LITERATURE SURVEY

Date	November 7, 2022
Team ID	PNT2022TMID06101
Project Name	Developing a flight delay prediction model using machine learning
Maximum Marks	4 Marks

1. Flight Delay Prediction System - Yogita Borse , Dhruvin Jain , Shreyash Sharma , Viral Vora,Aakash Zaveri (2020)

Abstract:

As discussed, considering the standard taxonomy of the flight delay and its problems, one will contemplate the scope of prediction to be one in every of these factors or combination of those factors[3]. The models developed during this system may be applied to predict the incidence of flight delay at airports. Such prognosticative capabilities would facilitate traffic managers and airline dispatchers to organize mitigation methods for reducing traffic disruptions.

This issue can be reduced by developing the flight delay prediction tool which can be developed using following methods.

Statistical analysis

Statistical model requires the use of correlation analysis, parametric and non parametric tests, multivariate analysis and econometric models. Government agencies have invested in these econometric models to understand the relationship between delay and Passenger demand, fare, size of aircraft etc

Probabilistic models

Probabilistic model requires analysis tools that estimates the probability of an event based on the historic data. The estimated outcome is given inform of a distribution function of the probability. The factor of randomness always makes an impact on the decision or the outcome produced by the probabilistic model.

2. A deep learning approach to flight delay prediction - Young Jin Kim; Sun Choi; Simon Briceno; Dimitri Mavris(2016)

Abstract:

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 airtraffic 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. Research on Flight Delay Prediction Based on Random Forest -

Peng Hu; Jianping Zhang; Ning Li(2021)

Abstract:

Based on the random forest model, this paper proposes a flight delay prediction model. By analyzing the departure flight data of Guangzhou Baiyun International Airport in June 2020, and selecting the data of ten landing airports, it analyzes the distribution of delayed, punctual, and early arrived. It studies the selection of features that impact on flight delays, and establishes random forest predictions model. Through case study, it researches the mean square error of different leaf sizes when the forest scale is 50 trees. The results show that the optimal leaf size is 5, and the minimum mean square error is 0.1096. And it analyzes the importance of features such as departure flight delay time, scheduled flight time, number of scheduled departure flights on the day, date, and landing airport. The research results also found that, when the forest size is 100 trees and the leaf size is 5, the out-of-bag mean square error is 0.1090, and the accuracyof the prediction model is high, which is close to 90%.