LITERATURE REVIEW ON FLIGHT DELAY PREDICTION MODEL USING MACHINE LEARNING

ABSTRACT

Flight delays hurt airlines, airports, and passengers. Their prediction is crucial during the decision making process for all players of commercial aviation. Moreover, the development of accurate prediction models for flight delays became cumbersome due to the complexity of air transportation system, the number of methods for prediction, and the deluge of flight data. In this context, this paper presents a thorough literature review of approaches used to build flight delay prediction models from the Data Science perspective.

We propose a taxonomy and summarize the initiatives used to address the flight delay prediction problem, according to scope, data, and computational methods, giving particular attention to an increased usage of machine learning methods. Besides, we also present a timeline of significant works that depicts relationships between flight delay prediction problems and research trends to address them.

INTRODUCTION

As the air travels have a significant role in economy of agencies and airports, it is necessary for them to increase quality of their services. One of the important modern life challenges of airports and airline agencies is flight delay. In addition, delay in flight makes passengers concerned and this matter causes extra expenses for the agency and the airport itself.

In 2007, U.S government had endured 31–40 billion dollar downsides due to flight delays. In 2017, 76% of the flights arrived on time. Where, in comparison to 2016, the percentage of on time flights decreased by 8.5%. As some of the reasons of flight delays the following can be mentioned: security, weather conditions, shortage of parts and technical and airplane equipment issues and fight crew delays. Delay in flight is inevitable, which has too much negative economic effects on passengers, agencies and airport. Furthermore, delay can damage the environment through fuel consumption increment and also leads to emission of pollutant gases. In addition, the delay affects the trade, because goods' transport is highly dependent to customer trust, which can increase or decrease the ticket sales, so that on time flight leads to customer's confidence. So that, flight prediction can cause a skillful decision and operation for agencies and airports, and also a good passenger

information system can relatively satisfy the customer.

According to abundant and diversity of reasons for flight delays, We are faced with a massive amount of data which is not possible to be processed through previous methods of data analysis like classification, or the decision tree and machine learning based methods to process this volume of data are not proper, because characteristics of older intelligent system has been designed by human and usually were personalized, also people rarely perceive some features and usually neglect these matters. On the other hand, in older learning process, as the number of categories available for classification increases, the level of difficulty increases and extraction of important and effective features becomes relatively impossible. Due to complexity and effect of parameters on each other, the problem of flight delay prediction is considered as NP Complete. Furthermore, the problem essentially is accompanied by oscillation and also these are considered as non-linear problems. On the other hand, applied data includes noise and error that should be handled to cope with the problem.

LITERATURE REVIEW

Nowadays, service quality plays an important role in attracting customers. Among these, air travels have their special customers and the most important matter in these travels is the flight time, on time arrival at destination for passengers such those who have an important meeting, that has been leading to high expenses for the passenger until they get to their destination on time. Flight delay has negative economic effects on the passenger, agencies and airport. Therefore, any reduction of these effect requires decreasing postponed flight price, so that prediction or estimation has a great significance and numerous studies has been to dedicated this subject. Correspondingly, all the scientists have tried to design a model that understands effective factors and computes effect of each factor and their relation. Overall, the prediction methods are classified into five groups including Statistical Methods, Probability methods, network-based methods, operational methods and machine learning methods .

In one of the best studies that has been performed based on statistics delay time has been considered to be reduced. Their study has investigated important factors before fly and those which occur on the ground. In the next step, it has predicted the delay at destination based on factors that occur in the vicinity of arrival time at destination. Eventually, results have shown that whenever, the delay is correctly predicted, passenger disaffection and fuel consumption decrease and consequently number of flight increases. Moreover, it is possible to increase the agencies benefits through reducing number of passengers who wrongly selected their routes or specifying the probabilities for some flights and optimizing delay time

prediction.

Another prominent investigation based on Probability has been done and the author believes that huge storm in U.S.A has highly affected the flight delay. This study has been devoted to predict delay based on mathematical calculations and through considering delay time duration of the flights that had been engaged to storm in the same day. Metrological reports have shown the effect of storm one hour before and after event cause ephemeral climate at the region. In the next step, Monte-Carlo simulation has been used to estimate the airport runway capacity, so that traffic of each runway would have been estimated. As the research has employed only one factor, the model has not enough accuracy, but it is possible to increase region air capacity path structure.

CONCLUSION

Predicting flight delays is on interesting research topic and required many attentions these years. Majority of research have tried to develop and expand their models in order to increase the precision and accuracy of predicting flight delays. Since the issue of flights being on-time is very important, flight delay prediction models must have high precision and accuracy.