

Proposed Solution

Proposed Solution

The dataset is analysed and preprocessed to remove/replace unnecessary entries using standard techniques.

Various classification models have been built to consider factors like arrival date, departure date, flight number, distance etc. Models like regression, KNN, Decision tree etc have been trained and an ensemble of the best performing models is built to further improve accuracy of the predictions. The models have been trained and tested successfully using the given dataset.

A UI layer is built using Python Flask application and details like Departure date, arrival date and distance is taken from the user and the data is fed into the final proposed ML model. The predictions are evaluated by the model and is returned to user.

Feasibility of idea

Ensembling of many base classifier models make the final model robust and more accurate. Thus all stakeholders can have their requirements met

Social impact

Accurate predictions of flight delays can lead to customer satisfaction, improves business for the airline and also improves productivity of aviation systems. This leads to a positive impact on a country's economy.

Scalability of solution

The model can be further scalable by training the proposed model on a larger dataset collected from one or more airports. Also, more powerful tools/computers can be used to predict results more effectively in lesser time.