Graduate Research Paper Analysis

Submitted by: - Parth Parashar
PSU ID: - 923928157

Topic: - Machine Learning for Air Transport Planning and Management.

Before starting off with this Research Paper Analysis, I would like to take this opportunity to thank my Prof. Anthony Rhodes for providing me with a chance to explore my interests in the fields of Machine Learning and consequently analyzing this research paper titled- “Machine Learning for air transport planning and management”. Additionally, I would also like to thank the original authors of this paper for providing me with the means to perform some additional research on their work.

In this research paper, the authors start off by discussing the general idea of aviation travel and how artificial intelligence can be used in the aviation industry. The aviation industry plays a very important role in today’s world as it facilitates global commerce, tourism and trade. The aviation industry is divided into two big group of businesses namely passenger airlines and Freight Airlines (also known as cargo airlines). The passenger airline industry consists of four major business models. These models are Full service network carrier, Lost cost carrier, Regional carrier and holiday carrier. Regardless of the business model chosen, forecasting is considered to be one of the most important management functions of the airline. This is because operating an airline is a very expensive business along with the fact that the profit is generally very

less and lies in the 1-2% range. The authors depict this point using a graph. The authors also depict the fact that even though the business of both passenger and freight airlines are growing quickly, the profits are not.

This makes every decision of the airline business very crucial. Forecasting is therefore very important to the aviation industry because this helps airlines plan routes according to the predicted number of passengers on those routes as well as how profitable that route can be.

The aviation industry till now employs the traditional methods of multiple linear regression models for passenger and aviation forecasting, which by themselves have evolved considerably over the years, but now can be challenged by more sophisticated and efficient machine learning models. The authors, by the medium of this paper, then compares the various machine learning models among themselves and with MLR to determine which could be the best methods for forecasting in the aviation industry.

The various machine learning algorithms that the author uses for comparison are Artificial Neural Networks, Genetic algorithms, Adaptive neuro fuzzy inference systems, Support Vector machines and regression trees.

The Artificial neural network based forecasting model was firstly developed for the Brazilian airlines and then extended. This was followed by the employment of more sophisticated neural networks along with back propagation techniques. Among these experiments performed by various researchers on small datasets (airspaces extending only in single countries), it was concluded that artificial neural networks performed better at predictions (forecasting) than the Multiple linear regression models. Among

other things, a lot of researchers also tried genetic algorithms as well as algorithms which were based on the combination of artificial neural networks and a group of data handling techniques and all these produced better results than MLR.

But to know which method performed the best and which worst, the authors decided to employ these methods to the data available from Australia's low cost carrier market. A low cost carrier is an airline which eliminates non-essential services in order to reduce the cost of airfare. This means that the tickets only include the cost of the fare and the rest of the other services cost extra. Also,LCC cuts costs in other operations as well like self-check in and bag drop so as to reduce the cost for staff. They do so to maximize profit. To compare the algorithms and their results, the same data points were used for training, testing and validation purposes.

There were a lot of variables used for getting the required results. Some of these variables are Australia's real best economy, discount fares, unemployment numbers and many more.

When all the above mentioned algorithms were run with the same inputs along with the same parameters , results for each of the six models in terms of the predicted RPKs relative to the actual RPKs for all 42 quarters of data are evaluated.According to the above evaluations, figures (graphs) are constructed. Each figure includes the equation (which would ideally give BETA= 1, ALPHA= 0), which gives an indication of the accuracy of the modelling tool, while the included R2 value is an indication of the precision of the modelling tool.

A parameter namely RPK (Total Revenue Passenger Kilometers) was used as a parameter for the axis of the graph/figure as well as served as a primary parameter for comparison of the various models. Also, RMSE (Root Mean Squared error) terms were calculated for each algorithm. The RMSE values determined which algorithm worked best and which worst.

To fully access the models, ANOVA (analysis of variance) was used to compare all 6 sets of MSE (mean squared error) terms for determining if they are statistically significantly different or not. Both the ANN (artificial neural networks) and ANFIS (adaptive neuro-fuzzy 4 inference systems) models result in MSE (mean squared error) terms that are statistically significantly better than all the other models. Relative to each other, the resultant errors indicate that the ANFIS model is not statistically significantly better than the ANN model. The largest error though is produced by the RT (Regression trees) model.

The authors, by using a real-world case study approach, have examined various machine learning and artificial intelligence based tools to model the algorithms for the aviation industry and its forecasting. Based on the resultant root mean square error to evaluate the models, the study found that the ANFIS approach had the lowest RMSE, whilst the Regression Tree approach had the highest RMSE. The results for the ANFIS were only marginally better than those for an ANN, and even those results were not statistically significantly different. The study also concludes that the application of artificial intelligence-based methods can be successfully applied in the airline industry to aid in operational planning and managing.

My Question to the author: -

I would like to ask the author as to why the Naive Bayes model was not included.

Suggestion: -

Although the research was clear, concise and thorough, a little more clarity on the data used for LCC airlines for determining the best and worst models would have been great.

If given an opportunity, I would like to extend this research paper to other methods of forecasting such as Naive Bayes models as well as other deep learning models could also help improve the forecasting saving cost for the airlines.