**Data-Driven Optimization for Airline Operations**

### **Introduction**

Airlines face a complex challenge in maximizing profitability while managing multiple operational areas, such as flight scheduling, ground crew efficiency, fuel management, and passenger services. Given strict regulations, optimizing these processes is essential for minimizing delays, improving turnaround times, and boosting overall performance.

The goal of this challenge is to apply advanced technologies—such as data analytics, IoT, machine learning, and process automation—to optimize key aspects of airline operations. Participants are tasked with building predictive models to improve efficiency in flight scheduling, ground operations, and fuel usage, leading to better performance and increased flight time utilization.

### **Problem Overview**

Participants are required to develop a predictive solution that enhances decision-making in various aspects of airline operations. The primary objectives are:

1. **Predict Flight Delays:** Use historical data to predict potential delays in scheduled flight operations, helping airlines preemptively adjust resources.
2. **Reschedule Flights:** Reschedule a given set of flights to minimize overall delays.

This challenge focuses on creating a scalable, data-driven model that works across different airports, aircraft, and routes, reflecting real-world variations in airline operations.

### **Dataset Description and Constraints**

Participants may use existing datasets or self-curated datasets to solve the problem but must ensure their dataset adheres to the following requirements:

#### **Training Data:**

* **Flight Schedule Data:**
  + Must include fields such as scheduled and actual departure/arrival times, and delay reasons.
  + Delay reasons should be categorized (e.g., weather-related, maintenance issues, or crew availability).

#### **Testing Data:**

* The testing data must include flight operations from different routes and different aircraft models not present in the training set to test model generalization.
* The dataset should have a minimum of 1000 flight records, ensuring model robustness across diverse scenarios.

### **Tasks and Evaluation**

Participants are required to develop predictive models addressing the following tasks:

1. **Flight Delay Prediction:**
   * Predict flight delays based on operational and external factors (e.g., crew scheduling, weather, previous delays).
   * Models will be evaluated on their accuracy in predicting delays and the impact of various delay factors.
2. **Flight Rescheduling:**
   * Reschedule a set of flights to minimize overall delays. Participants must consider constraints such as aircraft availability, crew scheduling, and airport regulations.

### **Evaluation Criteria:**

* **Predictive Accuracy:** Performance of the model in forecasting delays, ground task durations, and rescheduled flights.
* **Delay Reduction:** The model’s ability to minimize overall delays when rescheduling flights.
* **Efficiency:** The model’s ability to reduce overall turnaround time and fuel costs without compromising operations.
* **Generalization:** The capacity of the model to handle new routes, airports, and aircraft that were not present in the training data.
* **Innovation:** Originality in applying machine learning, IoT, and automation techniques to solve the problem.

### **Submission Requirements**

Participants must submit:

1. **Prediction Outputs:**
   * For flight delays: A CSV file listing predicted delays (in minutes) for each flight in the test set.
   * For flight rescheduling: A CSV file with the revised flight schedules and the expected reduction in delays.
2. **Model Overview:**
   * A brief report detailing the models used, the features considered, and how the predictions can be implemented in a real-world setting.