PROJECT REPORT

ON

Airline Satisfaction Prediction

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR SEMESTER V OF

T.E. (Information Technology)

SUBMITTED BY

Mr. Shubham Kamodkar (30)

Mr. Mikil Lalwani (37)

Ms. Aditi Miniyar (44)

UNDER THE GUIDANCE OF

PROF. Ms. Bincy Ivin

DEPARTMENT OF INFORMATION TECHNOLOGY
V.E.S. INSTITUTE OF TECHNOLOGY
2022-23

Certificate

This is to certify that project entitled

"Airline Satisfaction Prediction"

Group Members Names

Ms. Shubham Kamodkar (Roll No. 30)
Mr. Mikil Lalwani (Roll No. 37)
Ms. Aditi Miniyar (Roll No. 44)

In partial fulfillment of degree of T.E. (Sem V) in Information Technology for Project is approved.

Prof. Guide Name **Bincy Ivin**	**External Examiner**
Dr.(Mrs.)Shalu Chopra **H.O.D**	**Dr.(Mrs.)J.M.Nair** **Principal**

Date: 06/04/2022
Place: VESIT, Chembur

College Seal

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

- - - - - - - - - - -

(Signature)

Shubham Kamodkar (30)
Mikil Lalwani (37)
Aditi Miniyar (44)

ACKNOWLEDGEMENT

The project report on "Airline Satisfaction Prediction" is the outcome of the guidance, moral support and devotion bestowed on our group throughout our work. For this we acknowledge and express our profound sense of gratitude to everybody who has been the source of inspiration throughout project preparation. First and foremost we offer our sincere phrases of thanks and innate humility to Mrs. Shalu Chopra (HOD)", "Mrs. Smita Jangle (Deputy HOD)", "Ms. Bincy Ivin (Professor from Information Technology Department)" for providing the valuable inputs and the consistent guidance and support provided by them. We can say in words that we must at outset tender our intimacy for receipt of affectionate care to Vivekanand Education Society's Institute of Technology for providing such a stimulating atmosphere and conducive work environment.

Abstract

This project aims to develop a machine learning model for predicting airline customer satisfaction based on various factors such as flight duration, ticket price, seat comfort, in-flight entertainment, food quality, and overall customer service. The model is trained on a dataset of customer reviews and ratings from multiple airlines, and uses a variety of supervised learning techniques to predict customer satisfaction levels. The goal of this project is to provide airlines with insights into factors that influence customer satisfaction and to help them make data-driven decisions to improve their services and increase customer loyalty.Customer satisfaction questionnaires are a rich and strong source of information for companies to seek loyalty, customer and client retention, optimize resources, and repurchase products. Several advanced machine learning and statistical models have been employed to estimate the customer satisfaction score; however, there is not a single model that can yield the best result in all situations. Ensembles of regression techniques have demonstrated their effectiveness for various applications, where the success of these models lies in the construction of a set of single models.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

As with most industries, it can be a head-scratcher as to why customers choose a competing airline over yours and vice versa. For this reason, airline passenger satisfaction surveys can be a great asset to understanding consumer experiences, expectations, behaviors, and more.An airline passenger satisfaction survey refers to a type of market research that gathers feedback and information from passengers who have completed a trip with your airline or at your airport. Typically, a survey is sent a short while after the trip in an effort to learn more about the passenger's flying experience. Benefits of an airline passenger satisfaction survey:
Benchmarking:
To understand passenger satisfaction, you must first set a market research benchmark. Conducting your first passenger satisfaction survey sheds light on what you're doing right and what passengers think you could be doing better.With the first surveys being completed, you'll have data to compare to later on to see how things have improved.
A fresh perspective:
A third-party market research company, while executing a survey from beginning to end, will continuously provide you with unique, fresh, and unbiased recommendations. We don't have skin in the game, so we can share real, unfiltered insight on how you can improve your passenger experience.
Competition comparison:
Asking respondents how their experience compared to other competitors can shed light on how you stack up against others. The satisfaction survey can get valuable information from respondents on how they viewed the competition. These are just some of the benefits that conducting an airline passenger satisfaction survey can provide your organization.

1.2 Aim and Objectives

The objectives of this project are as follows:

The objective or goal of this project is to create a platform for airlines company to determine the important factors that influences the customer or passenger satisfaction.

These factors and surveys are important for the growth and reputation of a company.

This prediction would help the airline company to take right decisions towards customer service.

This will also help in analysing the results and taking right steps for company's betterment.

1.3 Motivation for the Work

The aviation industry is highly competitive, and airline companies are constantly striving to improve their services and customer satisfaction in order to gain a competitive edge. Customer satisfaction is a key factor that influences customer loyalty and retention, and it is therefore crucial for airlines to understand the factors that impact customer satisfaction.

The traditional methods of collecting customer feedback such as surveys and focus groups can be time-consuming, expensive, and often yield biased results. With the advent of machine learning and natural language processing techniques, it is now possible to analyze large volumes of customer data and extract meaningful insights from them.

The motivation for this project is to develop a machine learning model that can accurately predict customer satisfaction levels based on various factors, and provide insights that can help airlines make data-driven decisions to improve their services and customer experience. This can lead to increased customer loyalty, higher revenue, and a better reputation for the airline company.

1.4 Scope of Project

- Go to the website.
- Have a look at the various parameters on which whether you are satisfied or not will be decided.
- Mention the age, flight distance, gender and are you a loyal/disloyal customer.
- Then give the rating to the parameters on the basis of customer's experience. Ratings can be given between 0-5. (0- minimum satisfaction and 5- maximum satisfaction)
- Click on predict.
- The result will be displayed in the bottom. (The result could be either the customer is satisfied or the customer is not satisfied/neutral.)

1.5 Contribution

The airline satisfaction prediction project can contribute to society and an individual in several ways. It helps in improving customer experience by accurately predicting customer satisfaction levels and identifying factors that impact customer satisfaction, airlines can make data-driven decisions to improve their services and enhance the overall customer experience. This can lead to increased customer loyalty and retention, as well as a better reputation for the airline company.It also helps in increasing safety and security. Customer feedback and reviews can also provide valuable insights into safety and security issues, which can help airlines to identify and address potential safety hazards and security risks. It provides economic benefits too. The airline industry is a significant contributor to the global economy, and improving customer satisfaction levels can lead to increased revenue, job creation, and economic growth.Environmental benefits are also noticed by improving the overall customer experience, airlines may also encourage more people to travel by air, which can lead to a reduction in greenhouse gas emissions associated with other modes of transportation such as cars and buses.

1.6 Organization of the report

Title [1] - INTRODUCTION - It gives an overall idea of our project. It consists of Aim and objectives of our project,and motivation for our project and how we are contributing to the society.

Title [2] - LITERATURE SURVEY - In this, we have gone through multiple research papers such as Hosseini, S., Ansari, S. (2019). Airline customer satisfaction prediction using machine learning algorithms. Transportation research part C: emerging technologies, 107, 282-296. - "To study the service quality of flights and customer satisfaction." WE identified the problems in existing services and provided a solution for the same.

Title [3] - DESIGN IMPLEMENTATION - It consists an overall idea of our proposed system and how it works. Software and Hardware Requirements are specified. Idea of website is diagramatically represented such as:- Class diagram,Data flow diagram,Sequence diagram,Timeline chart. Algorithm ,Feasibility study and cost estimation is covered.

Title [4] - RESULTS AND DISCUSSION - Important code and Screenshots of GUI is provided here. Testing of accuracy and precision is done.

Title [5] - CONCLUSION - Summary of the entire project with future scope is covered here.

Chapter 2

Literature Survey

2.1 Introduction

The dataset for this project is obtained from Kaggle which contains the data sourced from a survey conducted by airlines on the satisfaction level of passengers/customers based on various factors. The dataset consists of 25 columns such as Age, Gender, Travel class, Arrival and Departure delays and also features that influences customer satisfaction level such as On-board service, Cleanliness, Seat comfort, Baggage handling etc. The dataset consists of a column or feature named 'satisfaction' which describes the overall satisfaction level of the customer. It has two values, 'neutral or dissatisfied' and 'satisfied'. This satisfaction feature is considered as the label feature. Airline satisfaction has been identified as an important topic for many airline companies. As well as it has been a important prediction prototype for academics.The airline satisfaction prediction focuses on understanding the factors that contribute to customer satisfaction and loyalty in the airline industry. There are multiple findings that have been done on these type of prediction models. The findings are the conclusion of using the model and the responses obtained from using the model. The findings are based on customer reviews, and these reviews help the company to identify whether they are doing good in their field and are the customers satisfied with what they provide.

2.2 Problem Definition

[1] Hosseini, S., Ansari, S. (2019). Airline customer satisfaction prediction using machine learning algorithms. Transportation research part C: emerging technologies, 107, 282-296.

It also helps the company to make any necessary changes in the type of services they provide or the service itself that they provide. Bad review and using the prediction model helps them identify any improvements that could be made and what the customers expect. One of the finding of this prediction model was service quality. Service quality has consistently been identified as a key factor in airline satisfaction. Studies have found that customers are most satisfied when they perceive the airline as providing high-quality service, including friendly and responsive staff, comfortable seating, and efficient operations. Another finding from the literature survey of airline satisfaction is Price. Price is another important factor in airline satisfaction, with many customers seeking affordable fares and value for money. This means, that customers

are always willing to pay a greater amount for best services and a comfortable travel. Studies have also found that customers are often willing to pay more for airlines that offer superior service and reliability. Thereby, customers don't choose airlines on basis of money but on the basis of services they enjoy and the type of travel they have.

[2] O'Neill, M., Palmer, A. (2020). Predicting passenger satisfaction: An examination of airline websites and online reviews. Journal of Air Transport Management, 82, 101774.

Literature survey also showed that loyalty programs have been found to be effective in increasing customer satisfaction and loyalty. Studies have found that customers who are members of loyalty programs are more likely to book with the same airline in the future and to recommend the airline to others.Thereby, as loyal the customers more he will prefer the same airline and is also expected to suggest this to others. Other findings include technology and customer service. Advances in technology have played a major role in airline satisfaction, with customers increasingly expecting convenient and efficient booking, check-in, and in-flight services. Studies have found that airlines that invest in technology and offer a seamless digital experience can significantly enhance customer satisfaction. Customers tend to choose the flight that have better digital space.

[3] Park, Y., Kim, J., Kim, K. (2018). Predicting passenger satisfaction with airline services: The case of low-cost carriers. Journal of Air Transport Management, 71, 34-41.

Customer service has also been identified as a key factor in airline satisfaction, with customers valuing helpful and friendly staff who are able to quickly resolve any issues or complaints. Customers loved their travel and experience when in flight attendees were good and respectful to them. The way the flight crew behaves with the customers plays an important role in the airline satisfaction review. Overall, these studies suggest that airline satisfaction is driven by a range of factors, including service quality, price, loyalty programs, technology, and customer service. Airlines that prioritize these factors and invest in enhancing the customer experience are more likely to achieve higher levels of satisfaction and loyalty among their customers.

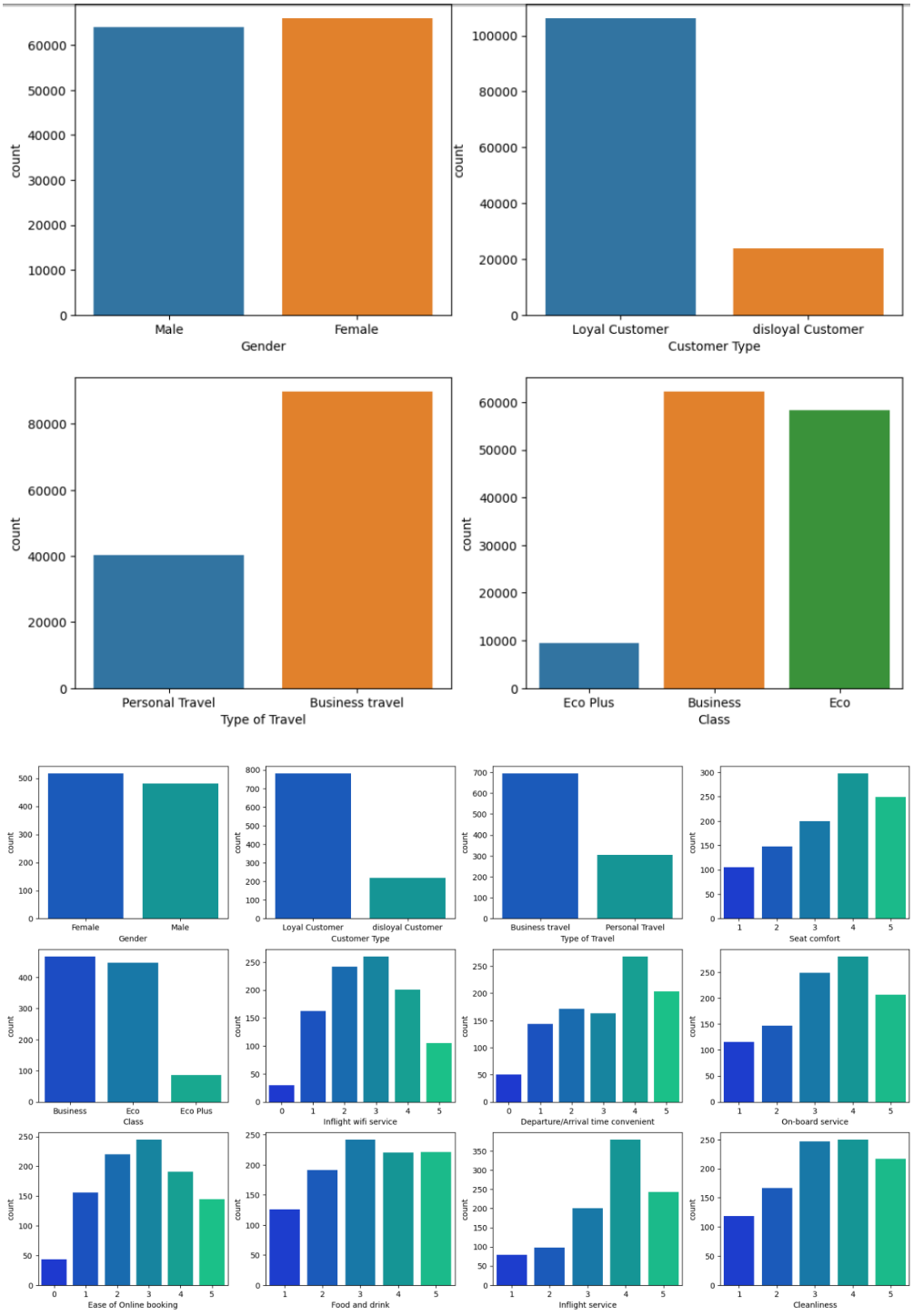
Chapter 3

Design Implementation

3.1 Proposed System

Our system provides a platform which includes basic yet necessary parameters such as Inflight wifi service, Departure/Arrival time convenient, Ease of Online booking, Gate location, Food and drink etc. Such parameters are not provided by any existing systems yet. The customer/traveller can give their feedback by giving ratings to each service. The customer can rate each service between 0-5. The rating 0 means that the service was worst while rating 5 would mean that the service provided was the best. We take into consideration multiple parameters. We also added an another parameter called as loyal/disloyal customer, which helps us in knowing if the customer has experienced any other flight company than itself. The system also takes into consideration any delay that has taken place while departure or arrival.

3.2 Requirement Gathering and Analysis



The requirement gathering and analysis of airline satisfaction prediction would involve the following steps:

Data collection and processing: Customer feedback and ratings data from multiple airlines would be collected and processed to create a dataset that can be used to train the machine learning model.

Feature selection and engineering: Relevant features and variables that can be

used to predict customer satisfaction levels would be selected and engineered to create new features that may improve the performance of the prediction model.

Model development and evaluation: Machine learning models using different algorithms and techniques would be developed and evaluated to identify the most accurate and effective model. The model would be trained on the dataset and validated using a test dataset.

Natural language processing: Natural language processing techniques would be used to process and analyze customer reviews and feedback to extract meaningful insights that can be used to improve customer satisfaction.

User interface development: A user-friendly interface would be developed to enable airlines to access and utilize the prediction model. The interface would provide insights on factors that impact customer satisfaction and recommendations on how to improve services.

Deployment and maintenance: The prediction model and user interface would be deployed, and regular maintenance and updates would be performed to ensure that the model continues to provide accurate predictions and insights.

3.3 Hardware Requirement

1. Processor - Intel Core i3 / Ryzen 5 3rd generation or above
2. Operating system - windows 7 or later,macOS, linux
3. SSD/HDD
4. 16 GB RAM

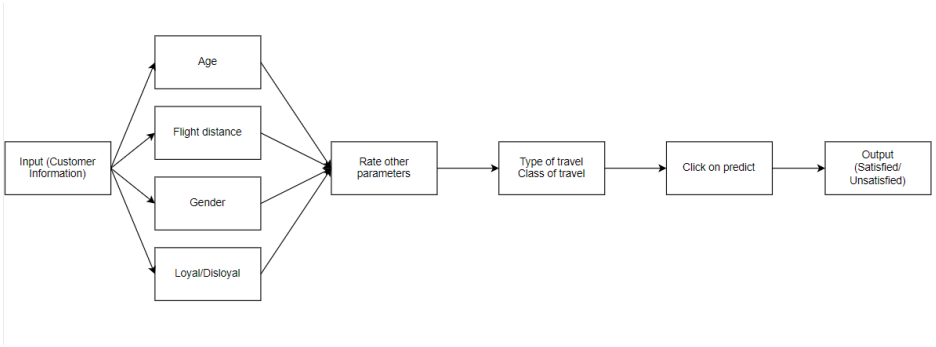
3.4 Software Requirement

1. Python 3.10.0
2. IDE - Colab notebook ver 6.4.12
3. Numpy version - 1.23.2
4. Pandas version - 1.4.3
5. Nltk version - 3.7
6. Tensorflow version - 2.9.1
7. Matplotlib version - 3.5.3

Figure 3.1: Block Diagram

3.5 UML Diagrams

3.5.1 Block Diagram



3.5.2 Architectural Diagram

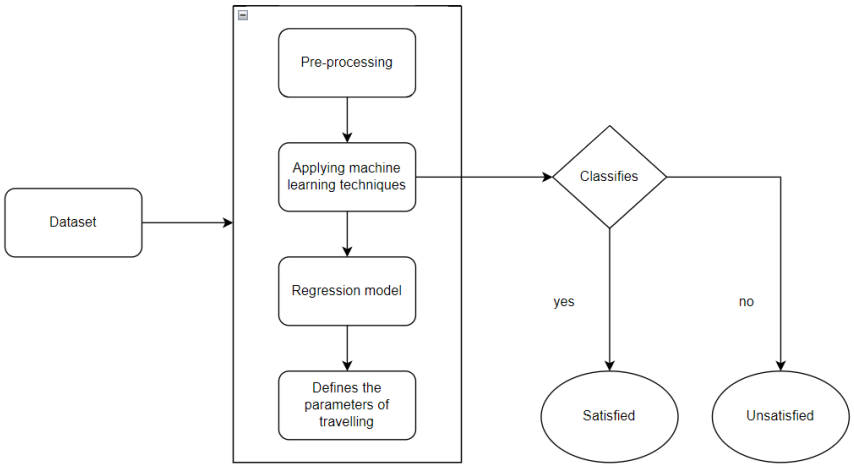


Figure 3.2: Architectural Diagram

3.5.3 Timeline Chart

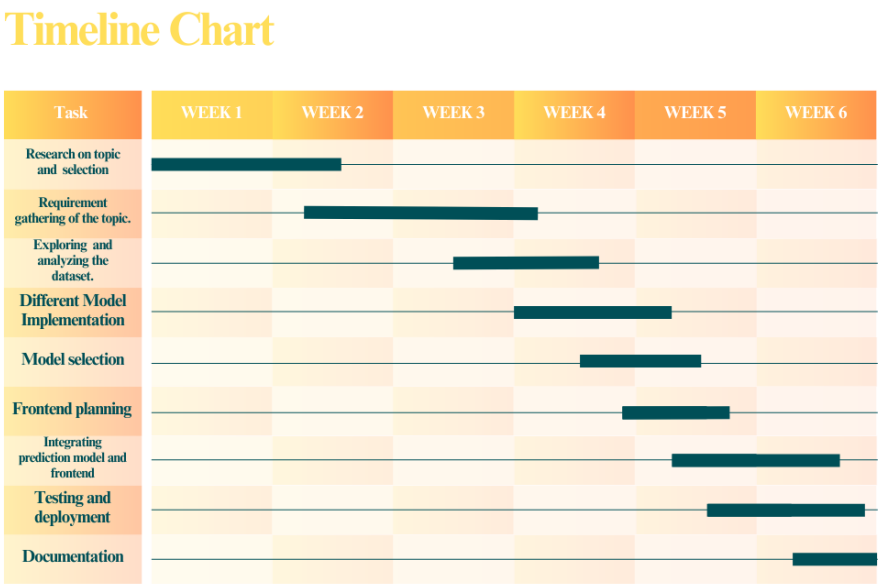


Figure 3.3: Timeline Chart

3.6 Feasibility Study

A feasibility study for an airline satisfaction prediction project would typically include an assessment of the technical, economic, and operational feasibility of the proposed system.

Here are some key considerations for each of these areas:

Technical feasibility:

Availability of customer feedback and ratings data from multiple airlines Availability of machine learning algorithms and natural language processing techniques for analyzing customer feedback Availability of suitable hardware and software infrastructure to support the development, deployment, and maintenance of the prediction model and user interface

Economic feasibility:

Cost of data collection and processing Cost of machine learning and natural language processing tools and technologies Cost of hardware and software infrastructure Potential benefits, such as increased customer loyalty and retention, improved reputation, and increased revenue
Operational feasibility:

Availability of skilled personnel to develop and maintain the prediction model and user interface Integration with existing airline systems and processes. Training and support for airline personnel to use the prediction model and user interface effectively.

Based on these considerations, an airline satisfaction prediction project is generally feasible, provided that there is access to customer feedback and ratings data from multiple airlines and that the costs and benefits of the project are carefully evaluated. Effective collaboration between technical experts, data scientists, and airline stakeholders is also essential for the success of the project.

Chapter 4

Results and Discussion

4.1 Code

```
[1] Importing necessary libraries

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import custom_functions as cf

[2] Visualizing the model

fig, axes = plt.subplots(3, 4, figsize = (20, 12))
sns.countplot(x = 'Gender', data = sample_data, palette= 'winter', ax = a
sns.countplot(x = 'Customer Type', data = sample_data, palette= 'winter',
sns.countplot(x = 'Type of Travel', data = sample_data, palette= 'winter'
sns.countplot(x = 'Seat comfort', data = sample_data, palette= 'winter',
sns.countplot(x = 'Class', data = sample_data, palette= 'winter', ax = ax
sns.countplot(x = 'Inflight wifi service', data = sample_data, palette= '
sns.countplot(x = 'Departure/Arrival time convenient', data = sample_data
sns.countplot(x = 'On-board service', data = sample_data, palette= 'winte
sns.countplot(x = 'Ease of Online booking', data = sample_data, palette=
sns.countplot(x = 'Food and drink', data = sample_data, palette= 'winter'
sns.countplot(x = 'Inflight service', data = sample_data, palette= 'winte
sns.countplot(x = 'Cleanliness', data = sample_data, palette= 'winter', a

[3] Training and Transforming the data

from sklearn.model_selection import train_test_split
X = df2.drop(columns = ['id','satisfaction'])
Y = df2['satisfaction']
train_X, test_X, train_y, test_y = train_test_split(X, Y, test_size=0.35,
\\
\\
\\
```

```
[4] Performance of the model

from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier(n_estimators=100)
rfc.fit(train_X,train_y)
#predicting test data
pred_y = rfc.predict(test_X)
#accuracy score
from sklearn.metrics import accuracy_score
score_rf = accuracy_score(test_y,pred_y)
score_rf

from sklearn.tree import DecisionTreeClassifier
dt = DecisionTreeClassifier(criterion = 'entropy', random_state = 0)
dt.fit(train_X, train_y)
y_pred = dt.predict(test_X)
score_dt = accuracy_score(test_y, y_pred)
print (score_dt)

import pickle
pickle_out = open('random_forest_classifier.pkl','wb')
pickle.dump(rfc,pickle_out)
pickle_out.close()

[5] Integrating the model with Streamlit

import streamlit as st
import pickle

with open('random_forest_classifier.pkl', 'rb') as file:
    clf = pickle.load(file)

def prediction(final_values):
    output = clf.predict([final_values])
    return output

def main():
    st.title('Passenger Satisfaction Prediction')

    slider_list = ['Inflight wifi service',
                'Departure/Arrival time convenient',
                'Ease of Online booking',
                'Gate location',
                'Food and drink',
                'Online boarding',
                'Seat comfort',
                'Inflight entertainment',
                'On-board service',
```

```
                'Leg room service',
                'Baggage handling',
                'Checkin service',
                'Inflight service',
                'Cleanliness']

rating = [-1]*14

col1, col2 = st.columns(2, gap='large')

with col1:
    age = st.number_input('Age', min_value=0,max_value=120, value=0)
    gender = st.radio('Gender',options=['Male','Female'])
    for i in range(0,len(slider_list),2):
        rating[i] = st.slider(slider_list[i], min_value=0, max_value
    departure_delay = st.number_input('Departure Delay(in Minutes)',
    travel_class = st.radio('Class',options=['Business','Economy', 'E

with col2:
    flight_distance = st.number_input('Flight Distance', min_value=0,
    cust_type = st.radio('Customer Type',options=['Loyal Customer','D
    for i in range(1,len(slider_list),2):
        rating[i] = st.slider(slider_list[i], min_value=0, max_value
    arrival_delay = st.number_input('Arrival Delay(in Minutes)', min_
    travel_type = st.radio('Type of Travel',options=['Business travel

if st.button('Predict'):
    final_values = []
    final_values += [age, flight_distance]
    final_values += rating
    final_values += [departure_delay, arrival_delay]

    if gender == 'Male':
        final_values += [0,1]
    else:
        final_values += [1,0]

    if cust_type == 'Loyal Customer':
        final_values += [1,0]
    else:
        final_values += [0,1]

    if travel_class == 'Business':
        final_values += [1,0,0]
    elif travel_class == 'Economy':
        final_values += [0,1,0]
    else:
        final_values += [0,0,1]
```

```
        if travel_type == 'Business travel':
            final_values += [1,0]
        else:
            final_values += [0,1]

        print(final_values)

        result = prediction(final_values)
        st.success('The customer is {}.'.format(result[0]))

if __name__ == '__main__':
    main()
```

4.2 Software Results

The implementation of project resulted in working of site in following way:

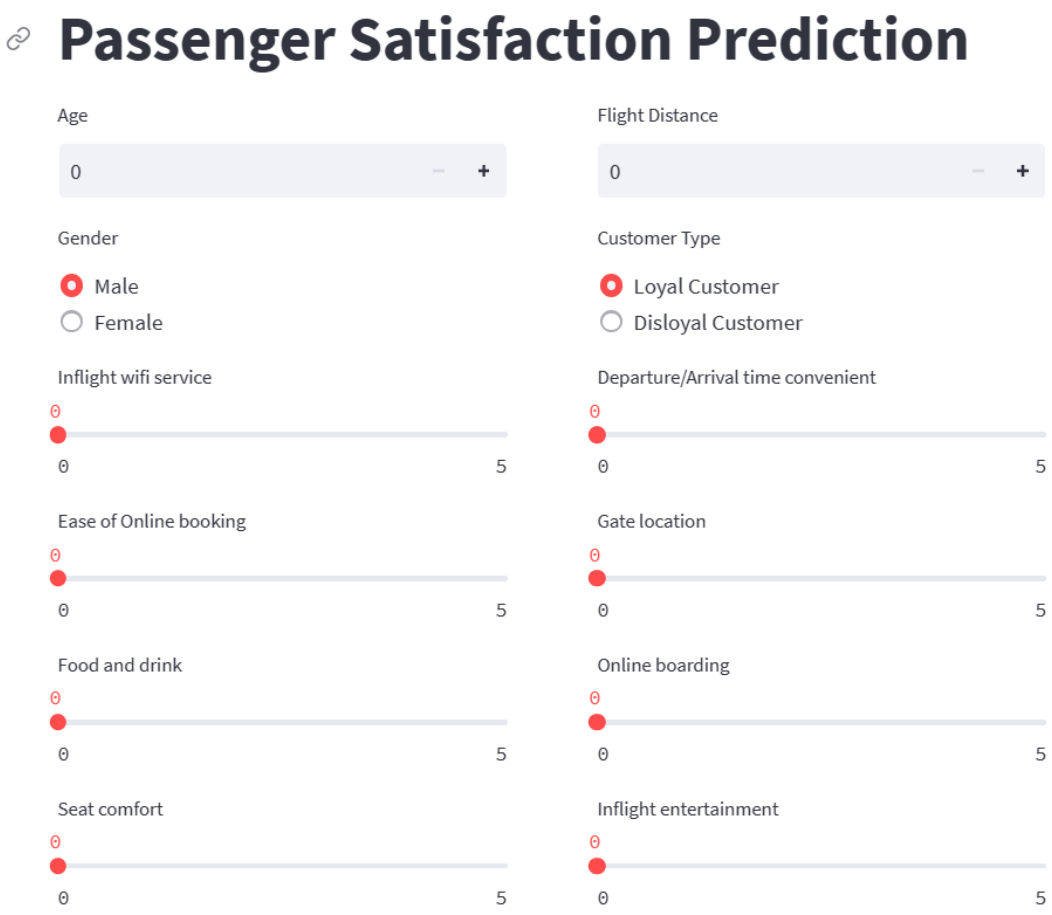
1) You put up all the ratings on the parameters according to your flight experience.

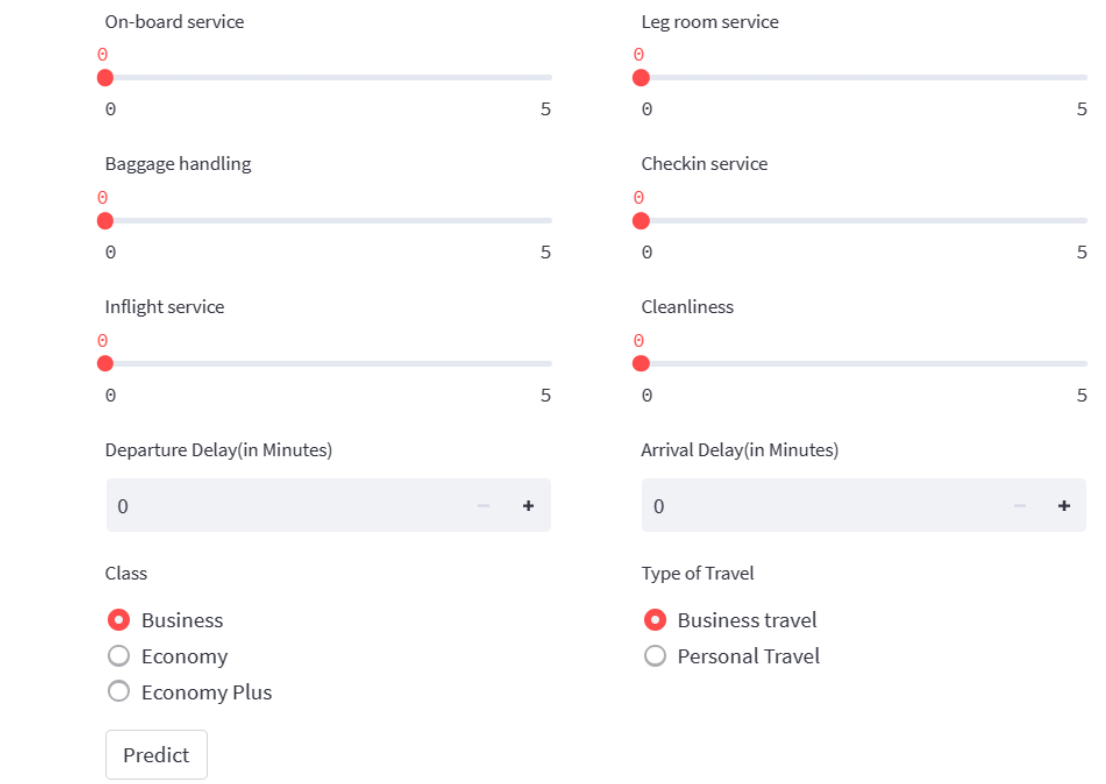
2) Submit the details by clicking on predict button.

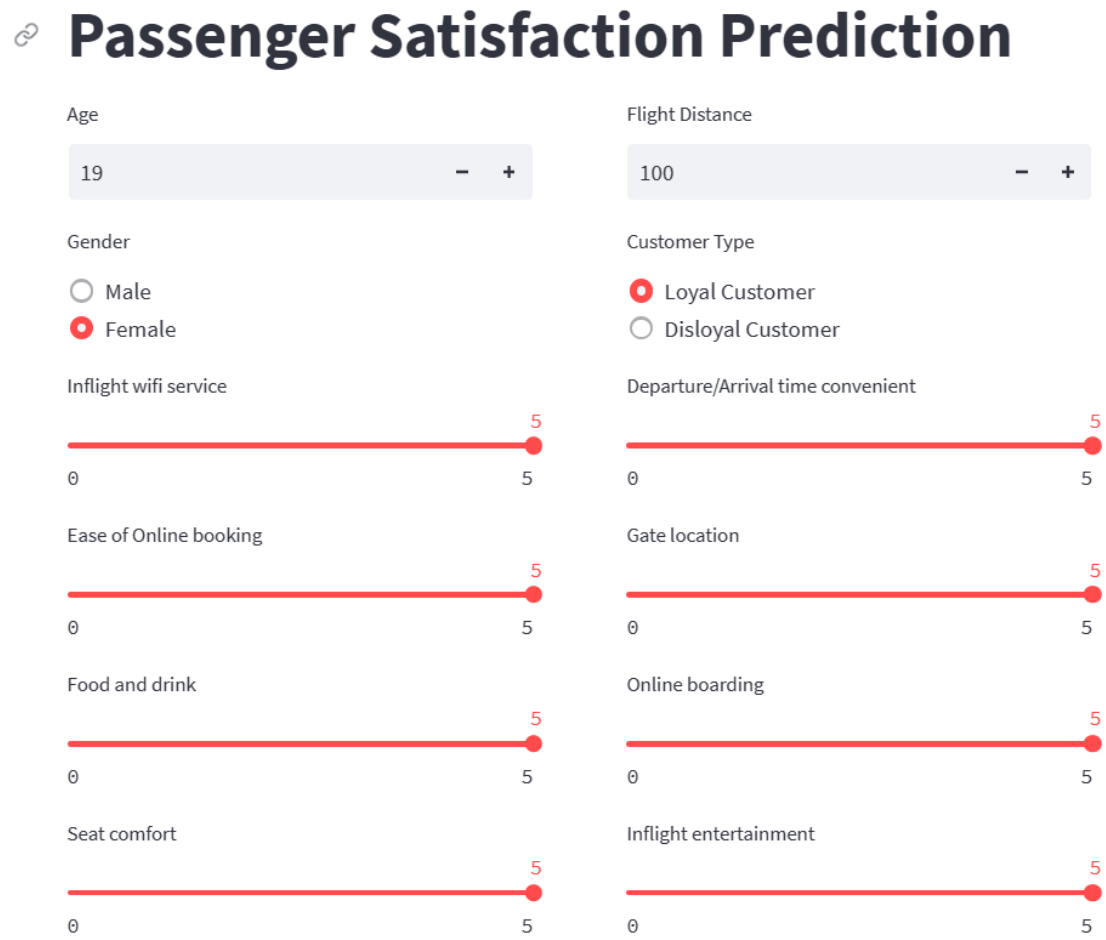
3) Model predicts whether the customer was satisfied or unsatisfied by the flight service.

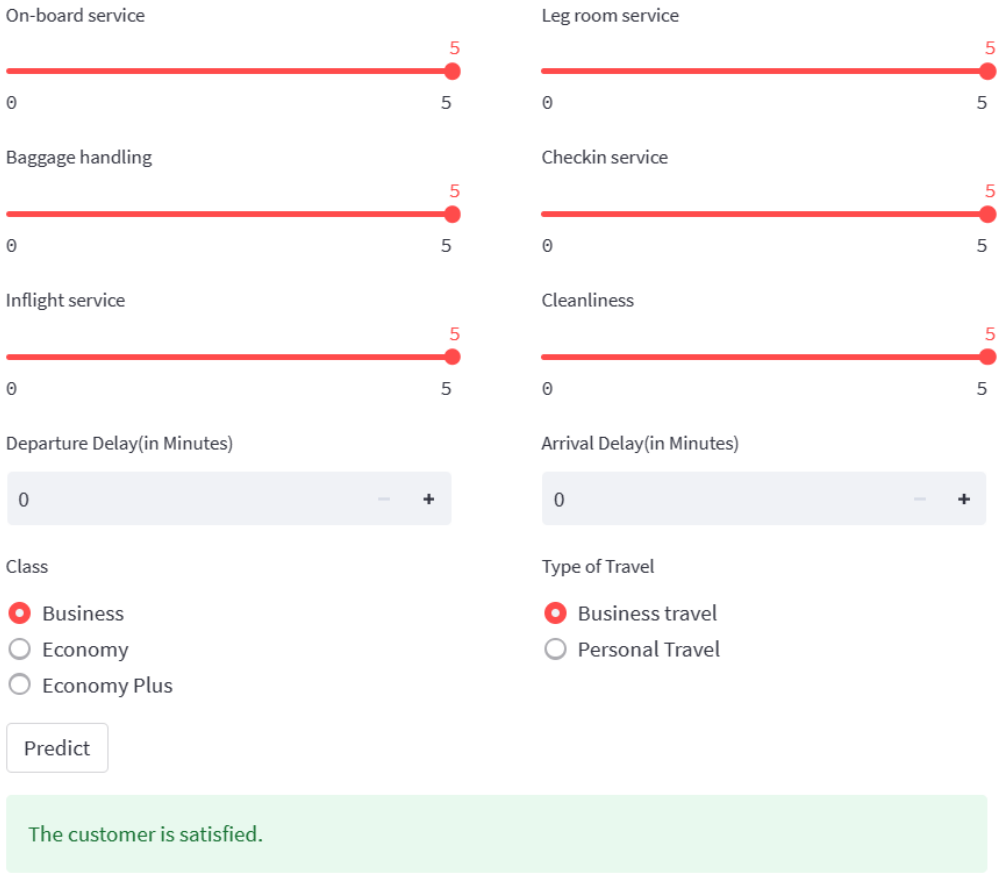
4) Based on these results, the company can improve their services and have a better insight about their customers.

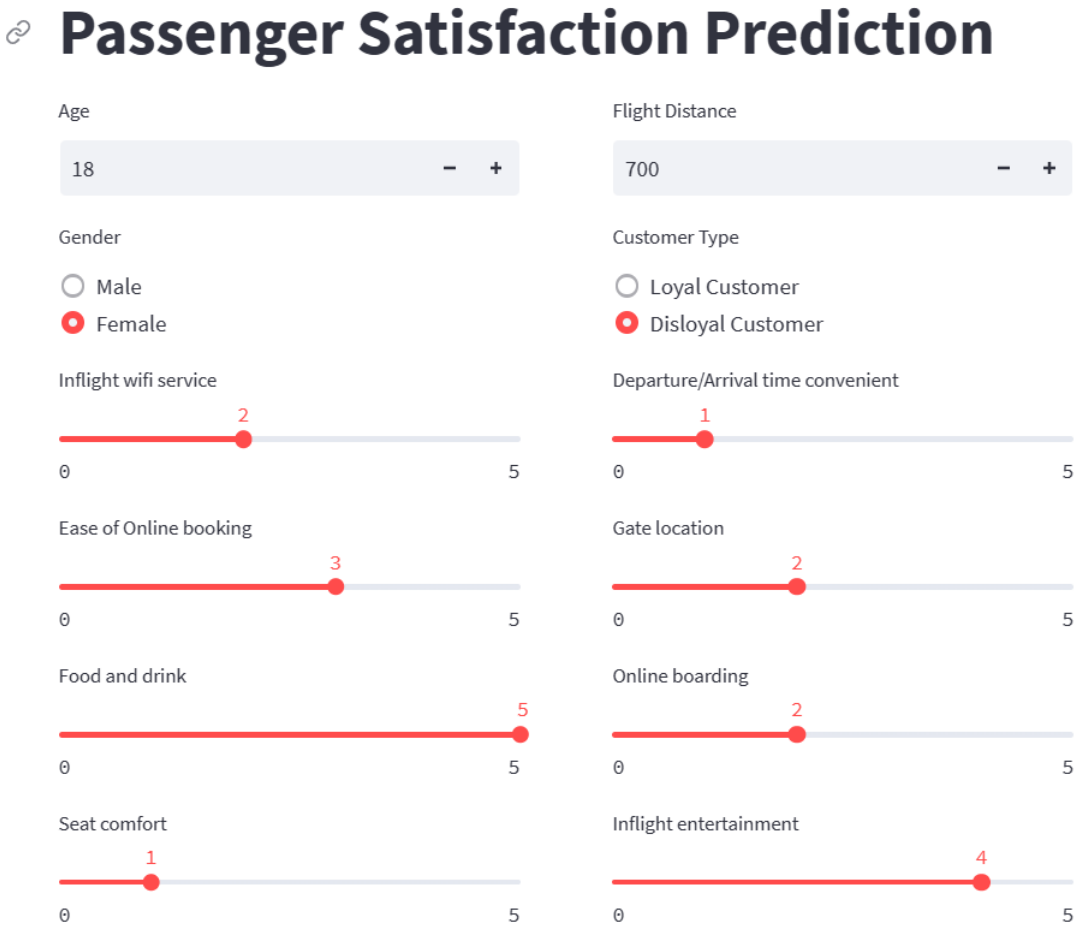
4.3 Screen Shots

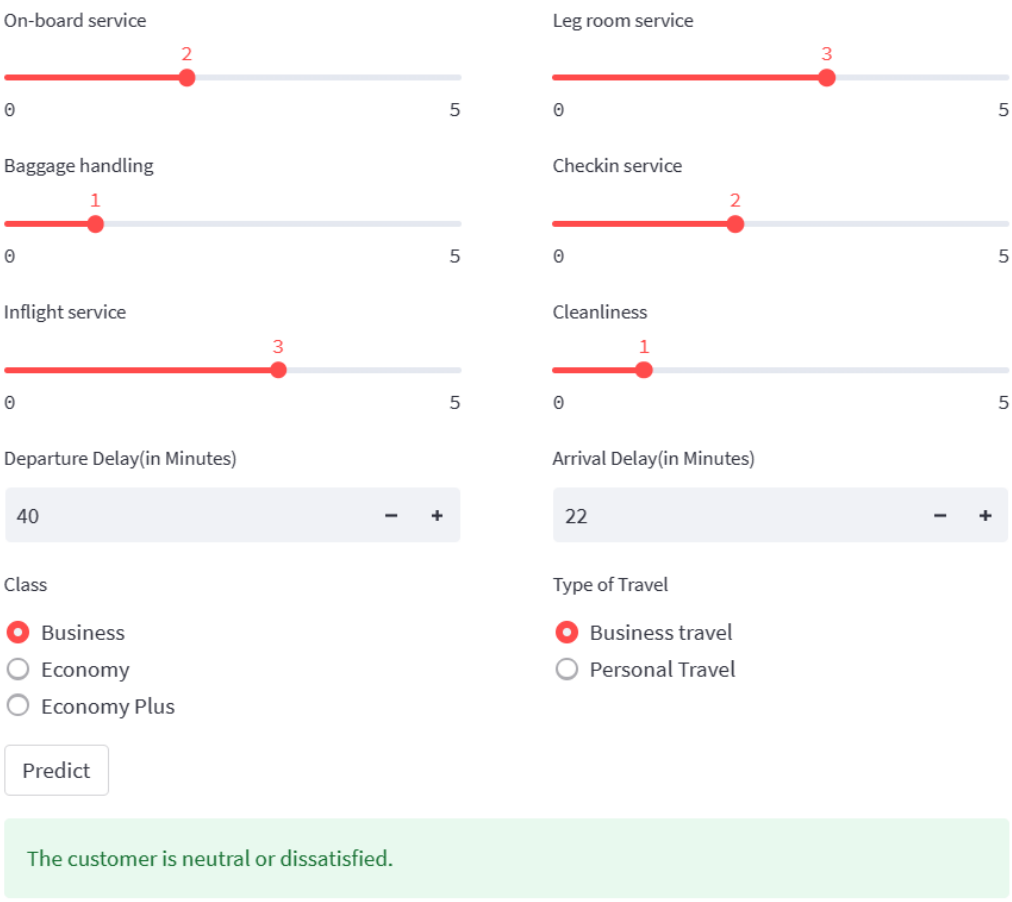



On-board service
0
0
5
Leg room service
0
0
5
Baggage handling
0
0
5
Checkin service
0
0
5
Inflight service
0
0
5
Cleanliness
0
0
5
Departure Delay(in Minutes)
0
Arrival Delay(in Minutes)
0
Class
Business
Economy
Economy Plus
Type of Travel
Business travel
Personal Travel
Predict


Passenger Satisfaction Prediction
Age
19
Flight Distance
100
Gender
Male
Female
Customer Type
Loyal Customer
Disloyal Customer
Inflight wifi service
5
0
5
Departure/Arrival time convenient
5
0
5
Ease of Online booking
5
0
5
Gate location
5
0
5
Food and drink
5
0
5
Online boarding
5
0
5
Seat comfort
5
0
5
Inflight entertainment
5
0
5


On-board service
5
0
5
Leg room service
5
0
5
Baggage handling
5
0
5
Checkin service
5
0
5
Inflight service
5
0
5
Cleanliness
5
0
5
Departure Delay(in Minutes)
0
Arrival Delay(in Minutes)
0
Class
Business
Economy
Economy Plus
Type of Travel
Business travel
Personal Travel
Predict
The customer is satisfied.


Passenger Satisfaction Prediction
Age
18
Flight Distance
700
Gender
Male
Female
Customer Type
Loyal Customer
Disloyal Customer
Inflight wifi service
2
0
5
Departure/Arrival time convenient
1
0
5
Ease of Online booking
3
0
5
Gate location
2
0
5
Food and drink
5
0
5
Online boarding
2
0
5
Seat comfort
1
0
5
Inflight entertainment
4
0
5


On-board service
2
0
5
Leg room service
3
0
5
Baggage handling
1
0
5
Checkin service
2
0
5
Inflight service
3
0
5
Cleanliness
1
0
5
Departure Delay(in Minutes)
40
Arrival Delay(in Minutes)
22
Class
Business
Economy
Economy Plus
Type of Travel
Business travel
Personal Travel
Predict
The customer is neutral or dissatisfied.

Chapter 5

Conclusion and Future Scope

Predicting airline satisfaction is a complex task that requires analyzing a variety of factors, including flight punctuality, onboard amenities, customer service, ticket prices, and overall travel experience. Machine learning algorithms can be trained on historical data to identify patterns and make predictions about future customer satisfaction levels. Additionally, the project might provide insights into the factors that contribute to airline satisfaction, which could be used by airlines to improve their services and increase customer satisfaction. It's important to note that customer satisfaction can be subjective and vary greatly between individuals. Therefore, predicting satisfaction levels with 100

The future scope of airline satisfaction prediction projects is vast, and there are several potential areas of development that could shape the future of this field. As more and more advanced machine learning algorithms are being developed, it is potentially possible that some machine algorithm would predict airline satisfaction to its maximum accuracy. This could include data from various data sources and making use of social media for better customer review. As the technology progresses, it is possible that new chatbots may be invented that would predict the airline satisfaction by using advanced features like natural language processing and sentiment analysis. This would definitely make the process faster and easier. Predictive analytics could also play a greater role in airline satisfaction prediction projects in the future, with airlines using data analytics to anticipate customer needs and preferences, and to provide personalized recommendations and offers.

References

[1] Hosseini, S., Ansari, S. (2019). Airline customer satisfaction prediction using machine learning algorithms. Transportation research part C: emerging technologies, 107, 282-296.

[2] O'Neill, M., Palmer, A. (2020). Predicting passenger satisfaction: An examination of airline websites and online reviews. Journal of Air Transport Management, 82, 101774.

[3] Park, Y., Kim, J., Kim, K. (2018). Predicting passenger satisfaction with airline services: The case of low-cost carriers. Journal of Air Transport Management, 71, 34-41.

[4] Sharma, A., Datta, S. (2018). A machine learning approach for predicting airline customer satisfaction. Journal of Air Transport Management, 67, 169-180.

[5] Wang, Y., Chen, K., Chen, M., Liu, C. (2019). Airline customer satisfaction prediction based on SVM and gradient boosting decision tree. Journal of Ambient Intelligence and Humanized Computing, 10(1), 253-264.