Applied GIS

ISSN: 1832-5505
Vol-05 Issue-03 Aug 2017

Using Deep Learning to Improve Flight Delay Prediction

Meenal Naranje[#1], Prof. Mandar Sohanil [2]
[#]*Department Of Computer Engineering, University Of Mumbai*

***Abstract*— Several machine learning tasks, such as image recognition, speech translation, and ground traffic flow prediction, have seen substantial improvements because to the use of deep learning. Our goal is to use deep learning to forecast when flights will be delayed, and then to provide reports on the situation at every given airport. This is done by assigning aircraft to delay categories using classifiers. Factors such as weather, airline, link delay, time of day, and holiday effects are used by the classifier. The number of Classes will be determined by the average wait times for events with the same set of attributes in the past. The model will be educated on this data and utilized for future forecasting. This method has the potential to provide a more powerful prediction model with better accuracy and efficiency.**

***Keywords*— Prediction, Multinomial Logistic Regression (MLR), Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM), and Recurrent Neural Networks (RNNs)**

INTRODUCTION

There are a number of people, places, and things that suffer when a flight is delayed. All commercial aviation stakeholders must factor in their forecast while making important decisions. In addition, the complexity of the air transportation system, the abundance of prediction techniques, and the flood of data connected to such systems made the creation of exact prediction models for flight delays difficult.

Even though India loses a lot of money due to delays, as mentioned in the introduction, little study is being done to improve delay prediction in the country. For the sake of study, academics have built learning and prediction models using a narrow collection of data or qualities. Since only a subset of relevant factors are taken into account, the resulting inconsistency reduces the predictive power of the models. There is currently no publically available technology that can forecast and indicate flight delays for passengers.

A. MOTIVATION

Figure 1 shows that Indian domestic flights have a dismal record of arriving on schedule. Data from four main airports in India—Delhi, Mumbai, Kolkata, and Chennai—show the percentage of flights that arrived on time for each carrier.

As can be seen, extremely few planes really arrive on schedule, which causes significant delays for passengers. In addition, airlines must offer amenities for customers while they wait, which drives up the price of flights.

Predicting flight delays means airlines can prepare for the possibility of delays and disseminate that information to passengers, who will therefore be more likely to show up at the airport at the scheduled departure time rather than waiting around.

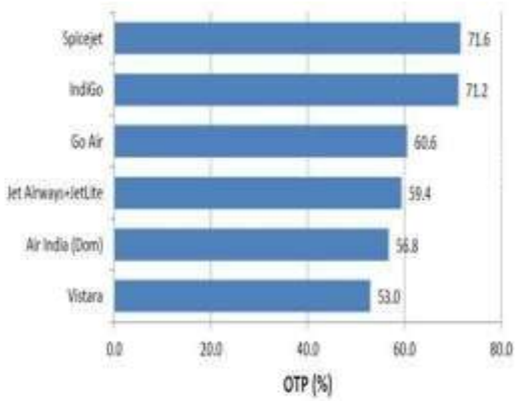


Fig 1. On Time Performance by Indian Flights at 4 Metro Airports

.

LITERATURE SURVEY

Approximately 275 airlines, or 83% of worldwide air traffic, are members of the International Air Transport group (IATA), the trade group for the world's airlines. They provide backing for a wide variety of aviation endeavors and contribute to the formation of industry policy on crucial aviation topics. With the goal of improving aviation safety in India, [1] the Directorate General of Civil Aviation's (DGCA) International Cooperation Group (ICG) coordinates all international initiatives with international organizations on policy, technical, and safety problems. [2]

Organizations from across the world demonstrate their commitment to working together to keep India's aviation system safe and make improvements where necessary. [2]

The first part of the ICG is a group of organizations working together voluntarily to improve aviation safety. These groups include international organizations, NGOs, national aviation authorities, operators, research organizations, and aircraft and equipment manufacturers.

A worldwide framework for information sharing and timely communication among the aviation community at national and international levels has been established by DGCA via ICG.

The ICG makes certain:

- Collaborating on questions of regulation.
- Data necessary for cooperation plans.

There will be a sharing of technical details.

Applied GIS

ISSN: 1832-5505
Vol-05 Issue-03 Aug 2017

• Harmonization and standardization of regulatory processes.

Include BASA and Technical Details.

The ICG's goal is to facilitate communication across different jurisdictions. This boosts the profile of the nations involved internationally, leading to greater market acceptability and new commercial opportunities.

.

A. An analysis method for flight delays based on Bayesian network:

In [3], Li Qianya, Wang Lei, and co-authors offer a Bayesian Network-based technique for analyzing and forecasting flight delays. When a delay occurs upstream, we are able to anticipate the circumstances that will cause delays for aircraft both upstream and downstream.

As shown by the series experiment, the solution has strong analytical and prediction dependability. In addition, it serves as a resource for the appropriate units so that they can anticipate the widespread flight delays. It exemplifies GMM-EM's use in tracing cascading effects from their underlying sources.

It uses Random Flight Points to forecast flight delays. RADAR is programmed to track flights to and from these airports. The RADAR system will track the plane's position and compare that to the time it was supposed to arrive. Therefore, we shall determine the Delay.

However, the placement of RADAR systems along the route of flight is very expensive for the system. Both on-ground delay and the effects of weather are ignored in the article. It is predicated entirely on a single flight's delay and its subsequent ripple effects. As a result, the precision drops, and future forecasts are made only with limited confidence.

.

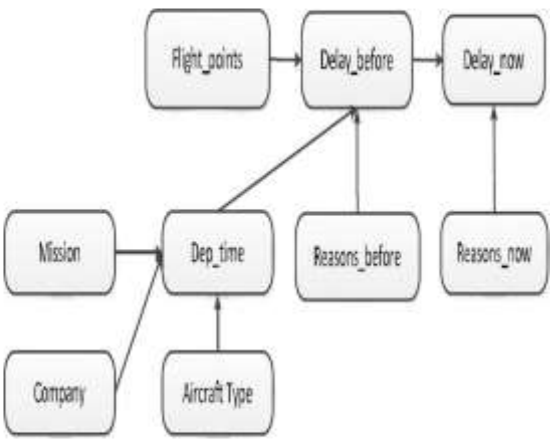
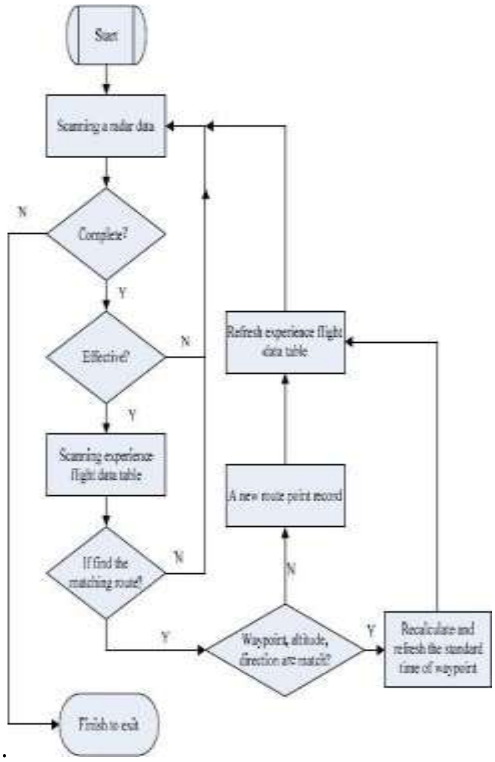


Fig 2. BN network Model [3]



.

Fig 3. Flowchart of obtaining Empirical Flight Data [3]

B. Risk Management Using Big Real Time Data:

C. As stated by Jie Cheng [4], After include all of the variables in the model, we have a complete model for predicting delays for individual flights. To emphasize the impact of the vacation effect, they ingeniously merged the weight function with a spline model. Second, they have successfully implemented the ARIMA model's robustness and efficiency in predicting the influence of random elements on non-stationary time series. The first and second-factor residuals are effectively eliminated. This means that the endeavor as a whole has been successful. The flight is monitored by RADAR systems, which also keep track of any delays. As a result, the System will function only in accordance with the existing delays. The priori probability is difficult to get, but we may use mathematical theories and years of historical data to produce it in preparation for our future work. If we can include a priori probability with actual data, this would be helpful. Airport business is not taken into account by the system (Real data)

D. A Flight Delay Prediction Model with consideration of Cross-Flight Plan Awaiting Resources:

E. Business of flight delay propagation due to cross-flight plan waiting for resources has been analyzed by Rong Yao, Wang Jiandong, and Xu Tao [5], who then provided a model and forecast method for flight delay propagation. The simulation findings demonstrate the efficacy of the model and algorithm in determining the flight delay propagation due to the sharing of airport resources across many flight plans, and in giving airports a transparent warning signal for any delays that may occur. In the future, it may be possible to include this model and algorithm into an RIA-based graphical platform for making informed predictions about flight delays. The

ISSN: 1832-5505
Vol-05 Issue-03 Aug 2017

article focuses only on resource reliance. The effects of things like bad weather and congested airways on flight delays are ignored.

F. Clusters and Communities in Air Traffic Delay Networks:

The results presented in this paper show the potential of clustering air traffic delay networks to identify typical delay states and typical types of days, accounting for both spatial and temporal patterns and connectivity. The studies aid in identifying major airports and restrictions that contribute to flight delays. According to the results of the community detection study, there are clusters of airports that share similar patterns of delays on specific days. Ongoing research into this problem involves a comparison of the typical day types discovered by clustering network time series (as was done in this publication) with the typical day types discovered by clustering aggregate delays seen at various airports. It is also possible to correlate the day-to-day changes in delay states to the interruptions and control measures (i.e., Traffic Management Initiatives) that occurred during that day. Finally, this study may be utilized to construct prediction models of air traffic network delays using the distinctive delay states, kinds of days, and major characteristics described here, which may lead to better decision support tools for stakeholders like air traffic management. Although it can produce the Delay network, it is unable to foretell the delay of any one aircraft.

G. Predicting flight departure delay at Porto Airport: A preliminary study:

H. Preliminary prediction findings were reported by Hugo Alonso and António Loureiro [7] for the challenge of anticipating airplane departure delays at Porto Airport. We predicted the delay by treating the issue as an ordinal classification job and used a method based on the so-called unimodal model. Using neural networks and trees, they've implemented the unimodal model and discovered, via experimentation, that arrival delay and ground operation time are the most important factors for departure delay prediction. The neural network implementation was less complicated and produced superior test-set performance. Interestingly, both implementations struggled to tell the difference between aircraft whose departure delays fell in the [-infinity,0] minute range and those whose departure delays were in the [0,15] minute range. This concern may be investigated at a later time. The article focuses only on arrival delays and ground operations. They have ignored the most important cause of aircraft delays: bad weather.

I. Prediction of weather-induced airline delays based on machine learning algorithms:

Researchers Sun Choi, Young Jin Kim, and colleagues [8] have investigated and suggested a prediction model that can categorize weather-related aircraft delays. In specifically, the model was developed using machine learning methods and prior information on the weather and traffic conditions for each OD pair. Random forest, AdaBoost, k-Nearest-Neighbors, and Decision Trees are only few of the supervised machine learning algorithms used in this research. SMOTE was used in conjunction with a random undersample because of the asymmetry in the data. The accuracy of the model's predictions were evaluated both in the validation set and the test set. There are certainly further avenues that might be explored to enhance the model. Taking into consideration the costs of a false positive and false negative would allow for the optimal performance of classifiers to be identified. Then, it might serve as the basis for a useful decision-making tool that helps forecast when planes will arrive. The predictive effectiveness of the model would also benefit from careful consideration of prediction uncertainty. It relies only on weather information and ignores all other variables. As a result, it becomes harder to make accurate predictions.

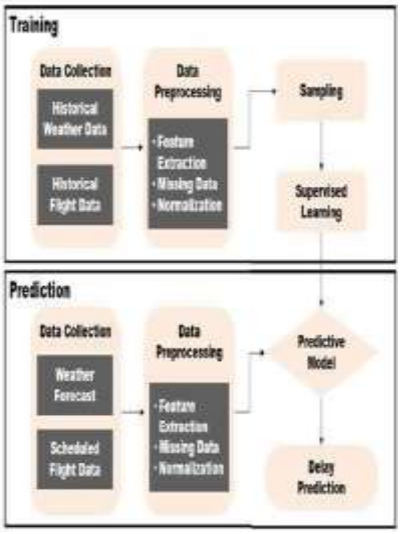


Fig 4. Summary of model Developed [8]

EXISTING SYSTEM

According to a prior research, the expenses associated with flight delays in the National Airspace System (NAS) are substantial. Costs to passengers, airlines, and other sectors of the NAS were over $33 billion in 2007. Numerous studies have been conducted to analyze and anticipate air traffic delays in an effort to save wasted expenses. The findings might be used to inform the development of air traffic management systems that are both more effective and less disruptive.

In many machine learning tasks, such as image identification, voice recognition, machine translation, etc., the accuracy of the classification and regression is greatly improved by using deep learning. It has also found use in the area of predicting the flow of traffic on the ground. One of the uses of air traffic data analytics is predicting when a flight will be delayed, thus

ISSN: 1832-5505
Vol-05 Issue-03 Aug 2017

it's important to assess whether or not a deep learning architecture can do a good job at that.

Detailed analyses of air traffic delay patterns are now possible thanks to a very accurate prediction model that was developed by merging numerous models based on the deep learning paradigm. Recurrent Neural Networks (RNNs) have shown to be quite effective in modeling time series data. A trustworthy one-day delay status might be obtained by implementing a deep LSTM RNN architecture into the prediction model. Then, the individual flight delay model has been fed the day's delay status, yielding the most accurate delay statuses for individual flights. Deeper network topologies have been demonstrated to increase RNN accuracy. [9]

PROBLEM DEFINITION

A. *PROBLEM STATEMENT*

We are planning to develop Flight Delay Prediction system which will predict delay of every individual flight which is scheduled in near future. A Multi-Layer Perceptron (MLP) and Multinomial Logistic Regression (MLR) algorithms are proposed for predicting flight delay with less time and more accuracy.

B. *OBJECTIVES*

To enhance performance of previously proposed models.
To provide predictions to airlines and passengers through simple interface. To evaluate performance of the model.

PROPOSED SYSTEM

Therefore, we are proposing a novel model for the Flight Delay Prediction to fill in some of the gaps found in the aforementioned literature review. Each component of the proposed system complements the others to form a powerful and precise whole. The courses consist of the following:

1. **Preparing the Data:**

Since the proposed system would draw from external datasets, these datasets may include more information than is necessary for the prediction task. This means the system will only save the data that is really useful for analysis.

(2) Identifier:

Each piece of the historical dataset will be segmented by the classifier module into one of many delay categories. Each flight will be assigned a delay category based on its specific characteristics. The trained dataset will include the results of this classification.

Third, a service that forecasts delays:

Every day, with no extra input from airport officials required, the system will function as intended. As adjustments are made, this will update the delay status of future flights. The database server will record the current state and use it to provide results for users.

Fourteenth, Application Host

Client queries will be handled by an application server. The query will be processed, and the result will be retrieved from the database server and returned. There will be no immediate action taken on your request.

Evaluation of Prediction Models

The assessment model will verify whether or not the sample was correctly categorised in the right delay class if the real Arrival and departure time of the aircraft is fed into the system. If there is a discrepancy, the classifier will adjust its parameters to account for it.The suggested system requires the following types of data sets:

First, the record of flight delays:

Attributes needed for classifier training for Flight Delay Prediction are Expected Arrival Time, Expected Departure Time, Actual Arrival Time, Actual Departure Time, Airline, and Flight Date. We may use this information to analyze patterns of delay across time, both in terms of airlines and dates.

Airport Timetable #2:

Predicting the outcome for certain flights requires knowing the schedule for such flights at a given airport. Included in this data collection should include the following: Flight Date, Airline, Expected Arrival Time, Actual Arrival Time, Actual Departure Time, and Actual Departure Time.

Classified samples using a delay classification system:

The classifier was used to build this database. This will be where we store our categorized samples and refer to when deciding how to handle future samples.

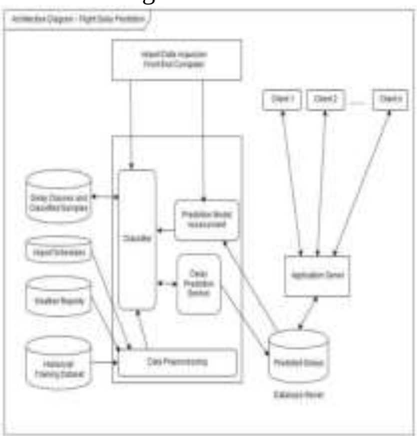


Fig 5: System Architecture

1. Predicted Delays:
   This is also generated database consisting of predictions of flight delays made by prediction

Applied GIS

ISSN: 1832-5505
Vol-05 Issue-03 Aug 2017

model and is intended to be used by application server to serve client requests for delay prediction.

CONCLUSION

Here we are going to use classification technique to classify flight details in a class, we call it as delay class as it reflects delay to that flight. The classifier will be trained with historical data of previous flights and their details with or without delay. Based on this system will predict delay for individual flight as well as it will generate some facts and figures for the on-time performance of the airport.

Conclusions

The version of this template is V2. Most of the formatting instructions in this document have been compiled by Causal Productions from the IEEE LaTeX style files. Causal Productions offers both A4 templates and US Letter templates for LaTeX and Microsoft Word. The LaTeX templates depend on the official IEEEtran.cls and IEEEtran.bst files, whereas the Microsoft Word templates are self-contained. Causal Productions has used its best efforts to ensure that the templates have the same appearance.

Causal Productions permits the distribution and revision of these templates on the condition that Causal Productions is credited in the revised template as follows: “original version of this template was provided by courtesy of Causal Productions (www.causalproductions.com)”.

Acknowledgment

Since I have worked hard to get this far in my dissertation. However, without the kind backing and assistance of numerous people and our university, none of this would have been possible. I'd want to give everyone our deepest appreciation.
I'd want to express my gratitude to Professor Mandar Sohani for all the help he's given me during this research. Under his watchful eye, I've been able to gather all the knowledge I needed to succeed.
The special theme seminar would not have been possible without the helpful cooperation and support of our project coordinator, Professor Vipul Dalal, and Head of Department, Dr. Arun Chavan.

For all of Dr. Sunil A. Patekar's support and inspiration as our school's principal, please accept my deepest appreciation.

I'd also want to express my gratitude to the individuals who have voluntarily contributed their time and skills to the project's development, as well as the colleagues who have assisted me along the way.The first part of my dissertation would not have been possible without the support and cooperation of my parents and the faculty and staff at Vidyalankar Institute of Technology.

References

[1] (See [1] http://www.iata.org)

[2] [2]http://dgca.nic.in

[3] [3]An Analysis Method for Flight Delays Based on Bayesian Network, Li Qianya, Wang Lei, Fei Rong, Wang Bin, Hei Xinhong, The 27th Chinese Control and Decision Conference (2015 CCDC), 2015.

[4] 2014 Master's Thesis by Jie Cheng titled "Risk Management Using Big Real Time Data"

[5] "A Flight Delay Prediction Model taking Into Account Cross-Flight Plan Awaiting Resources," 2nd International Conference on Advanced Computer Control, Rong Yao, Wang Jiandong, and Xu Tao, 2010.

[6]

[7] "Clusters and Communities in Air Traffic Delay Networks," 2016 American Control Conference (ACC), authors Karthik Gopalkrishnan, Hamsa Balakrishnan, and Richard Jordan.

[8] Alonso, Hugo, and Loureiro, António [7] 2015, 7th International Joint Conference on Computational Intelligence (IJCCI) "Predicting flight departure delay at Porto Airport: A preliminary study"

[9] According to "Prediction of weather-induced airline delays based on machine learning algorithms" by Sun Choi, Young Jin Kim, Simon Briceno, and Dimitri Mavris (cited in [8]),

[10] Conference on Digital Avionics Systems (DASC), 2016 IEEE/AIAA 35th Annual

[11] 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), 2016; Young Jin Kim, Sun Choi, et al., "A deep learning approach to flight delay prediction."

[12] For example, see [10] Simon Serrarens's "COMPARING CONVOLUTIONAL AND LSTM NEURAL NETWORKS FOR INTRADAY FOREX DIRECTION PREDICTION" (Master's Thesis, 2017–2018).