PROJECT REPORT

1. Introduction

1.1 Overview

There has been a remarkable expansion in commercial aviation in the past decade, with more people preferring air travel for a fast and comfortable journey. However, flight delays have become quite common across the world with the growth of the aviation sector. Besides inconvenience to the travellers, flight delays have a negative impact on the economy. The airline companies incur substantial monetary losses and observe a fall in their reputation if their flights are delayed often. The unforeseen delays also have a cascading effect on various other sectors. According to a report by the Joint Economic Committee of United States Congress, the total cost of flight delays to the US economy was over $40 billion with $19 billion to the airlines, $12 billion to the passengers and around $10 billion to other industries. The delayed flights also pose certain environmental concerns. Delayed flights consumed an additional 740 million units of jet fuel and released over 7 million metric tonnes of additional Carbon Dioxide[1].

1.2 Purpose

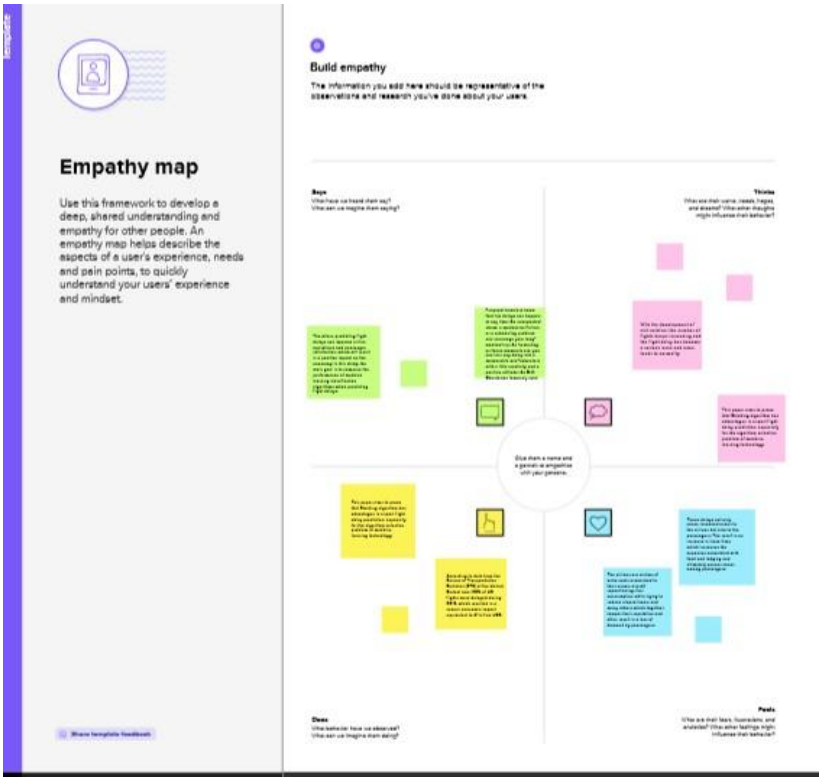
- Flight delays are gradually increasing and bring more financial difficulties and customer dissatisfaction to airline companies. To resolve this situation, supervised machine learning models were implemented to predict flight delays.

- Therefore, predicting flight delays can improve airline operations and passenger satisfaction, which will result in a positive impact on the economy. In this study, the main goal is to compare the performance of machine learning classification algorithms when predicting flight delays.

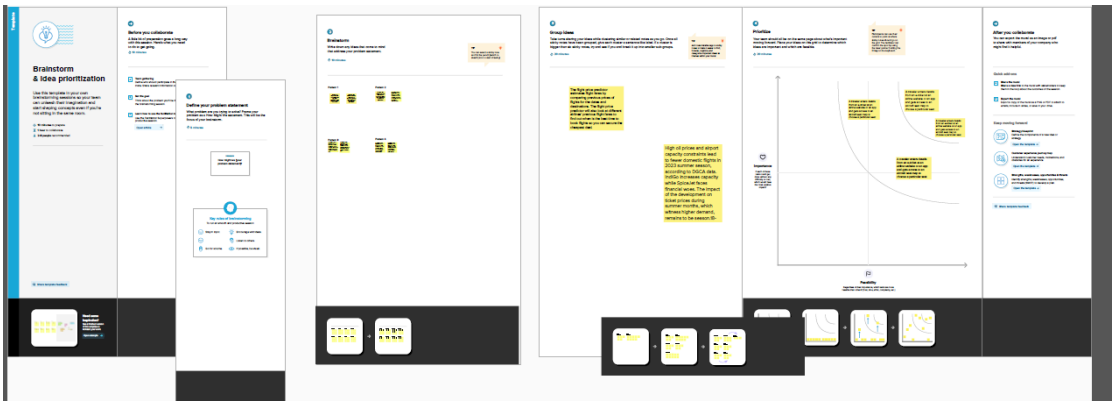
- Flight delays not only irritate air passengers and disrupt their schedules but also cause a decrease in efficiency, an increase in capital costs, reallocation of flight crews and aircraft, and additional crew expenses

2.Problem Definition &Design thinking

2.1 Empathy Map



2.2 Brainstorm



3.Result

3.1 Date Model

2. Advantages & Disadvantages

*** Effective prediction of flight delays will not only reduce the economic losses caused by airline delays, but also reduce the stay time and improve the traveling satisfaction of passengers.**

*** Carriers attribute flight delays to several causes such as bad weather conditions, airport congestion, airspace congestion, and use of smaller aircraft by airlines. These delays and cancellations tarnish the airlines' reputation, often resulting in loss of demand by passengers.**

6 . Appendix

```
#!/usr/bin/env python
# coding: utf-8


# In[1]:
```

```
from flask import Flask, request, jsonify, render_template,
url_for , request
import pickle
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
import pandas as pd

# Import dataset
df = pd.read_csv('Data/Processed_data15.csv')

# Label Encoding
le_carrier = LabelEncoder()
df['carrier'] = le_carrier.fit_transform(df['carrier'])

le_dest = LabelEncoder()
df['dest'] = le_dest.fit_transform(df['dest'])

le_origin = LabelEncoder()
df['origin'] = le_origin.fit_transform(df['origin'])
```

```
# Converting Pandas DataFrame into a Numpy array
X = df.iloc[:, 0:6].values # from column(years) to column(distance)
y = df['delayed']

X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.25,random_state=61) # 75% training and 25% test

app = Flask(__name__)
model = pickle.load(open('model.pkl', 'rb'))

@app.route('/')
def home():
    return render_template('index.html')

@app.route('/predict',methods=['POST'])
def predict():
    year = request.form['year']
    month = request.form['month']
    day = request.form['day']
```

```
            carrier = request.form['carrier']
            origin = request.form['origin']
            dest = request.form['dest']
            year = int(year)
            month = int(month)
            day = int(day)
            carrier = str(carrier)
            origin = str(origin)
            dest = str(dest)

            if year >= 2013:
               x1 = [year,month,day]
               x2 = [carrier, origin, dest]
               x1.extend(x2)
               df1 = pd.DataFrame(data = [x1], columns = ['year',
'month', 'date', 'carrier', 'origin', 'dest'])

               df1['carrier'] = le_carrier.transform(df1['carrier'])
               df1['origin'] = le_origin.transform(df1['origin'])
               df1['dest'] = le_dest.transform(df1['dest'])
```

```
    x = df1.iloc[:, :6].values

    ans = model.predict(x)

    output = ans


  return render_template('index.html',
prediction_text=output)


if__name__== '__main__':

    app.run(debug=False)

# For mac, make 'app.run(debug=True)'
```

TEAM PROFILE

- **Team leader: YOGESHWARI.R**
- **Member 1: PAVITHRA.P**
- **Member 2: UDHAYA KUMAR.R**
- **Member 3: NAVEEN.P**