

## FLIGHT DELAY PREDICTION USING MACHINE LEARNING.

### Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

#### Activity 1: Define Problem Statement

Problem Statement:

The objective is to create a machine learning model that can accurately predict whether a flight will be delayed based on a variety of features. These features may include flight details, weather conditions, airport traffic, and other relevant factors. The model's predictions can be used by airlines to optimize scheduling, manage resources, and provide timely information to passengers

Ref. template: [Click Here](#)

Flight Delay Problem Statement Report: [Click Here](#)

#### Activity 2: Project Proposal (Proposed Solution)

Flight delays are a common and frustrating issue for airlines and passengers alike. Predicting flight delays can significantly improve operational efficiency for airlines and enhance the travel experience for passengers. This proposal aims to develop a predictive model to forecast flight delays using historical flight data and various external factors

Ref. template: [Click Here](#)

Flight Delay Project Proposal Report: [Click Here](#)

#### Activity 3: Initial Project Planning

Initial project plan, the team can systematically approach the development and deployment of a flight delay prediction model, ensuring that all key aspects are addressed and managed effectively.

Ref. template: [Click Here](#)

Flight Delay Project Planning Report: [Click Here](#)

## Milestone 2: Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant flight data from Kaggle, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

### Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

The dataset for "Flight Delay Prediction dataset " is sourced from Kaggle. It includes applicant details and financial metrics. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive.

Ref. template: [Click Here](#)

**Flight Data Prediction Data Collection Report:** [Click Here](#)

### Activity 2: Data Quality Report

The dataset for "Flight Delay Prediction " is sourced from Kaggle. It includes applicant details and financial metrics. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Ref. template: [Click Here](#)

**Flight Delay Prediction Data Quality Report:** [Click Here](#)

### Activity 3: Data Exploration and Preprocessing

Data Exploration involves analyzing the loan applicant dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These crucial steps enhance data quality, ensuring the reliability and effectiveness of subsequent analyses in the loan approval project.

Ref. template: [Click Here](#)

**Flight Delay Prediction Data Exploration and Preprocessing Report:** [Click Here](#)

## Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for flight delay. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, Linear Regression, Extra tree Classifier) , initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

### **Activity 1: Feature Selection Report**

The Feature Selection Report outlines the rationale behind choosing specific features for the flight delay model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to discern flight delays.

Ref. template: [Click Here](#)

Flight Delay Prediction Feature Selection Report: [Click Here](#)

### **Activity 2: Model Selection Report**

The Model Selection Report details the rationale behind choosing Random Forest, Decision Tree, Linear Regression, Extra tree Classifier models for flight delay prediction. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Ref. template: [Click Here](#)

Flight Delay Prediction Model Selection Report: [Click Here](#)

### **Activity 3: Initial Model Training Code, Model Validation and Evaluation Report**

The Initial Model Training Code employs selected algorithms on the flight delay dataset, setting the foundation for predictive modeling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and in predicting loan outcomes.

Ref. template: [Click Here](#)

Flight Delay Prediction Model Development Phase Template: [Click Here](#)

## **Milestone 4: Model Optimization and Tuning Phase**

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

### **Activity 1: Hyperparameter Tuning Documentation**

The Gradient Boosting model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

### **Activity 2: Performance Metrics Comparison Report**

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the Gradient Boosting

model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

### **Activity 3: Final Model Selection Justification**

The Final Model Selection Justification articulates the rationale for choosing Gradient Boosting as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal loan approval predictions.

Ref. template: [Click Here](#)

Flight Delay Prediction Model Optimization and Tuning Phase Report: [Click Here](#)

## **Milestone 5: Project Files Submission and Documentation**

For project file submission in Github, Kindly click the link and refer to the flow. [Click Here](#)

For the documentation, Kindly refer to the link. [Click Here](#)

## **Milestone 6: Project Demonstration**

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.