

Literature Survey

Developing Flight Delay Prediction Model using Machine Learning

Analysis of flight delay and causal factors is crucial in maintaining airspace efficiency and safety. However, delay samples are not independent since they always show a certain aggregation pattern. Therefore, this study develops a novel spatial analysis approach to explore the delay and causal factors which is able to take dependence and the possible problem involved including error correlation and variable lag effect of causal factors on delay into account. The study first explores the delay aggregation pattern by measuring and quantifying the spatial dependence of delay. The spatial error model (SEM) and spatial lag model (SLM) are then established to solve the error correlation and the variable lag effect, respectively. Results show that the SEM and SLM achieve better fit than ordinary least square (OLS) regression, which indicates the effectiveness of considering dependence by employing spatial analysis. Moreover, the outcomes suggest that, aside from the well-known weather and flow control factors, delay-reduction strategies also need to pay more attention to reducing the impact of delay at the previous airport.

Paper 1: Sensor dynamics in high dimensional phase spaces via nonlinear transformations: Application to helicopter loads monitoring

Author: Julio J. Valdés; Catherine Cheung; Matthew Li

Year: 2014

Publisher: IEEE - 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)

Summary: By considering on loads in critical helicopter components like the main rotor blade normal bending in level flight, steady turn and rolling pull-out flight conditions exhibits a differential behaviour. Several techniques from the fields of computational intelligence and

nonlinear systems analysis coincide in identifying an increasing degree of complexity in the sequence level flight, steady turn and rolling pull-out manoeuvres. More flight records and more flight conditions must be analysed in order to have a better characterization of the flight dynamics.

References: <https://ieeexplore.ieee.org/document/7008691>

Paper 2: A Machine Learning approach for prediction of on time performance of flights.

Author: Balasubramanian Thiagarajan ; Lakshminarasimhan Srinivasan; Aditya Vikram Sharma ;Dinesh Sreekanthan ;Vineeth Vijayaraghavan

Year: 2017

Publisher: IEEE - 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)

Summary: A two-stage predictive model was developed to efficiently predict the departure and arrival delays of flights using flight schedule and weather features. Various supervised machine learning algorithms were implemented. It was found that the departure delay prediction had comparatively higher error rates due to a weak feature set. Furthermore, a Decision Support Tool was developed using the model to predict real-time flight delays. In the future, more data can be extracted by considering a larger number of airports over a longer time frame to improve the model and other deep architectures can also be implemented.

References: <https://ieeexplore.ieee.org/document/8102138>

Paper 3: Identification, Characterization, prediction of Traffic flow patterns in multi airport systems.

Author: Mayara Condé Rocha Murça; Robert John Hansman

Year: 2019

Publisher: IEEE - IEEE Transactions on Intelligent Transportation Systems

Summary: The data-driven framework is based on a sequential application of machine learning methods on historical flight tracks, weather forecasts and airport operational data. A multi-layer clustering analysis is performed to mine spatial and temporal trends in flight trajectory data for traffic flow pattern identification. The results revealed significant variability in throughput and delay performance for different metroplex configurations, emphasizing the importance of anticipating the behavior of the metroplex as a system when forecasting individual airport capacity. Future research goes along this direction by exploring the development of higher-fidelity models for airport capacity prediction that take as input detailed weather information and metroplex configuration forecasts in order to deliver probabilistic capacity forecasts for strategic TMI planning.

References: <https://ieeexplore.ieee.org/document/8373742>

Paper 4: A Novel Integration Platform to Reduce Flight Delays in the National Airspace System.

Author: Chuyang Yang; Zachary A. Marshall; John H. Mott

Year: 2020

Publisher: IEEE - 2020 Systems and Information Engineering Design Symposium (SIEDS)

Summary: By integrating various databases with existing NextGen's SWIM and FAA CDM and GDP programs and harnessing remote cloud computing of deep learning algorithms, precise and accurate flight delay forecasts are generated. Allowing for the full realization of potentials in schedule optimization, emission reduction, and resource utilization, the delay predictions provided by the proposed system could significantly grow airports' capabilities through improved operational efficiency. System principles, safety-risk, impacts, and sustainability assessments are presented in this paper. The swift implementation of the proposed system would effectively enhance the capacity of the NAS and further the goals of the FAA, connecting more people and ideas from all corners of the continent.

References: <https://ieeexplore.ieee.org/document/9106657>

Paper 5: Machine Learning Model - based Prediction of Flight Delay

Author: N Lakshmi Kalyani;G. Jeshmitha;Bindu Sri Sai U.;M. Samanvitha;J. Mahesh;B.V. Kiranmayee

Year: 2020

Publisher: IEEE - 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)

Summary: The flight's delay is estimated along with the estimation of delay time in minutes using machine learning algorithms namely Decision Tree Algorithm (XGBoost) and Linear regression. Data set of both flights and weather will be taken to compare with the given inputs and validate them by applying classification and Regression concepts of Machine Learning.

References: <https://ieeexplore.ieee.org/document/9243339>

Paper 6:

Author: Yushan Jiang; Yongxin Liu; Dahai Liu; Houbing Song

Year: 2020

Publisher: IEEE - 2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)

Summary: QCLCD and AOTP data are used to construct a new dataset with both flight information and weather condition. Then this dataset is further explored about some useful pattern toward flight delay. For the future work, we suggest to pre-design the arrangement of features when using encoding methods increasing the feature dimensions in order to leverage the advantage of convolution neural network on high dimensional feature space

References: <https://ieeexplore.ieee.org/document/9251206>

Paper 7: Flight Delay Prediction: Data Analysis and Model Development

Author: Azib Anees; Wei Huang

Year: 2021

Publisher: IEEE - 2021 26th International Conference on Automation and Computing (ICAC)

Summary: The proposed model gains insight into factors causing flight delays, cancellations and the relationship between departure and arrival delay using exploratory data analysis. In addition, Random Forest (RF) algorithm is used to train and test the big dataset to help the model development. A web application has also been developed to implement the model and the testing results are presented with the limitation discussed.

References: <https://ieeexplore.ieee.org/document/9594260>

Paper 8: Flight delay predictions and the study of its causal factors using machine learning algorithms.

Author: Cho Yin Yiu; Kam K. H. Ng; Kin Chung Kwok; Wing Tung Lee; Ho Tung Mo

Year: 2021

Publisher: IEEE - 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)

Summary: Several machine learning approaches were adopted in this research to predict flight delay, including the decision tree, random forest, k-nearest neighbour, Naïve Bayes, and artificial neural networks. The results show that all algorithms achieved more than 80% of accuracy and artificial neural networks perform the best among the alternatives. While Naïve Bayes is the least accurate, k-nearest neighbour have the lowest F 1 score.

References: <https://ieeexplore.ieee.org/document/9633571>

Paper 9: Flight Delay Prediction Using Machine Learning Algorithm XGBoost

Author: Subhani Shaik, KP Surya Tej

Year: 2019

Publisher: Researchgate

Summary: We are proposing machine learning algorithms like XGBoost regressed, Linear regression Techniques. The aim of this research work is to predict Flight Delay, which is highest economy producing field for many

countries and among many transportation this one is fastest and comfort, so to identify and reduce flight delay.

References:

https://www.researchgate.net/publication/344227817_Flight_Delay_Prediction_Using_Machine_Learning_Algorithm_XGBoost

Paper 10: Flight delay prediction for commercial air transport: A deep learning approach

Author: BinYu^a, ZhenGuo^a, SobhanAsian^a, HuaizhuWang^b, GangChen^{de}

Year: 2019

Publisher: ScienceDirect

Summary: The proposed method has proven to be highly capable of handling the challenges of large datasets and capturing the key factors influencing delays. This ultimately enables connected airports to collectively alleviate delay propagation within their network through collaborative efforts

References:

<https://www.sciencedirect.com/science/article/abs/pii/S136655451831197>

Paper 11:

Author: H. Khaksar ,A. Sheikholeslami

Year: 2019

Publisher: Scientia Iranica

Summary: 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 occurrence 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.

References: http://scientiairanica.sharif.edu/article_20020_0.html

Paper 12:

Author: Hatipoğlu, Irmak; Tosun, Ömür; Tosun, Nedret

Year: 2022

Publisher: Ebscohost

Summary: There is only one way to prevent these problems before they occur, and that is to know which flights will be delayed. The aim of this study is to predict delayed flights. For this, the use of machine learning techniques, which have become widespread with the development of computer capacities and data storage systems, is preferred. Methods: Estimations are made with three up-to-date techniques XGBoost, LightGBM, and CatBoost techniques based on Gradient Boosting from machine learning techniques. The bayesian technique is used for hyper-parameter settings. In addition, the Synthetic Minority Over-Sampling Technique (SMOTE) technique is also used, as the majority of flights are on time and delayed flights, which constitute a minority class, may adversely affect the results.

References:

<https://web.p.ebscohost.com/abstract?direct=true&profile=ehost&scope=sit&authtype=crawler&jrnl=18952038&AN=155638803&h=rIVjuOvl2enh%2fgJqjEbsfoJUOWN%2b9%2fBfNYeMhXFOPwEJEd5wCaXw%2fgmele3q%2fiC5PglX6PqXV9I66qDvnAbb%2fw%3d%3d&crl=c&resultNs=AdminWebAuth&resultLocal=ErrCrlNotAuth&crlhashurl=login.aspx%3fdirect%3dtrue%26profile%3dehost%26scope%3dsite%26authtype%3dcrawler%26jrnl%3d18952038%26AN%3d155638803>