Optimizing Flight Booking Decisions through Machine Learning Price Predictions

1) INTRODUCTION

To create Optimizing Flight Booking Decisions through Machine Learning Price Predictions

1.1 Overview

People who work frequently travel through flight will have better knowledge on best discount and right time to buy the ticket. For the business purpose many airline companies change prices according to the seasons or time duration. They will increase the price when people travel more. Estimating the highest prices of the airlines data for the route is collected with features such as Duration, Source, Destination, Arrival and Departure. Features are taken from chosen dataset and in the price wherein the airline price ticket costs vary overtime. we have implemented flight price prediction for users by using KNN, decision tree and random forest algorithms. Random Forest shows the best accuracy of 80% for predicting the flight price. also, we have done correlation tests and metrics for the statistical analysis.

1.2 Purpose

- Entered input is analysed by the model which is integrated.
- Once model analyses the input the prediction is showcased on the UI

To accomplish this, we have to complete all the activities listed below,

- Define Problem / Problem Understanding
  - Specify the business problem
  - Business requirements
  - Literature Survey
  - Social or Business Impact.
- Data Collection & Preparation

- ○ Collect the dataset
- ○ Data Preparation
- ● Exploratory Data Analysis
- ○ Descriptive statistical
- ○ Visual Analysis
- ● Model Building
- ○ Training the model in multiple algorithms
- ○ Testing the model
- ● Performance Testing & Hyperparameter Tuning
- ○ Testing model with multiple evaluation metrics
- ○ Comparing model accuracy before & after applying hyperparameter tuning
- ● Model Deployment
- ○ Save the best model
- ○ Integrate with Web Framework
- ● Project Demonstration & Documentation
- ○ Record explanation Video for project end to end solution
- ○ Project Documentation-Step by step project development procedure

2 **Problem Definition & Design Thinking**

1: Business requirements

The business requirements for a machine learning model to predict personal loan approval include the
ability to accurately predict loan approval based on applicant information, Minimise the number of false
positives (approved loans that default) and false negatives (rejected loans that would have been successful). Provide an explanation for the model's decision, to comply with regulations and improve
transparency.

2: Literature Survey (Student Will Write)

As the data is increasing daily due to digitization in the banking sector, people want to apply for loans through
the internet. Machine Learning (ML), as a typical method for information investigation, has gotten more
consideration increasingly. Individuals of various businesses are utilising ML calculations to take care of the
issues dependent on their industry information. Banks are facing a significant problem in the approval of the
loan. Daily there are so many applications that are challenging to manage by the bank employees, and also
the chances of some mistakes are high.Most banks earn profit from the loan, but it is risky to

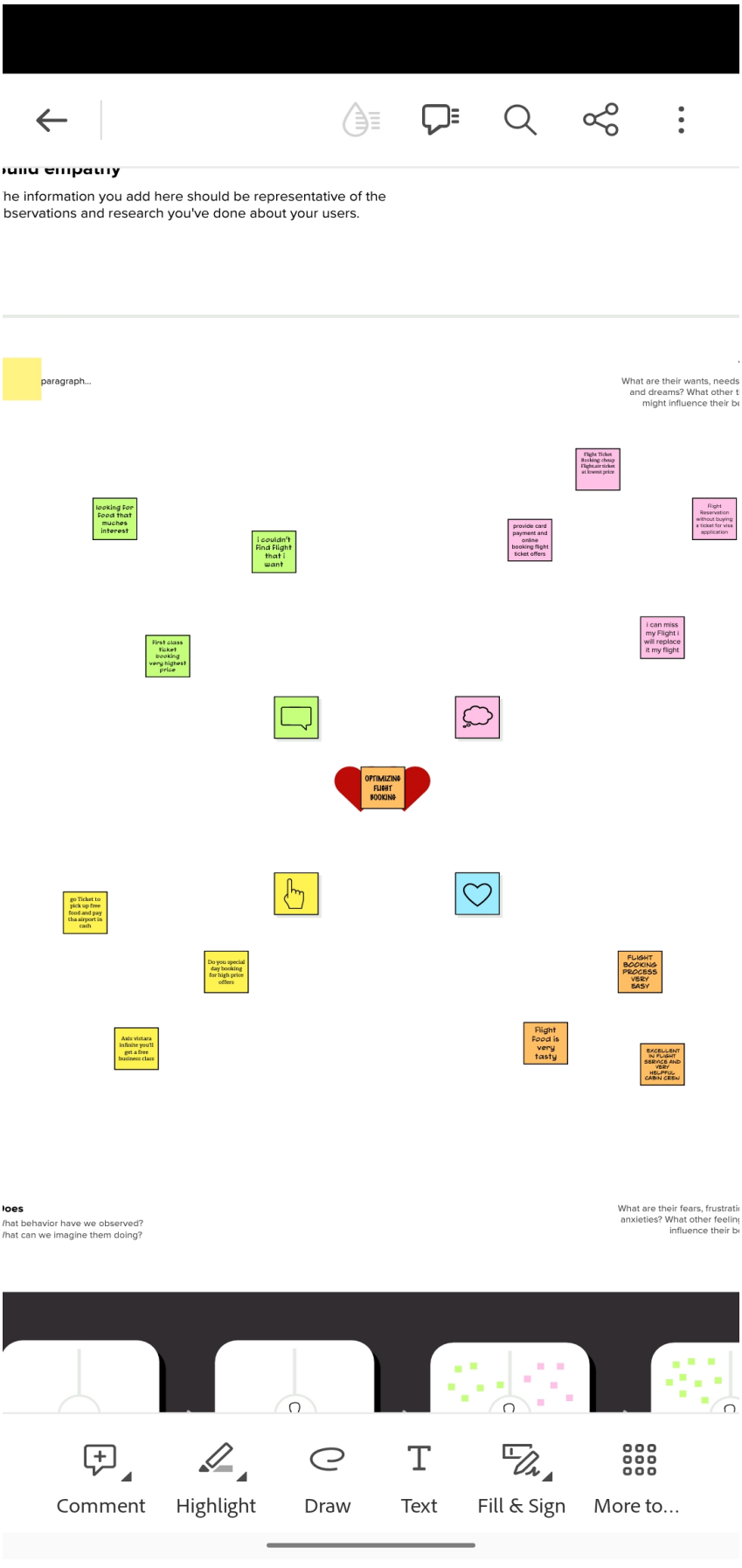
choose
deserving customers from the number of applications.There are various algorithms that have been used with
varying levels of success. Logistic regression, decision tree, random forest, and neural networks have all
been used and have been able to accurately predict loan defaults. Commonly used features in these studies
include credit score, income, and employment history, sometimes also other features like age, occupation,
and education level.

3: Social or Business Impact.

Social Impact: - Personal loans can stimulate economic growth by providing individuals with the funds they
need to make major purchases, start businesses, or invest in their education.

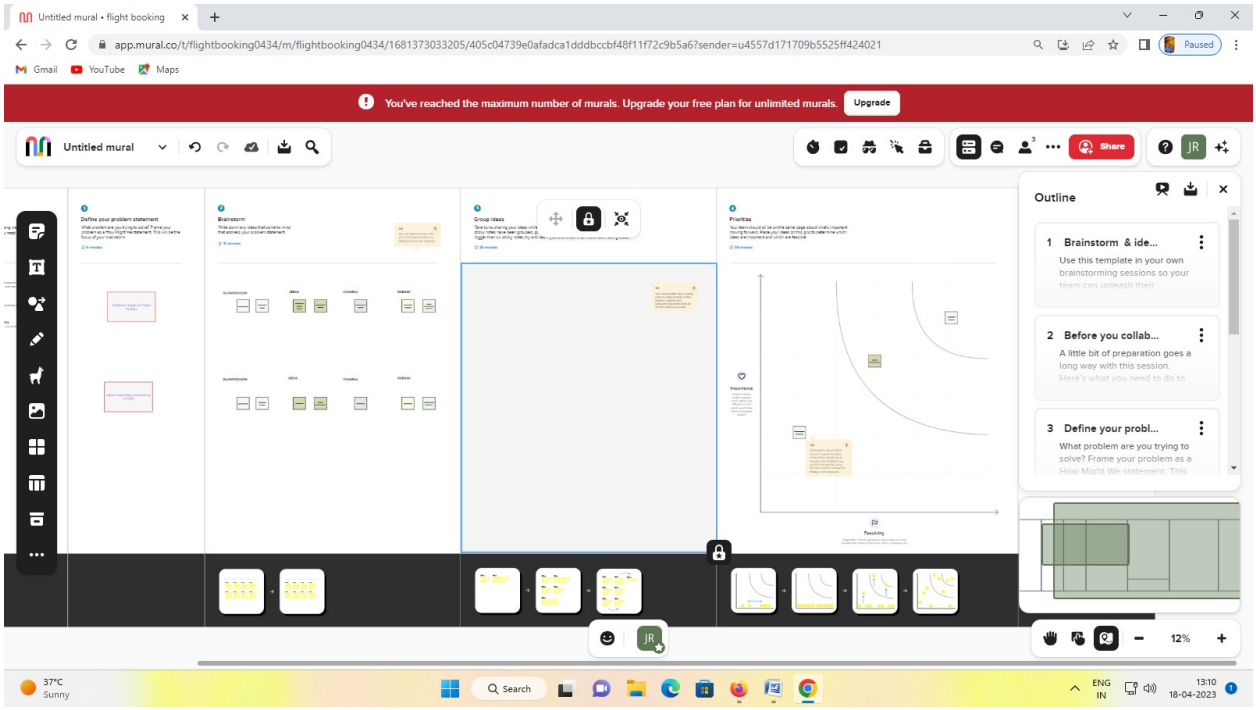
Business Model/Impact: - Personal loan providers may charge fees for services such as loan origination,
processing, and late payments. Advertising the brand awareness and marketing to reach out to potential
to generate revenue borrowers

Empathy Map:

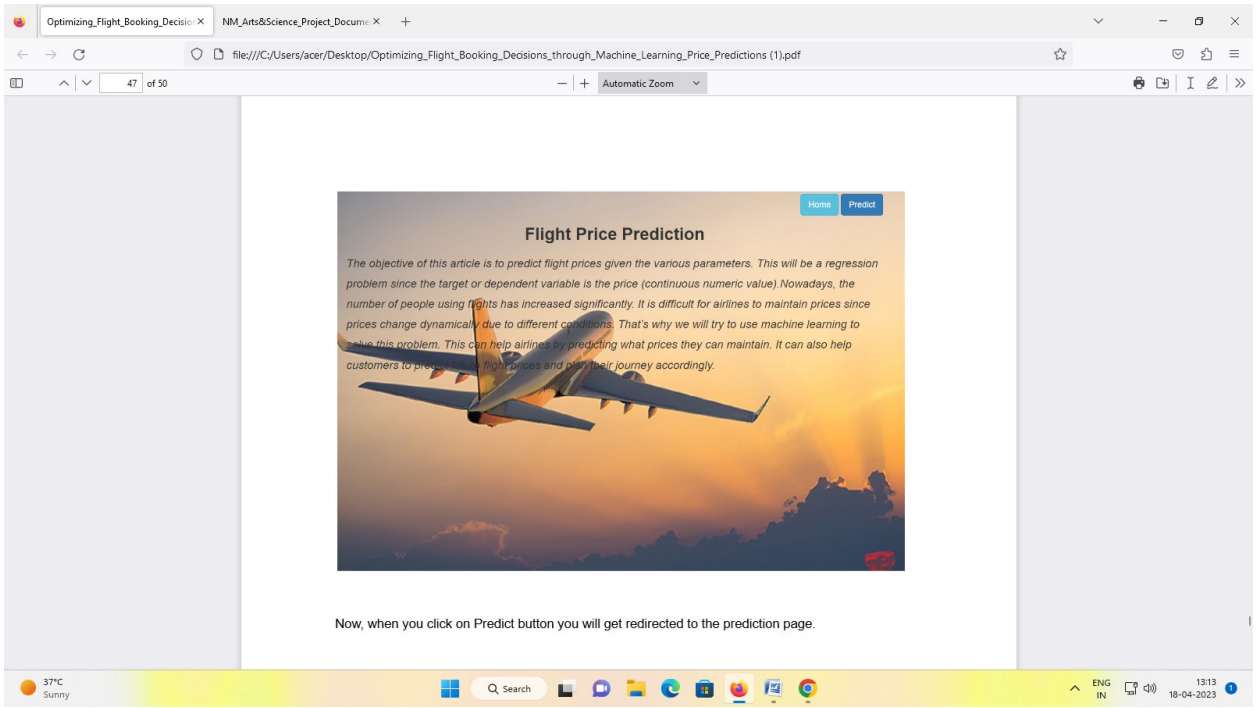

he information you add here should be representative of the
bservations and research you've done about your users.
paragraph...
What are their wants, needs
and dreams? What other t
might influence their be
looking for food that muches interest
I couldn't Find Flight that i want
provide card payment and online booking flight ticket offers
Flight Reservation without buying a ticket for visa application
i can miss my Flight i will replace it my flight
First class ticket booking very highest price
OPTIMIZING FLIGHT BOOKING
go Ticket to pick up free food and pay the airport in cash
Do you special day booking for high price offers
FLIGHT BOOKING PROCESS VERY EASY
Axis victara infinite you'll get a free business class
Flight Food is very tasty
oes
/hat behavior have we observed?
/hat can we imagine them doing?
What are their fears, frustrati
anxieties? What other feelin
influence their be
Comment
Highlight
Draw
Text
Fill & Sign
More to...

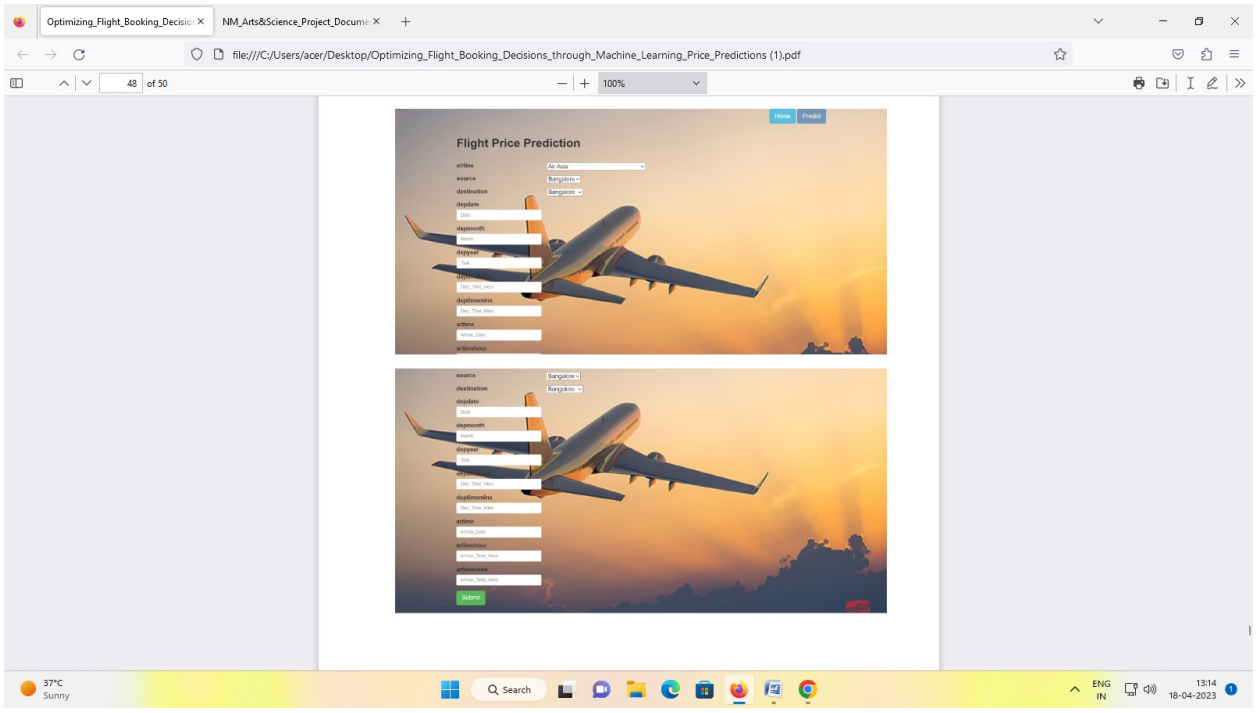
Ideation & Br

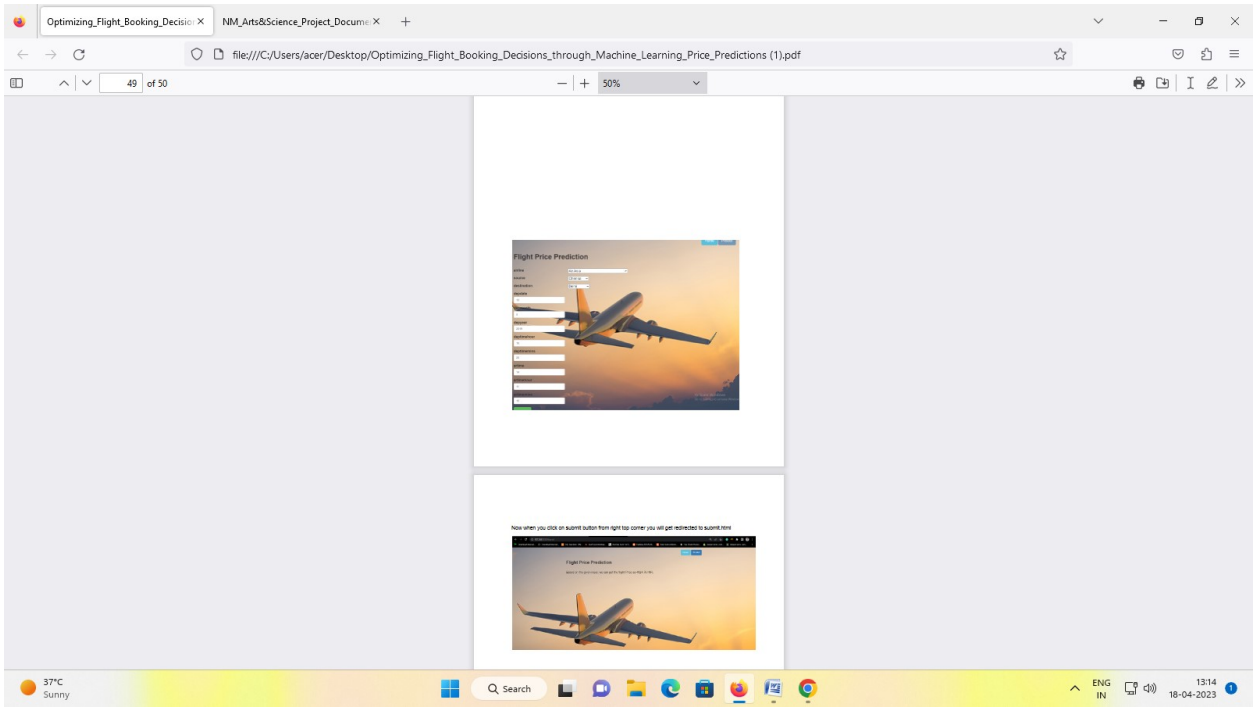
Ideation & Brainstorming Map:



2)RESULT







4) **ADVANTAGES & DISADVANTAGES**

ADVANTAGES:

Your Business will be available 24/7.

- **Reservation can be minimized anytime.**

- **Easier to manage the calendar.**
- **Conveniently manage everything through a single screen.**
- **Get Valuable insight about your business.**
- **An online Booking system saves a lot of time.**
- **Online Booking System is a modern approach. BOOK**
- **This helps in increasing sell-through upselling. ONLINE**
- **You will learn more about your customers.**

You get paid before providing services.

- **Cuts the overloads of the staff's members.**

DISADVANTAGES

Need a good internet connection.

High competition on the market.

Not every online booking system is equal, so you have to choose the best one as per your business.

5)CONCLUSION

Recommendations and ConclusionOnline ticket booking system is an application where the customer can book a ticketonline and 24*7 hours a day from anyplace in the world. Customers can also interact with theticket booking website to know any other details they want. Online ticket booking system hasbeen developed successfully. System performance is also found to be satisfactory. This is a user-friendly application. Through this application, the cost can be reduced and efficiency isincreased. There are several procedures that can be selected by customers. With the help of thisapplication customers can book tickets, can know the status of a flight, bus or trains, a Sourcestation and destination can be chosen according to their choice, can select seats, can choose than.

6) FUTURE SCOPE

The research report includes specific segments by region (country), company, Type, and Application. This study provides information about the sales and revenue during the historical and forecasted period. Understanding the segments helps identify the factors that aid the market growth. The Online Airline Reservation System research report

provides information about the market area, which is further subdivided into sub-regions and countries/regions. In addition to the market share in each country and sub-region, this report chapter also contains information on profit opportunities.

The use of online airline reservation systems has seen significant growth in recent years as more consumers prefer to book their flights online. According to a study by the International Air Transport Association, over 80% of air tickets were purchased through digital channels in 2020. Additionally, the study found that mobile bookings accounted for over 50% of all digital ticket sales. The trend towards digital booking is expected to continue, as travelers seek convenient and efficient ways to plan their trips. In response, airlines and travel companies have invested in the development of online airline reservation systems that provide a seamless booking experience for their customers. With the growing demand for online booking, the online airline reservation system market is poised for continued growth in the coming years.

The Online Airline Reservation System research report is an expert's analysis that mainly includes companies, types, applications, regions, countries, etc. Also, the reports analyse sales, revenue, trade, competition, investment, and forecasts. Industrial Analytics market research covers COVID-19 impacts on the upstream, midstream, and downstream industries. Also, this study offers detailed market estimates by emphasising statistics on several aspects covering market dynamics like drivers, barriers, opportunities, threats, and industry news & trends.

7)**APPENDIX**

A. **Source Code**

```
[ ]import numpy as np


[ ]import pandas as pd

[ ]import matplotlib.pyplot as plt

[ ]import seaborn as sns

[ ]from sklearn.model_selection import train_test_split

[ ]from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier

[ ]from sklearn.tree import DecisionTreeClassifier
```

[ ]from sklearn.neighbors import KNeighborsClassifier

[ ]from sklearn.metrics import f1_score

[ ]from sklearn.metrics import classification_report, confusion_matrix

[ ]import warnings

[ ]import pickle

[ ]from scipy import stats

[ ]warnings.filterwarnings('ignore'

[ ]plt.style.use('fivethirtyeight')

[ ]data=pd.read_csv("Data_Train.csv")

[ ]data.head()

[ ]for i in category:

[ ]print(i,data{i},unique())

[ ]data.Data_of_journey=data,data_of_journey.str.split('/')

[ ]data.Data_of_journey

[ ]data.Total_stops.unique

[ ]data.rount=data.Route.str.split('+')

[ ]data.Rount

[ ]data.Dep_Time=data.Dep_Time.str.split(:)

[ ]data['Dep_Time_Hour']=data.Dep_Time.str[0]

[ ]data['Dep_Time_Hour']=data.Dep_Time.str[1]

[ ]data.Arrival_Time=data.Arrival_Time.str.split('')

[ ]data['Arrival_data']=data.Arrival_Time.str[1]

[ ]data['Time_of_Arrival']=data.Arrival_Time.str[0]

[ ]data['Time-of-arrival']=data.Time_of_Arrival.str.split[':']

[ ]data['Arrival_Time_Hour']=data.Time_of_Arrival.str[0]

[ ]data['Arrival_Time_Hour']=data.Time_of_Arrival.str[1]

```
[ ]data.Duration=data.Duration.str.split('')

[ ]data['Travel_Hours']=data.Duration.str[0]

[ ]data['Travel_Hours']=data['Travel_Hours'.str.split('h')]

[ ]data['Travel_Hours']=data['Travel Hours'].str[0]

[ ]data.Travel_Hours=data.Travel_Hours

[ ]data.['travel_mins']=data.duration.str[1]

[ ]data.Travel_mins=data.Duration.str.split('m')

[ ]data.Travel_mins=data.Duration.str[0]

[ ]data.Total_stops.replace('non stop',0,inplace=True)

[ ]data.Total_stops=data.Total_stops.str[o]

[ ]#we also treat the 'total_stops'column,and replace non stop-Flight with 0 value and
extract the integer part of the 'Total_

[ ]ta.Total_stops.replace('non_stop',0,inplace=True)

[ ]data.Total_stop=data.Total_stops.str.splite('')

[ ]data.Total_stops=data.Total_stops.str[0]

[ ]data.Additional_lnfo.unique()

[ ]array('no','ln flight meal not include',

    no cheak-in baggage included', 'i short loyover','No lnfo',

    'i long layover','change airports','Businees class',

    'Red-eye flight','2 long layover'], dtype=object)

[ ]data.Additional_lnfo.replace('No lnfo','No info',inplace=True)

[ ]data.isnull().sum()

[ ]data.drop(['Cily4','City5','City6'],axis=1,inplace=true=True)

[ ]data.drop(['data_of_journey','Rount','Dep_Time','Arrival_Time','Duration'],axis=1,
inplace=True)

[ ]data.drop(['Time_of_Arrival'],axis=1,inplace=True)

[ ]data.isnull().sum()
```

[ ]data['City'].fillna('None',inplace=True)

[ ]data['Arrival_data'].Fillna(data['Date'].inplace=True)

[ ]data['travel_Nins'].Fillna(0,inplace=True)

[ ]data.info()

[ ]data.Date=data.Date.astype('int64')

[ ]data.Month=data.Month.astype('int64')

[ ]data.Year=data.Year.astype('int64')

[ ]data.Dep_Time_Hour=data.Dep_Time_Hour.astype('int64')

[ ]data.Dep_Time_Hour=data.Dep_Time_Hour.astype('int64')

[ ]data.Dep_Time.Mins=data.Dep_Time_Mins.astype('int64')

[ ]data.Arrival_data=data.Arrival_date.astype('int64')

[ ]data.Arrival_Time_Hour=data.Arrival_Time_Hour.astype('int64')

[ ]data.Arrival_Time_Hour=data.Arrival_Time_Mine.astype('int64')

[ ]data.Travel_Mine=data.Traval_Mins.astype('int64')

[ ]data[data['Traval_Hours']=='sm']

[ ]data.Travel_Hours-data.Travel_Hours.astype('int64')

[ ]categorical=['Airline', 'Source', 'Destination', 'Additional_Info'',''City'']

[ ]numerical=['Total Stops', 'Date', 'Month', 'Year', 'Dep_Time_Hour', 'Dep_Time_Mins', 'Arrival_date', 'Arrival_Time_Hour', ''Arrival Time_Mins', 'Travel_Hours', 'Travel_Mins']

[ ]from sklearn.preprocessing import LabelEncoder le-LabelEncoder()

[ ] data.Airline-le.fit_transform(data.Airline)

[ ] data.Source…