



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 4, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Survey on Flight Delay Prediction by Machine Learning

Prof. Swati Dhadake, Tuljai Kadam, Amanoddin Shaikh, Sumit Sabale, Bhagyashri Shinde

Department of Computer Engineering, Smt Kashibai Navale College of Engineering (SKNCOE), India

ABSTRACT: Today, aviation plays a vital role in global transportation, with 4,444 businesses relying on various airlines to connect them to the rest of the world. But extreme weather conditions can directly affect air services due to flight delays. To address this issue, accurate prediction of these flight delays could prepare passengers for the deterrent effect would have on their journey and enable airlines to address potential causes of flight delays in advance. flight to reduce the negative impact of. The goal of this project is to create a web application powered by the machine learning algorithm to predict flight delays. Let's look at the implementation of the Random Forest Regression algorithm for predictive models. The main idea behind the web application is also to allow the public to interact with the platform to retrieve the scheduled delays for their flights.

KEYWORDS: Flight delay prediction, machine learning, random forest, web app

I. INTRODUCTION

In the modern world, the key components of any transportation system are passenger aircraft, cargo aircraft, and air traffic control systems. Since then, countries around the world have tried to develop ways to improve the air transport system. This fundamentally changed airline operations. Flight delays periodically affect passengers in modern vehicles. An estimate 20% of airline flights are cancelled or delayed each year, costing passengers over \$20 billion in costs and time. Flight delays are unavoidable and have too many negative economic impacts on passengers, agencies and airports. In addition, delays harm the environment by increasing fuel consumption and causing polluting gas emissions. In addition, delays affect commerce, because the transportation of goods is highly dependent on customer trust, which can increase or decrease ticket sales, so flying on time leads to customer trust. Only in this way can flight forecasts enable agencies and airports to make skillful decisions and operations, and a good passenger information system can satisfy customers.

II. RELATED WORK

MingdaoLu, Peng Wei, Mingshu Heand Yinglei Teng:

This paper presents a flight delay prediction algorithm based on a gradient machine learning classifier. Fully dig into the influence of the previous flight, the situation of departure and landing airports, and the overall situation of flights on the same route, etc., to analyse and predict flight delays. A series of experimental results show that the algorithm can accurately classify and predict flight delays, and has good analysis and prediction reliability. Future research includes the application of more advanced and appropriate preprocessing techniques and artificial intelligence algorithms to achieve better performance. In this article, we use flight data covering all airports in China, which has been tested to have good generalization prediction ability. Therefore, the model can also be trained in the future using data from other countries.

SuvojitManna, SanketBiswas, RiyankaKundu, SomnathRakshi, PritiGupta andSubhasBarman:

This paper presents, a gradient enhanced decision tree model was used to predict flight delays using six important attributes. The study here concludes that, for the given data, the model reaches the highest coefficients of determination of 92.3185% for the arrival case and 94.8523% for the departure case. The model can be used to accurately predict flight delays at US airports. Individuals and airlines can also use the model to accurately predict flight delays. It is also useful for tourists before choosing a one-time airline when traveling. A limitation of the model is that it can predict flight delays for the 70 airports on which it was trained. Extending the model to other airports requires historical data for those airports. This could be the next step for logistics in air transport

**Lu Hao , Mark Hansen , Yu Zhang , Joseph Post:**

In this study, two different approaches were used to study the interaction between the three New York airports and the rest of the NAS, in particular the impact of delays at New York airports on the system. By applying econometric models, our results confirm the existence of dependencies. A one minute delay on each flight at the three New York airports would result in a 0.08 minute delay on each flight at the other major NAS airports. The rest of the NAS will also generate 0 for each minute of flight delay. New York was 22 minutes behind. These different values are reasonable given the size of the operations of the three New York airports compared to the 32 other airports considered in our model. Using coefficients from econometric models, different hypothetical counterfactual scenarios involving congestion relief in New York City to quantify how delays at other airports and the system as a whole would be affected. In the most extreme case, once the flight has flown into all three New York airports, all delays disappear, reducing overall delays by 3 minutes/flight, or 24.5% of total NAS delays. More realistically, if only "local" delay sources in NYC were eliminated, the NAS delay reduction would be 2.4-2.8 minutes/flight, or 19.22%, depending on how we implement the concept. These estimates are reasonable. Table 6 Latency and latency reduction results for the SWAC simulation model, 12 days.

Daniel Alberto Pamplona, Li Weigang, Alexandre De Barros, Elcio Hideiti Shiguemori :

The purpose of this study is to present the application of artificial neural networks to a flight delay prediction model for the route between Sao Paulo (Congonhas) - Rio de Janeiro (Santos Dumont). Since the main contribution of this study is to configure the ANN using the random search technique to realize the hyper parameterization of the network.

Juan Jose Rebollo, Hamsa Balakrishnan:

This article presents a new model to predict air traffic delays. The proposed model considers both temporal and spatial (i.e. network) delay conditions as explanatory variables and uses a random forest algorithm to predict departure delays from 2 to 24 hours in the future. In addition to the local delay variables describing arrival or departure delays, indicate the most influential airports and links in the network (origin-destination pairs), the new network delay variables characterize the global delay state of the entire national airspace system in the forecast. offered at the time. This paper analyses the performance of the proposed predictive model to classify latency as above or below a certain threshold and predict latency values.

Leonardo Moreira, Christofer Dantas, Leonardo Oliveira:

This paper focuses on the unbalanced distribution of delay classes (presence and absence), experimentally evaluating various preprocessing methods to develop machine learning flight delay classification models. The models are built from a dataset that combines domestic flight operations with weather conditions at airports. Our results show that models applying the balancing technique performed better at predicting the occurrence of lags, obtaining a success rate of approximately 60%.

Juan Jose Rebollo, Hamsa Balakrishnan:

This article presents a new model to predict air traffic delays. The proposed model considers both temporal and spatial (i.e. network) delay conditions as explanatory variables and uses a random forest algorithm to predict departure delays from 2 to 24 hours in the future. In addition to the local delay variables describing arrival or departure delays, indicate the most influential airports and links in the network (origin-destination pairs), the new network delay variables characterize the global delay state of the entire national airspace system in the forecast. offered at the time. This paper analyses the performance of the proposed predictive model to classify latency as above or below a certain threshold and predict latency values. The models were trained and validated on operational data from 2007 and 2008, and evaluated using the 100 worst latency links of the system.

Young Jin Kim, Sun Choi, Simon Briceno and Dimitri Mavris:

From this study, we demonstrated that the deep architecture can improve the accuracy of airport delay prediction models. In particular, by applying a deep LSTM RNN architecture to a predictive model, a single reliable latency state of days can be obtained. The most accurate delay state for an individual flight is then obtained by introducing the delay state by one day into the individual flight delay model. It gives peak results to predict individual flight delays. 4044 The next step is to apply other deep architectures to the flight delay 4044 prediction and analysis task.



Shichao Zhang, Senior Member, IEEE, Xuelong Li, Fellow, IEEE, Ming Zong, Xiaofeng Zhu*, and Ruili Wang:

In this paper, we propose two new kNN classification algorithms, kTree and k*Tree methods, to select optimal k values for each test sample for effective and efficient kNN classification. The key idea of our proposed method is to design a training stage to reduce the running cost of the test stage and improve the classification performance.

Daniel Alberto Pamplona, Li Weigang, Alexandre Gomes de Barros, Elcio Hideiti Shiguemori:

The purpose of this study is to present the application of artificial neural networks to a flight delay prediction model for the route between Sao Paulo (Congonhas) - Rio de Janeiro (Santos Dumont). Since the main contribution of this study is the constructing an ANN for the using a random search technique, hyper parameterization of the network occurs.

III. EXISTING SYSTEM/OPEN ISSUES

Random forest is an ensemble learning method that uses decision trees as sub classifiers and introduces a random selection of features into the decision tree. The algorithm for constructing a random forest-based classifier is as follows: Note that D_i is generated by randomly subsampling the full training dataset, while A_j is randomly selected from the full feature set for construction. After the random forest builds the framework classification, the set classifier uses the result with the most votes among the N sub classifiers as its prediction. The power of each sub classifier and the independence of the sub classifiers together improve the accuracy of the model. The randomness of the under sampled training dataset to generate subsets, select features, and create sub classifiers largely avoids the probability of overfitting.

IV. CONCLUSION

In this paper, a random forest-based architecture was implemented to predict individual flight delays. The experimental results show that random forest-based methods can achieve good performance for binary classification tasks and that there is still room for improvement in multi-category classification tasks. In summary, the Random Forest-based architecture offers better adaptability at the cost of training accuracy when dealing with limited datasets. Future work will include training the model on the latest aeronautical data at selected airports. Test the model on real data provided by the aeronautical data server. Include other factors contributing to the delay in Assessment of Anticipated Delays.

REFERENCES

- [1]. M. Leonardi, "Intrusion Detection with Ads-b Anomaly and Sensor Surveillance Tracking," IEEE Trans. Aerosp. former. system, in press, doi:10.1109/TAES.2018.2886616.
- [2]. Yu. A. Nijssour, G. Caddum, G. Gagnon, F. Gagnon, K. Yuen and R. Ma hapatra, "Adaptive air-to-ground communication systems based on ads-b and global multilateralation", IEEE Trans. Milestone Technologies, Vol.65,5, p. 3150–3165, 2015. IJARSCT ISSN (Online) 2581-9429 International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume 3, Issue 2, February 2023 Copyright to IJARSCT DOI: 10.48175/IJARSCT-8384 236 www.ijarsct.co.in Impact Factor: 7.301
- [3]. J.A. F. Zuluaga, J.F.V. Bonya, J.D. O. Pabon and K. M. S. Rios, "Radar Error Calculation and Correction System Based on ads-b and Smart Business Tools", Proc. International Qanahan Conference. security technology., seconds. 1–5, IEEE, 2018.
- [4]. D.A. Pamplona, L. Weigang, A.G. (1999). Wright, E.H. Shiguemori, и c. J.P. Alves, "Supervised Neural Networks with Multi-Input Layers for Air Traffic Delay Prediction", Proc. International Jt. Meeting Neural Networks , pp. 1–6, IEEE, 2018.
- [5]. S. Manna, S. Biswas, R. Kundu, S. Rakshit, P. Gupta, and S. Barman, "Statistical Approach to Flight Delay Prediction Using the Gradient Boosted Decision Tree", Proc. International Conference Intel Account. data science , str. 101-1 1–5, IEEE, 2017.
- [6]. L. Moreira, C. Dantas, L.D., et al. Oliveira, J. Soares and E. Ogasawara, "A Evaluation of Preprocessing Techniques in Machine Learning Models for Flight Delay", Proc. International Jt. Meeting Neural Networks, pp. 1–8, IEEE, 2018.
- [7]. J. J. Rebollo and H. Balakrishnan, "Characterization and Prediction of Air Traffic Delay", Transp. Resolution Part C Technology, Vol. 44, s. 231–241, 2014.
- [8]. L. Hao, M. Hansen, Y. Zhang and J. Post, "New York, New York: Two Methods to Assess the Impact of New York Airport Delays", Transp. Part of resolution ELogist. Transportation Obr., Vol. 70, p. 245–260, 2014.



[9]. ANAC, "National Civil Aviation Authority of Brazil". anac.gov, 2017. [Online] Available at: <http://www.anac.gov.br/>.

[10]. S. Zhang, X. Li, M.Zong, X. Zhu and R. Wang, "Efficient knnclassification with different number of nearest neighbors", IEEE Trans.Neural Networks. study.sist., vol. 29, no. 5, p. 1774–1785, 2017.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details