Assignment V– Individual Take Home Assignment

Rahul S. Kalubowila

Department of Statistics, University of Colombo

IS 3005

2016S16026 | S13581

Contents

Acknowledgment 4

Executive Summary 5

Introduction 6

Outline 6

Data 6

Methodology 6

Objectives 6

Model Fitting Procedure for TS1 7

1. Stationarity 7

2. ACF and PACF 7

3. ARMA Model Results 8

4. Forecast & Performance 9

Model Fitting Procedure for TS2 9

1. Stationarity 9

2. ACF and PACF 10

3. ARMA (1,1) Model Results 10

4. Residual Analysis 11

5. Forecast & Performance 12

Comparison 15

Mean Squared Error (MSE) 15

Runtime 15

Conclusion 16

Acknowledgment

I would like to express my deepest appreciation to all those who provided me the possibility to complete this report. A special gratitude I give to the course lecturer, Dr. Nadeeka Basnayake, whose contribution in stimulating suggestions and encouragement, helped me to coordinate my project especially in writing this report. Furthermore, I would also like to acknowledge with much appreciation the crucial role of the staff who gave the permission to use all required equipment and the necessary materials to complete the task. I have to appreciate the guidance given by supervisors especially in our analysis that has improved our analysis skills and thank their comment and advices.

Executive Summary

Following is the executive summary from of 500 simulated time series data points to compare classical model fitting procedure & ANN model fitting procedure for time series data.

Factors like runtime and interpretability the classical models are better. But when comparing the RMSE values we found that LSTM performed better.

We can improve RMSE for classical models also if we utilized advance forecasting methods using the classical approach. But due to the time constraints basic forecasting was used for classical models.

Factors like amount of data, number of layers, Units in layers and number of iterations effects performance significantly for ANN models.

Introduction

Outline

The purpose of this study is to compare empirically classical model fitting procedure & ANN model fitting procedure for time series data. For this study we are simulating the data for our time series. Additionally, we are forecasting on validation data to make comparisons.

Data

- TS1 – ARMA (2,2) with AR coefficients 0.75, -0.25, and MA coefficients 0.65, 0.35 specification was used to simulate the time series.

- TS2 – Introduced 25 extreme observations to TS1 data series such that the number of observations in both the series are the same and the extreme values (1*np.random.randn(25)+10) are replaced in a random manner.

Methodology

Analysis was carried out using Python programming language.

Objectives

Classical model fitting procedure of TS1 & TS2

ANN model fitting procedure for TS1 and TS2

Compare RMSE (Root Mean Square Error)

Compare Runtime

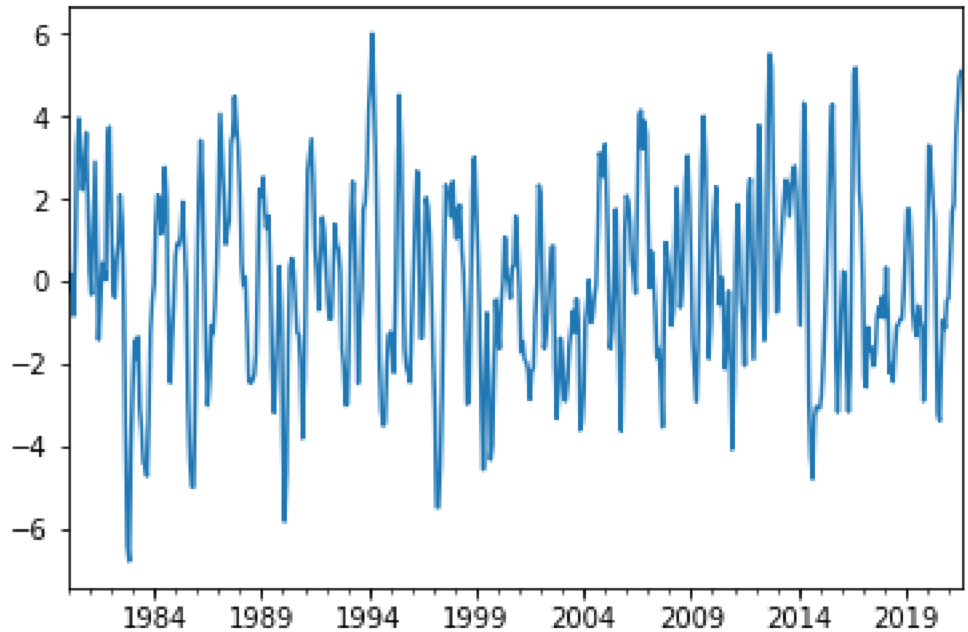
Compare Interpretability

Model Fitting Procedure for TS1

Classical Time Series Model Fitting for Known Artificial Data

1. Stationarity

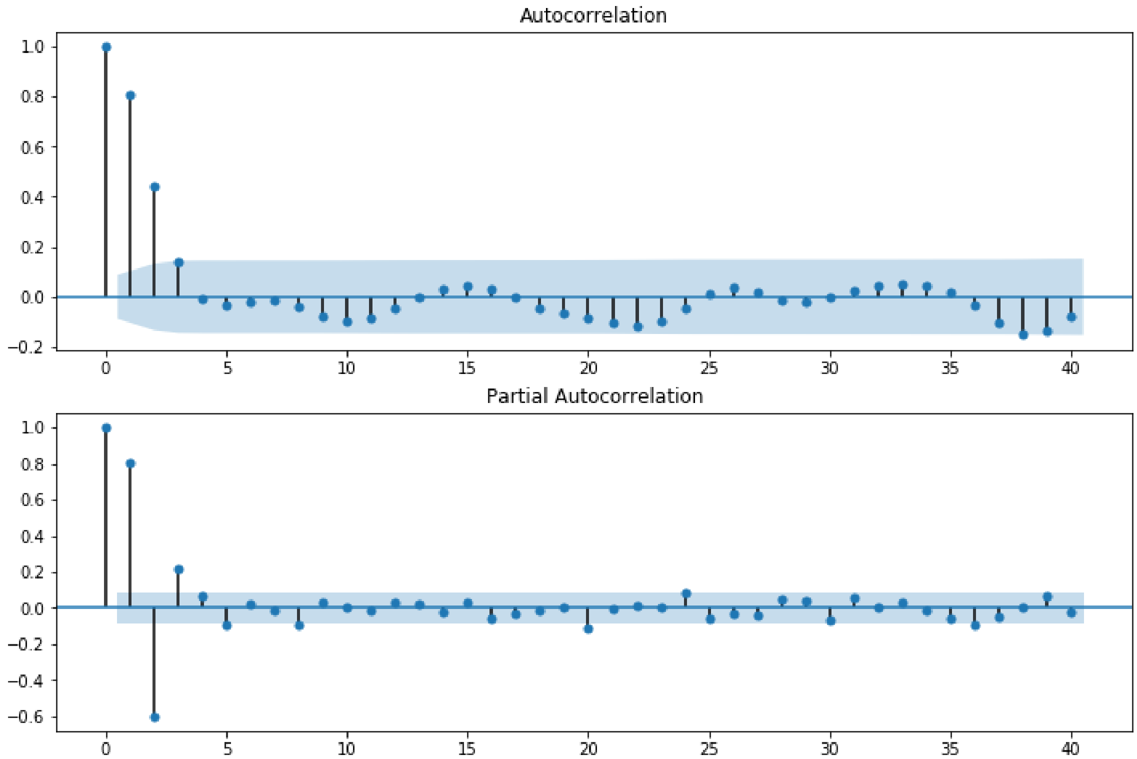
A stationary process has the property that the mean, variance and autocorrelation structure do not change over time.



Running the Augmented Dickey-Fuller unit root test on data we obtain a p-value of 4.5774564652911975e-12 is obtained through regression surface approximation from MacKinnon 1994, but using the updated 2010 tables. Thus, we can reject the null hypothesis and confirm stationarity.

2. ACF and PACF

The ACF and PACF should be considered together. It can sometimes be tricky going, but a few combined patterns do stand out. Looking at the following plots we can deduce the following.



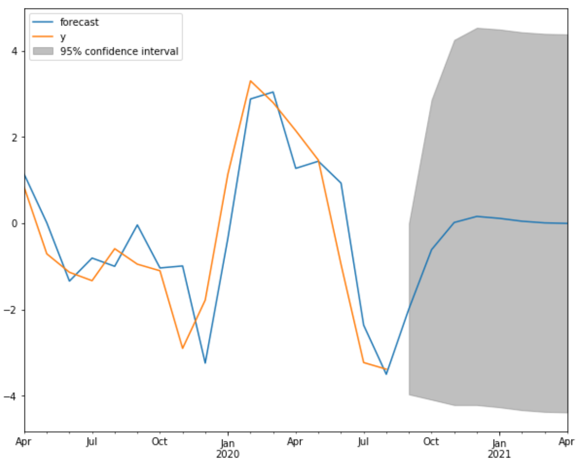
3. ARMA Model Results

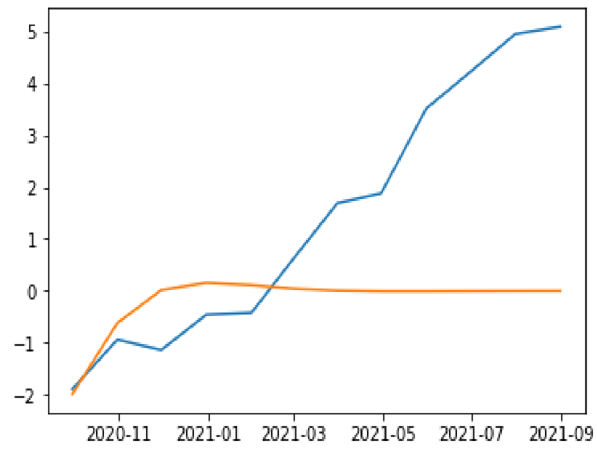
```
                              ARMA Model Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  488
Model:                     ARMA(2, 2)   Log Likelihood                -699.765
Method:                       css-mle   S.D. of innovations              1.013
Date:                Sun, 01 Mar 2020   AIC                           1409.530
Time:                        18:29:52   BIC                           1430.481
Sample:                    01-31-1980   HQIC                          1417.760
                         - 08-31-2020
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1.y        0.7352      0.110      6.702      0.000       0.520       0.950
ar.L2.y       -0.2357      0.088     -2.683      0.008      -0.408      -0.063
ma.L1.y        0.7038      0.106      6.653      0.000       0.496       0.911
ma.L2.y        0.3942      0.077      5.132      0.000       0.244       0.545
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.5599           -1.3454j            2.0599           -0.1133
AR.2            1.5599           +1.3454j            2.0599            0.1133
MA.1           -0.8926           -1.3191j            1.5927           -0.3447
MA.2           -0.8926           +1.3191j            1.5927            0.3447
-----------------------------------------------------------------------------
Runtime: 0.14106345176696777
```

It can be observed that the ARMA coefficients are very close to the coefficients we used for simulation.

4. Forecast & Performance

After the model was fitted on the training set. We predicted values on the remaining year that was taken as the validation set and compared results. We can see that ARMA model values converge at 0.

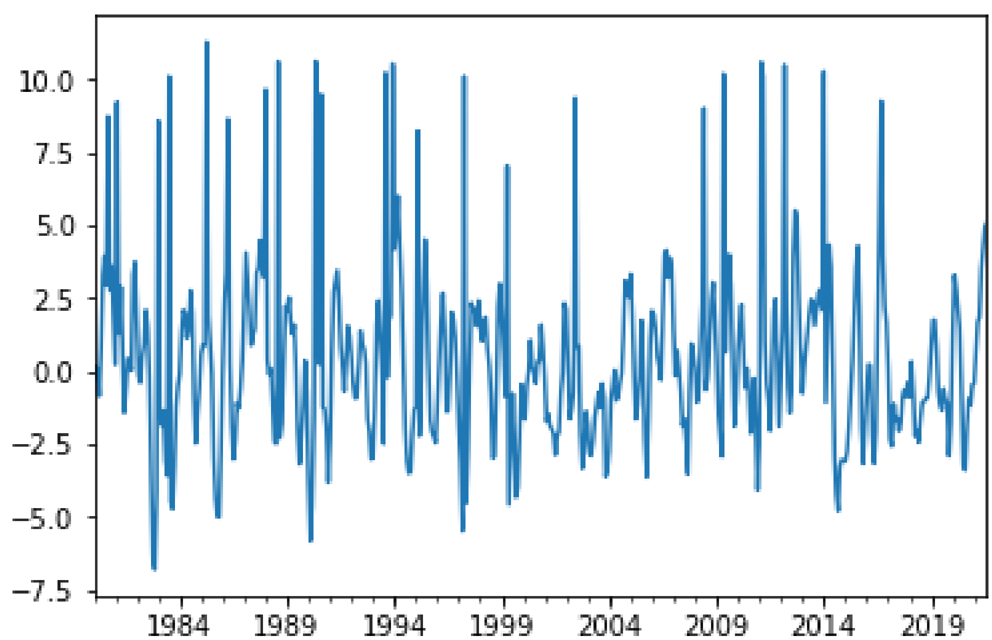




Model Fitting Procedure for TS2

Classical Time Series Model Fitting for Unknown Artificial Data

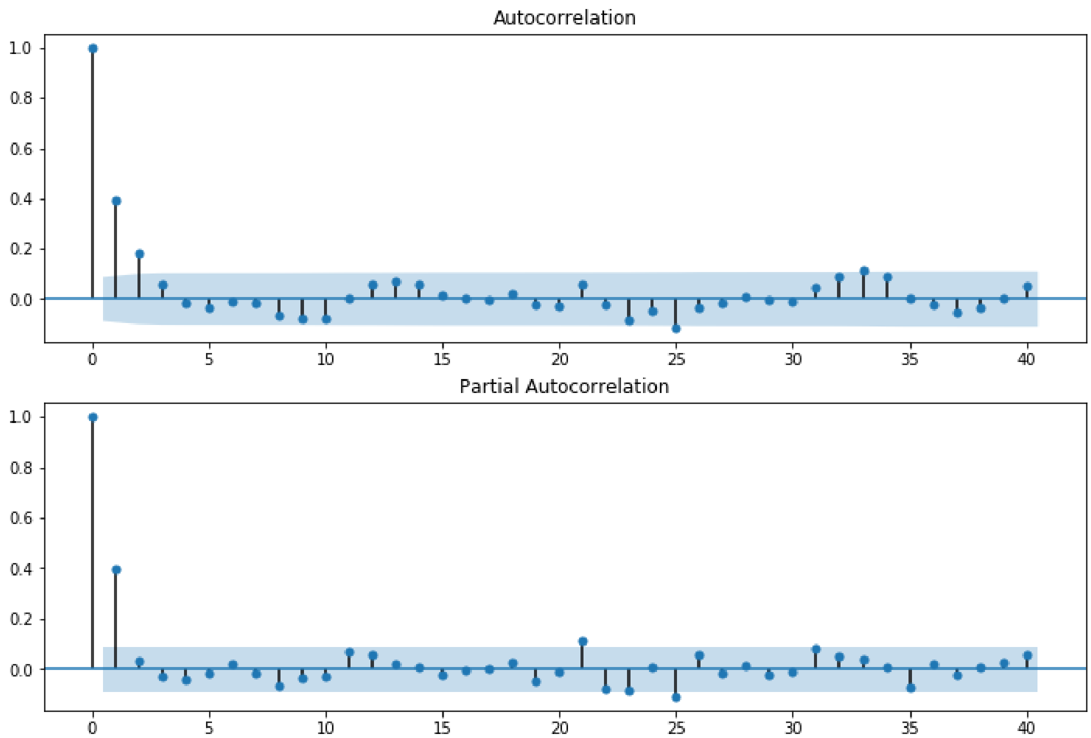
1. Stationarity



Running the Augmented Dickey-Fuller unit root test on data we obtain a p-value of 5.812867264889022e-26 is obtained through regression surface approximation from MacKinnon 1994, but using the updated 2010 tables. Thus, we can reject the null hypothesis and confirm stationarity.

2. ACF and PACF

The ACF and PACF should be considered together. It can sometimes be tricky going, but a few combined patterns do stand out. ARMA models (including both AR and MA terms) have ACFs and PACFs that both tail off to 0. These are the trickiest because the order will not be particularly obvious.



Since, we have high values in both ACF & PACF plots, we can assume the ARMA (1,1) to be a good fit. But further analysis is needed to confirm this hypothesis.

Afterwards we ran a loop trying different combinations to arrive to the best model. This model was also the ARMA (1,1) model. Thus, we used that model as the baseline for this analysis.

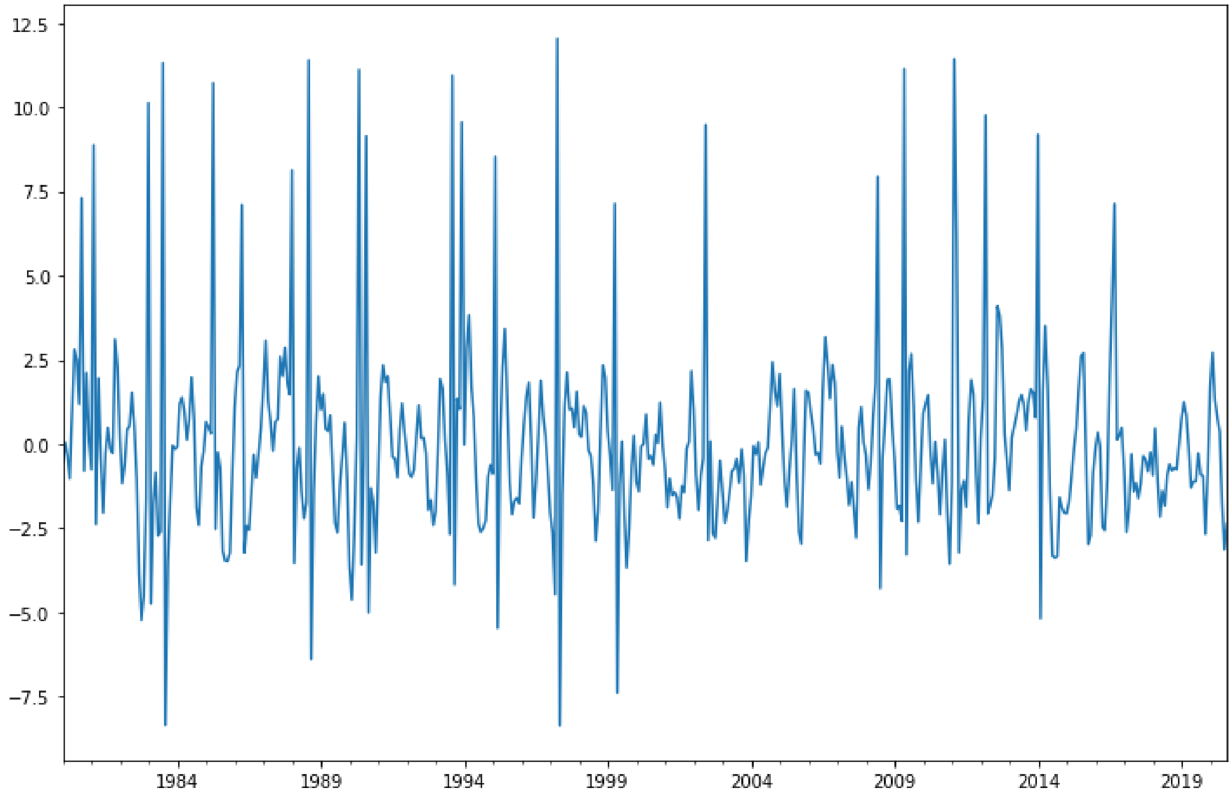
3. ARMA (1,1) Model Results

```
const       0.361525
ar.L1.y     0.446071
ma.L1.y    -0.072266
```

Above are the parameters for the model. Following are AIC, BIC & HQIC values for the model respectively.

```
2389.1571357425864 2405.918397365999 2395.74102405854
```

4. Residual Analysis



Residual Plot 1

Durbin Watson Test

The Durbin Watson Test is a measure of autocorrelation (also called serial correlation) in residuals from regression analysis.

Durbin Watson Test Statistic - 1.9993581572867287

Since, the value is very close to 2, we can say that there is no autocorrelation in the model.

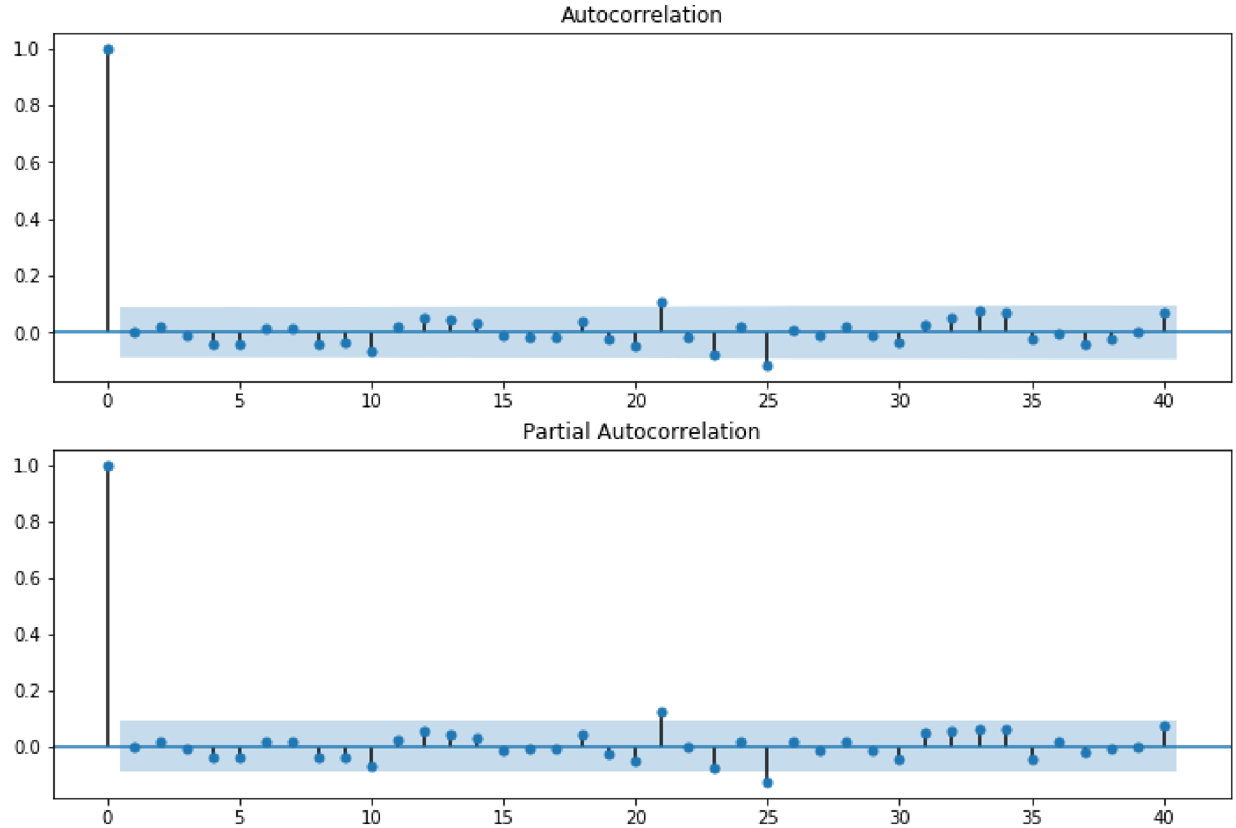
Ljung Box Test

It is a test of lack of fit: if the autocorrelations of the residuals are very small, we say that the model doesn't show 'significant lack of fit'.

```
Ljung Box Test Statistic for 20 lags - (array([11.49613543]), array([0.93232825]))
```

Hence, our model does not show lack of fit.

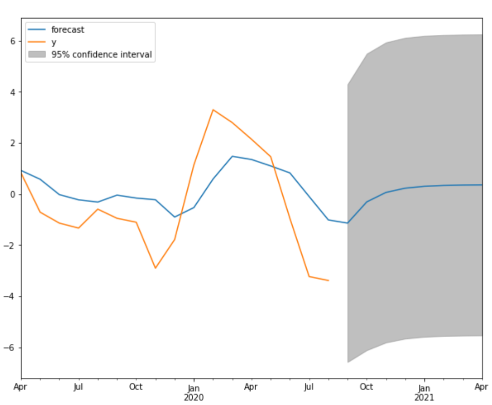
Residual ACF & PACF

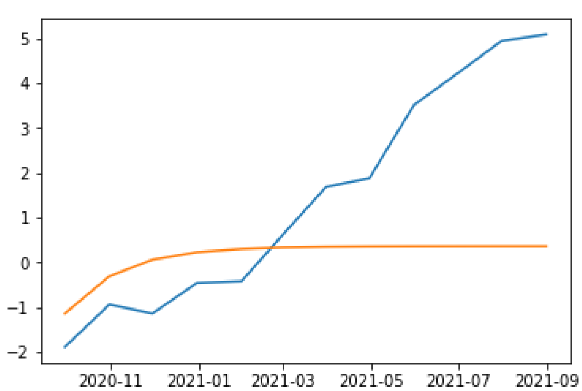


It can be observed that there ae no significant spikes in both ACF & PACF functions of residuals for our model. That's a good sign and proves our residuals are stationary.

5. Forecast & Performance

After the model was fitted on the training set. We predicted values on the remaining year that was taken as the validation set and compared results. We can see that ARMA model values converge at 0.



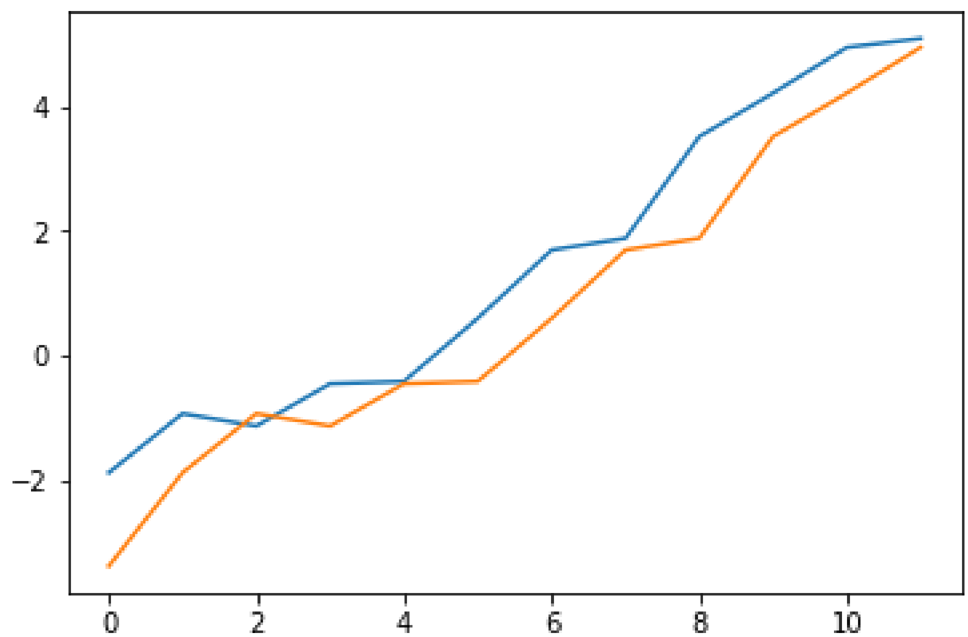


ANN model fitting procedure for TS1 and TS2

1. Persistence Model Forecast – Baseline Model

A good baseline forecast for a time series with a linear increasing trend is a persistence forecast.

The persistence forecast is where the observation from the prior time step (t-1) is used to predict the observation at the current time step (t).



2. Long Short-Term Memory Model

The Long Short-Term Memory network (LSTM) is a type of Recurrent Neural Network (RNN).

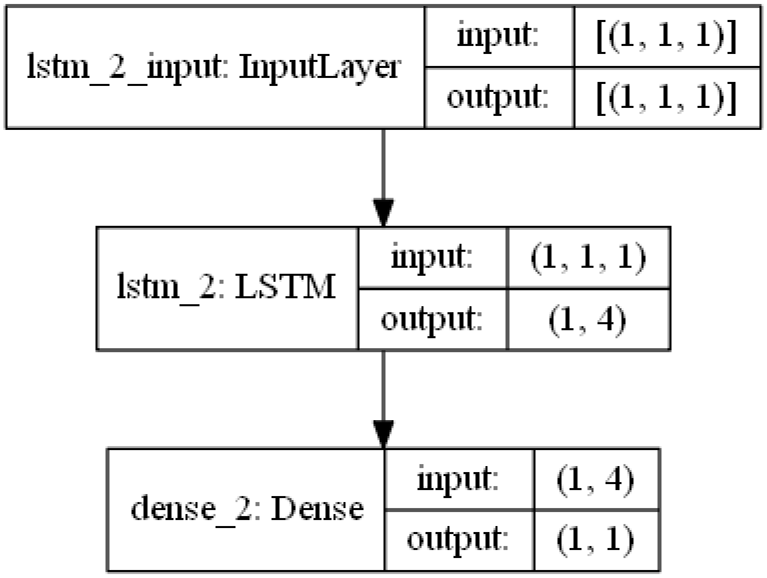
A benefit of this type of network is that it can learn and remember over long sequences and does not rely on a pre-specified window lagged observation as input

Before we can fit an LSTM model to the dataset, we must transform the data.

1. Transform the time series into a supervised learning problem
2. Transform the time series data so that it is stationary.
3. Transform the observations to have a specific scale.

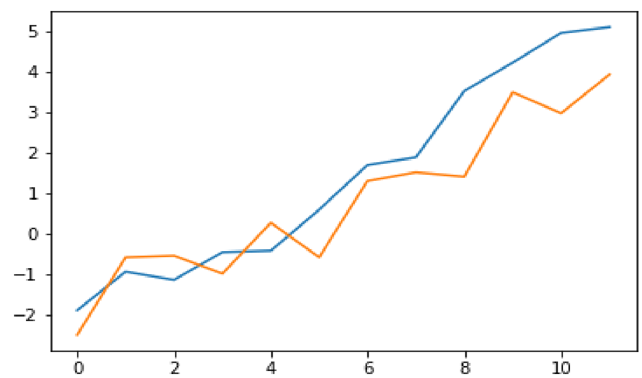
The required steps are done in the code attached. All functions are executed as separate python functions for interpretability. Following are the results from the 2 models.

For both time series data, we used a LSTM model with 2 dense layers with 4 neurons each and 20 iterations.



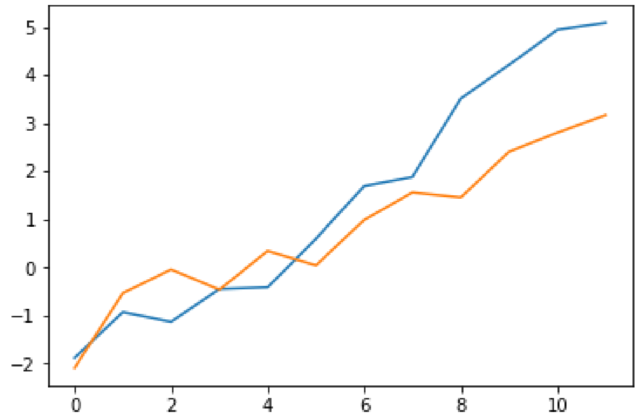
TS1

```
Month=1, Predicted=-2.445661, Expected=-1.887990
Month=2, Predicted=-0.607034, Expected=-0.936715
Month=3, Predicted=-0.460770, Expected=-1.137193
Month=4, Predicted=-0.941245, Expected=-0.456617
Month=5, Predicted=0.311034, Expected=-0.419927
Month=6, Predicted=-0.599351, Expected=0.591934
Month=7, Predicted=1.372153, Expected=1.689647
Month=8, Predicted=1.510329, Expected=1.877750
Month=9, Predicted=1.430099, Expected=3.512217
Month=10, Predicted=3.399665, Expected=4.209811
Month=11, Predicted=3.012022, Expected=4.944123
Month=12, Predicted=3.693534, Expected=5.087330
```



TS2

```
Month=1, Predicted=-1.712033, Expected=-1.887990
Month=2, Predicted=-0.600970, Expected=-0.936715
Month=3, Predicted=-0.247514, Expected=-1.137193
Month=4, Predicted=-0.574822, Expected=-0.456617
Month=5, Predicted=0.140141, Expected=-0.419927
Month=6, Predicted=-0.106984, Expected=0.591934
Month=7, Predicted=0.766190, Expected=1.689647
Month=8, Predicted=1.290435, Expected=1.877750
Month=9, Predicted=1.297783, Expected=3.512217
Month=10, Predicted=2.203537, Expected=4.209811
Month=11, Predicted=2.529807, Expected=4.944123
Month=12, Predicted=2.836607, Expected=5.087330
Test RMSE: 1.378
```



Comparison

	Classical		LSTM	
	TS1	TS2	TS1	TS2
RMSE	2.727	2.511	1.062	1.247
Runtime	0.1411 Seconds	0.0998 Seconds	22.9848 Seconds	22.6476 Seconds

Root Mean Squared Error (RMSE)

The root mean squared error tells you how close a regression line is to a set of points. It does this by taking the distances from the points to the regression line (these distances are the “errors”). We can see that LSTM model has lower RSME values for both the datasets.

Runtime

In computer science, runtime, run time or execution time is the time when the CPU is executing the machine code. Here, we want to see how much time is needed to run the algorithm and relate that to resource consumption. We can see that LSTM takes far more time compared to the classical models.

Interpretability

Interpretability is closely connected with the ability of users to understand the model. From the above two types of model we were able to run better and more insightful diagnostics on the classical models. Also, more resources and literature are published on the classical models as well. Thus, when it comes to interpretability classical models are superior.

Conclusion

In conclusion, for factors like runtime and interpretability the classical models are better. But when comparing the RMSE values we found that LSTM performed better.