Developing Flight Delay Prediction Model Using Machine Learning

Team Leader: K.Manishankar
Team Members: Anishvikram.B, Kiran Kannan, Aswin.S

Problem Statement

A flight delay is when an airline flight takes off and/or lands later than its scheduled time. The Federal Aviation Administration (FAA) considers a flight to be delayed when it is 15 minutes later than its scheduled time. A cancellation occurs when the airline does not operate the flight at all for a certain reason. This has led to phenomenal growth in air traffic and on the ground. An increase in air traffic growth has also resulted in massive levels of aircraft delays on the ground and in the air. These delays are responsible for large economic and environmental losses. The main objective of the model is to predict flight delays accurately in order to optimize flight operations and minimize delays. The problem of flight delay prediction is approached most often by predicting a delay class or value. However, the aviation industry can benefit greatly from probabilistic delay predictions on an individual flight basis, as these give insight into the uncertainty of the delay predictions.

Therefore, in this study, two probabilistic forecasting algorithms, Mixture Density Networks and Random Forest regression are applied to predict flight delays at a European airport. Techniques such as Decision Trees, AdaBoost, and KNearest Neighbors are used for predicting individual flight delays. Binary classification performed by the model helps predict the scheduled flight delay. These algorithms estimate well the distribution of arrival and departure flight delays with a Mean Absolute Error of less than 15 min. To illustrate the utility of the estimated delay distributions, integration of these probabilistic predictions into a probabilistic flight-to-gate assignment problem is done. The objective of this problem is to increase the robustness of flight-to-gate assignments. Considering probabilistic delay predictions, the proposed flight-to-gate assignment model reduces the number of conflicted aircraft by up to 74% when compared to a deterministic flight-to-gate assignment model. To summarize, the results illustrate the utility of considering probabilistic forecasting for robust airport operations' optimization.