APPLIED DATA SCIENCE

DEVELOPING A FLIGHT DELAY PREDICTION MODEL USING MACHINE LEARNING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

UNIVERSITY COLLEGE OF ENGINEERING BIT CAMPUS ANNA UNIVERSITY TIRUCHIRAPPALLI

PRESENTED BY

SUNDARESAN M - 810019106083

SUNDARESVAR P - 810019106084

SUREN GOPAL D - 810019106085

THULASI VEERARAGAVAN L -

810019106089

Khaksar, H., & Sheikholeslami, A. (2017). Airline delay prediction by machine learning algorithms. Scientia Iranica.

Flight planning, as one of the challenging issue in the industrial world, is faced with many uncertain conditions. One such condition is delay occurrence, which stems from various factors and imposes considerable costs on airlines, operators, and travelers. With these considerations in mind, we implemented flight delay prediction through proposed approaches that are based on machine learning algorithms. Parameters that enable the effective estimation of delay are identified, after which Bayesian modeling, decision tree, cluster classification, random forest, and hybrid method are applied to estimate the occurrences and magnitude of delay in a network. These methods were tested on a U.S. flight dataset and then refined for a large Iranian airline network. Results showed that the parameters affecting delay in US networks are visibility, wind, and departure time, whereas those affecting delay in Iranian airline flights are fleet age and aircraft type. The proposed approaches exhibited an accuracy of more than 70% in calculating delay occurance and magnitude in both the whole-network US and Iranian. It is hoped that the techniques put forward in this work will enable airline companies to accurately predict delays, improve flight planning, and prevent delay propagation.

Sun Choi, Young Jin Kim, Simon Briceno , Dimitri Mavris.
Prediction of weather-induced airline delays based on machine learning algorithms.

The primary goal of the model proposed in this paper is to predict airline delays caused by inclement weather conditions using data mining and supervised machine learning algorithms. US domestic flight data and the weather data from 2005 to 2015 were extracted and used to train the model. To overcome the effects of imbalanced training data, sampling techniques are applied. Decision trees, random forest, the AdaBoost and the k-Nearest-Neighbors were implemented to build models which can predict delays of individual flights. Then, each of the algorithms' prediction accuracy and the receiver operating characteristic (ROC) curve were compared. In the prediction step, flight schedule and weather forecast were gathered and fed into the model. Using those data, the trained model performed a binary classification to predicted whether a scheduled flight will be delayed or on-time.

Juan JoseRebollo , HamsaBalakrishnan. Characterization and prediction of air traffic delays.

This paper presents a new class of models for predicting air traffic delays. The proposed models consider both temporal and spatial (that is, network) delay states as explanatory variables, and use Random Forest algorithms to predict departure delays 2–24 h in the future. In addition to local delay variables that describe the arrival or departure delay states of the most influential airports and links (origin–destination pairs) in the network, new network delay variables that characterize the global delay state of the entire National Airspace System at the time of prediction are proposed. The paper analyzes the performance of the proposed prediction models in both classifying delays as above or below a certain threshold, as well as predicting delay values. The models are trained and validated on operational data from 2007 and 2008, and are evaluated using the 100 most-delayed links in the system. The results show that for a 2-h forecast horizon, the average test error over these 100 links is 19% when classifying delays as above or below 60 min. Similarly, the average over these 100 links of the median test error is found to be 21 min when predicting departure delays for a 2-h forecast horizon. The effects of changes in the classification threshold and forecast horizon on prediction performance are studied.

REFERENCES

Thiagarajan B, et al. A machine learning approach for prediction of on-time performance of flights. In 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC). New York: IEEE. 2017.

Esmaeilzadeh, E., & Mokhtarimousavi, S. (2020). Machine learning approach for flight departure delay prediction and analysis. Transportation Research Record: Journal of the Transportation Research Board, *2674*(8), 145–159.

Yu, B., Guo, Z., Asian, S., Wang, H., & Chen, G. (2019). Flight delay prediction for commercial air transport: A deep learning approach. Transportation Research Part E: Logistics and Transportation Review, *125*, 203–221.

Ye, B., Liu, B., Tian, Y., & Wan, L. (2020). A methodology for predicting aggregate flight departure delays in airports based on supervised learning. Sustainability, *12*(7), 2749.

Liu YJ, Cao WD, Ma S. Estimation of arrival flight delay and delay propagation in a busy hub-airport. In 2008 Fourth International Conference on Natural Computation. New York: IEEE. 2008.