Aston Business School
Birmingham

PORTFOLIO TASK 6: TIME SERIES USING ANN

INTRODUCTION

A UK/US exchange rate dataset is provided where the value of one pound sterling in US dollars is reported as a time series variable. The data set includes the daily exchange rate from January 4, 2010 until August 7, 2020. This study forecasts the exchange rate for August 8, 2020 using an artificial neural network in IBM SPSS.

Using available historical inputs, forecasting of one step ahead will be done. For a two step ahead forecasting, available historical input and the one-step forecast will be used. This process repeats until we forecast all the required values.

ANALYSIS OF DATA

For the data to be ready for analysis, ensure that there are missing values. In case of missing values, replace the blanks with the average of previous four values. In the same way, the missing values are handled in this case for the final dataset. Next, a line chart is plotted showing the values over time.

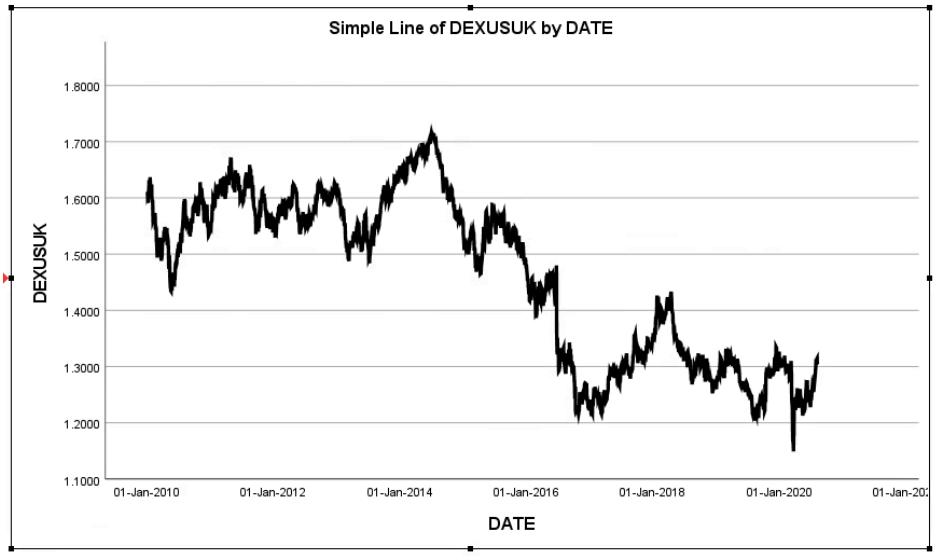


Fig: Line chart showing exchange rate values over time

From the graph, it can be seen that there is a downward trend, i.e., the data is non stationary.

For further analysis, ACF and PACF plots will be required.

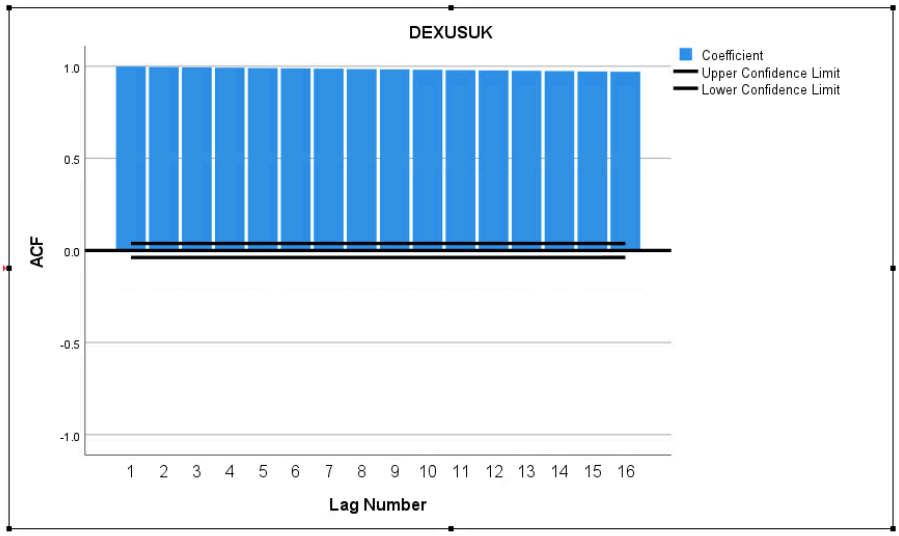


Fig: Image showing ACF plot of initial non-stationary data

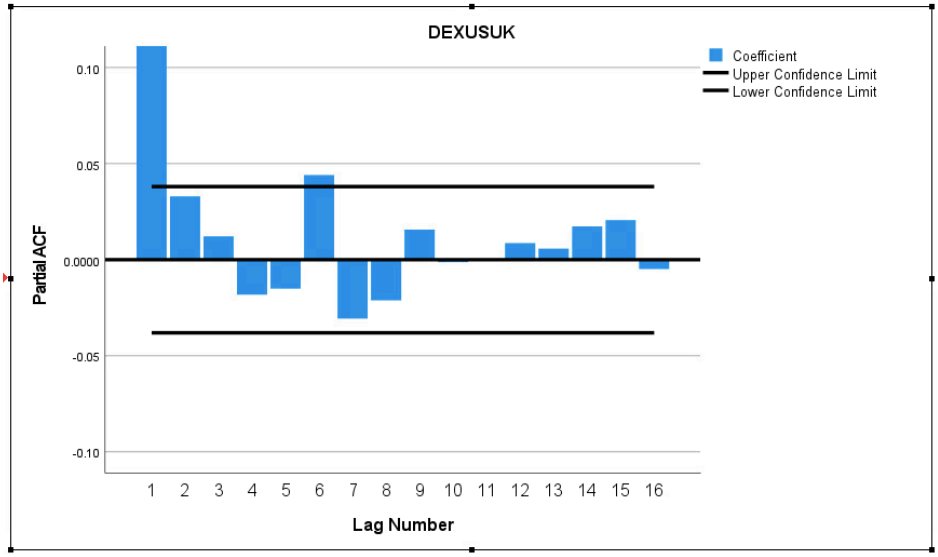


Fig: Image showing PACF plot of initial non-stationary data

There are 2 lags in the PACF plot, one at t=1 and the other at t=6.

The above information is for non-stationary data. It can be converted to stationary by differencing. The following results are obtained after first differencing.

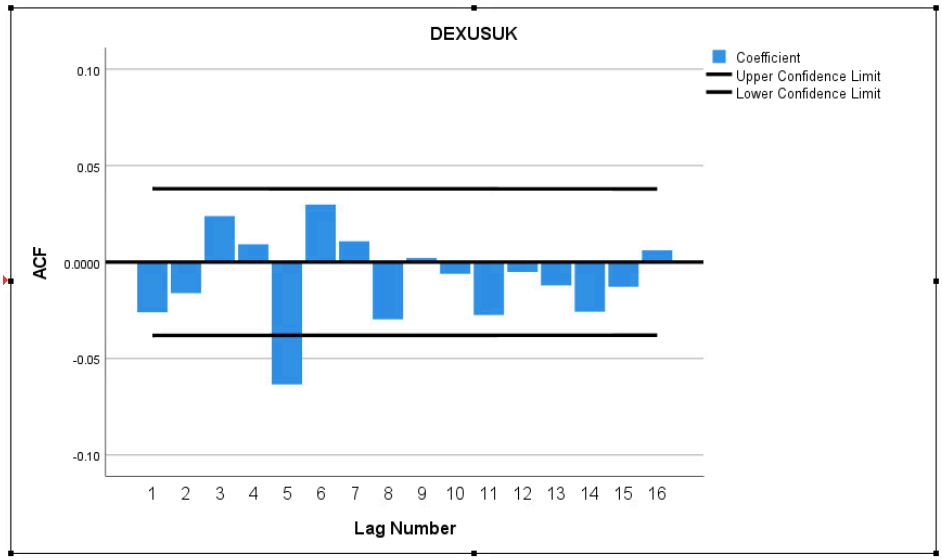


Fig: Image showing ACF plot after first differencing

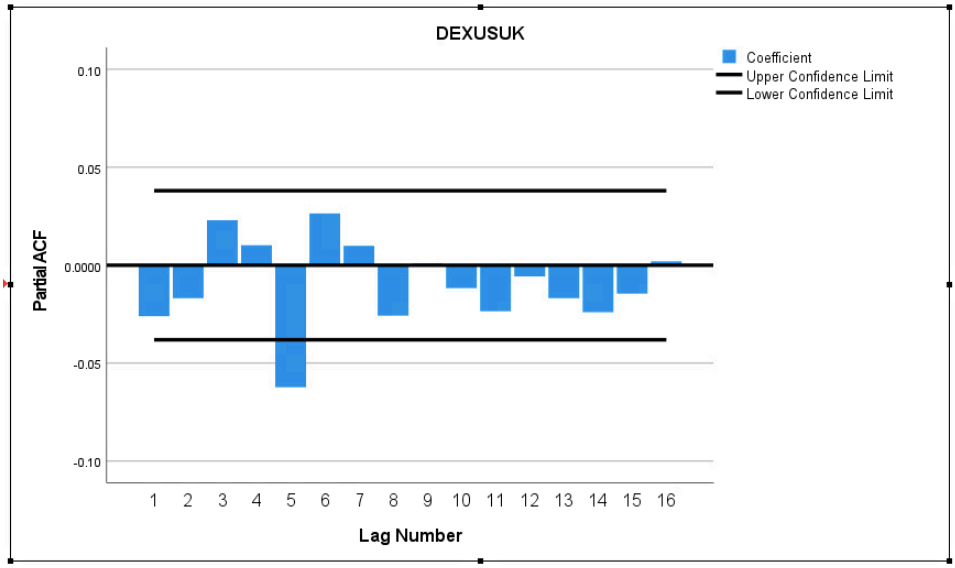


Fig: Image showing PACF plot after first differencing

It can be observed from the PACF Plot above that there is one significant lag at t=5. Therefore, 5 inputs can be fed to the neural networks. Now, create a time series with 6 variables in the data, Y(t), Y(t-1), Y(t-2), Y(t-3), Y(t-4) and Y(t-5). The following table showing lags at these respective times will be obtained.

	DATE	DEXUSUK	yt5	yt4	yt3	yt2	yt1	yt
1	04-Jan-10	1.6109	.	.	.	.	.	1.61
2	05-Jan-10	1.6009	.	.	.	.	1.61	1.60
3	06-Jan-10	1.6016	.	.	.	1.61	1.60	1.60
4	07-Jan-10	1.5912	.	.	1.61	1.60	1.60	1.59
5	08-Jan-10	1.5993	.	1.61	1.60	1.60	1.59	1.60
6	11-Jan-10	1.6146	1.61	1.60	1.60	1.59	1.60	1.61
7	12-Jan-10	1.6181	1.60	1.60	1.59	1.60	1.61	1.62
8	13-Jan-10	1.6288	1.60	1.59	1.60	1.61	1.62	1.63
9	14-Jan-10	1.6320	1.59	1.60	1.61	1.62	1.63	1.63
10	15-Jan-10	1.6240	1.60	1.61	1.62	1.63	1.63	1.62
11	18-Jan-10	1.6257	1.61	1.62	1.63	1.63	1.62	1.63
12	19-Jan-10	1.6370	1.62	1.63	1.63	1.62	1.63	1.64
13	20-Jan-10	1.6287	1.63	1.63	1.62	1.63	1.64	1.63
14	21-Jan-10	1.6214	1.63	1.62	1.63	1.64	1.63	1.62
15	22-Jan-10	1.6120	1.62	1.63	1.64	1.63	1.62	1.61
16	25-Jan-10	1.6236	1.63	1.64	1.63	1.62	1.61	1.62
17	26-Jan-10	1.6109	1.64	1.63	1.62	1.61	1.62	1.61
18	27-Jan-10	1.6204	1.63	1.62	1.61	1.62	1.61	1.62
19	28-Jan-10	1.6242	1.62	1.61	1.62	1.61	1.62	1.62
20	29-Jan-10	1.6009	1.61	1.62	1.61	1.62	1.62	1.60
21	01-Feb-10	1.5924	1.62	1.61	1.62	1.62	1.60	1.59
22	02-Feb-10	1.5968	1.61	1.62	1.62	1.60	1.59	1.60
23	03-Feb-10	1.5915	1.62	1.62	1.60	1.59	1.60	1.59
24	04-Feb-10	1.5763	1.62	1.60	1.59	1.60	1.59	1.58

Data View | Variable View

Fig: Image showing the final dataset for analysis

Neural Networks:

Analyzing the above data, using 'Multilayer Perceptron' of Neural Networks a model is built with Yt as the dependent variable.

Network Information

Input Layer	Covariates	1	yt-5
		2	yt-4
		3	yt-3
		4	yt-2
		5	yt-1
	Number of Units[a]		5
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		3
	Activation Function		Sigmoid
Output Layer	Dependent Variables	1	yt
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares

a. Excluding the bias unit

Fig: Image showing network information

Case Processing Summary

		N	Percent
Sample	Training	1363	49.4%
	Testing	726	26.3%
	Holdout	671	24.3%
Valid		2760	100.0%
Excluded		10	
Total		2770	

Fig: Image showing case processing summary

From the case process summary above it can be seen that 10 variables were excluded.

The following image shows the multilayer perceptron visualization. The hidden layer activation function is a Hyperbolic tangent.

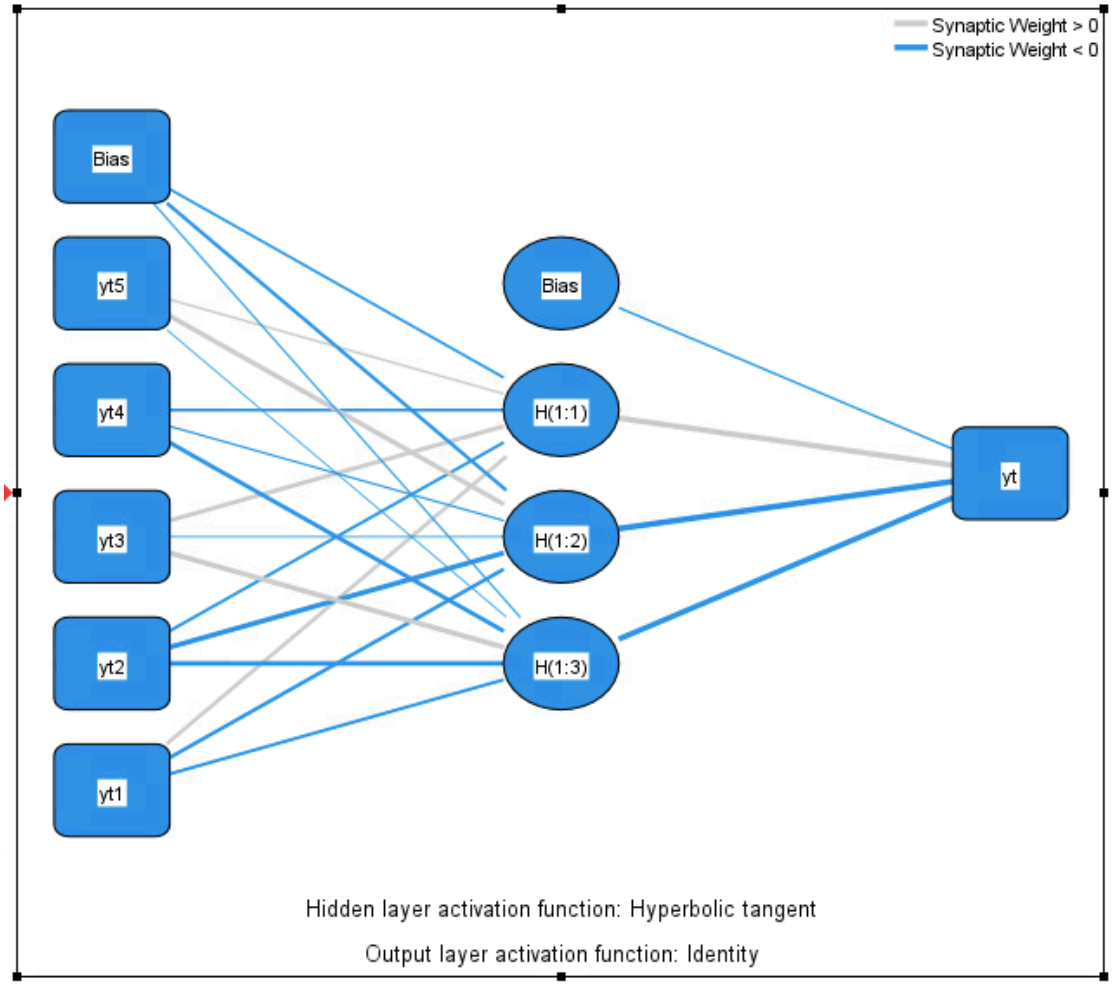


Fig: ANN model

Model Summary

Training	Sum of Squares Error	2.733
	Relative Error	.004
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.02
Testing	Sum of Squares Error	1.223
	Relative Error	.003
Holdout	Relative Error	.004

Dependent Variable: yt

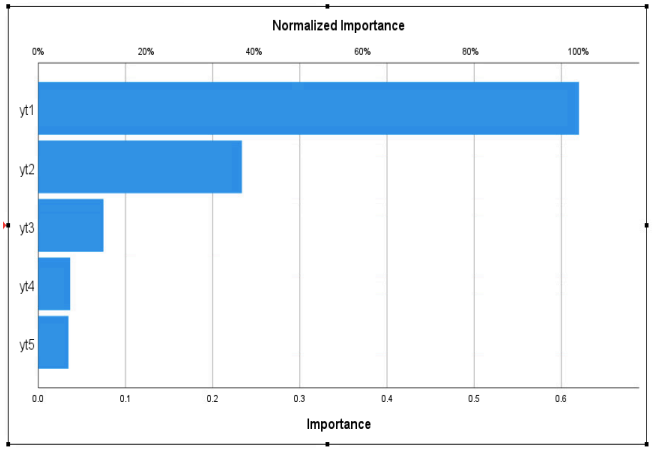
a. Error computations are based on the testing sample.

Parameter Estimates

Predictor		Predicted: Hidden Layer 1 H(1:1)	H(1:2)	H(1:3)	Output Layer yt
Input Layer	(Bias)	-.374	.117	-.383	
	yt5	.469	.056	.080	
	yt4	.001	.114	.060	
	yt3	.223	.303	-.304	
	yt2	-.282	-.144	-.250	
	yt1	.350	-.966	-.265	
Hidden Layer 1	(Bias)				1.756
	H(1:1)				1.350
	H(1:2)				-2.495
	H(1:3)				-2.448

Fig: Tables showing model summary and parameter estimates

The parameter estimates represent the values in the hidden layers of the neural networks.



Independent Variable Importance

	Importance	Normalized Importance
yt-5	.035	5.6%
yt-4	.037	5.9%
yt-3	.075	12.1%
yt-2	.234	37.7%
yt-1	.620	100.0%

Fig: Graph and table showing variable importance

Yt1 is the most important variable and Yt5 is the least important variable.

Forecast:

It can be observed that the model is good, since, the predicted and forecasted values coincide with each other in the graph below.

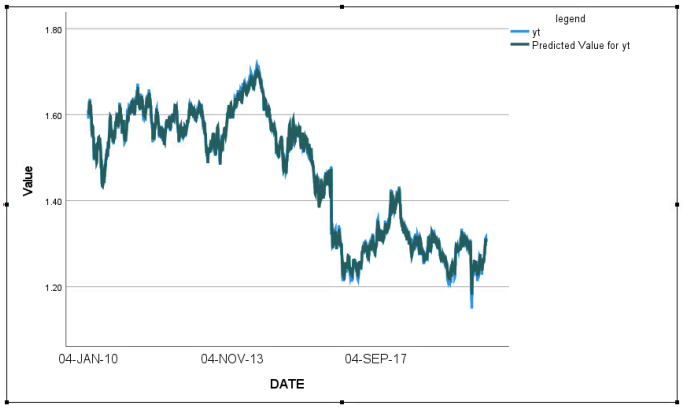
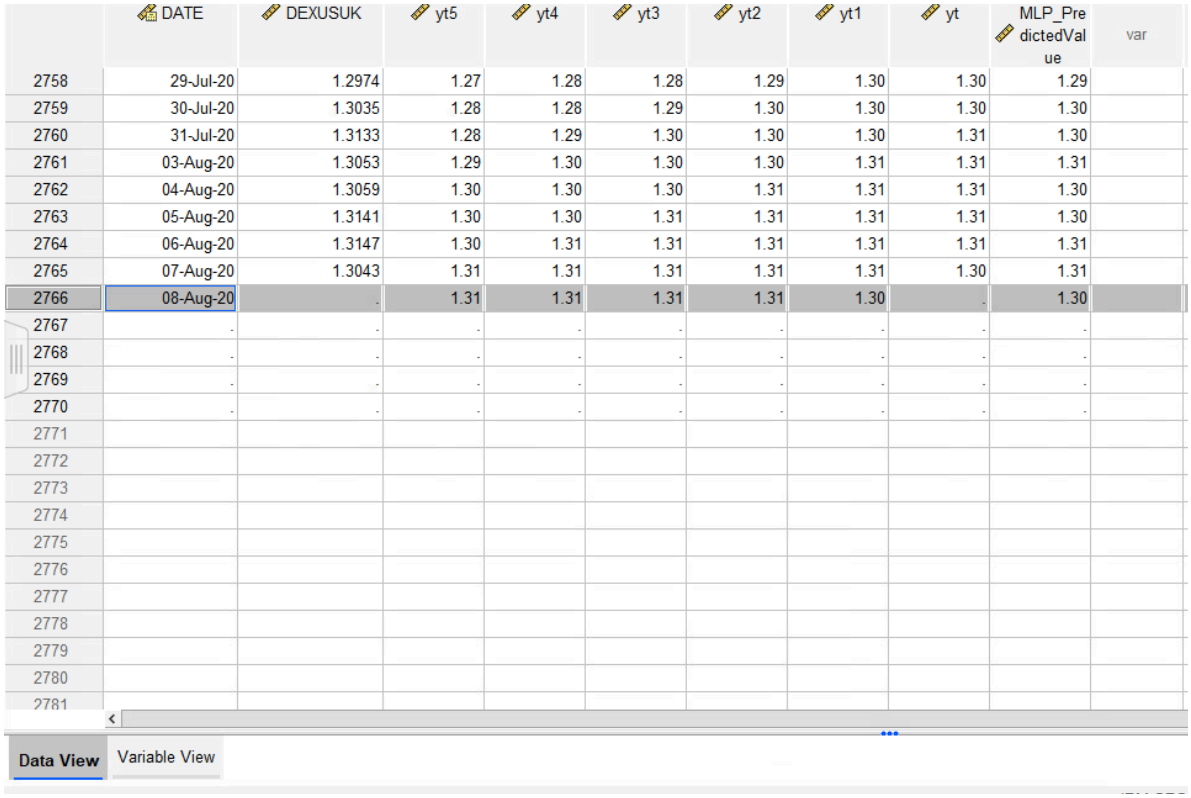


Fig: Graph depicting predicted and forecasted values



	DATE	DEXUSUK	yt5	yt4	yt3	yt2	yt1	yt	MLP_PredictedValue	var
2758	29-Jul-20	1.2974	1.27	1.28	1.28	1.29	1.30	1.30	1.29	
2759	30-Jul-20	1.3035	1.28	1.28	1.29	1.30	1.30	1.30	1.30	
2760	31-Jul-20	1.3133	1.28	1.29	1.30	1.30	1.30	1.31	1.30	
2761	03-Aug-20	1.3053	1.29	1.30	1.30	1.30	1.31	1.31	1.31	
2762	04-Aug-20	1.3059	1.30	1.30	1.30	1.31	1.31	1.31	1.30	
2763	05-Aug-20	1.3141	1.30	1.30	1.31	1.31	1.31	1.31	1.30	
2764	06-Aug-20	1.3147	1.30	1.31	1.31	1.31	1.31	1.31	1.31	
2765	07-Aug-20	1.3043	1.31	1.31	1.31	1.31	1.31	1.30	1.31	
2766	08-Aug-20	.	1.31	1.31	1.31	1.31	1.30	.	1.30	
2767	.	.	.	.	.	.	.	.	.	
2768	.	.	.	.	.	.	.	.	.	
2769	.	.	.	.	.	.	.	.	.	
2770	.	.	.	.	.	.	.	.	.	
2771										
2772										
2773										
2774										
2775										
2776										
2777										
2778										
2779										
2780										
2781										

Data View | Variable View

Fig: Image showing sav file with predicted values

The value for 08-Aug-20 is predicted as 1.30.

Conclusion:

Using ANN, a good model was obtained as evident from the graph. On comparing it with the original value from the website, it was found that the rate on 8th August was 1.3082 and the one obtained here is 1.30. Therefore, it can be concluded that the most parsimonious model is obtained.