AIRLINES PASSENGER FORECASTING USING LSTM BASED RECURRENT NEURAL NETWORKS

Varun Gupta, Kanupriya Sharma, Mohit Singh Sangwan

Abstract: *Investors take huge risk in making investments because they lack the information about the future prospects of the company or we can also say that they lack forecasting. Due to which they may have to face loss in future. Similar is the case with transportation, especially when it comes to airline mode because of huge increase in investment and risk factor. Thus, this paper proposes a system using Long Short Term Memory (LSTM) based Recurrent Neural Networks for forecasting the number of airline passengers. Unlike the existing schemes, the proposed scheme is capable enough to minimize the Root Mean Square Error to maximum extent which ultimately would lead to better forecasting results. The proposed scheme consist of four steps: (i) LSTM model construction, (ii) data processing and making data suitable for LSTM model, (iii) fitting a stateful LSTM network model to the training data and (iv) evaluation of the static LSTM model on the test data and report the performance of forecast. Furthermore, the obtained results indicate that the proposed scheme reduces the chances of false positive to great extent and is practical enough to be implemented in real-time scenario.*

Keywords: *Airline passenger forecasting; time series prediction; LSTM; recurrent neural networks; deep learning, artificial intelligence*

ITHEA Keywords: *I.2.6 [Artificial Intelligence]: Learning – Connectionism and neural nets*

Introduction

Time series modelling is a dynamic research area which has attracted attentions of researcher’s community over last few decades. The main aim of time series modelling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model is then used to generate future values for the series that is used to make forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the past [Raicharoen et al, 2003].

Now a days travelling is playing most important role in life of people because of which huge amount of investment is made by various people. Faster you travel from one place to another better it is for user and because of which airlines are major targets to investors. This paper deals with the airline dataset which gives the current and previous records of number of passengers travelled using airline. With help of the theses records this paper will analyse the number of passengers in the upcoming months and years. This information will be the key factor for investors to decide whether they should do investment in the airline business or not depending on the change in rate of the number of customers whether it's positive or negative.

For solving the above problem we have used LSTM with memory between the batches. LSTM is a type of Recurrent Neural Networks (RNN). The 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. The dataset used contains information of the number of passengers who had travelled through the airline in particular month and year. The dataset has been transformed to make it more suitable for LSTM model.

The rest of the paper is organized as follows. Section II presents the relative works which have been done in this field. Section III gives the brief description about the working of the proposed scheme. Section IV elaborates the methodology and implementation of the proposed scheme. The results and discussion are presented in Section V. The paper is finally concluded in Section VI.

Related Work

The model presented in the next section is the result of inspiration we have taken from prior works on prediction on airline passengers by various authors discussed below.

M.M. Mohie El-Din et al. [El-Din et al, 2017] used the back-propagation neural network and genetic algorithm to forecast the air passenger demand in Egypt (International and Domestic). The factors that influence air passenger were identified, evaluated and analyzed by applying the back-propagation neural network on the monthly data from 1970 to 2013. Lawrence R. Weatherford et al. [Weatherford et al, 2003] did comparative study of new method and traditional forecasting technique such as moving average, exponential smoothing, regression etc. All methods were compared on the basis of a standard error measure- mean absolute percentage error. Yukun Bao et al. [Bao et al, 2012] proposed an ensemble empirical mode decomposition (EEMD) based support vector machines (SVM) modeling framework incorporating a slope-based method to restrain the end effect issue occurring during the shifting process of EEMD. Along with above research papers, one of the most popular and frequently used stochastic time series models is the Autoregressive Integrated Moving Average (ARIMA) model by G.P. Zhang [Zhang, 2003] [Zhang, 2007], K.W. Hipel and McLeod [Hipel and McLeod, 1994]. The basic assumption made to implement this model is that the considered time series is linear and follows a particular known statistical distribution, such as the normal distribution. ARIMA model has subclasses of other models, such as the Autoregressive (AR) given by J. Lee [Lee], G.E.P. Box et al. [Box et al, 1970], Moving Average (MA) by K.W. Hipel and McLeod [Hipel and McLeod, 1994] and Autoregressive Moving Average (ARMA) by John H. Cochrane et al. [Cochrane, 1997] models.

For seasonal time series forecasting, Box and Jenkins [Box et al, 1970] had proposed a quite successful variation of ARIMA model, viz. the Seasonal ARIMA (SARIMA) by K.W. Hipel and McLeod [Hipel and McLeod, 1994], Box and Jenkins [Box et al, 1970]. The popularity of the ARIMA model is mainly due to its flexibility to represent several varieties of time series with simplicity as well as the associated Box-Jenkins methodology by G.P. Zhang [Zhang, 2003], K.W. Hipel and McLeod [Hipel and McLeod, 1994] for optimal model building process. But the severe limitation of these models is the pre-assumed linear form of the associated time series which becomes inadequate in many practical situations.

Proposed Scheme

This section presents the proposed scheme for forecasting of airlines passengers which consists of a sequential layer, two LSTM layers and one dense layer. Figure 1 shows the architecture of the proposed network.

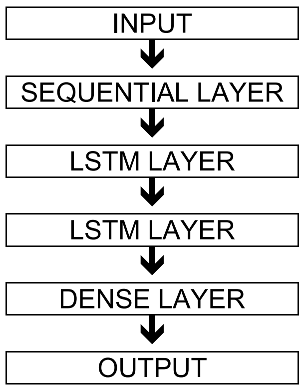

Figure 1. Proposed Scheme

SEQUENTIAL LAYER:

It is used to create a sequential model. It acts as a linear stack of layers. All the other layers like LSTM and Dense are added to it to create a model.

LSTM LAYER:

Long Short Term Memory networks – usually just called “LSTMs” – a special kind of RNN, capable of learning long term dependencies in the given time-series data. They were first introduced by Hochreiter & Schmidhuber in 1997 [Hochreiter and Schmidhuber, 1997 a]. [Hochreiter and Schmidhuber, 1997 b]. LSTMs are explicitly designed to solve the long-term dependency problem. All recurrent neural networks have the form of a chain of repeating modules of neural networks. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer. Figure 2 shows simple RNN architecture.

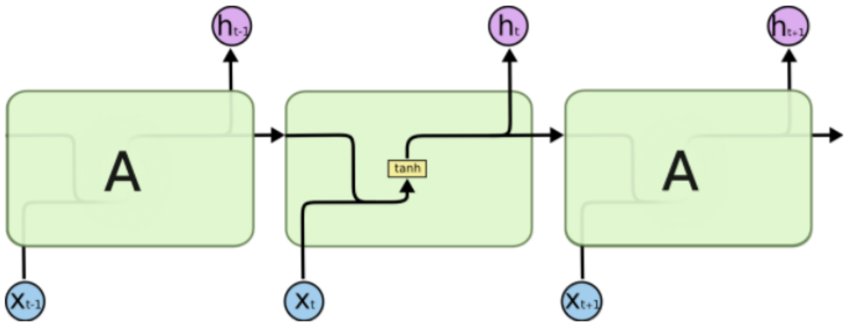

Figure 2. Simple RNN Architecure [Cohah 2015]

LSTMs also have this chain like structure, but the repeating module has a different structure. Instead of having a single neural network layer, there are four, interacting in a very special way. Figure 3 depicts LSTM based RNN architecture.

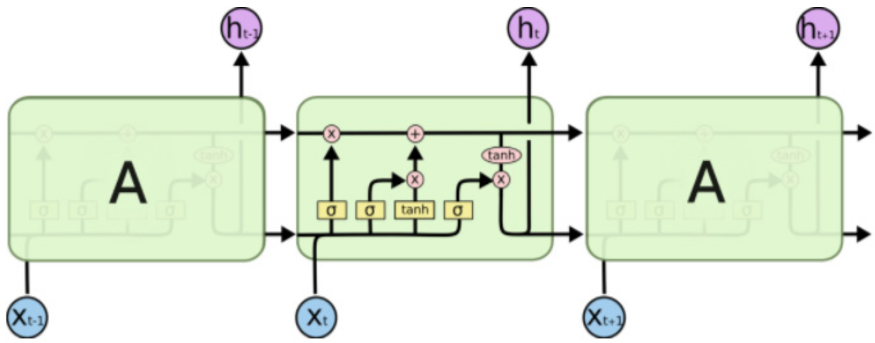

Figure 3. LSTM based RNN Architecture [Cohah 2015]

DENSE LAYER:

A dense layer is simply a layer where each unit or neuron is connected to each neuron in the next layer.

Methodology and Implementation

The steps taken to implement the proposed scheme were:

1. The dataset “international-airlines-passenger.csv” was loaded using pandas read_csv() function.

2. The data was normalised using MinMaxscaler() function of sklearn module .
3. The dataset was divided into two parts: one for training (67%) and the other for testing (33%) and the divided data chunks were reshaped so that they could be feed into the model.
4. The model was trained and fitted.
5. Results obtained for prediction on the training and testing dataset.

Results and Discussions

The results obtained after implementation of the proposed scheme are summarised using a graph shown in Figure 4 which shows how accurately the proposed model predicts the actual data points.

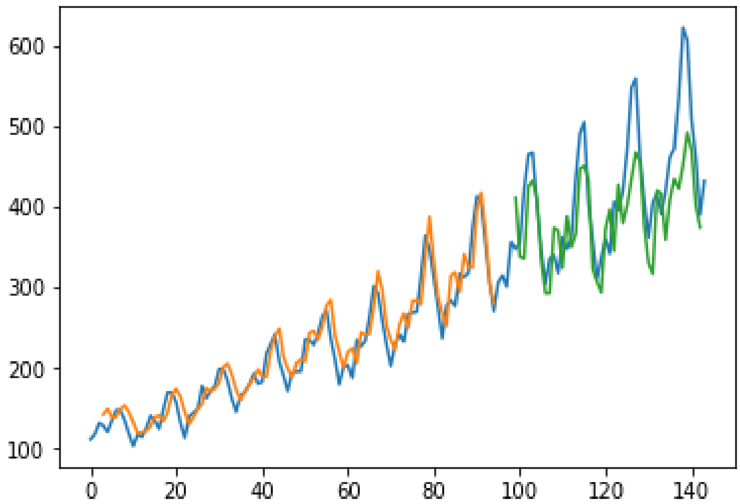

Figure 4. Actual Vs Predicted Results

The keys to interpret the graph shown in Figure 4 are presented in Table 1.

Table 1. Keys for Figure 4

COLOUR	DATA POINTS
BLUE	ACTUAL DATA POINTS
ORANGE	PREDICTED PONTS OF TRAINING DATA
GREEN	PREDICTED POINTS OF TESTING DATA

Figure 5 shows the MAPE (Mean Absolute Percentage Error) during training and testing. The error obtained during training phase was 8.79% and the error obtained during testing phase was 10.8%.

Figure 5. Error during training and testing phase

Comparison with the existing schemes

This segment elaborates the comparison of the proposed scheme with the existing schemes. The comparison results reveal that the proposed scheme is the most effective scheme relative to others [Weatherford et al, 2003].

Table 2. Comparison with the existing schemes

Reference	Model used	MAPE
Larry R Weatherford et al. [Weatherford et al, 2003]	Single Layer Multiple Inputs	Train : 62% Test : 63%

Lawrence R. Weatherford et al. [Weatherford et al, 2003]	Holt-Winters	7.4413%
M.M. Mohie El-Din et al. [El-Din et al, 2017]	Back-propagation neural network and genetic algorithm	NOT REPORTED
Proposed Scheme	Two Layer LSTM	Test : 8.79% Train : 10.8%

In reference [Weatherford et al, 2003], the MAPE (Mean Absolute Percentage Error) value of the best trained neural network is 62% for training data and 63% for testing data. In reference [Weatherford et al, 2003], the MAPE value is 7.4413% because of larger train dataset. While the MAPE value of the proposed model is 8.79% for training data and 10.8% for testing data.

Thus, the proposed model performs better than the other existing schemes for forecasting of airline passenger.

Conclusion and Future Scope

The forecasting of airline passengers is crucial to strategic decision making by the promoters and investors of airlines. In this paper, we have proposed a system for forecasting of number of airline passengers using LSTM based recurrent neural networks. Results obtained indicate that the proposed scheme performs better than the other existing schemes.

In future, the proposed work can further be extended by incorporating convolution layers in the architecture and by trying to make the model even deeper using multiple LSTM layers.

Bibliography

[Raicharoen et al, 2003] T. Raicharoen, C. Lursinsap, P. Sanguanbhoki, Application of critical support vector machine to time series prediction, Circuits and Systems, 2003. In: ISCAS '03.Proceedings of the 2003 International Symposium on Vol. 5, 25-28 May, 2003. pp. V-741-V-744

[Zhang, 2003] G.P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neurocomputing, Vol. 50, 2003, pp. 159–175.

[Zhang, 2007] G.P. Zhang, A neural network ensemble method with jittered training data for time series forecasting. Information Sciences, Vol. 177 Issue 23, 2007. pp. 5329–5346.

[Cochrane, 1997] John H. Cochrane, Time Series for Macroeconomics and Finance, Graduate School of Business, University of Chicago, Spring 1997.

[Hipel and McLeod, 1994] K.W. Hipel, A.I. McLeod, Time Series Modelling of Water Resources and Environmental Systems. Amsterdam, Vol. 45, 1st Edition, Elsevier 1994.

[Lee] J. Lee, Univariate time series modeling and forecasting (Box-Jenkins Method), Econ 413, lecture 4.

[Box et al, 1970] G.E.P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, CA, 1970.

[Hochreiter and Schmidhuber, 1997 a] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation, Vol. 9, Issue 8, 1997, pp.1735–80.

[Hochreiter and Schmidhuber, 1997 b] Sepp Hochreiter, Jürgen Schmidhuber, Neural Computation, Vol. 9, Issue 8, 1997, pp. 1735-1780.

[El-Din et al, 2017] M. M. Mohie El-Din, M. S. Farag and A. A. Abouzeid, Airline Passenger Forecasting in EGYPT (Domestic and International), International Journal of Computer Application (0975-8887), Vol. 165, Issue.6, May 2017.

[Bao et al, 2012] Yukun Bao, Tao Xiong, and Zhongyi Hu, Forecasting Air Passenger Traffic by Support Vector Machines with Ensemble Empirical

Mode Decomposition and Slope-Based Method, Discrete Dynamics in Nature and Society, Vol. 2012, 2012, pp. 1-12.

[Cohah 2015] Cohah's Blog: Understanding LSTM Networks. Accessed from http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[Weatherford et al, 2003] Larry R Weatherford, Boddan Wilamowski and T W Gentry, Neural Network forecasting for airlines: A comparative analysis, Journal of Revenue and Pricing Management, Volume 1, Issue 4, 2003, pp 319–331.

Authors' Information

	***Varun Gupta*** *– Associate Professor in Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India;* *e-mail: varun3dec@gmail.com* ***Major Fields of Scientific Research****: Deep learning, Machine learning, Software Engineering*
	***Kanupriya Sharma****- Student in Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India;* *e-mail: kanupriyasharma18@yahoo.in*
	***Mohit Singh Sangwan****- Student in Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India;* *e-mail: rtr.mohitsangwan@gmail.com*