**CHAPTER 1**

**INTRODUCTION**

* 1. **OVERVIEW**

OVER the last twenty years, air travel has been increasingly preferred among travelers, mainly because of its speed and in some cases comfort. This has led to phenomenal growth in air traffic and on the ground. An increase in air traffic growth has also resulted in massive levels of aircraft delays on the ground and in the air. These delays are responsible for large economic and environmental losses. According to, taxi-out operations are responsible for 4,000 tons of hydrocarbons, 8,000 tons of nitrogen oxides and 45,000 tons of carbon monoxide emissions in the United States in 2007. Moreover, the economic impact of flight delays for domestic flights in the US is estimated to be more than $19 Billion per year to the airlines and over $41 Billion per year to the national economy In response to growing concerns of fuel emissions and their negative impact on health, there is active research in the aviation industry for finding techniques to predict flight delays accurately in order to optimize flight operations and minimize delays.

Using a machine learning model, we can predict flight arrival delays. The input to our algorithm is rows of feature vector like departure date, departure delay, distance between the two airports, scheduled arrival time etc. We then use decision tree classifier to predict if the flight arrival will be delayed or not. A flight is delayed when difference between scheduled and actual arrival times is greater than 15 minutes. Furthermore, we compare decision tree classifier with logistic regression and a simple neural network for various figures of merit. Finally, it will be integrated to web based application.

* 1. **PURPOSE**

In the present world, the major components of any transportation system include passenger airline, cargo airline, and air traffic control system. With the passage of time, nations around the world have tried to evolve numerous techniques of improving the airline transportation system. This has brought drastic change in the airline operations. Flight delays occasionally cause inconvenience to the modern passengers. Every year approximately 20% of airline flights are canceled or delayed, costing passengers more than 20 billion dollars in money and their time. A few factors responsible for the flight delays like runway construction to excessive traffic are rare, but bad weather seems to be a common cause. Some flights are delayed because of the reactionary delays, due to the late arrival of the previous flight. It hurts airports, airlines, and affects a company's marketing strategies as companies rely on customer loyalty to support their frequent flying programs.

**CHAPTER 2**

**PROBLEM DEFINITION AND DESIGN THINKING**

**2.1 EMPATHY MAP**

An empathy map is a square divided into four quadrants with the user or client in the middle. Each of the four quadrants comprises a category that helps us delve into the mind of the user. The four empathy map quadrants look at what the user says, thinks, feels, and does.

WHAT BRAND SHOULD I CHOOSE?

WASTING TOO MUCH OF TIME

I WAS EXPECTING SOMETHING

I WANT SOMETHING AWESOME

**THINKS**

**SAYS**

**FEELS**

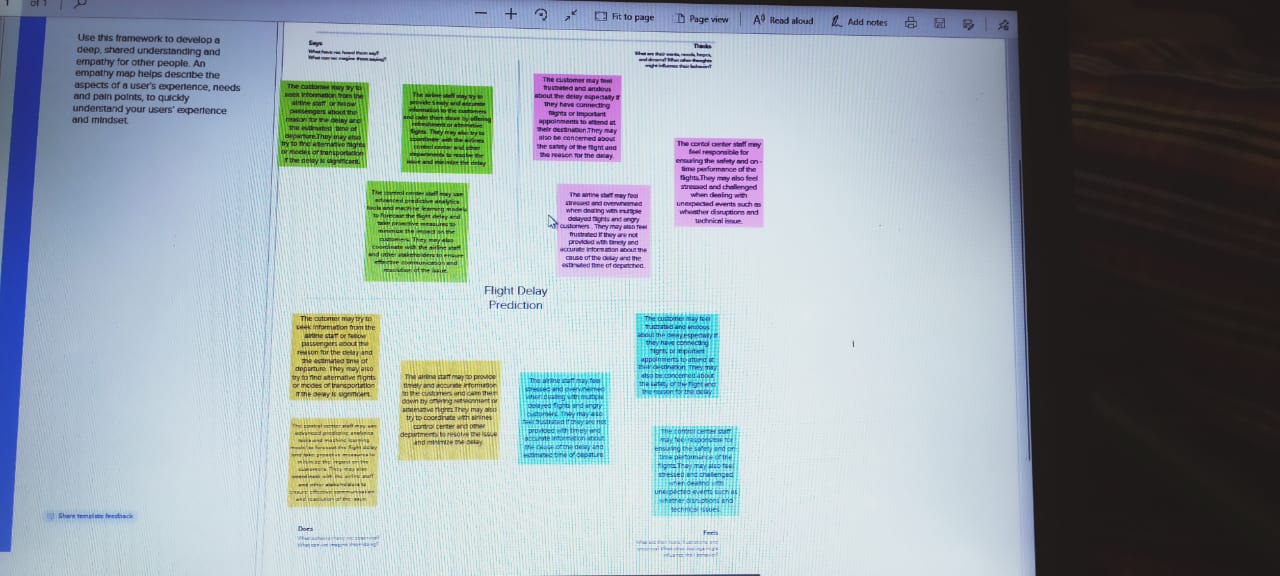
**DOES**

EXITED

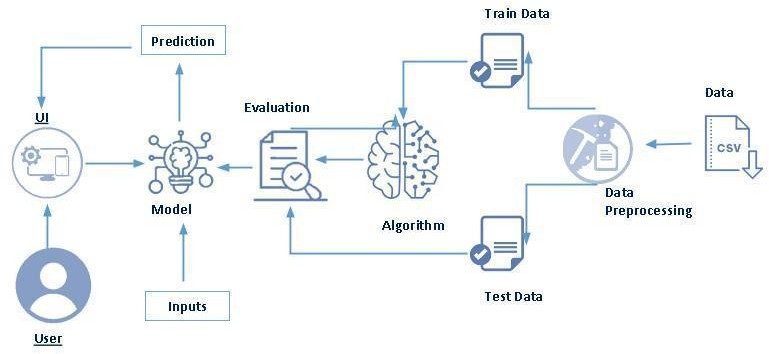
ASK FRIENDS

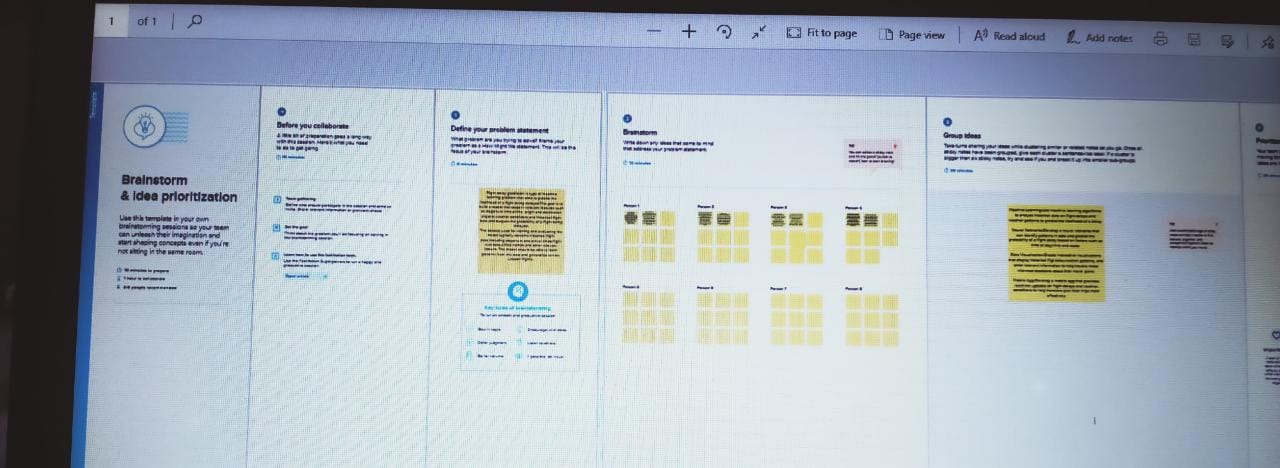
FEAR

COMPARE PRODUCTS



**2.2 IDEATION MAP**

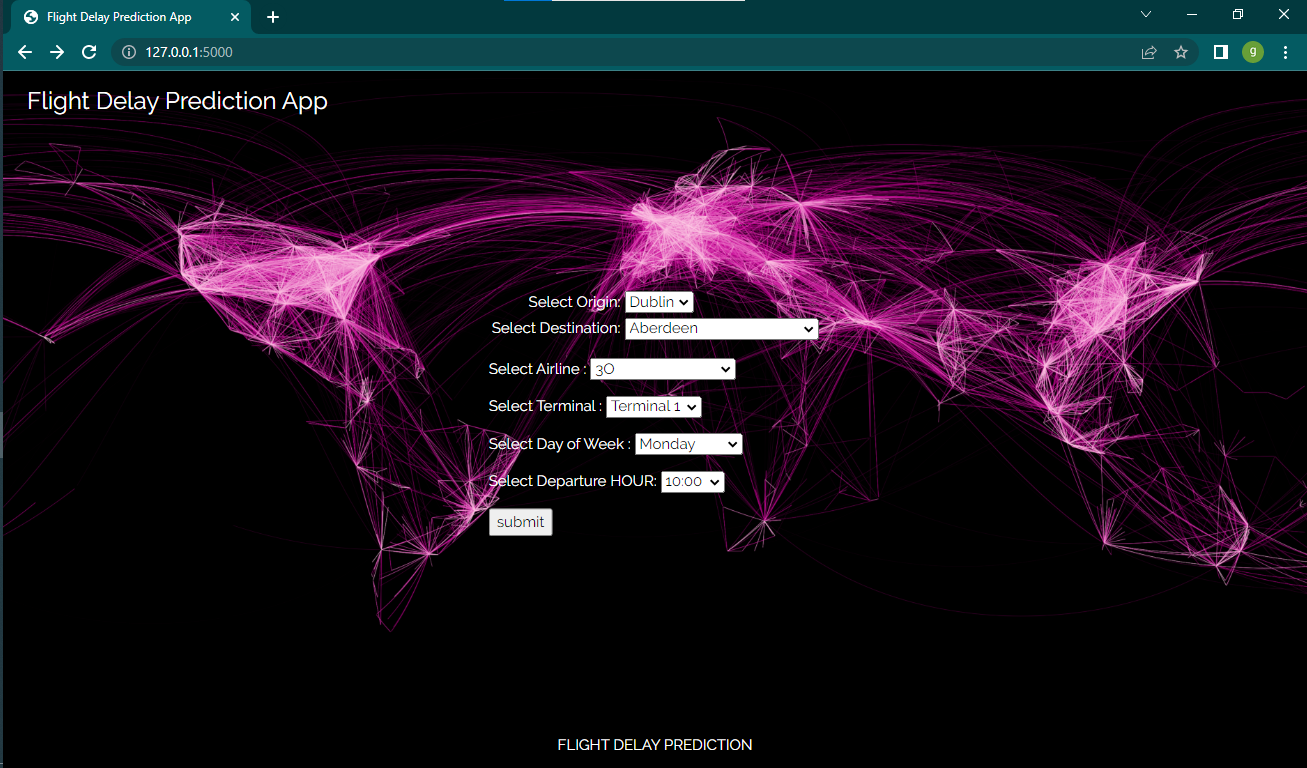


****

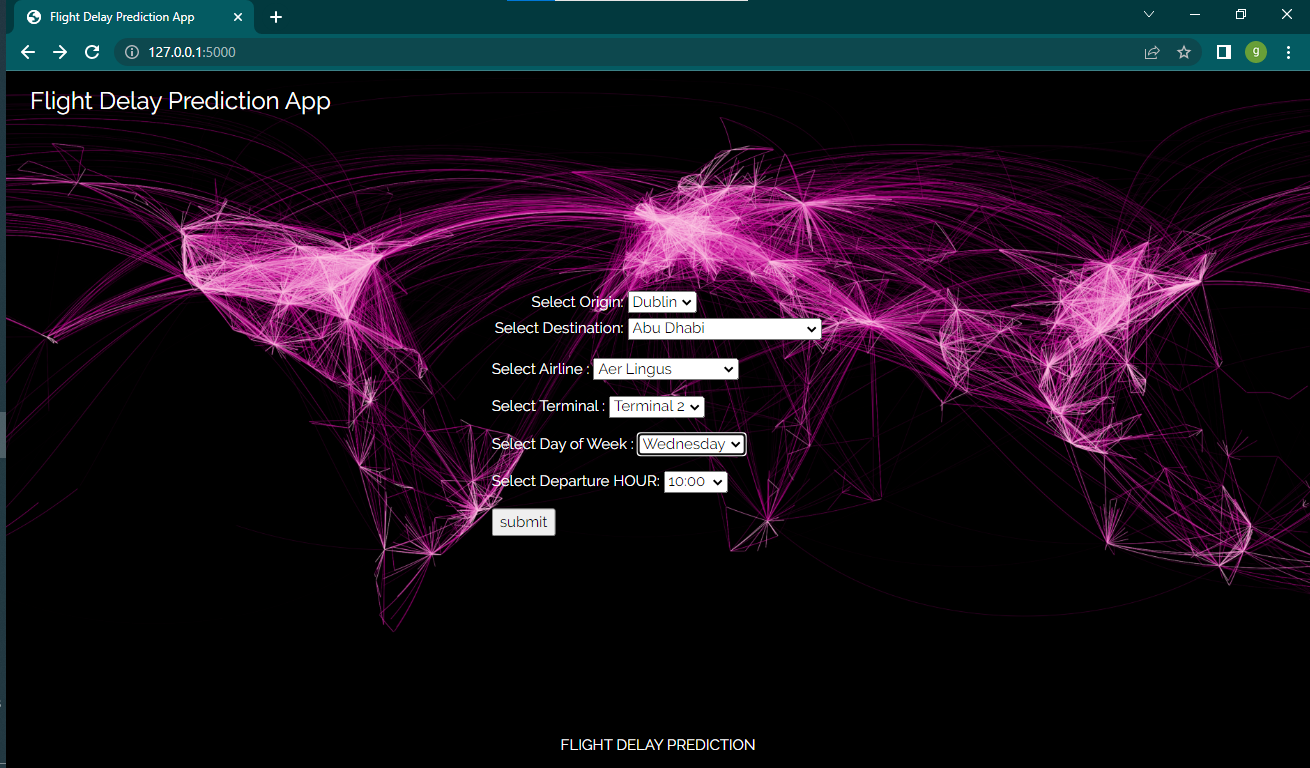
**CHAPTER 3**

**RESULT**

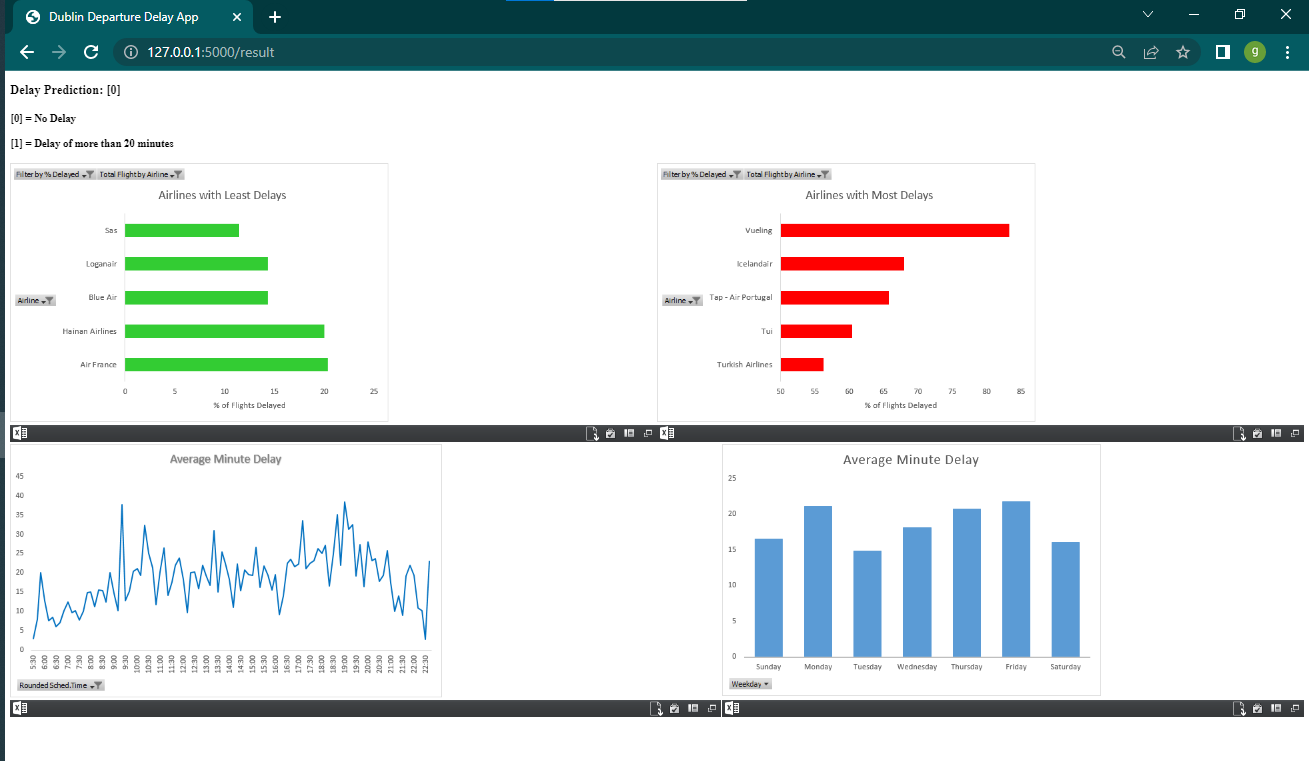
**3.1 RESULT**

****

**HOME PAGE**

****

**INPUT PAGE**

****

**OUTPUT (PREDICTION)**

**CHAPTER 4**

**ADVANTAGES AND DISADVANTAGES**

**4.1 ADVANTAGES**

* Predicting flight delays can improve airline operations
* Increases passenger satisfaction
* This will result in a positive impact on the economy.
* It is widely used by aircraft operators throughout the world to inform and facilitate corrective actions in a range of operational areas by offering the ability to track and evaluate flight operations trends, identify risk precursors, and take the appropriate remedial action.
* Reduces stress for passengers associated with air travel.
* Improves communication between airlines and passengers.
* This helps airlines and airports better manage their operations.

**4.2 DISADVANTAGES**

* Airlines aren't required to compensate passengers when flights are delayed or canceled due to problems deemed beyond the company's control, like bad weather.
* As flight delay cost a lot to the airlines as well as passangers in financial and environmental terms, flight delay is a the talk of the hour.
* Flight delay causes surging of prices by costing a lot on operational purpose They may increase prices to customers and operational prices to airlines.
* As the outcome is directly associated with the passanger and the airlines which inturn is liked to another set of airline and pasaangers it is very crucial to get real time delay for each player within the air transport system
* They also aren't required to provide a refund when the passenger initiates the cancellation or flight change.
* Leads to decrease customer satisfication.
* 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.

**CHAPTER 5**

**APPLICATION**

**5.1 APPLICATION**

* Flight booking applications help the airline industry automate the booking process.
* Users worldwide can book flights on the go using the simple apps, which include features such as quick flight search, download tickets, check and modify booking details, one-tap check-in, and many more.
* This system can be used by all the common people.

**CHAPTER 6**

**CONCLUSION**

**6.1 CONCLUSION**

Predicting flight delays is on interesting research topic and required many attentions these years. Majority of research have tried to develop and expand their models in order to increase the precision and accuracy of predicting flight delays. Since the issue of flights being on-time is very important, flight delay prediction models must have high precision and accuracy. A flight is delayed when difference between scheduled and actual arrival times is greater than 15 minutes. Using a machine learning model, we can predict flight arrival delays. The input to our algorithm is rows of feature vector like departure date, departure delay, distance between the two airports, scheduled arrival time etc. Finally, it will be integrated to web based application.

**CHAPTER 7**

**FUTURE SCOPE**

**7.1 FUTURE SCOPE**

Further supportive study is required to correlate all the problem, scope and method for getting most accurate result. Although weather conditions are the major reasons for flight delay, other unprecedented events such as major calamities, natural or man-made can cause major delay in flight. the future work of this project includes incorporating a larger dataset. There are many different ways to preprocess a larger dataset like running a Spark cluster over a server or using a cloud-based services like AWS and Azure to process the data. With the new advancement in the field of deep learning, we can use Neural Networks algorithm on the flight and weather data.

**CHAPTER 8**

**APPENDIX**

**8.1 SOURCE CODE**

from flask import Flask, request

from flask import render\_template

import pickle

import numpy as np

app = Flask(\_\_name\_\_)

@app.route('/')

def home():

return render\_template('home.html')

@app.route('/result',methods=['POST','GET'])

def get\_delay():

if request.method=='POST':

result=request.form

#Prepare the feature vector for prediction

pkl\_file = open('x\_data.txt', 'rb')

index\_dict = pickle.load(pkl\_file)

new\_vector = np.zeros(len(index\_dict))

try:

new\_vector[index\_dict['Terminal\_'+str(result['Terminal'])]] = 1

except:

pass

try:

new\_vector[index\_dict['Weekday\_'+str(result['Weekday'])]] = 1

except:

pass

try:

new\_vector[index\_dict['Airline\_'+str(result['Airline'])]] = 1

except:

pass

try:

new\_vector[index\_dict['Origin\_'+str(result['Origin'])]] = 1

except:

pass

try:

new\_vector[index\_dict['Destination\_'+str(result['Destination'])]] = 1

except:

pass

try:

new\_vector[index\_dict['Scheduled Time\_'+str(result['Scheduled Time'])]] = 1

except:

pass

pkl\_file = open('model.pkl', 'rb')

adamodel = pickle.load(pkl\_file)

new\_vector = new\_vector.reshape(1, -1)

Prediction = adamodel.predict(new\_vector)

return render\_template('result.html',prediction=Prediction)

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

app.debug = True

app.run()