Literature Review

1. Flight Delay Prediction Based on Aviation Big Data and Machine Learning

Authors: Guan Gui, Fan Liu, Jinlong Sun, Jie Yang, Ziqi Zhou

Date of Publication: 18 November 2019

Accurate flight delay prediction is fundamental to establish the more efficient airline business. Recent studies have been focused on applying machine learning methods to predict the flight delay. Most of the previous prediction methods are conducted in a single route or airport. This paper explores a broader scope of factors which may potentially influence the flight delay, and compares several machine learning-based models in designed generalized flight delay prediction tasks. To build a dataset for the proposed scheme, automatic dependent surveillance-broadcast (ADS-B) messages are received, pre-processed, and integrated with other information such as weather condition, flight schedule, and airport information. The designed prediction tasks contain different classification tasks and a regression task. Experimental results show that long short-term memory (LSTM) is capable of handling the obtained aviation sequence data, but overfitting problem occurs in our limited dataset. Compared with the previous schemes, the proposed random forest-based model can obtain higher prediction accuracy (90.2% for the binary classification) and can overcome the overfitting problem.

2. A deep learning approach to flight delay prediction

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

Date of publication: 12 December 2016

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.

3. Airline delay prediction by machine learning algorithms

Authors: H. Khaksar, A. Sheikholeslami

Date of publication: 28 January 2017

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.

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

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

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 predict whether a scheduled flight will be delayed or on-time.