## OPTIMIZING FLIGHT BOOKING DECISIONS THROUGH MACHINE

# 1.INTRODUCTION

# 1.1 OVERVIEW

In this article, 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.

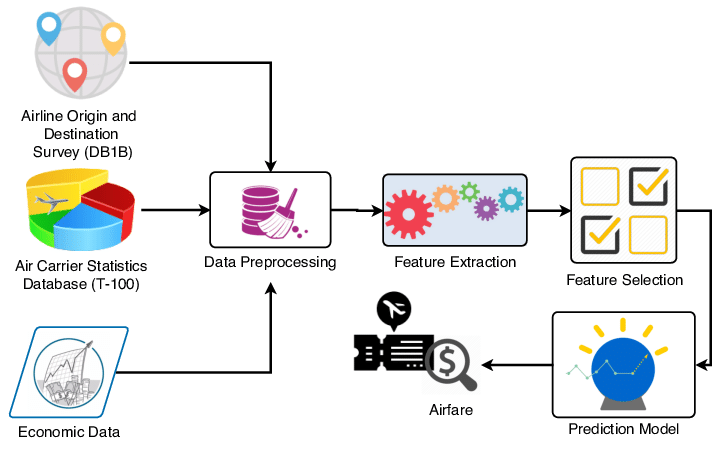
# 1.2 PURPOSE

This can be a regression problem since the target or label is the price (continuous numeric value). Airline companies use complex algorithms to calculate flight prices given various conditions present at that individual time. These methods take financial, marketing, and various social factors into consideration to predict flight prices. Nowadays, the amount of individuals using flights has increased significantly. It's difficult for airlines to keep up with prices since prices change dynamically because of different conditions. That’s why we are going to attempt to use machine learning to unravel this problem. This will help airlines by predicting what prices they will maintain. It may also help customers to predict future flight prices and plan their journey accordingly.

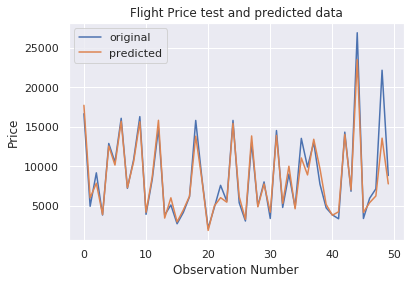
We will analyze the flight fare prediction using Machine Learning, in the dataset, we will be using a few necessary features to draw some predictions about the price of the flight like what type of airline it is, what is the arrival time, what is departure time, what is the duration of the flight, source, destination, and more.

# 2.PROBLEM DEFINITION & DESIGN THINKING

## 2.1EMPATHY MAP



# 2.2 IDEATION &BRAINSTROMING MAP



# 3.RESULT

# OUTPUT 1

# 

# OUTPUT 2

Flight number : G8 134 GoAir

Jay Prakash Narayan International Airport

Scheduled : 21:00 IST

Terminal : N/A

Estimated : 21:00 IST

Gate : N/A

Indira Gandhi International Airport

Scheduled : 22:40 IST

Terminal : T2

Estimated : 22:40 IST

Gate : 205

Terminal : N/A

Gate : N/A

Terminal : T2

Gate : 205

1

# 4.Advantages

* Traveler get the fare prediction handy using which it’s easy to decide the airlines.
* Saves time in searching / deciding for airlines.

# Disadvantages

* Improper data will result in incorrect fare predictions.

# 5.APPLICATIONS:

# Skyscanner

### Kayak

### Google Flights

### Hitlist

### Hipmunk

### Momondo

### Hopper

These are the applications for get a solution for flight price prediction

## 6.CONCLUSION

Machine learning models were examined in this case study to forecast the average flight price at the business segment level. We used training data to train the training data and test data to test it. These records were used to extract a number of characteristics. Our suggested model can estimate the quarterly average flight price using attribute selection strategies.To the highest possible standard, much prior studies into flight price prediction using the large dataset depended on standard statistical approaches, which have their own limitations in terms of underlying issue estimates and hypotheses. To our knowledge, no other research have included statistics from holidays, celebrations, stock market price fluctuations, depression, fuel price, and socioeconomic information to estimate the air transport market sector; nonetheless, there are numerous restrictions.As example, neither of the databases provide precise information about ticket revenue, including such departing and arrival times and days of the week. This framework may be expanded in the future to also include airline tickets payment details, that can offer more detail about each area, such as timestamp of entry and exit, seat placement, covered auxiliary items, and so on. By merging such data, it is feasible to create a more robust and complete daily and even daily flight price forecast model. Furthermore, a huge surge of big commuters triggered by some unique events might alter flight costs in a market sector. Thus, incident data will be gathered from a variety of sources, including social media sites and media organizations, to supplement our forecasting models. We will also examine specific technological Models, such as Deeper Learning methods, meanwhile striving to enhance existing models by modifying their hyper-parameters to get the optimum design for airline price prediction

# 7.FUTURE SCOPE

The scope of Machine Learning is not limited to the investment sector. Rather, it is expanding across all fields such as banking and finance, information technology, media & entertainment, gaming, and the automotive industry. As the Machine Learning scope is very high, there are some areas where researchers are working toward revolutionizing the world for the future

# 8.APPENDIX

# SOURCE CODE 1 flight price predection

import numpy as np *# linear algebra*

port pandas as pd *# data processing, CSV file I/O (e.g. pd.read\_csv)*

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.ensemble import ExtraTreesRegressor

from sklearn.model\_selection import GridSearchCV

from sklearn.tree import DecisionTreeRegressor

from sklearn.pipeline import make\_pipeline

from sklearn.preprocessing import StandardScaler

from sklearn.model\_selection import train\_test\_split, cross\_val\_score

from sklearn.preprocessing import MinMaxScaler

from sklearn import metrics

sns.set()

import os

for dirname, \_, filenames **in** os.walk('/kaggle/input'):

for filename **in** filenames:

print(os.path.join(dirname, filename))

numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.preprocessing import StandardScaler

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import mean\_squared\_error as mse

from sklearn.metrics import r2\_score

from math import sqrt

from sklearn.linear\_model import Ridge

from sklearn.linear\_model import Lasso

from sklearn.tree import DecisionTreeRegressor

from sklearn.ensemble import RandomForestRegressor

from sklearn.preprocessing import LabelEncoder

from sklearn.model\_selection import KFold

from sklearn.model\_selection import train\_test\_split

from sklearn.model\_selection import GridSearchCV

from sklearn.model\_selection import RandomizedSearchCV

from prettytable import PrettyTable

## SOURCE CODE 2 Flight booking

# import module

import requests

from bs4 import BeautifulSoup

# UDF for get HTML code

# from URL

def get\_html(Airline\_code, Flight\_number, Date, Month, Year):

    def getdata(url):

        r = requests.get(url)

        return r.text

    # url

    url = "+Airline\_code + \

        "/"+Flight\_number+"?year="+Year+"&month="+Month+"&date="+Date

    # pass the url

    # into getdata function

    htmldata = getdata(url)

    soup = BeautifulSoup(htmldata, 'html.parser')

    return(soup)

# Get Flight number

# from Html code

def flight\_no(soup):

    Flight\_no = ""

    # Find div tag with

    # unique class name

    for i in soup.find("div", class\_="ticket\_\_FlightNumberContainer-s1rrbl5o-4 hgbvHg"):

        Flight\_no = Flight\_no + (i.get\_text()) + " "

    return (Flight\_no)

# Get Airport name

# from HTML code

def airport(soup):

    Airport\_name = []

    # Find div tag with

    # unique class name

    for i in soup.find\_all("div", class\_="text-helper\_\_TextHelper-s8bko4a-0 CPamx"):

        Airport\_name.append(i.get\_text())

    return (Airport\_name)

# get status

# from HTML code

def status(soup, Airport\_list):

    Time\_status = []

    Airport\_List = []

    Status\_str = []

    Gate = []

    Gate\_no = []

    # Find div tag with

    # unique class name

    # to get Gate number

    for data in soup.find\_all("div", class\_="ticket\_\_TGBLabel-s1rrbl5o-15 gcbyEH text-helper\_\_TextHelper-s8bko4a-0 dfeqpK"):

        Gate.append(data.get\_text())

    for data in soup.find\_all("div", class\_="ticket\_\_TGBValue-s1rrbl5o-16 icyRae text-helper\_\_TextHelper-s8bko4a-0 cCfBRT"):

        Gate\_no.append(data.get\_text())

    # Get status from

    # html code

    for i in soup.find\_all("div", class\_="text-helper\_\_TextHelper-s8bko4a-0 bcmzUJ"):

        Status\_str.append(i.get\_text())

    for i in soup.find\_all("div", class\_="text-helper\_\_TextHelper-s8bko4a-0 cCfBRT"):

        Time\_status.append(i.get\_text())

    # traverse the Data

    # from scraping data

    for item in range(4):

        if item == 0:

            print(Airport\_list[0])

        if item == 2:

            print("")

            print(Airport\_list[1])

        print(Status\_str[item] + " : " + Time\_status[item])

        print(Gate[item] + " : " + Gate\_no[item])

    for item in range(len(Gate)):

        print(Gate[item] + " : " + Gate\_no[item])

# Driver code

if \_\_name\_\_ == '\_\_main\_\_':

    # Input Data from geek

    Airline\_code = 'G8'

    Flight\_number = '134'

    Date = '23'

    Month = '10'

    Year = '2020'

    # Calling the get\_html

    # with argument

    # function calling

    soup = get\_html(Airline\_code, Flight\_number, Date, Month, Year)

    print("Flight number : ", flight\_no(soup))

    Airport\_list = airport(soup)

    status(soup, Airport\_list)

## REFERENCE

[1] Rajankar, Supriya, and Neha Sakharkar. "A Survey on Flight Pricing Prediction using Machine Learning." Internatıonal Journal Of Engıneerıng Research & Technology (Ijert) 8.6 (2019): 1281- 1284.

[2] Smith, Barry C., John F. Leimkuhler, and Ross M. Darrow. "Yield management at American airlines." interfaces 22.1 (1992): 8-31.

[3] Groves, William, and Maria Gini. "An agent for optimizing airline ticket purchasing." Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems. 2013.

[4] Janssen, Tim, et al. "A linear quantile mixed regression model for prediction of airline ticket prices." Radboud University (2014).

[5] Wohlfarth, Till, et al. "A data-mining approach to travel price forecasting." 2011 10th International Conference on Machine Learning and Applications and Workshops. Vol. 1. IEEE, 2011.

[6] Papadakis, Manolis. "Predicting Airfare Prices." (2014)