Airport Passenger Traffic Data Prediction

Harsha Vardhan Miryala | Sanju Prabhath Reddy | Sri Harsha Majeti | Dr. Durga Toshniwal
Indian Institute of Technology Roorkee

Abstract

We develop a machine learning approach using Regression Techniques, Neural Networks and Long Short Term Memory networks to forecast the airport passenger traffic. This project aims to implement time series data prediction on the data obtained from official websites and also to try out simulation and econometric modelling.
Typically, the forecasts are not final objectives in and of themselves. An essential ingredient to preparing forecasts is to understand the purpose for which they will be used. In economic terms, the forecasts of activity are usually meant to reflect the demand for aviation services. It is the demand for services that drives the forecasts and therefore helps airport planners provide the appropriate supply in terms of the infrastructure needed.

Introduction

Air transportation is an integral part of the global economy. Scheduled commercial passenger transport alone has grown over the past half century to more than 27 million flights in 2005 across the world, and more than 11.6 million flights to or from the United States alone. Other sectors, including cargo, air taxi, and general aviation, also have expanded rapidly. Airports are obviously a fundamental part of modern air transportation systems—some serve as hubs for extensive national and international transportation networks, whereas others function as the complementary spokes; still others serve as parts of multi-airport regional systems; and some serve as important connections to the outside world for otherwise isolated populations. Airports serve both commercial and private aviation - often referred to as air carrier and general aviation.
The construction, operation, and future expansion of airports can require substantial initial and ongoing investments, a large share of which is usually paid for with public money. Consequently, for any individual airport, it is important to be able to forecast future demands for aviation services to assess the potential need for further investments in capacity or services to meet those demands. Accurate forecasts are essential for effective airport planning and decision making, and for the efficient provision of capacity.
The type of forecast and level of effort required to produce it depends importantly on the purpose for which the forecast is being made. Short-term aviation forecasts are needed to support operational planning and often are used to assess personnel requirements at an airport or the need for incremental improvements or expansions of landside facilities and terminal areas, air cargo facilities, general aviation hangar space, etc.

Methodology

These type of situations are usually approached in four methods namely -

- Time Series Modelling
- Econometric Modelling
- Simulation Modelling
- Market Share forecasting

Market share analysis involves measuring current activity at an airport as a share of some other aggregate measure, and then assuming that the share will remain constant (or perhaps change in some prespecified way) so that airport activity will grow along with the projected growth in the aggregate activity.
Simulation models can be used when one needs to obtain a particular estimate, so each model has its own rules and corresponding inputs.
Econometric modelling involves statistical estimation of a regression equation between a dependent variable and a set of explanatory variables. The possible variables are -

- Macroeconomic and demographic factors such as the level of and growth in the economy, population, incomes, etc.;
- Airline market factors, including fares, flight frequency, and schedules;
- Air transport production costs and technology;
- Regulatory factors;
- Infrastructure constraints and improvements;
- Substitutes for air travel.

Time Series modelling involves some form of extrapolating existing data out into the future. In its simplest form it is based only on values of the variable being forecast and projects the future based on current or past trends.
Three types are tried in this - ARIMA,LSTM and MLP.
ARIMA[3] is an acronym that stands for AutoRegressive Integrated Moving Average. It is a generalization of the simpler AutoRegressive Moving Average and adds the notion of integration.
The parameters of the ARIMA model are defined as follows:

- **p**: The number of lag observations included in the model, also called the lag order.
- **d**: The number of times that the raw observations are differenced, also called the degree of differencing.
- **q**: The size of the moving average window, also called the order of moving average.

Data it is divided into trend part, seasonal part and residual part in fig 1.

MLP[4] is a supervised backpropagation learning neural network. The data of the past 12 months acted as the input for the current month prediction.

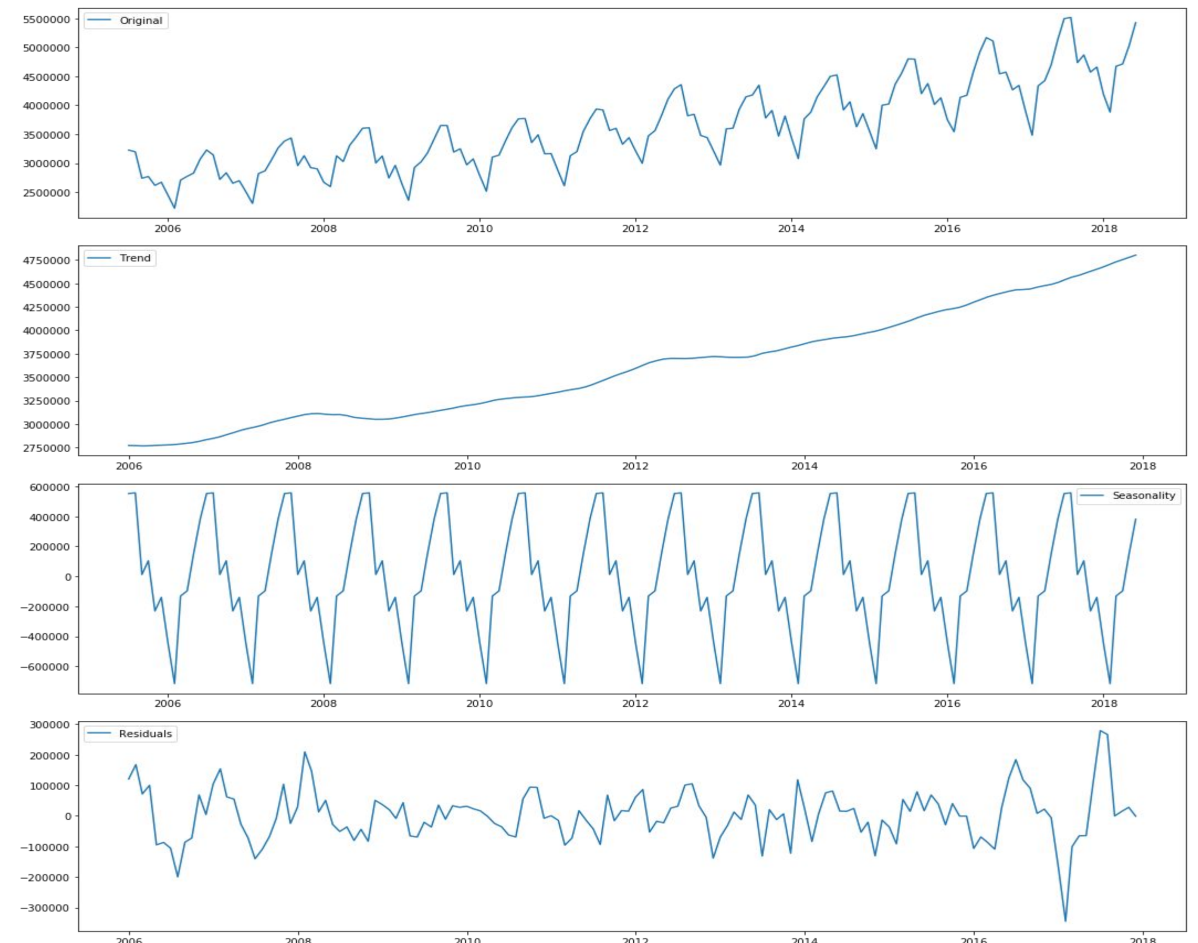
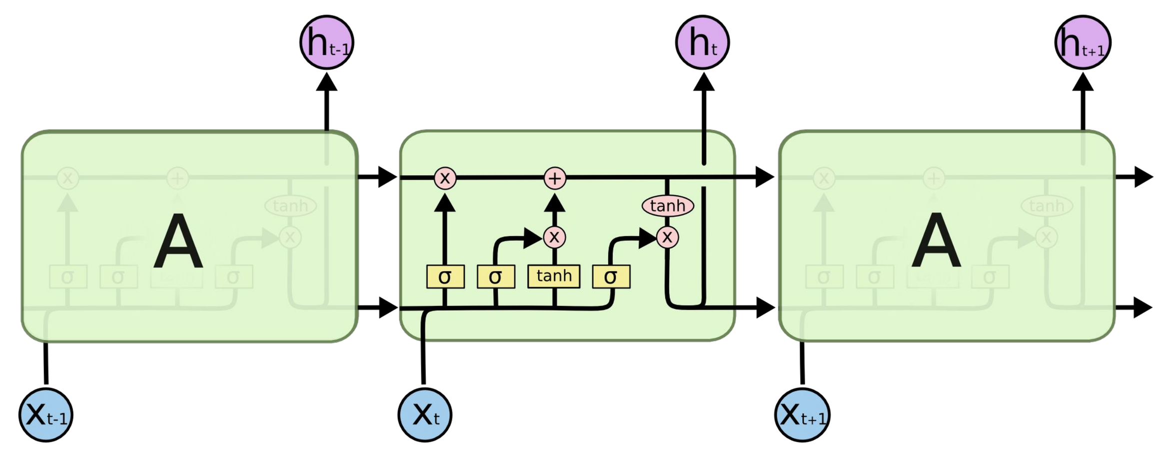


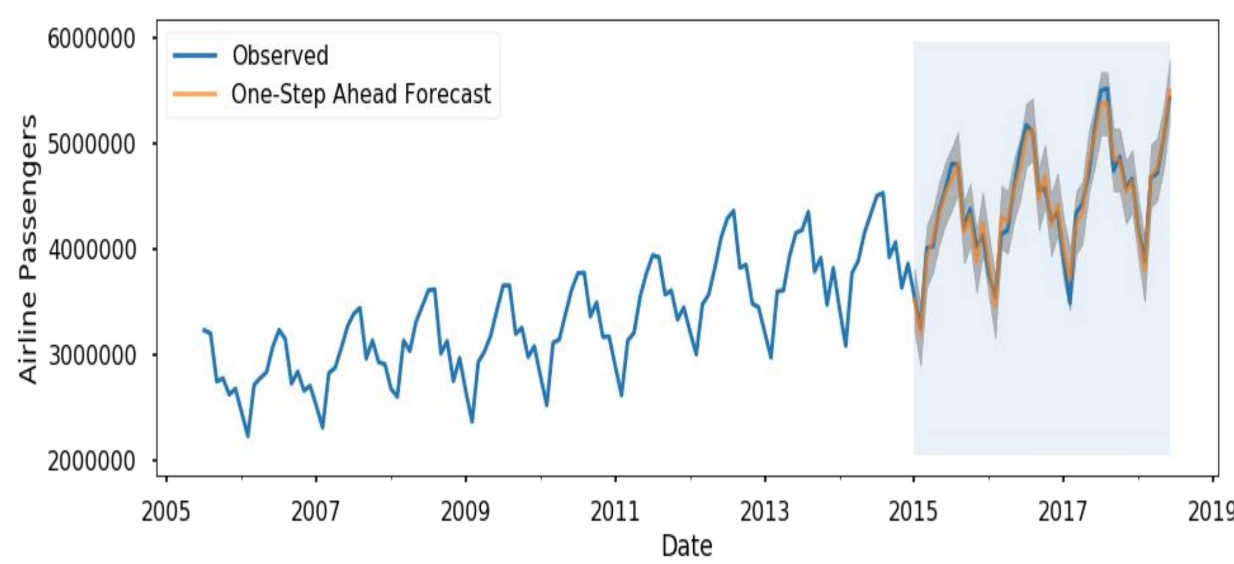
Fig 1

Instead of neurons, LSTM[2] networks have memory blocks that are connected through layers. A block has components that make it smarter than a classical neuron and a memory for recent sequences. A block contains gates that manage the block's state and output. A block operates upon an input sequence and each gate within a block uses the sigmoid activation units to control whether they are triggered or not, making the change of state and addition of information flowing through the block conditional.

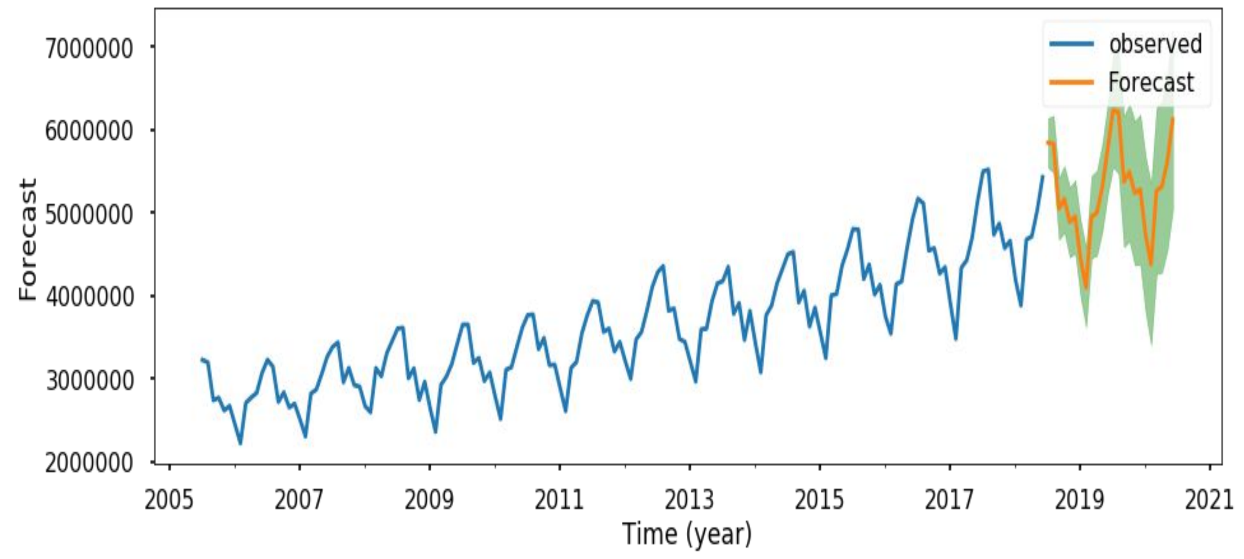


Results

From the above mentioned three methods, ARIMA gave better results and the below diagram shows the actual data and the forecasted one.



Now, through the model, the prediction for the next two years, showing the upper and lower limit.



Conclusion

- Among the devoloped models, ARIMA gave better results.
- The selection of an appropriate forecasting method may depend on both technical and budgetary factors.
- From a technical standpoint, two primary drivers in determining which may be the most appropriate method are the purpose for which a forecast is being made and how the metrics being forecast relate to available historical data.

Future Works

- Improving the existing ARIMA model and trying for a better time series model.
- Collect different possible data as mentioned that can affect the traffic and try to build an efficient econometric model.
- Implement a simulation model and the possibility of a market share forecast.

Sound of work

- Long-Term Airport Planning and Capacity Needs
  - Airside and Lanside facility expansion
- Short-Term Operational Planning
  - Aiport Personnel and FAA tower staffing requirements
- Financial Planning
  - Bond issues ans use of public funds and annual budget

References

[1] AIRPORT COOPERATIVE RESEARCH PROGRAM by FAA - http://www.trb.org/Publications/Blurbs/158684.aspx
[2] LSTM by Schmidhuber et al. https://www.bioinf.jku.at/publications/older/2604.pdf
[3] ARIMA by Jenkins et al.
[4] MLP by Hinton et al https://www.iro.umontreal.ca/~vincentp/ift3395/lectures/backprop_old.pdf