Flight Delay Prediction Based on Aviation Big Data and Machine Learning

ABSTRACT

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 over fitting 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 over fitting problem.

**EXISTING SYSTEM**

* Nowadays, aircrafts have become a necessity because they easy life. They are efficient in carrying goods and passengers around the world. It also supplies emergencies in warfare and takes a vital role in carrying medical necessities. Hence, advent of airplanes is considered important. Delays in aircrafts can affect thousands of people across the globe either directly or indirectly. There are a lot of reasons of delays in aircrafts such as critical weather, security issues, traffic and many more.
* There are several methods implemented in the existing system to predict the flight delays but due to various complexities of the ATFM and the huge datasets involved, it has become very difficult to find an accurate solution for this complication. Many algorithms have been implemented to forecast flight delays. We are using Python in Visual Studio Code. We implement Binary Classification to prepare a model that can predict the delays.

Disadvantages

* + In the existing system, the system is not using Data Transformation and Balancing.
  + This system is less performance due to lack of Data Cleaning and Data Integration.

**PROPOSED SYSTEM**

* The proposed work benefits from considering as many factors as possible that may potentially influence the flight delay. For instance, airports information, weather of airports, traffic flow of airports, traffic flow of routes. The contributions of this paper can be summarized as follows:
* The system explores a broader scope of factors which may potentially influence the flight delay and quantize those selected factors. Thus we obtain an integrated aviation dataset. Our experimental results indicate that the multiple factors can be effectively used to predict whether a flight will delay.
* Several machine learning based-network architectures are proposed and are matched with the established aviation dataset. Traditional flight prediction problem is a binary classification task. To comprehensively evaluate the performance of the architectures, several prediction tasks covering classification and regression are designed.Conventional schemes mostly focused on a single route or a single airport [4], [6], [12]. However, our work covers all routes and airports which are within our ADSB platform.

**Advantages**

* Proposed methods implementing ADS-B Message Based Aviation Big Data Platform which is more effective and fast.
* ADS-B system is a communication and surveillance integrated system for air traffic management (ATM) where flights periodically broadcast location and other information on the same frequency band.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).