

**Project Design Phase-I**  
**Proposed Solution Template**

Date	24 September 2022
Team ID	PNT2022TMID40519
Project Name	Project – Developing a Flight Delay Prediction Model using Machine Learning
Maximum Marks	2 Marks

**Proposed Solution Template:**

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	<p>Flight delays can be very annoying to airlines, airports, and passengers. Moreover, the development of accurate prediction models for flight delays became very difficult due to the complexity of air transportation flight data. In this project, we try to resolve this problem with approaches used to build flight delay prediction models using BPN and Radial Basis Function. 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. The data set that records information of flights departing from JFK airport during one year was used for the prediction. Seven algorithms (Logistic Regression, K-Nearest Neighbor, Gaussian Naïve Bayes, Decision Tree, Support Vector Machine, Random Forest, and Gradient Boosted Tree) were trained and tested to complete the binary classification of flight delays. The evaluation of algorithms was fulfilled by comparing the values of four measures: accuracy, precision, recall, and f1-score. These measures were weighted to adjust the imbalance of the selected data set. The comparative analysis showed that the Decision Tree algorithm has the best performance with an accuracy of 0.9777, and the KNN algorithm has the worst performance with an f1-score of 0.8039. Tree-based ensemble classifiers generally have better performance over other base classifiers.</p>
2.	Idea / Solution description	<p>As discussed, weather condition plays an important role in proper and timely functioning of flights. We propose a flight delay prediction system which focuses mainly on predicting delay of a flight based on the weather situation. To make the system more scalable it is necessary to</p>

		choose an algorithm which considers all the parameters to be independent.
3.	Novelty / Uniqueness	predicting flight delays can improve airline operations and passenger satisfaction, which will result in a positive impact on the economy. In this study, the main goal is to compare the performance of machine learning classification algorithms when predicting flight delays.
4.	Social Impact / Customer Satisfaction	Flight delays not only irritate air passengers and disrupt their schedules but also cause a decrease in efficiency, an increase in capital costs, reallocation of flight crews and aircraft, and additional crew expenses. So using this flight delay prediction will be used for peoples or passengers
5.	Business Model (Revenue Model)	Business model requires analysis tools that estimates the probability of an event based on the historic data. The estimated outcome is given in form of a distribution function of the probability. The factor of randomness always makes an impact on the decision or the outcome produced by the business model.
6.	Scalability of the Solution	The results show a high accuracy in prediction of delays above a given threshold. For instance, with a delay threshold of 60 minutes we achieve an accuracy of 85.8% and a delay recall of 86.9%. We also consider the effects on performance of varying model parameters, such as the classification threshold or the number of weather observations used. The goal of this work is to implement a predictor of the arrival delay of a scheduled flight due to weather conditions.