. Consequently, the most optimistic approach for our model would be to rely on 1-ahead predictions. However, it's essential to note that even 1-ahead predictions may not

be highly accurate due to the inherent volatility of stock markets, influenced by various global factors on a day-to-day basis. Considering all circumstances, 1-ahead prediction emerges as the best approach among the available options.

RNN

For RNN We will also use the same dataset and the same preprocessing steps, as we see from Fig.6 & Fig.7 the performance of the RNN model is not bad compared to LSTM.

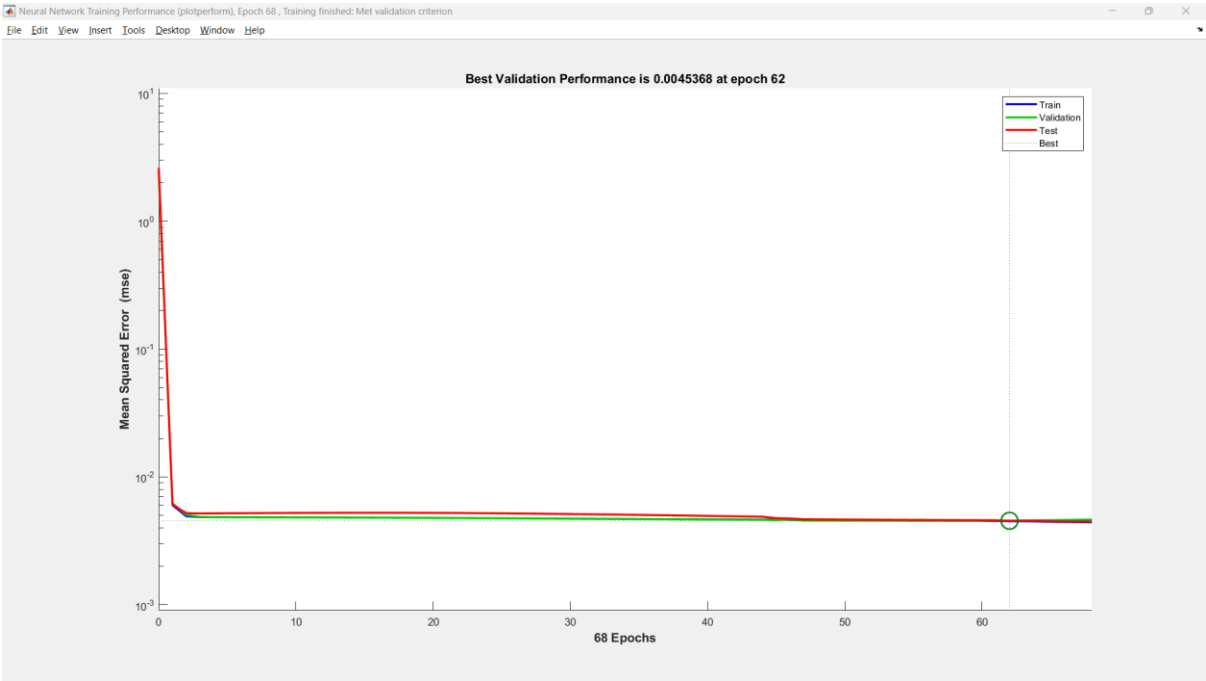

Fig.6

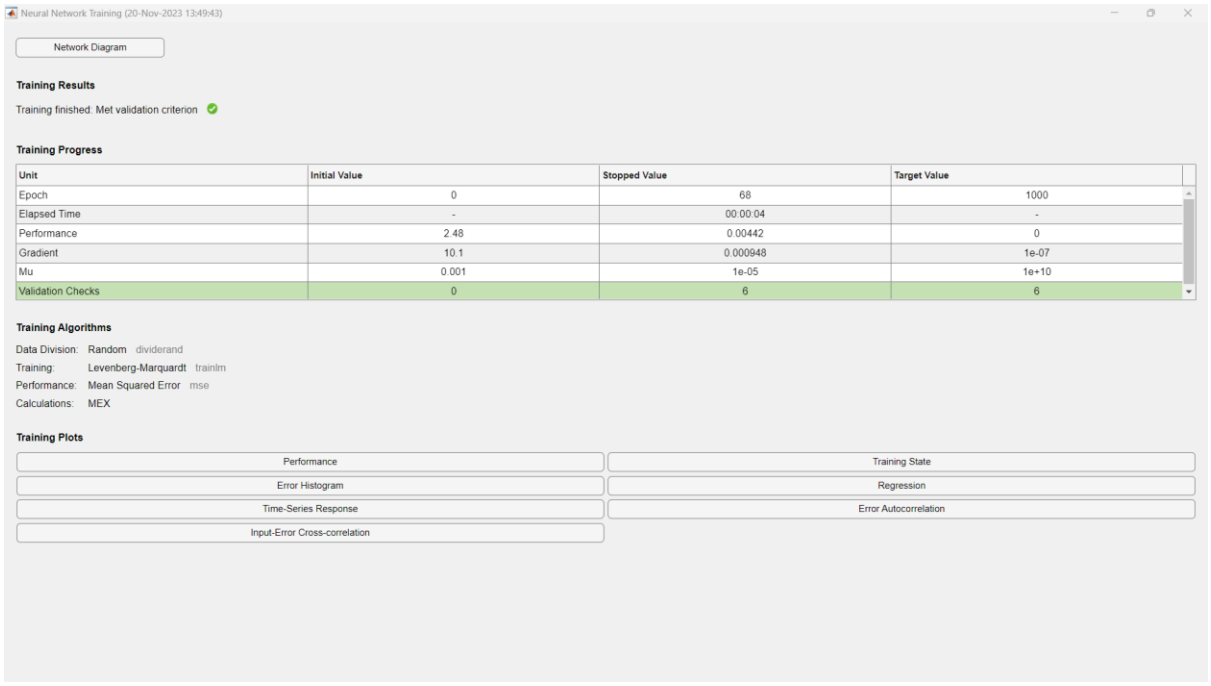

Fig.7

From Fig.8 and Fig.9 we can see that the prediction of the RNN Model is very accurate as compared to LSTM Model.

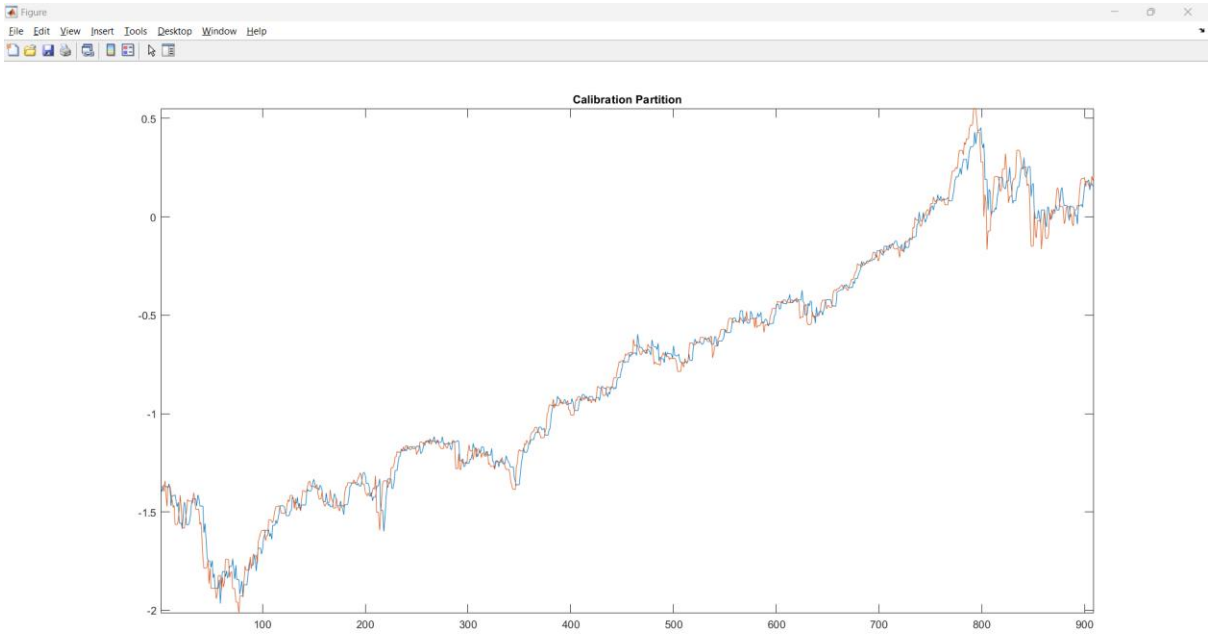

Fig.8

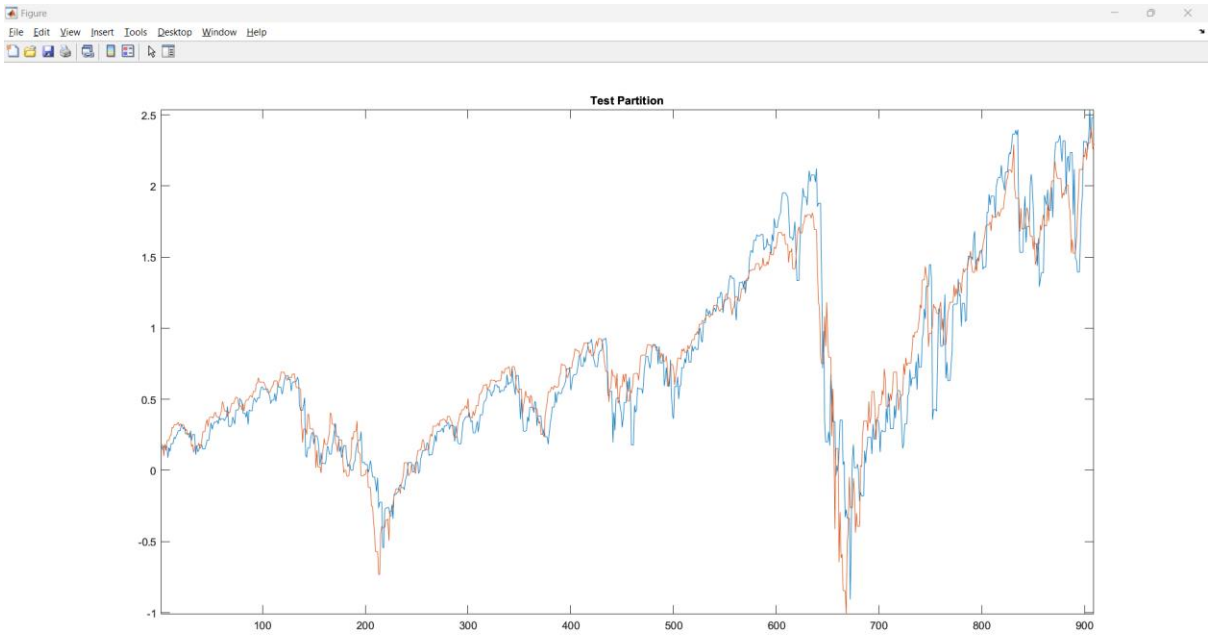

Fig.9

Conclusion:

Both the LSTM and RNN models demonstrate the viability of time-series forecasting, particularly when dealing with relatively stable datasets. Yahoo stock, in comparison to some other stocks, exhibits a generally stable pattern, with occasional deviations such as during the COVID outbreak and the recent economic crisis. It's essential to acknowledge that the inherent volatility in time-series data can introduce deviations in our neural network models.