# Architecture

# **Flight Fare Prediction**

#### **Abstract**

The airline industry has always been the preferred mode of transport for those who would like to reach their destination quickly. The pricing of these airline tickets is considered to be a tug of war between the airline carriers and the passengers where each side wants the best rates. Passengers prefer to travel by flights that offer them the best rate at the lowest fare available while the airlines would like to bank in more money for their tickets. Flight prices are also affected by seasons, particular occasions when passengers are expected to travel more. This makes it harder to guess the flight prices. This project provides a method to predict flight fares using machine learning models like Linear regression, Random Forest, Decision tree, Extra trees regression, Gradient boosting and Ridge regression.

# **Table of contents**

1. Introduction	1
1.1 Scope	
2. Technical specifications	1
2.1 Dataset	1
2.2 Prediction	
2.3 Logging	2
2.4 Deployment	2
3. Proposed Solution	
4. Model Training workflow	

#### 1. Introduction

The purpose of this document is to present a detailed description of the Flight fare prediction system. It covers the purposes and features of the system and how the system performs the actions

The objective of the project is to create a solution to find the ticket price of the user's preferred airline using machine learning models which will make it easy for users to plan their trips

Predicting the fare of flight tickets depends upon a variety of factors like the their preferred airline company, source and destination of travel, number of stops etc

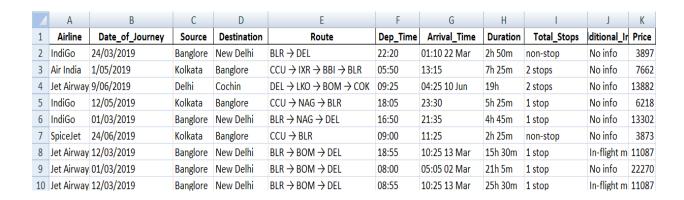
### 1.1 Scope

The flight fare default prediction system is a machine learning based model that will help us to predict flight fares. The ability to predict or guess the ticket prices of their preferred flights can help users to make calculated decisions about which airline to fly and when to make the travel

# 2. Technical specifications

#### 2.1 Dataset

The dataset used for the study consists of 10683 rows of data and 11 columns which relate the flight and its details.



#### 2.2 Prediction

The system prompts the user to input the data related to the flight. After the information has been loaded, the system will output the prediction of their preferred flight route.

# 2.3 Logging

The system logs all the steps so that the user will be aware of the processes working internally. The type of logging chosen for each depends upon the type of operation that is carried out like info, critical, debug or warning logging. Logging helps to debug issues, so it is a mandatory step

## 2.4 Deployment

The machine learning model will be deployed as an API using FastAPI

# 3. Proposed Solution

The proposed solution is to build various machine learning models like Linear regression, Random Forest, Decision tree, Extra trees regression, Gradient boosting and Ridge regression and find which model is able to best predict the fare of a flight given a set of training data.

# 4. Model Training workflow

The different models used in the project will be used to predict the fare of a flight based on the data used in the training stage. The goal is to identify the model which is able to make predictions with the highest R2 score so that the fares can be predicted easily.

