# KARAMVEER BHAURAV PATIL COLLEGE

**BSC IN INFORMATION TECHNOLOGY**

# Project name: Flight Price Prediction Using Flask App In ML

**Project By: Sandhya Shivanad Ghuge**

**Seat No.: 237787**

**2023-2024**

Introduction

we will be analyzing the flight fare prediction using Machine Learning dataset using essential exploratory data analysis techniques then will draw some predictions about the price of the flight based on some features such as what type of airline it is, what is the arrival time, what is the departure time, what is the duration of the flight, source, destination and more.

The Airline Flight Fare Prediction project is to predict airline flight fares across the Indian cities. The dataset for the project is taken from Kaggle, and it is a time-stamped dataset so, while building the model, extensive pre-processing was done on the dataset especially on the date-time columns to finally come up with a ML model which could effectively predict airline fares across various Indian Cities. The dataset had many features which had to pre-processed and transformed into new parameters for a cleaner and simple web application layout to predict the fares. The various independent features in the dataset were:

Airline: The name of the airline.

Date\_of\_Journey: The date of the journey

Source: The source from which the service begins.

Destination: The destination where the service ends.

Route: The route taken by the flight to reach the destination.

Dep\_Time: The time when the journey starts from the source.

Arrival\_Time: Time of arrival at the destination.

Duration: Total duration of the flight.

Total\_Stops: Total stops between the source and destination.

Additional\_Info: Additional information about the flight

Price: The price of the ticket

# Popular flight Price Predictor

## Airhint

## The **airfare predictor** also shows the price range and the confidence level regarding the best moment to buy your flight tickets. Moreover, we will tell you if it's best to wait or book your tickets now and if not what are the chances of price drop. AirHint tells you how to fly cheap and is all about great flight predictions and deals!

## Hopper

## Hopper is an accredited travel agency. Hopper partner with airlines, hotels, homes, and car rental providers across the globe so you can feel confident you’re booking the perfect vacation at the best price. Hopper sells over $4.5 billion of travel every year. We’ve raised over $600 million of private capital and are backed by some of the world's largest institutional investors and banks.

## Skyscanner

## Skyscanner is one of the popular apps when it comes to airfare tracking. And a large part of that has to do with its simplified and easy-to-use interface, which allows for the ability to search and compare different flights in one place. Besides, it also simplifies the experience further by helping you with finding the cheapest months and dates to fly, in addition to a price alert that can be set accordingly.

## Kayak

Kayak is a veteran when it comes to flight search and airfare prediction. The service curates flight information from various travel sites to offer you the best deals for flights, hotels, or rental cars. Its app includes Kayak Trip, which is a sweet feature that allows you to organize all your travel plans in one place, such as flight status updates, security wait time, airport terminal maps, etc.

Apart from offering deals for booking flights and hotels, the app provides a price forecast to help you determine if the airfares are likely to go up in the future. In addition to this, it also offers over price history charts for the last 90 days, which helps you get an idea of the airfare over the period. Among other things, Kayak offers real-time notifications for flight status, gate change, and baggage claim, and also allows you to navigate through the airport by offering step-by-step directions with offline access to your travel itinerary and airport maps.

# Proposed system

We plan to predict ticket prices for upcoming flights to help customers in selecting the optimum time for travel and the cheapest flight to the desired destination.

Throughout this project, a linear regression model will be applied to forecast the cheapest flight tickets to a specific destination based on data scraped from a variety of travel websites such as Momondo, Kayak, and Expedia.

The prediction provisionally will be based on the following features:

previous flight prices

destination

departure

class

airline

Who benefits from exploring this question or building this model/system?

Flight customers.

Data Description:

The dataset will be created by scraping web content from different travel websites.

The Search Engine Results - Flights & Tickets Keywords Dataset will also be used as it provides Rankings for world top destinations on Google.

If modeling, what will you predict as your target?

Flight prices.

# Existing popular flight Price Predictor disadvantages

Early prediction of the demand along a given route could help an airline company preplan the flights and determine appropriate pricing for the route. Existing demand prediction models generally try to predict passenger demand for a single flight/route and market share of an individual airline. Price discrimination allows an airline company to categorize customers based on their willingness to pay and thus charge them different prices. Customers could be categorized into different groups based on various criteria such as business vs leisure, tourist vs normal traveler, profession etc. For example, business customers are willing to pay more as compared to leisure customers as they rather focus on service quality than price.

Despite the fact that there are several studies conducted on both sides, customer and airlines, no attempt has been made to present a literature survey and review of existing work. Therefore, the main goal of this paper is to present a comprehensive literature review of existing studies related to this topic which can be utilized by future researchers. We first classify and present existing studies into two categories based on their desired goal (customer side models and airline side models). We then group existing work based on the specific problem being addressed. Several issues have been discussed including data sources, features and various techniques employed for prediction. We believe that this is an important contribution for researchers who are aiming to work on this exciting area of research.

# Advantage of proposed system

The Flight ticket prices increase or decrease every now and then depending on various factors like timing of the flights, destination, duration of flights. In the proposed system a predictive model will be created by applying machine learning algorithms to the collected historical data of flights. Optimal timing for airline ticket purchasing from the consumer’s perspective is challenging principally because buyers have insufficient information for reasoning about future price movements. In this project we majorly targeted to uncover underlying trends of flight prices in India using historical data and also to suggest the best time to buy a flight ticket. The project implements the validations or contradictions towards myths regarding the airline industry, a comparison study among various models in predicting the optimal time to buy the flight ticket and the amount that can be saved if done so. Remarkably, the trends of the prices are highly sensitive to the route, month of departure, day of departure, time of departure, whether the day of departure is a holiday and airline carrier. Highly competitive routes like most business routes (tier 1 to tier 1 cities like Mumbai-Delhi) had a non-decreasing trend where prices increased as days to departure decreased, however other routes (tier 1 to tier 2 cities like Delhi - Guwahati) had a specific time frame where the prices are minimum. Moreover, the data also uncovered two basic categories of airline carriers operating in India – the economical group and the luxurious group, and in most cases, the minimum priced flight was a member of the economical group. The data also validated the fact that, there are certain time-periods of the day where the prices are expected to be maximum. The scope of the project can be extensively extended across the various routes to make significant savings on the purchase of flight prices across the Indian Domestic Airline market..

# Technology used

Python : Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. ... Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse

ML: Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase. They will be required to help identify the most relevant business questions and the data to answer them.

Libraries

Pandas

seaborn

numpy

matplotlib

datetime

pickle

sklearn

# Discussion

Dynamic pricing is one of the most common pricing strategies implemented by the airline industry to adjust ticket prices in response to various internal and external factors such as changes in demand, competitor promotions, ability of users to buy, availability of seats and others. Airlines need to predict changes in these factors to implement a dynamic pricing scheme that dynamically adjusts ticket prices to increase their profit. On the other hand, customers are also interested to forecast how ticket prices would change in the future to be able to buy tickets at lower prices. Therefore, researchers have developed various prediction models both for airlines and customers to help them deal with dynamic pricing. The two most common methods proposed for airlines are demand prediction and price discrimination which we collectively refer to as Airlines side models. Customer side modes involve optimal ticket purchase time prediction models and ticket price prediction models. There is a tradeoff between money saving by customer and increasing revenue by companies. As customers become more strategic by using customer side tools, it becomes more difficult for the airlines to apply dynamic pricing and to generate profit and vice versa. Therefore, there is a need for a prediction model that can predict the optimal ticket prices that can bring mutual benefit both for customers and airlines.