Flight Delay Prediction For Aviation Industry Using Machine Learning

Milestone 1: define problem/problem understanding

Business problem;

Flight delay is a significant problem that negatively impacts the aviation industry and costs billions of dollars each year. Most existing studies investigated this issue using various methods based on applying machine learning methods to predict the flight delay. However, due to the highly dynamic environments of the aviation industry, relying only on single route of airport may not be sufficient and applicable to forecast the future of flights. The purpose of this project is to analyze a broader scope of factors which may potentially influence the flight delay it compares several machine learning-based models in designed generalized flight delay prediction tasks. In this project we have used flight delay dataset from US Department of Transportation (DOT) to predict flight delays. We have used supervised learning algorithms to predict flight departure delay and then model evaluation is done to get best model and our model can identify which features were more important when predicting flight delays.

Business requirement;

Consistently a significant number of flights are delayed or crossed out because of numerous of reasons. These delays

bother travelers. These delays also cost a lot to the aircraft organization. Flight delays negatively affect carriers, air terminals and travelers. There are different methodologies used to manufacture flight delays expectation models from the Data Science point of view. The key resource of a flight includes aircraft, cockpit crew and cabin crew. For purposes of dispatching resources effectively, the three resources may be distributed independently. If the initial flight of a flight plan is delayed due to bad weather or other factors, it may result in the delays of the directly downstream flights that need to await its resources. If the delays continue to spread to the lower downstream flights, it may result in large area delay propagation. The method proposed here introduces and summarizes the initiatives used to address the flight delay prediction problem, according to scope, data and computational methods, giving special attention to an increasing usage of machine learning methods. 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.

Literature survey;

Mohammed Ayaz Hussain Khan, BE, Department of CSE, ayazhussaink1@gmail.com. Mohammed Farhan Uddin, BE,

Department of CSE, Mohdfarhan2001624@gmail.com. Growth in aviation industries has resulted in air-traffic jamming causing flight delays. Flight delays not only have economic impact but also injurious environmental properties. Air-traffic supervision is becoming increasingly challenging. Airlines delays make immense loss for business field as well as in budget loss for a country, there are so many reasons for impede in flights some of them are, some of them are due to security issues, mechanical problems, due to weather conditions, Airport congestion etc. we are proposing machine learning algorithms like Random Forest, Decision Tree, MLP Classifier, Naive Bayes, KNN, Gradient Boosting Classifier, Voting Classifier, SVM, Logistic Regression, Ridge Regression and Neural Network Techniques. The aim of this research work is to predict Flight Delay, Which is highest economy producing field for many countries and among many transportation this one is fastest and comfort, so to identify and reduce flight delays, can dramatically reduce the flight delays to saves huge amount of turnovers, using machine-learning algorithms. random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regeression.* Rebollo JJ, Balakrishnan H. Characterization and prediction of air traffic delays. Transportation Res Part C Emerg Technol. 2014;44:231–41. * Thiagarajan B, et al. A machine learning approach for prediction of on-time performance of flights. In

2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC). New York: IEEE. 2017

Social or business impact;

This study is carried by a long-term prediction approach that can be implemented long before a flight's departure. So, features as a result of events occurring close to the flight time, as the arrival time of the aircraft, did not include in the models. Also, while similar studies in flight delay prediction tend to benefit from data related to flights, in this study it is attempted to enrich the models with data related to aircraft and passengers and cargo. Hopefully, this study with other increasing numbers of studies in this subject contributes to improving on-time performances of flights so that both airline customers, airline personnel, airlines itself can benefit from time left to themselves saved from delays. Also shrinking extra fuel consumption would be important for the environment.

Business impact:

Finally rebound of opportunity costs of delays can add value to the economy. Flight delays are gradually increasing and bring more financial difficulties and customer dissatisfaction to airline companies. To resolve this situation, supervised machine learning models were implemented to predict flight delays.