LITERATURE SURVEY

1. Applying Machine Learning to Aviation Big Data for Flight Delay Prediction

Authors: Yushan Jiang, Liu Yongxin, Dahai Liu, Houbing Song

Year: 2021

Link: https://ieeexplore.ieee.org/document/9389876

Flight delay has been a serious and widespread problem that needs to be solved. One promising solution is the flight delay prediction. Although big data analytics and machine learning have been applied successfully in many domains, their applications in aviation are limited. This paper presents a comprehensive study of flight delay spanning data pre-processing, data visualization and data mining, in which we develop several machine learning models to predict flight arrival delays. Two data sets were used, namely Airline On-Time Performance (AOTP) Data and Quality Controlled Local Climatological Data (QCLCD). This paper aims to recognize useful patterns of the flight delay from aviation data and perform accurate delay prediction. The best result for flight delay prediction (five classes) using machine learning models is 89.07% (Multilayer Perceptron). A Convolution neural network model is also built which is enlightened by the idea of pattern recognition and success of neural network method, showing a slightly better result with 89.32% prediction accuracy.

2. Learning Machine Model - based Prediction of Flight Delay

Authors: Lakshmi Kalyani, G.Jeshmitha, Bindu Sri Sai U, M.Samanvitha, J.Mahesh, B.V.Kiranmayee

Year: 2020

Link: https://ieeexplore.ieee.org/document/9243339

Prior prediction of flight arrival delays is necessary for both travelers and airlines because delays in flights not only trigger huge economic loss but also airlines end up losing their reputation that was built for several years and passengers lose their valuable time. Our paper aims at predicting the arrival delay of a scheduled individual flight at the destination airport by utilizing available data. The predictive model presented in this work is to foresee airline arrival delays by employing supervised machine learning algorithms. US domestic flight data along with the weather data from July 2019 to December 2019 were acquired and are used while training the predictive model. XG Boost and linear regression algorithms were applied to develop the predictive model that aims at predicting flight delays. The performance of each algorithm was analyzed. Flight data along with the weather data was given to the model. Using this data, binary classification was carried out by the XG Boost trained model to predict whether there would be any arrival delay or not, and then linear regression model predicts the delay time of the flight.

3. Multi-Task Local-Global Graph Network for Flight Delay Prediction

Authors: Tianyi Wang, Shu-Ching Chen

Year: 2022

Link: https://ieeexplore.ieee.org/document/9874045/authors#authors

Airline on-time performance has always been a key factor in evaluating the punctuality of the civil aviation industry and has a profound impact on airlines, airports, and passengers. As a result, there have been increasing demands for the systematic analysis of flight delays and the development of accurate and efficient tools for flight delay prediction. In this paper, a deep learning framework based on graph convolutional networks and multi-task learning is proposed for flight delay prediction. We first use graph convolutional networks to capture the local and global spatial dependencies among the airports. A multi-decoder sequence-to-sequence model is developed to extract the temporal correlation from the data. We further apply a hierarchical graph fusion approach to combine features at different levels of the network to exploit their cross-modality correlations. The model is trained using a dynamic multi-task learning strategy to predict flight arrival and departure delays at the same time to boost the model's generalization and performance. The proposed model is evaluated on a large-scale public flight record dataset against several state-of-the-art methods. The experimental results demonstrate that our model can outperform all baseline methods in predicting short to medium-term flight delays.

4. Outlier Analysis of Airport Delay Distributions in US and China

Author: Jingyi Qu, Jinjie Zhang

Year: 2021

Link: https://ieeexplore.ieee.org/document/9389876/authors#authors

By analyzing the historical data of flight delay, the delay problem can be effectively prevented and controlled. The spread of delay propagation in continuous flights often leads to large-scale flight delays. Aiming at the problems of flight delay propagation, a flight chain model is proposed to evaluate flight delays. Because flight data has the characteristics of time and space, the flight delay visualization analytics system is put forward to directly interpret the spatiotemporal pattern of flight delay. The time-varying delay trend is shown by the time line chart and sequence histogram. At the same time, it analysis the time sequence of the impact of flight delays on the consecutive flights through the thermal sequence diagram of flight and parallel coordinate chart. The spatial model reflects the spatial distribution rule and spatial clustering results of delay in airports in different regions by thermal map and rectangular tree diagram. Simultaneously, through the tree diagram of flight propagation delay from the perspective of spatial spread, the delay sweep problems are intuitively analyzed. According to the results of the delay spread assessment and visualization, it can provide corresponding suggestions for various departments to control the delay propagation.

5. State-Of-Art Machine Learning Techniques to Predict Airlines Delay

Authors: Sai Sharan, M.Sriniketh, Harsha Vardhan, Dannana Jayanth

Year: 2021

Link: https://ieeexplore.ieee.org/document/9702590/authors#author

Nowadays everyone is becoming extremely busy that makes them follow the time very precisely. In the commercial aviation sector, flight delays are a significant cause of dissatisfaction with customers. So, the prediction of flight delays plays a pivotal role in travelers comfort and alleviates the airline's economic losses. This paper analyzes the performance of the machine learning algorithms such as Random Forest, Ada Boost, and XG Boost classifier to handle the delay time prediction of flight by considering multiple parameters such as weather conditions, flight schedule, etc., that are responsible for flight delay. The paper does a detailed comparative analysis of the algorithms used. Our study can also be applied to various other applications, such as predicting demand-based airline fares.

6. Predicting Flight Delay Risk Using a Random Forest Classifier Based on Air Traffic Scenarios and Environmental Conditions

Authors: Markus Bardach, Eduard Gringinger, Michael Schrefl, Christoph G. Schuetz

Year: 2020

Link: https://ieeexplore.ieee.org/document/9256474/authors#authors

A reduction of delay costs can be achieved through more adaptable flight planning, which hinges on accurate prediction of delays. In order to counteract the expected delay of flights, air traffic control may adapt flight plans through slot swapping, opening another runway, or changing the runway configuration, for example. Environmental conditions and external events such as runway and airspace closures may render a flight plan obsolete, which must be taken into account when aiming to reduce delay. Air traffic control must recognize changes in the environment and external events such as runway and airspace closures as early as possible in order to adapt flight plans accordingly and avoid delays. Current systems employed by air traffic control do not sufficiently leverage the multitude of available data for the detection of upcoming congestion and, consequently, flight delays. Therefore, flight plans are not adapted fast enough in air traffic scenarios with potentially high delay. In this paper, we aim to predict the risk class of an air traffic scenario based on the expected cost of the delays, and considering information about environmental conditions and external events. In particular, we present a random forest classifier for Atlanta International Airport, which achieves an accuracy of 82.5% for the highest and thus most important risk classes.

7. A Computational Intelligence-Based Prediction Model for Flight Departure Delays

Authors: Johanna Hopane, Barnabas Gatsheni

Year: 2019

Link: https://ieeexplore.ieee.org/document/9071393/authors#authors

Flight departure delays are a major problem at OR Tambo International airport (ORTIA) located in Johannesburg in South Africa. These delays are more pronounced at the beginning and end of the month. Flight delays at ORTIA do impact negatively on business, on job opportunities and on tourists. Machine learning algorithms namely Decision Trees (J48), Support Vector Machine (SVM), KMeans Clustering (K-Means) and Multi Layered Perceptron (MLP) were used to construct the flight departure delays prediction models. Cross-validation (CV) was used for evaluating the models. The best prediction model was selected by using a confusion matrix and the ROC curve. The results show that the models constructed using data and the Decision Trees is suited for flight departure delay prediction as it gave the best prediction of 67.144%. The implications of the model is that travellers wishing to travel from ORTIA can foretell the flight departure delays using the tool. The tool will allow the travellers to enter variables such as month, week of month, day of week and time of day.

8. A deep learning approach to flight delay prediction

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

Year: 2016

Link: https://ieeexplore.ieee.org/document/7778092/authors#authors

Deep learning has achieved significant improvement in various machine learning tasks including image recognition, speech recognition, machine translation and etc. Inspired by the huge success of the paradigm, there have been lots of tries to apply deep learning algorithms to data analytics problems with big data including traffic flow prediction. However, there has been no attempt to apply the deep learning algorithms to the analysis of air traffic data. This paper investigates the effectiveness of the deep learning models in the air traffic delay prediction tasks. By combining multiple models based on the deep learning paradigm, an accurate and robust prediction model has been built which enables an elaborate analysis of the patterns in air traffic delays. In particular, Recurrent Neural Networks (RNN) has shown its great accuracy in modeling sequential data. Day-to-day sequences of the departure and arrival flight delays of an individual airport have been modeled by the Long Short-Term Memory RNN architecture. It has been shown that the accuracy of RNN improves with deeper architectures. In this study, four different ways of building deep RNN architecture are also discussed. Finally, the accuracy of the proposed prediction model was measured, analyzed and compared with previous prediction methods. It shows best accuracy compared with all other methods