Flight Delay Prediction For Aviation Industry Using Machine Learning

(NAAN MUDHALVAN PROJECT WORK)

PG & RESEARCH DEPARTMENT OF COMPUTER SCIENCE

THIRUVALLUVAR GOVT. ARTS COLLEGE, RASIPURAM

(Affiliated to Periyar University, Salem-11)

APRIL 2023

THIRUVALLUVAR GOVT ARTS COLLEGE, RASIPURAM

(Affiliated to Periyar university, salem-11)

P.G. AND RESEARCH DEPARTMENT OF COMPUTER SCIENCE

(NAAN MUDHALVAN PROJECT WORK)

TITLE :

Flight Delay Prediction For Aviation Industry Using Machine Learning

By

Team ID : NM2023TMID24468

Team Leader : KARTHIKEYAN V (20UCS2301)

Team Member : BALAJI S (20UCS2298)

Team Member : BHARATHAN P S (20UCS2299)

Team Member : GOKULAKRISHNAN P (20UCS2300)

Team Member : PANDI DURAI R (20UCS2305)

GITHUP LINK :

https://github.com/Karthikeyan-Vaiyapuri/flight-delay-prediction-for-aviation-industry-using-machine-learning.git

INTRODUCTION

1.1 OVERVIEW

Binary Classification:

Predicting whether a flight will be delayed or not. This is a common approach where the problem is framed as a binary classification task, where the model predicts whether a flight will be delayed (1) or not delayed (0) based on historical data, such as weather conditions, airline schedules, and previous flight delays.

Multiclass Classification:

Predicting the severity of flight delays. This approach involves categorizing flights into multiple classes based on the severity of delays, such as mild delay, moderate delay, and severe delay. This can provide more detailed insights and allow airlines to take appropriate actions based on the predicted severity of delays.

1.2 PURPOSE

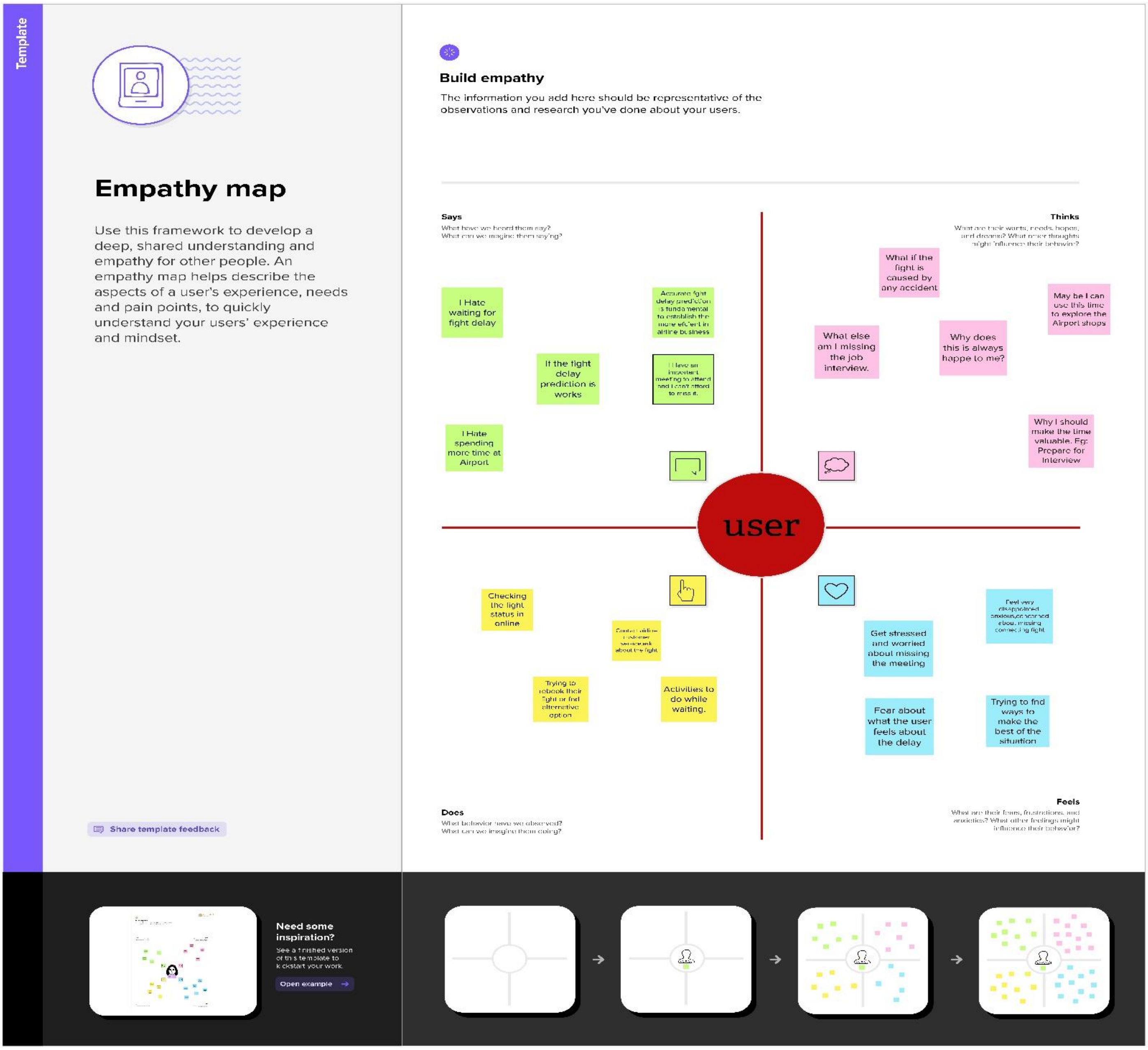
Develop a model that predicts flight delays based on weather conditions, such as temperature, precipitation, wind speed, and visibility. The model can use historical weather data in combination with flight data to identify patterns and correlations between weather conditions and flight delays.Build a model that predicts flight delays based on the historical performance of airlines, such as on-time performance, previous delays, and cancellations. The model can leverage data from various airlines to identify trends and patterns in airline performance that may impact flight delay

PROBLEM DEFINITION & DESIGN THINKING

2.1 EMPATHY MAP

An empathy map is a collaborative visualization used to articulate what we know about a particular type of user. It externalizes knowledge about users in order to

- create a shared understanding of user needs
- aid in decision makein



Template

Empathy map

Use this framework to develop a deep, shared understanding and empathy for other people. An empathy map helps describe the aspects of a user's experience, needs and pain points, to quickly understand your users' experience and mindset.

Share template feedback

Build empathy

The information you add here should be representative of the observations and research you've done about your users.

Says

What have we heard them say? What can we imagine them saying?

I Hate waiting for fight delay

Accurate fight delay prediction is fundamental to establish the more efcient in airline business

If the fight delay prediction is works

I have an important meeting to attend and I can't afford to miss it.

I Hate spending more time at Airport

Thinks

What are their wants, needs, hopes, and dreams? What other thoughts might influence their behavior?

What if the fight is caused by any accident

May be I can use this time to explore the Airport shops

What else am I missing the job interview.

Why does this is always happe to me?

Why I should make the time valuable. Eg: Prepare for Interview

user

Checking the light status in online

Contact airline customer service ask about the fight

Trying to rebook their fight or find alternative option

Activities to do while waiting.

Feel very disappointed anxious,concerned about missing connecting fight

Get stressed and worried about missing the meeting

Fear about what the user feels about the delay

Trying to find ways to make the best of the situation

Does

What behavior have we observed? What can we imagine them doing?

Feels

What are their fears, frustrations, and anxieties? What other feelings might influence their behavior?

Need some inspiration?

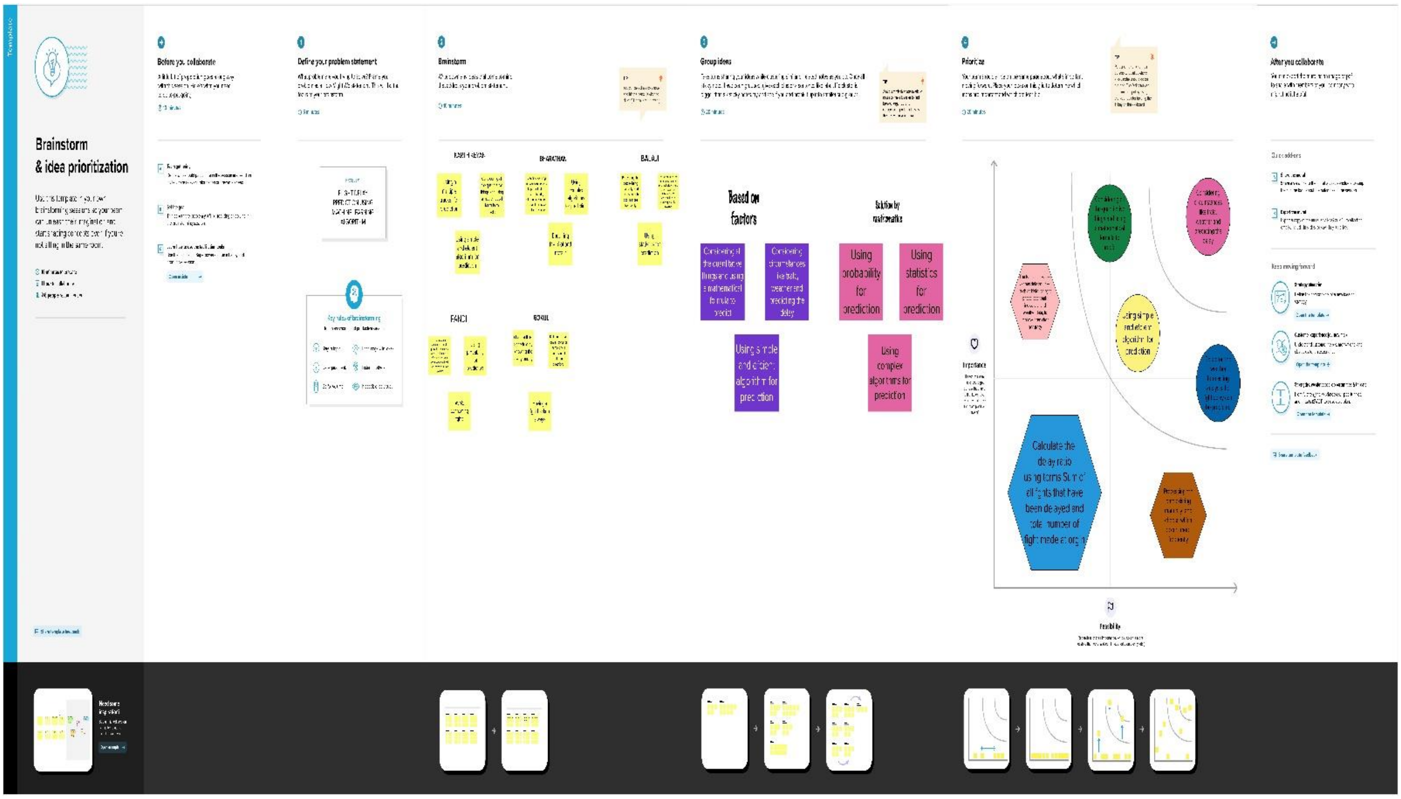
See a finished version of this template to kickstart your work.

Open example

2.2 IDEATION AND BRAINSTROMING

- **Time of Day and Seasonal Patterns:** Develop a model that captures time of day and seasonal patterns in flight delays, such as peak travel times, holiday seasons, and weather-related patterns. The model can help airlines proactively manage operations during high-impact periods to reduce delays.

- **Airport Congestion Prediction:** Build a model that predicts flight delays based on airport congestion levels, such as runway capacity, air traffic control delays, and gate availability. The model can help airlines anticipate potential congestion-related delays and make operational adjustments according

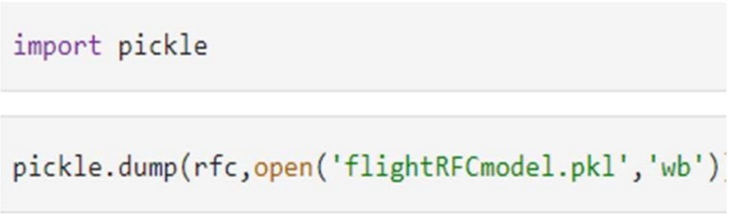


Brainstorm
& idea prioritization
Before you collaborate
Define your problem statement
Brainstorm
Group ideas
Prioritize
After you collaborate
Based on factors
Solution by mathematics
Considering all the quantitative things and using a mathematical formula to predict
Considering circumstances like traffic, weather and predicting the delay
Using probability for prediction
Using statistics for prediction
Using simple and efficient algorithm for prediction
Using complex algorithms for prediction
Using simple and efficient algorithm for prediction
Calculate the delay ratio using terms Sum of all flights that have been delayed and total number of flight made at origin
Importance
Feasibility

MODEL DEPLOYMENT

3.1 SAVE THE BEST MODEL

Saving the best model after comparing its performance using different evaluation metrics means selecting the model with the highest performance and saving its weights and configuration. This can be useful in avoiding the need to retrain the model every time it is needed and also to be able to use it in the future.



```
import pickle
```

```
pickle.dump(rfc,open('flightRFCmodel.pkl','wb')
```

3.2 INTEGRATE WITH WEB FRAMEWORK

In this section, we will be building a web application that is integrated to the model we built. A UI is provided for the uses where he has to enter the values for predictions. The enter values are given to the saved model and prediction is showcased on the UI.

This section has the following tasks

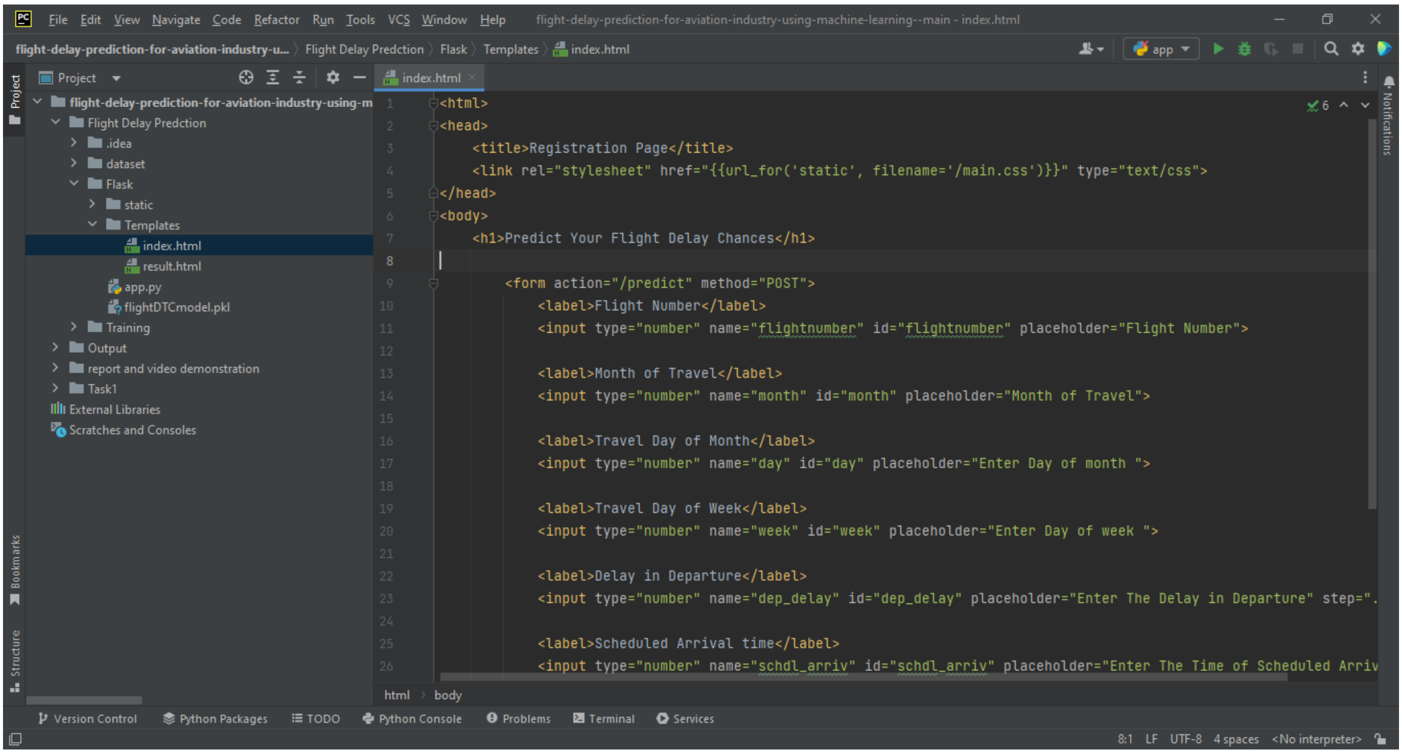
- **Building HTML Pages**
- **Building server side script**
- **Run the web application**

3.2.1 Building html pages

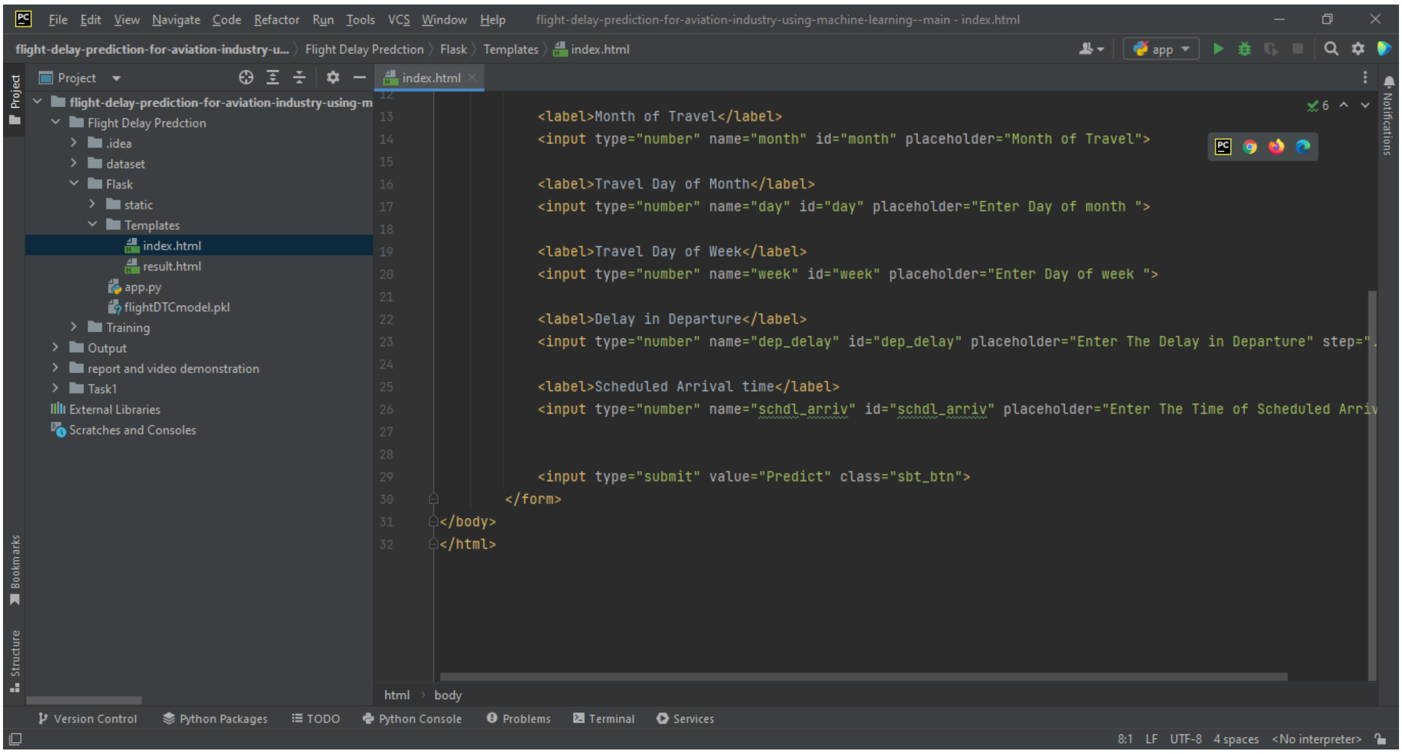
In this project we have created two html pages

- Index.html
- Result.html

Index.html :

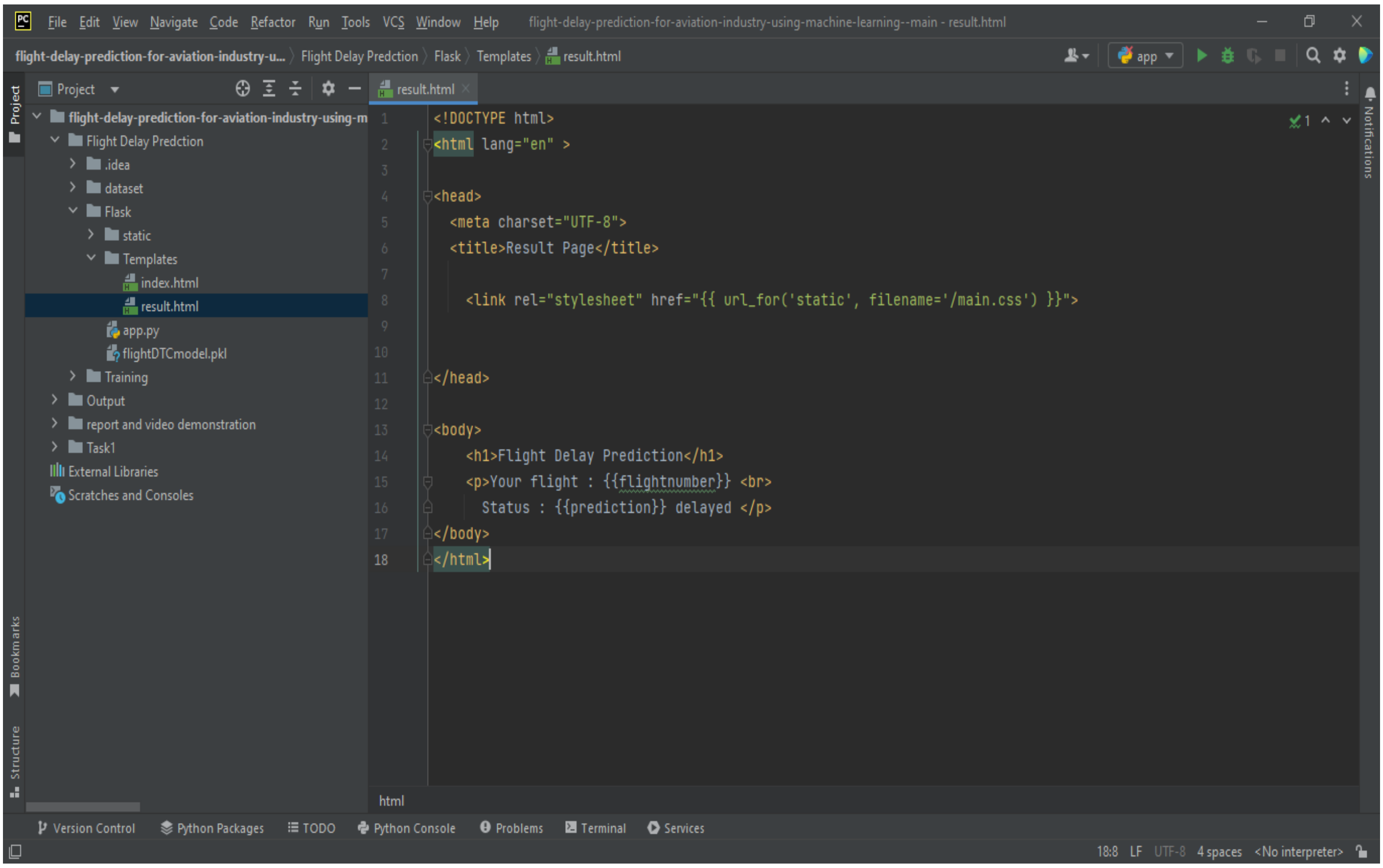


```
<html>
<head>
    <title>Registration Page</title>
    <link rel="stylesheet" href="{{url_for('static', filename='/main.css')}}" type="text/css">
</head>
<body>
    <h1>Predict Your Flight Delay Chances</h1>

        <form action="/predict" method="POST">
            <label>Flight Number</label>
            <input type="number" name="flightnumber" id="flightnumber" placeholder="Flight Number">

            <label>Month of Travel</label>
            <input type="number" name="month" id="month" placeholder="Month of Travel">

            <label>Travel Day of Month</label>
            <input type="number" name="day" id="day" placeholder="Enter Day of month ">

            <label>Travel Day of Week</label>
            <input type="number" name="week" id="week" placeholder="Enter Day of week ">

            <label>Delay in Departure</label>
            <input type="number" name="dep_delay" id="dep_delay" placeholder="Enter The Delay in Departure" step=".

            <label>Scheduled Arrival time</label>
            <input type="number" name="schdl_arriv" id="schdl_arriv" placeholder="Enter The Time of Scheduled Arriv
```



```
            <label>Month of Travel</label>
            <input type="number" name="month" id="month" placeholder="Month of Travel">

            <label>Travel Day of Month</label>
            <input type="number" name="day" id="day" placeholder="Enter Day of month ">

            <label>Travel Day of Week</label>
            <input type="number" name="week" id="week" placeholder="Enter Day of week ">

            <label>Delay in Departure</label>
            <input type="number" name="dep_delay" id="dep_delay" placeholder="Enter The Delay in Departure" step=".

            <label>Scheduled Arrival time</label>
            <input type="number" name="schdl_arriv" id="schdl_arriv" placeholder="Enter The Time of Scheduled Arriv


            <input type="submit" value="Predict" class="sbt_btn">
        </form>
</body>
</html>
```

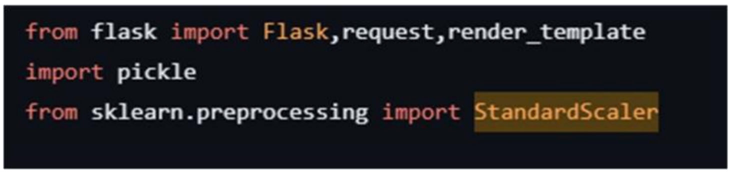
Result.html


File Edit View Navigate Code Refactor Run Tools VCS Window Help
flight-delay-prediction-for-aviation-industry-using-machine-learning--main - result.html
flight-delay-prediction-for-aviation-industry-u... 〉 Flight Delay Predction 〉 Flask 〉 Templates 〉 result.html
app
Project
result.html
flight-delay-prediction-for-aviation-industry-using-m
Flight Delay Predction
.idea
dataset
Flask
static
Templates
index.html
result.html
app.py
flightDTCmodel.pkl
Training
Output
report and video demonstration
Task1
External Libraries
Scratches and Consoles
Bookmarks
Structure
Notifications
html
Version Control
Python Packages
TODO
Python Console
Problems
Terminal
Services
18:8 LF UTF-8 4 spaces <No interpreter>

```html
<!DOCTYPE html>
<html lang="en" >

<head>
  <meta charset="UTF-8">
  <title>Result Page</title>

    <link rel="stylesheet" href="{{ url_for('static', filename='/main.css') }}">


</head>

<body>
    <h1>Flight Delay Prediction</h1>
    <p>Your flight : {{flightnumber}} <br>
      Status : {{prediction}} delayed </p>
</body>
</html>
```

3.2.2 Building server side script

Import the libraries



```
from flask import Flask,request,render_template
import pickle
from sklearn.preprocessing import StandardScaler
```

Load the saved model. Importing the flask module in the project is mandatory. An object of Flask class is our WSGI application. Flask constructor takes the name of the current module (__name__) as argument.

Main Function:

```
if __name__=='__main__':
        app.run(debug=True)
```

ADVANTAGES & DISADVANTAGES

ADVANTAGES :

- Conduct feature engineering to identify and extract relevant features from the collected data that can be used as input features for the machine learning model.
- This may involve selecting relevant features, normalizing or scaling features, and creating new features that may capture important information related to flight delays.
- Evaluate the performance of the developed model using appropriate evaluation metrics, such as accuracy, precision, recall, F1-score, and AUC-ROC, depending on the problem type (binary classification, multi- class classification, regression, etc.).
- Validate the model's performance on a separate test set to ensure its generalization ability.
- Perform thorough model analysis, interpretability, and sensitivity analysis to gain insights into its behavior and limitations.

DISADVANTAGES :

- Refine the selected model(s) based on the evaluation results and feedback.
- Fine-tune the model hyperparameters, adjust the feature engineering techniques, or explore ensemble methods to improve the model performance.
- Repeat the model training and evaluation process iteratively until the desired performance level is Achieved.
- Deploy the trained and refined machine learning model(s) in a production environment, such as a web application, mobile app, or cloud- based service, depending on the project requirements.
- Ensure that the model is integrated seamlessly into the operational workflow and follows the necessary security and privacy measures.
- Conduct thorough testing and validation of the deployed model(s) to ensure its accuracy, reliability, and stability in real-world scenarios

CONCLUSION

The paper performed a prediction of the occurrence of flight delays by adapting it into a machine learning problem. A supervised machine learning approach in the form of binary classification was used for the prediction.

Seven algorithms were used for delay prediction, and four measures were used for algorithms performance evaluation. Due to the imbalanced nature of the data set, evaluation measures were weighted to eliminate the dominant effect of non-delayed flights over delayed flights. After applying classifiers to the delay prediction, the values of their four measures were compared to evaluate the performance of each model.The data set selected for this paper is imbalanced distributed, which may cause significant variation in the performance of each algorithm. In this paper, this problem was solved by the use of weighted evaluation measures. For future studies, using techniques such as SMOTE can better resolve this imbalance and improve the prediction. The result of algorithm comparison shows that tree-based ensemble algorithms tend to better predict flight delays of this data set. It will be valuable to repeat similar experimental processes using more tree-based ensemble algorithms to discover their significance in flight delay prediction.

FUTURE SCOPE

In conclusion, machine learning can be a useful tool in predicting flight delays

These delays and cancellations tarnish the airlines reutation, often resulting in loss of demand by passenger

Further research and development are needed to improve the accuracy and reliability of machine learning models in predicting flight delays. Carriers attribute fight to several causes such as bad weather conditions,airspace congestion and use of smaller aircraft by airlines.

APPENDIX

Flight Delay Prediction For Aviation IndustryUsing Machine Learning

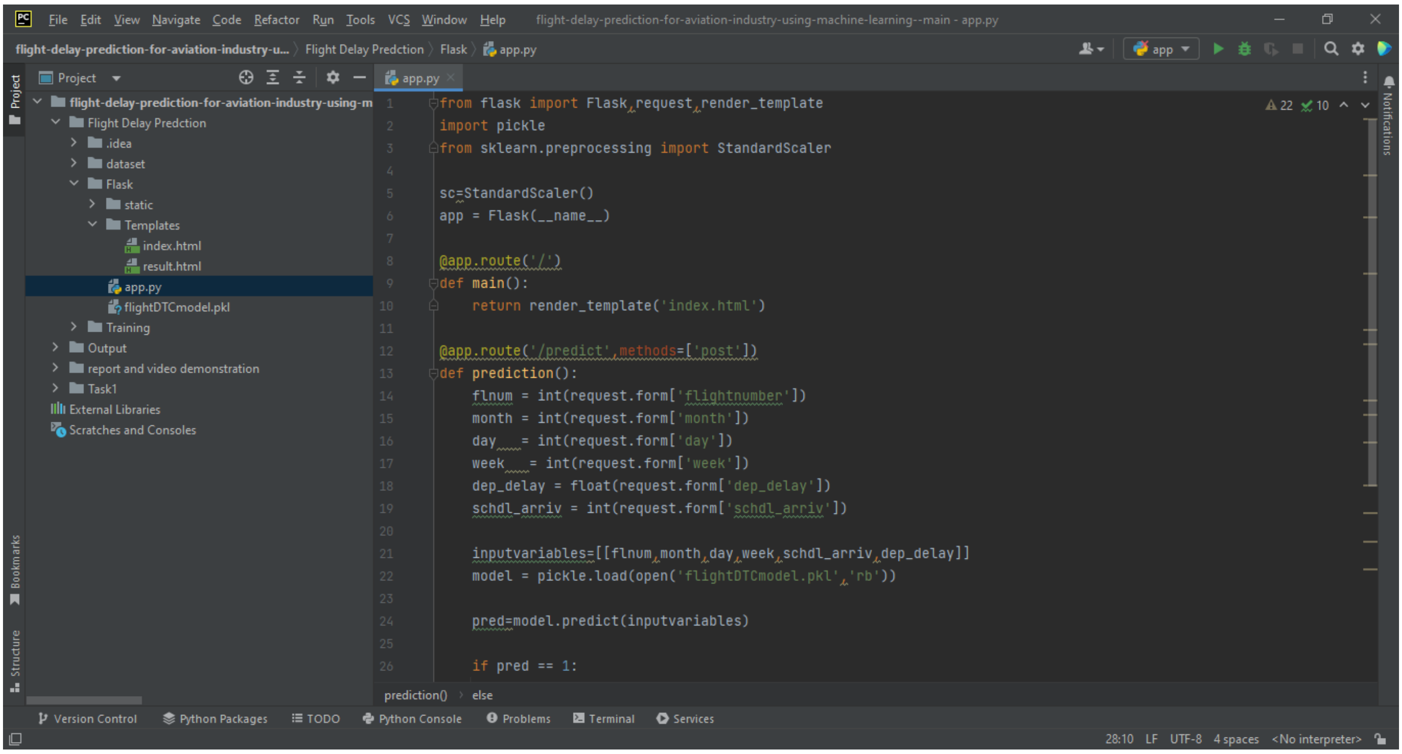
GitHub Link :

https://github.com/Karthikeyan-Vaiyapuri/flight-delay-prediction-for-aviation-industry-using-machine-learning.git

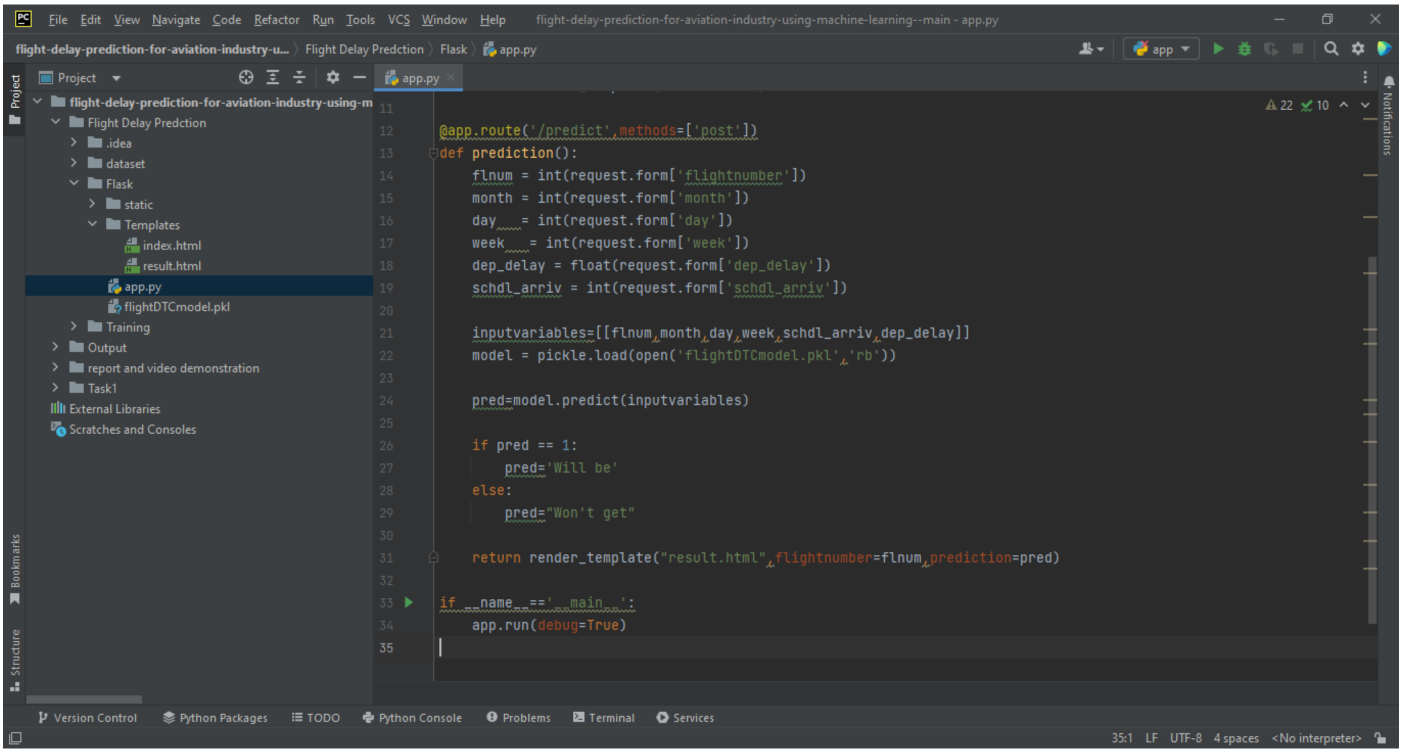
Video Link :

https://youtu.be/AyGWXjxzr2E

SAMPLE CODING

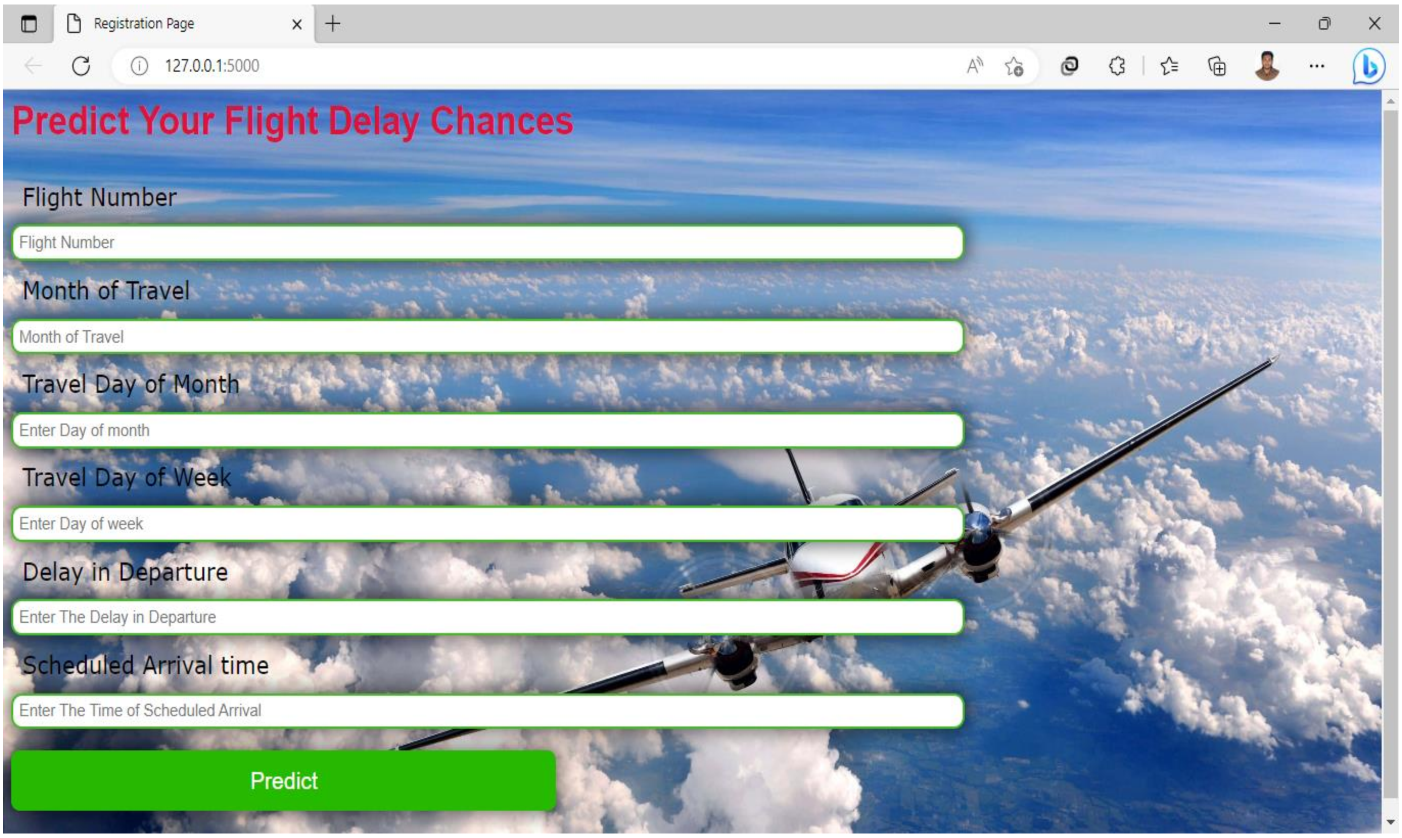

File Edit View Navigate Code Refactor Run Tools VCS Window Help
flight-delay-prediction-for-aviation-industry-using-machine-learning--main - app.py
flight-delay-prediction-for-aviation-industry-u... 〉 Flight Delay Predction 〉 Flask 〉 app.py
Project
flight-delay-prediction-for-aviation-industry-using-m
Flight Delay Predction
.idea
dataset
Flask
static
Templates
index.html
result.html
app.py
flightDTCmodel.pkl
Training
Output
report and video demonstration
Task1
External Libraries
Scratches and Consoles
prediction() 〉 else
Version Control
Python Packages
TODO
Python Console
Problems
Terminal
Services
28:10 LF UTF-8 4 spaces <No interpreter>

```python
from flask import Flask,request,render_template
import pickle
from sklearn.preprocessing import StandardScaler

sc=StandardScaler()
app = Flask(__name__)

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

@app.route('/predict',methods=['post'])
def prediction():
    flnum = int(request.form['flightnumber'])
    month = int(request.form['month'])
    day  = int(request.form['day'])
    week  = int(request.form['week'])
    dep_delay = float(request.form['dep_delay'])
    schdl_arriv = int(request.form['schdl_arriv'])

    inputvariables=[[flnum,month,day,week,schdl_arriv,dep_delay]]
    model = pickle.load(open('flightDTCmodel.pkl','rb'))

    pred=model.predict(inputvariables)

    if pred == 1:
```


File Edit View Navigate Code Refactor Run Tools VCS Window Help
flight-delay-prediction-for-aviation-industry-using-machine-learning--main - app.py
flight-delay-prediction-for-aviation-industry-u... 〉 Flight Delay Predction 〉 Flask 〉 app.py
Project
Version Control
Python Packages
TODO
Python Console
Problems
Terminal
Services
35:1 LF UTF-8 4 spaces <No interpreter>

```python

@app.route('/predict',methods=['post'])
def prediction():
    flnum = int(request.form['flightnumber'])
    month = int(request.form['month'])
    day  = int(request.form['day'])
    week  = int(request.form['week'])
    dep_delay = float(request.form['dep_delay'])
    schdl_arriv = int(request.form['schdl_arriv'])

    inputvariables=[[flnum,month,day,week,schdl_arriv,dep_delay]]
    model = pickle.load(open('flightDTCmodel.pkl','rb'))

    pred=model.predict(inputvariables)

    if pred == 1:
        pred='Will be'
    else:
        pred="Won't get"

    return render_template("result.html",flightnumber=flnum,prediction=pred)

if __name__=='__main__':
    app.run(debug=True)

```

OUTPUT



Predict Your Flight Delay Chances

Flight Number

Flight Number

Month of Travel

Month of Travel

Travel Day of Month

Enter Day of month

Travel Day of Week

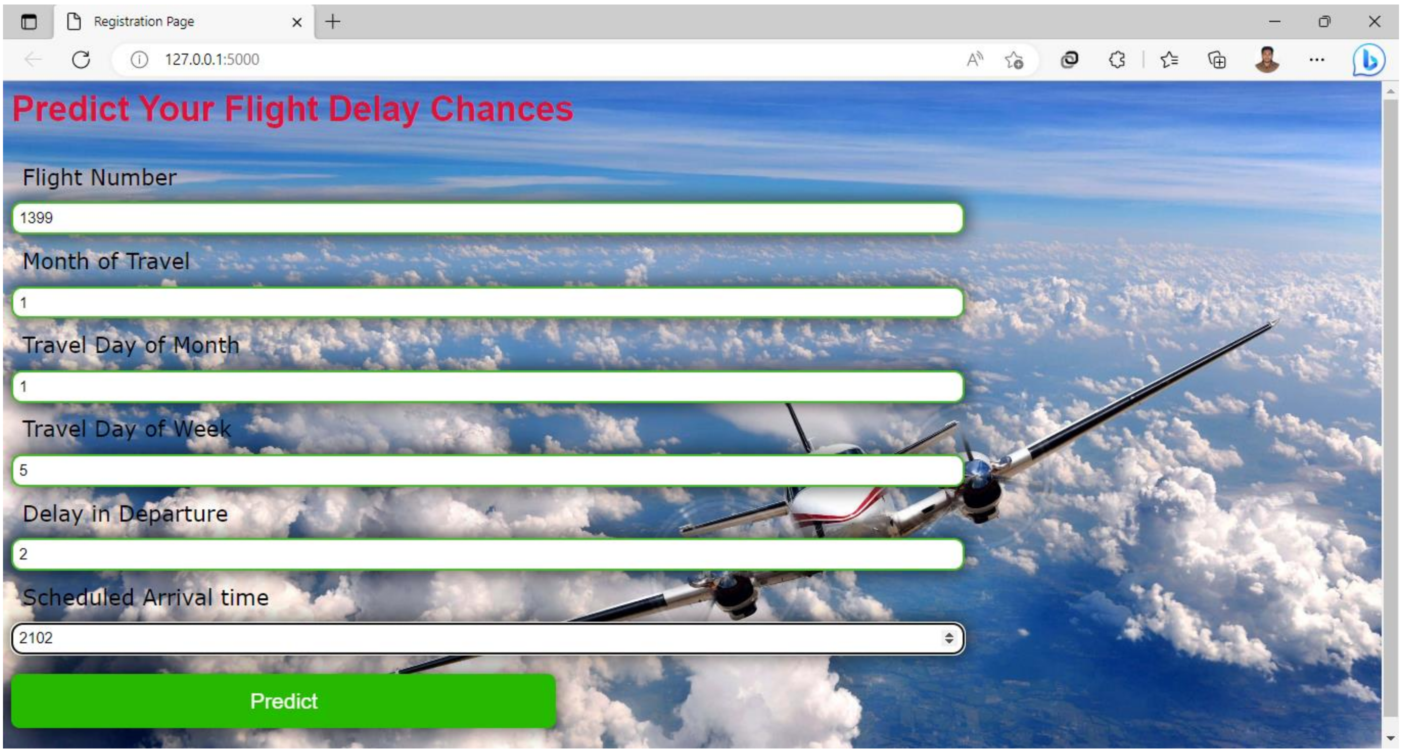
Enter Day of week

Delay in Departure

Enter The Delay in Departure

Scheduled Arrival time

Enter The Time of Scheduled Arrival



Predict Your Flight Delay Chances

Flight Number

1399

Month of Travel

1

Travel Day of Month

1

Travel Day of Week

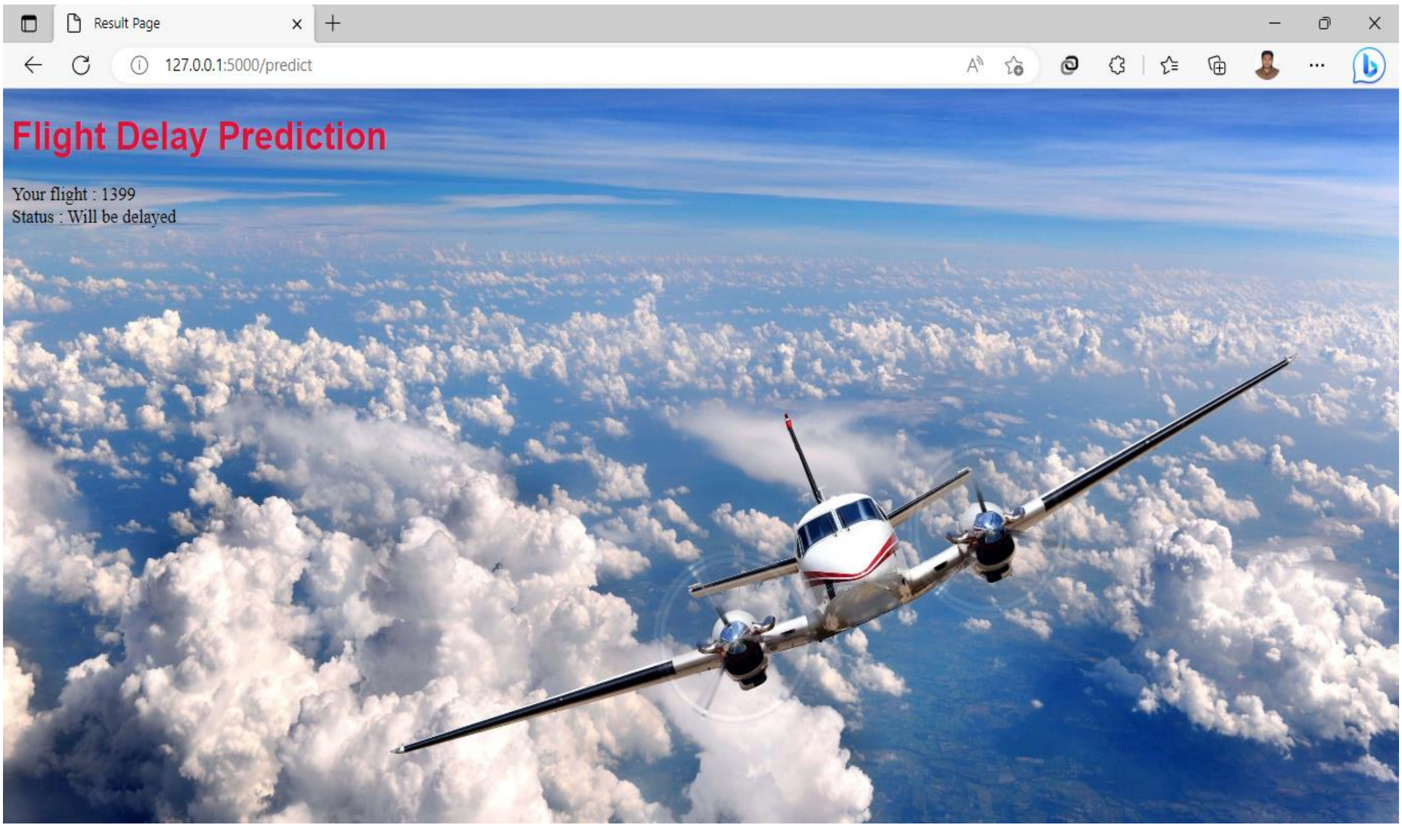
5

Delay in Departure

2

Scheduled Arrival time

2102



Flight Delay Prediction

Your flight : 1399
Status : Will be delayed