FLIGHT DELAY PREDICTION FOR AVIATION INDUSTRY USING MACHINE LEARNING

MILESTONE 1: define problem / problem understanding

ACTIVITY 1: specify the business problem

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.

KEYWORDS:
Air transportFlight delay predictionDelay propagationAdvanced transportation informationDeep belief learning

ACTIVITY 2: Business requirements

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**.**

ACTIVITY 3: Literature survey (student will write)

Varsha Venkatesh, Arti Arya, Pooja Agarwal, S Lakshmi, Sanjay Balana

2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), 562-567, 2017

The aviation industry is one of the largest industries in the transportation sector.
One of the major problems faced in this industry is the prevalence of delays.These delays not only cause dissatisfaction among the passengers but also huge losses for airlines. In this study, an approach using gated recurrent neural networks is explored for the purpose of predicting delays encountered in the aviation industry.
The proposed approach uses Long Short-Term Memory(LSTM) architectures such as vanilla LSTM and Bi-directional LSTM to empirically evaluate its efficiency in predicting the delay for a future flight. While previous works have explored this approach for classifying the delay status as either delay or no-delay, we take this a step further by predicting the amount of delay that will be experienced. The trained models were tested on an unseen test data set and were able to predict the delay with an error value of 33 minutes.In the aviation industry, flight arrival delays cause approximately 18 billion of loss to customers as stated in the literature. So, it becomes inevitable on the part of the aviation authorities to predict such delays and take necessary action to fix this loss for customer satisfaction.
In this paper, an approach based on machine learning techniques is proposed that predicts the flight arrival delays considering input parameters ranging from distance to their corresponding weather details to make a decision of whether the specific flight is delayed or not. It makes use of neural networks and deep learning concepts to estimate flight delay.The proposed approach is tested on real world flight big dataset that gives an accuracy of 77% using deep nets and 89% using neural nets. This approach can achieve reliable prediction with respect to if flight arrival delay is to be expected or not, moving forward the use of such a model can come in handy not only for airline administrators but also the passengers who can rearrange their schedules and arrange accommodation.

ACTIVITY 4: social or business impact

The use of machine learning to predict flight delays can have significant social and business impacts in the aviation industry.

From a social perspective, flight delays can cause a great deal of inconvenience for passengers. Passengers may miss connecting flights, important appointments, or events, resulting in frustration and potentially monetary losses. By predicting flight delays, airlines can proactively inform passengers and allow them to make alternative arrangements, minimizing disruptions and improving the overall travel experience.

From a business perspective, predicting flight delays can help airlines optimize their operations, reduce costs, and improve customer satisfaction. By identifying potential delays in advance, airlines can adjust schedules, allocate resources more efficiently, and minimize the impact of delays on their bottom line. Additionally, by improving on-time performance, airlines can improve their reputation and gain a competitive edge in the market.

Overall, the use of machine learning to predict flight delays has the potential to improve both the customer experience and the operational efficiency of the aviation industry, leading to positive social and business impacts.