PREDICTION OF TOURISM FLOW IN INDIA USING LSTM MODEL

[1] Dipali Koshti,[2]Supriya Kamoji,[3]Scarlet Lopes
[1]Assistant Professor,[2]Assistant Professor,[3]Student
[1]Department of Computer Engineering,
[1]Fr. Conceicao Rodrigues College of Engineering, Mumbai, India

Abstract: In today's world, travel and tourism is one of the major sources of foreign exchange for both developed and developing economies. According to the information given in statista.com, approximate contribution of this industry in the world's GDP is 8 to 10 percent. Therefore, increasing the revenue generated from tourism in countries has become a serious concerned for most of the countries, especially for developing countries like India. The proposed system predicts the tourist arrival in India from 56 different countries. Given the name of the county, the year and the month, it predicts percentage of tourist expected to arrive in India from the given country. With the current advancement in the neural networks, Deep learning has proven its results for vast number of applications. We have compared two models namely Support Vector Regression and Long Short Term Memory for predicting tourism flow in India. Our experimental results show that LSTM model performs better than SVR model for time series data.

Index Terms **-** Tourism prediction, Long Short Term Memory, Support Vector Regression, Deep learning Model

1. INTRODUCTION

Tourism is a leading industry worldwide, representing approximately 10 percent of the worldwide GDP according to the World Travel & Tourism Council's tourism satellite account method. Furthermore, increasing tourism also increases employment directly or indirectly which is one of the major concerned in the growing country's economy. Tourism represents a cross sector (umbrella) industry, including many related economic sectors such as culture, sports, and agriculture, where over 30 different industrial components have been identified that serve travelers. In addition, tourism greatly influences regional development, owing to its SME (small- and medium-sized enterprises) structure and relatively small entrance barriers.

2. LITERATURE SURVEY

Autoregressive Integrated Moving Averages ARIMA have been popularly used for solving forecasting problems. The authors (A. Ariyo, Adewumi, Ayo, 2014) used ARIMA to forecast stock prices from the historical daily stock prices obtained from two different countries stock exchange namely New York Stock Exchange (NYSE) and Nigeria Stock Exchange (NSE). They used Nokia stock data covering the period from 25th April, 1995 to 25th February, 2011. The data set contained a total number of 3990 observations. The data included four elements namely: open price, low price, high price and close price. The stock data of Zenith bank was used in this study covered the period from 3rd January, 2006 to 25th February, 2011 with total of 1296 observations. The author suggested the best values of (p,d,q) for both the data sets. Biljana Petrevska (2017) suggested the use of Box–Jenkins methodology (ARIMA) for forecasting international tourism in F.Y.R Macedonia and analyzed the trends and patterns of the tourist arrival using ARIMA. Choden and Suntaree Unhapipat (2018) modelled the monthly international visitors in Bumthang, Bhutan from January 2012 to December 2016 using SARIMA, the seasonal ARIMA model. The drawback of ARIMA model is that it is linear statistic time series model, which may not fit well for out of sample predictability tests due to model uncertainty and parameter instability. (Z. Hu, J. Zhu and K. Tse, 2013)

Paula Fernandes and Tiexeira (2017) used artificial neural networks for analyzing the tourism time series: "Monthly Guest Nights in Hotéis" in Northem Portugal recorded between January 1987 and December 2006. Authors used the model with 4 neurons in the hidden layer with the logistic activation function and was trained using the Resilient Backpropagation algorithm. B. Gui, X. Wei, Q. Shen, J. Qi and L. Guo (2014) combined support vector machine with Fuzzy logic for forecasting financial time series market.

In 1997 Sepp Hochreiter and Jurgen Schmidhuber introduced Long Short Term Memory neural networks to solves complex, artificial long time lag tasks that have never been solved by previous recurrent network algorithms. According to Yanjie Duan, Yisheng Lv, Fei-Yue Wang (2016) Deep learning models considering sequence relation are promising in traffic series data prediction. They used LSTM for predicting travel time for the travelers based on the data provided by London Highway.

YiFei Li1, Han Cao (2018) have predicted the flow of tourism for travel data obtained from wild goose pagoda which cover the travel data from January 2013 to December 2015, provided by Xi'an Museum. Authors compared performance of two LSTM model namely Simplex LSTM and stacked LSTM.

The rest of this paper is organized as follows. Section III presents the theoretical description of the methods we used. Section IV presents the two models SVR and LSTM that we used for prediction. Section V discusses the experimental results and performance of the two models. Finally, we conclude the paper in section VI.

3. THEORETICAL FRAMEWORK

We have implemented two models for prediction of tourism flow. The first model uses Support Vector Regression for prediction and the second is LSTM based model.

3.1 Model 1: Support Vector Regression based model

Support vector machine (SVM) is a machine learning algorithm which is used for solving classification and regression problems. SVM was first introduced by Vladimir Naumovich Vapnik and his colleagues in 1992. SVM regression is considered a nonparametric technique because it relies on kernel functions. It uses the maximum margin algorithm: a non-linear function is leaned by linear learning machine mapping into high dimensional kernel induced feature space.

3.2 Model 2: LSTM based Model

First we discuss the Recurrent Neural Networks (RNN), the predecessor of LSTM. Recurrent Neural Networks (RNN), is the predecessor of LSTM. A usual **RNN** has a short-term memory. In combination with **LSTM** they also have a long-term memory, which is discussed in detail below.

3.2.1 Recurrent Neural Network:

The major drawback of traditional neural networks is they cannot relate the present input with the previous input. Recurrent neural networks introduce the concept of internal memory to address this issue. Because of this internal memory, RNN can store the previous information and can predict what is coming next in a sequential data. Thus, RNN is suitable for the data that is sequential in nature and have temporal dynamics. As shown in Fig. 1, in RNN information cycles through a loop. It takes into consideration the current input and also what it has learned from the previous inputs while making any decisions.

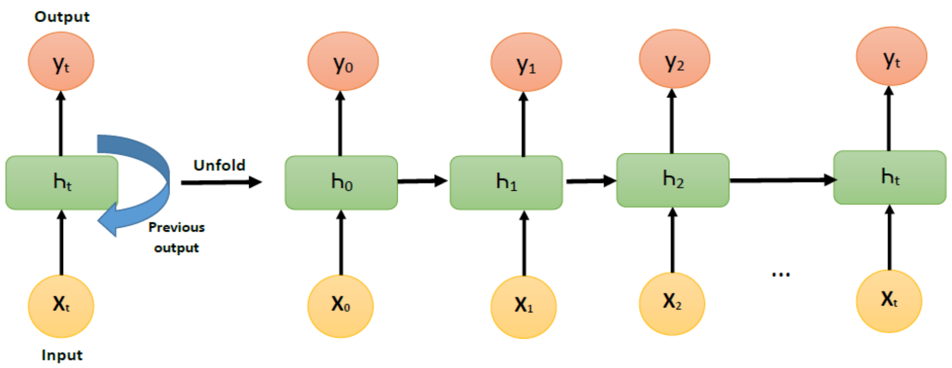


Fig. 1 Architecture of RNN

Fig. 1 shows a simple RNN with one input unit, one output unit, and one recurrent hidden unit when unfolded into a full network. Here x_t is the input, y_t is the output and h_t is the hidden state at time step t. Here h_t, is the 'memory' of the network. One problem with Recurrent neural networks is, they work fine when dealing with short-term dependencies but fail when applied to long term dependencies. This is because RNN suffers from a vanishing gradient problem. Vanishing Gradient problem occurs when gradients become too small and the model stops learning. RNN can remember things for just small durations of time. This issue can be resolved by applying a slightly tweaked version of RNNs – the Long Short-Term Memory Networks.

3.2.2 LSTM Network

LSTM was first introduced by Sepp Hochreiter and Jurgen Schmidhuber (1997), which solves the vanishing gradient problem of RNN. LSTM is an extension of RNN which basically extends their memory. Because of the extended memory, LSTM can remember their inputs over a long period of time. This makes them more suitable for the sequences that have very long time lags in between.

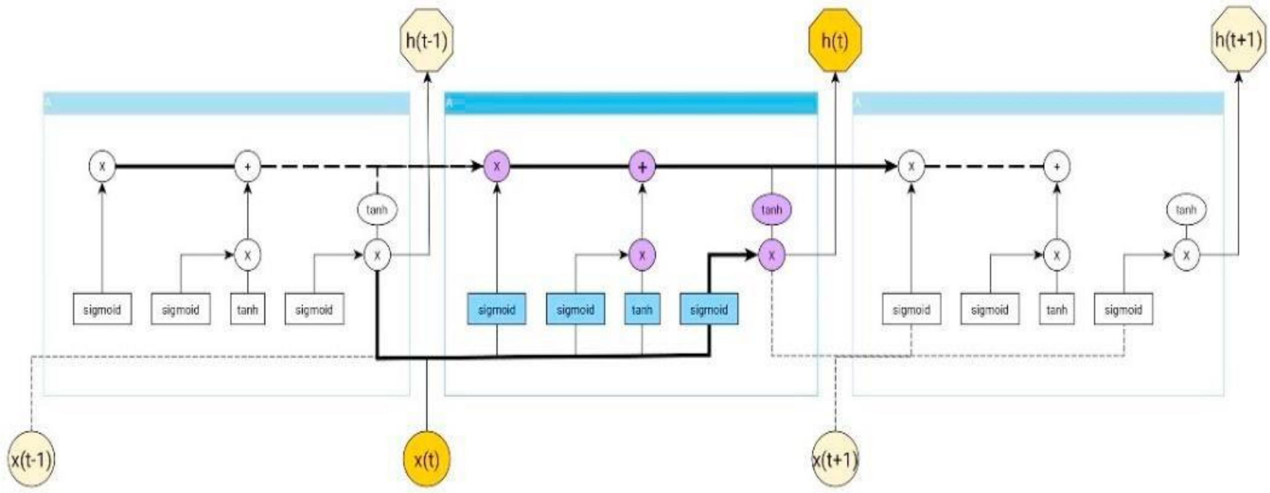


Fig 2: LSTM Network

The LSTM uses gated cells which act like a computer's memory. Each cell has three gates: input gate, Forget gate and output gate. Fig 2 shows the general form of LSTM cell. Unlike digital gates these gates are analog implemented as sigmoid and they range from 0 to 1. The problematic issue of vanishing gradients is solved through LSTM because it keeps the gradients steep enough and therefore the training relatively short and the accuracy high.

Let us denote the input time series as $X = (x_1, x_2, \dots \dots, x_T)$, hidden state cells as $H = (h_1, h_2, \dots \dots, h_T)$,output sequence as $Y = (y_1, y_2, \dots \dots, y_T)$. LSTM NNs do the computation as follows:

$$h_t = H\left(W_{hy}\, x_t + W_{hh}\, h_{t-1} + b_n\right) \quad (1)$$

$$y_t = \left(W_{hy}\, h_t + b_y\right) \quad (2)$$

The LSTM structure depicts that above is implemented through the following functions.

$$i_t = \sigma(W_{xi}\, x_t + W_{hi}\, h_{t-1} + W_{ci}\, c_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma(W_{xf}\, x_t + W_{hf}\, h_{t-1} + W_{cf}\, c_{t-1} + b_f) \quad (4)$$

$$c_t = f_t\, c_{t-1} + i_t * \tanh h\,(W_{xc}\, x_t + W_{hc}\, h_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma(W_{xo}\, x_t + W_{ho}\, h_{t-1} + W_{co}\, c_t + b_o) \quad (6)$$

$$h_t = o_t \tanh h\,(c_t) \quad (7)$$

All above equation holds for t = 1,2,……..,T.

The first step in LSTM is to decide which information is not important and to be thrown away. This is done by 'forget gate' which is a sigmoid layer. Based on the values of $h_{t-1}$and x_t it outputs a number between 0 and 1. Value 1 means completely keep the value and 0 means completely forget the value. Next, the input gate decides what new information to be stored in the cell state. It decides which value is to be updated. Next, the *tanh* layer creates a vector of new candidate values, Ct, that could be added to the state. The old cell state c_{t-1} is updated in to the new cell state c_t. We multiply the old state by f_t so that things that are to be forgotten are not considered. Then we add $(i_t * c_t)$. This is the new cell state c_t. Finally, the output gate decides what is to be output. This output will be the filtered version of the cell state. First, we run a sigmoid layer which decides what parts of the cell state to be output. Then, we pass the cell state through *tanh* and multiply it by the output of the sigmoid gate.

4. THE PROPOSED SYSTEM

4.1 Dataset

The international travel data of 56 counties from January 2003 to December 2017 is obtained from the website of Ministry of Tourism, Government of India. The data contains the number of foreign travelers traveling to India month wise from the year 2003 to 2017. It basically contains four fields. 1) Country 2) Total number of tourists visiting yearly 3) they year of visit 4) the month of visit. The dataset contains total of 10,453 records.

4.2 Data Preprocessing

The following data transforms are performed on the dataset prior to fitting a model and making a forecast.

1. As the ministry had data of 2 years was missing we impute the missing data points using historical average value. The following experiments are done on imputed data.
2. We normalize the data to the range of 0 to -1.
3. We transformed the time series data so that it is stationary. Specifically, a lag=1 differencing to remove the increasing trend in the data.
4. Transform the time series into a supervised learning problem. We organized the data into input and output patterns where the observation at the previous time step is used as an input to forecast the observation at the current time step.

4.3 Data Splitting

In our experiment, data is divided into two parts. The training datasets and validation dataset. Data from the year 2003 to 2013 has been used for training the LSTM model. And data from year 2014 to 2017 has been used for testing the model accuracy.

4.4 The Model Description

For SVR model, we used RBF kernel, the penalty parameter c was set to 0.001 and gamma was set to default –0.002. The LSTM network has a visible layer with 1 input, a hidden layer with 5 LSTM blocks or neurons an output layer that makes a single value prediction. The default sigmoid activation function is used for the LSTM blocks. We set a dropout 0.2 to avoid over-fitting. Our model uses '**adam**' optimizer. The network is trained for 100 epochs and a batch size of 1 is used.

4.5 Evaluation Criteria

In order to evaluate the performance of two models, three performance evaluation metric have been used: MAPE (Mean Absolute Percent Error), MAE (Mean Absolute Error) and RMSE (Root mean square error).

$$MAPE = \frac{100}{n} \times \sum_{i=1}^{n} \left| \frac{y_i - y'_i}{y_i} \right| \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - y'_i| \quad (9)$$

$$RMSE = \sqrt{\left[\frac{1}{n} \sum_{i=1}^{n} (|y_i - y'_i|)^2 \right]} \quad (10)$$

Where n is the total number of test samples , y_i is the target output and y'_t is the predicted output.

5. RESULTS AND DISCUSSION

We have predicted the tourism flow (in percentage) from several countries for a specific year from the test data set and then compared it with the actual tourist flow. Table 1.1 shows the comparison of SVR and LSTM based on three metrics discussed above.

Table 1.1 : Comparison of SVR and LSTM model for tourism prediction

Model	Accuracy	RMSE	MAPE	MAE
SVR	0.9774	0.0273	40.00	0.0225
LSTM	0.9859	0.0202	18.75	0.0140

Our experimental results show that LSTM performs better than SVR for prediction. There is the significant difference between the MAPE values. According to the Criteria of MAPE for Model Evaluation in Lewis (1982), the predicted data with the LSTM model has a good accurate forecast.

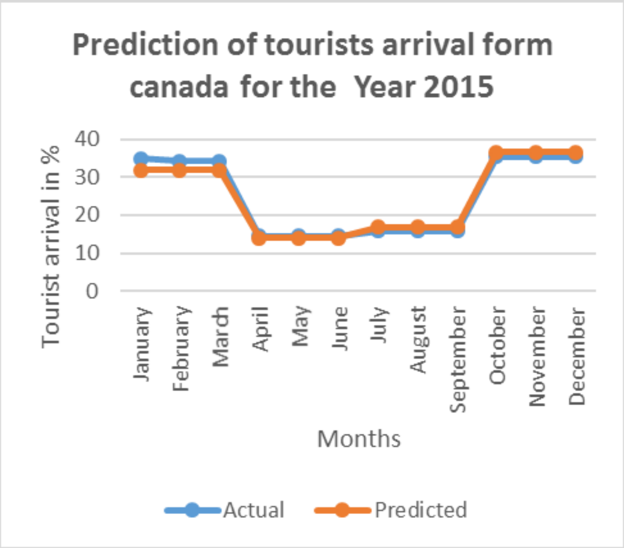
Table 1.2 : Model evaluation based on MAPE values

MAPE values (%)	Model evaluation
<10	Highly accurate
10-20	Good Forecasting
20-50	Reasonable forecasting
>50	Inaccurate forecasting

Hence we finalize our LSTM model for tourism prediction for India. The Table 1.3 shows actual tourist arrival and prediction month wise from Canada for the year 2015,2016 and 2017 along with the MPAE values.

Table 1.3: Month wise actual and predicted values of tourist arrival from Canada for the year 2015,2016 and 2017

Year→	2015			2016			2017		
Month	Actual Tourist (%)	Predicted Tourist (%)	MAPE	Actual Tourist (%)	Predicted Tourist (%)	MAPE	Actual Tourist (%)	Predicted Tourist (%)	MAPE
January	34.98	31.98	8.576329	34.5	31.98	7.304348	34.6	31.98	7.572254
February	34.2	31.98	6.491228	34.5	31.98	7.304348	34.6	31.98	7.572254
March	34.2	31.98	6.491228	34.5	31.98	7.304348	34.6	31.98	7.572254
April	14.5	14	3.448276	14.1	14	0.70922	14.1	14	0.70922
May	14.5	14	3.448276	14.1	14	0.70922	14.1	14	0.70922
June	14.5	14	3.448276	14.1	14	0.70922	14.1	14	0.70922
July	15.9	16.88	6.163522	16.5	16.88	2.30303	16.3	16.88	3.558282
August	15.9	16.88	6.163522	16.5	16.88	2.30303	16.3	16.88	3.558282
September	15.9	16.88	6.163522	16.5	16.88	2.30303	16.3	16.88	3.558282
October	35.4	36.71	3.700565	34.9	36.71	5.186246	35	36.71	4.885714
November	35.4	36.71	3.700565	34.9	36.71	5.186246	35	36.71	4.885714
December	35.4	36.71	3.700565	34.5	31.98	7.304348	35	36.71	4.885714



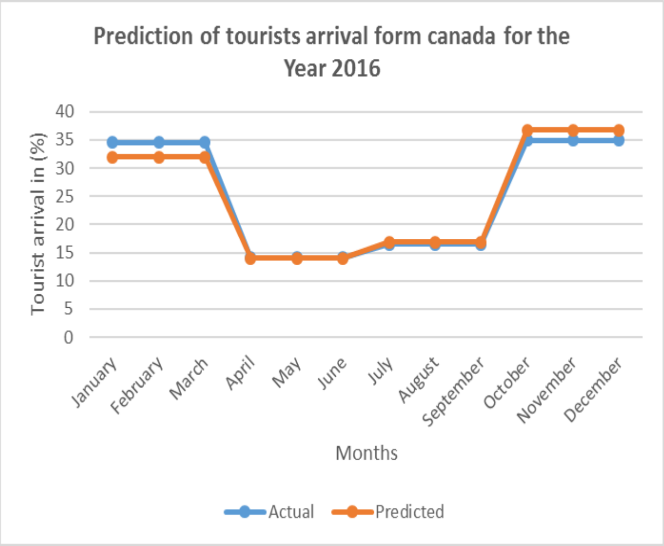


Fig. 3: Plot of actual and predicted values of tourism flow from Canada for the year 2015 and 2016.

6. CONCLUSION

In this paper we have exploited the use of deep learning model LSTM for prediction of tourism flow in India from 56 countries. We have compared SVR model with LSTM model for prediction of tourism flow in India. Experimental results show that LSTM model outperforms SVR model with the accuracy of 98.5% and MAPE 18.5% which is considered to be a good forecasting model.

7. FUTURE WORK

Prediction of tourism can be further enriched by considering other parameters that may affect the tourism flow. This would help the tourism ministry of India to understand the cause for growth or decline in the number of tourists arriving in India.

REFERENCES

[1] Statista, the statistic portal Retrieved from https://www.statista.com/topics/962/global-tourism

[2] World Travel Tourism Council, "World Travel & Tourism Council" retrieved from https://www.wttc.org/

[3] A. A. Ariyo, A. O. Adewumi and C. K. Ayo, "Stock Price Prediction Using the ARIMA Model," *2014* UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, *Cambridge*, 2014, pp. 106-112. doi: 10.1109/UKSim.2014.67

[4] **Biljana Petrevska**," Predicting tourism demand by A.R.I.M.A. models", Economic Research – Ekonomska Itstrazivanja, 2017 Vol. 30, No. 1, 939- 950, https://doi.org/10.1080/1331677X.2017.1314822

[5] **Choden and Suntaree Unhapipat**."ARIMA model to forecast international tourist visit in Bumthang, Bhutan", IOP Conf. Series: Journal of Physics: Conf. Series 1039 (2018) 012023 doi :10.1088/1742-6596/1039/1/012023.

[6] Z. Hu, J. Zhu and K. Tse, "Stocks market prediction using Support Vector Machine," 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering, Xi'an, 2013, pp. 115-118. doi: 10.1109/ICIII.2013.6703096

[7] **Paula Fernandes and Joao Paulo Teixeira**, "A new approach of Modelling and Forecasting monthly overnights in the Northern region of Portugal", IFC 4 ,4th International Finance Conference, 15th , 16th and 17th March 2007, Tunisia.

[8] B. Gui, X. Wei, Q. Shen, J. Qi and L. Guo, "Financial Time Series Forecasting Using Support Vector Machine," 2014 Tenth International Conference on Computational Intelligence and Security, Kunming, 2014, pp. 39-43.doi: 10.1109/CIS.2014.22

[9] **Sepp Hochreiter, J•urgen Schmidhuber**,"LONG SHORT-TERM MEMORY", Journal of Neural Computation, Volume 9 , issue 8 , November 15 1997,1735-1780 , doi/10.1162/neco.1997.9.8.1735

[10] **Yanjie Duan, Yisheng Lv and Fei-Yue Wang**, "Travel time prediction with LSTM neural network," 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, 2016, pp. 1053-1058.doi: 10.1109/ITSC.2016.7795686

[11] **YiFei , Han Cao**,"Prediction for Tourism Flow based on LSTM Neural Network", Procedia Computer Science, Volume 129, 2018, Pages 277-283, https://doi.org/10.1016/j.procs.2018.03.076.

[12] **Corinna Cortes, Vladimir Vapnik** "Support Vector Network" , Machine Learning, 20, 273-297 (1995) https://doi.org/10.1007/BF00994018

[13] Ministry of Tourism Govt. of India , Retrieved from http://www.tourism.gov.in/market-research-and-statistics, last updated 05 April 2019

[14] **Lewis, C.D**.; (1982); "Industrial and Business Forecasting Method"; Butterworth Scientific; London