SONA COLLEGE OF TECHNOLOGY
(Autonomous Institution) **SALEM-636005**

DEVELOPING A FLIGHT DELAY PREDICTION MODEL USING MACHINE LEARNING

TEAM ID:PNT2022TMID18451

Bonafide record of work done by

PASUPULETI MIDHILESH

PALLAVARAPU CHANDU

PUVVADA VENKATA NAGA SAI UDAY CHANDRA

SHAIK SULEMAN

CONTENTS

CHAPTER 1

INTRODUCTION

Travelers have begun to favour air travel more and more over the past 20 years, primarily due to its quickness and occasional comfort. Both on the ground and in the air, as a result, have experienced amazing growth. Massive amounts of ground and airborne aircraft delays have also been brought on by an increase in air traffic. Large economic and environmental losses are the result of these delays. The model's primary goal is to correctly forecast flight delays in order to improve aircraft operations and reduce delays.

1.1. PROJECT OVERVIEW

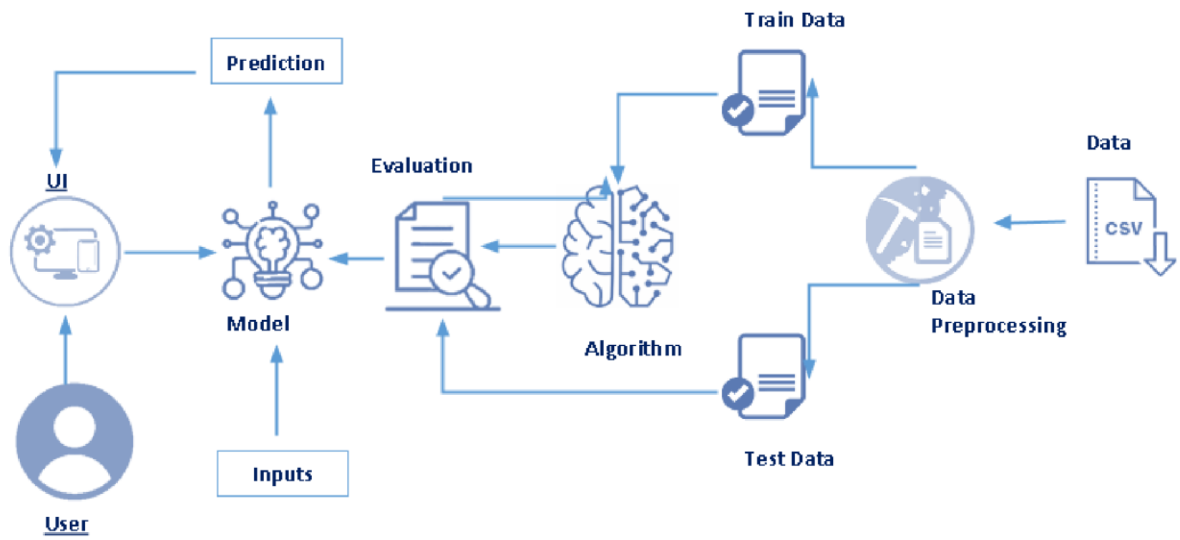


Figure 1.1. Technical Architecture

Flight arrival delays can be predicted using a machine learning algorithm. Rows of feature vectors, such as departure date, delay, travel time between the two airports, and scheduled arrival time, provide the input to our algorithm. The decision tree classifier is then used to determine whether or not the flight arrival will be delayed. When there is more than a 15-minute gap between the scheduled and actual arrival timings, a flight is deemed to be delayed. For various figures of merit, we contrast the decision tree classifier with logistic regression and a straightforward neural network.

1.2. PURPOSE

The main goal of this project is to predict the flight delay using machine learning algorithms. Flight planning is one of the difficulties in the industrial environment because there are many unpredictabilities. One such condition is the incidence of delays, which can result from a variety of causes and impose significant expenses on airlines, operators, and passengers. Delays in departure can be brought on by inclement weather, seasonal and holiday demands, airline policies, technical issues with airport infrastructure, baggage handling, and mechanical equipment, and a buildup of delays from earlier flights. Hence Predicting flight delays can improve airline operations and passenger satisfaction, which will result in a positive impact on the economy.

CHAPTER 2

LITERATURE SURVEY

2.1 Existing problem

Commercial aviation is a complex distributed transportation system. It deals with valuable resources, demand fluctuations, and a sophisticated origin destination matrix that needs orchestration to provide smooth and safety operations. Furthermore, individual passengers follow her itineraries while airlines plan various schedules for aircrafts, pilots and flight attendants. Stages can take place at terminal boundaries, airports, runways, and airspace, being susceptible to different kinds of delays. Some examples include mechanical problems, weather conditions, ground delays, air traffic control, runway queues and capacity constraints.

Airports are significant nodes of air transportation. The number of airport flight delays has been on increase in recent years. Delayed flights are defined by the Federal Aviation Administration when they arrive or depart more than 15 minutes later than scheduled. In 2019, the arrival delay rate is 19.2% and the departure delay rate is 18.18% in the United States. Flight delays can cause many negative effects, such as passengers' inconvenience, increased airport pressure, and airline losses. Effective flight delay prediction could provide support for flight plan and emergency plan formulation, reduce the economic loss, and alleviate the negative impact (1). Hence, a delay prediction turns out very important. International Civil Aviation Organization (ICAO) has enabled a program called Air Traffic Flow Management (ATFM) with the objective of ensuring that the traffic volume is compatible with the capacities declared by aviation authorities in order to reduce ground and en-route delays. Another application of traffic management is the Free Route Airspace (FRA) concept which consists of using the shortest routes possible in order to reduce flight time, CO2 emissions, and fuel waste. Moreover, several other models have been developed to solve delays problem based on probability, statistics, graph and network representations, operational research studies, and so on (2). The Related Work gives the recent works and research on this topic.

RELATED WORKS

the principle of the Stacking classification algorithm is introduced, the SMOTE algorithm is selected to process imbalanced datasets, and the Boruta algorithm is utilized for feature selection. There are five supervised machine learning algorithms in the first-level learner of Stacking including KNN, Random Forest, Logistic Regression, Decision Tree, and Gaussian Naive Bayes. The second-level learner is Logistic Regression. To verify the effectiveness of the proposed method, comparative experiments are carried out based on Boston Logan International Airport flight datasets from January to December 2019. Multiple indexes are used to comprehensively evaluate the prediction results, such as Accuracy, Precision, Recall, F1 Score, ROC curve, and AUC Score. The results show that the Stacking algorithm not only could improve the prediction accuracy but also maintains great stability (1). Multi Layer Perceptron (MLP) to train and test data applied. The neural network MLP was able to predict flight arrival delay with a coefficient of determination R2 of 0.9048, and the selective procedure achieved a time saving and a better R2 score of 0.9560. To enhance the reliability of the proposed method, the performance of the MLP was compared with that of Gradient Boosting (GB) and Decision Trees (DT). The result is that the MLP outperformed all existing benchmark methods (2). Here present the first data driven systemic study of air transport delays in China, of their evolution and causes, based on 11 million flights between 2016 and 2018. A significant fraction of the delays can be explained by a few variables, e.g., weather conditions and traffic levels, the most important factors being the presence of thunderstorms and the season of the year. Remaining delays can often be explained by en-route weather phenomena or by reactionary delays. This study contributes towards a better understanding of delays and their prediction through a data-driven methodology, leveraging on statistics and data mining concepts (3). Recognize useful patterns of the flight delay from aviation data and perform accurate delay prediction. The best result for flight delay prediction (five classes) using machine learning models is 89.07% (Multilayer Perceptron). A Convolution neural network model is also built which is enlightened by the idea of pattern recognition and success of neural network method, showing a slightly better result with 89.32% prediction accuracy (4). Explores a broader scope of factors which may potentially influence the flight delay, and compares several machine learning based models in designed generalized flight delay prediction tasks. Compared with the previous schemes, the proposed random forest-based model can obtain higher prediction accuracy (90.2% for the binary classification) and can overcome the overfitting problem (5). The model demonstrated to reduce by 30% the take-off time prediction errors of the current system one hour before the time that flight is scheduled to depart from the parking position and presents an extension of the model, which overcomes this look-ahead time constraint and allows to improve take-off time predictions as early as the initial flight plan is received. In addition, a subset of the original set of input features has been meticulously selected to facilitate the implementation of the solution in an operational air traffic flow and capacity management system, while minimising the loss of predictive power. Finally, the importance and interactions of the input features are thoroughly analysed with additive feature attribution methods (6). The designed prediction tasks contain different classification tasks and a regression task. Experimental results show that long shortterm memory (LSTM) is capable of handling the obtained aviation sequence data, but overfitting problem occurs in our limited dataset. Compared with the previous schemes, the proposed random forest-based model can obtain higher prediction accuracy (90.2% for the binary classification) and can overcome the overfitting problem (7).

References

1. *Flight Delay Classification Prediction Based on.* **Jia Yi, 1 Honghai Zhang ,.** [ed.] Chi-Hua Chen. Honghai Zhang; zhh0913@163.com : Wiley, 2021. p. 10.
2. *A Multilayer Perceptron Neural Network with Selective-Data.* **Hajar Alla, Lahcen Moumoun , and Youssef Balouki.** [ed.] Jianping Gou. Settat, Morocco : Hindawi, 2021. Hindawi Scientific Programming. p. 12.
3. *Characterization and Prediction of Air Transport.* **Massimiliano Zanin 1, * , Yanbo Zhu 2,3, Ran Yan 3.** 2020, MDPI Journals, p. 15.
4. *Applying Machine Learning to Aviation Big Data.* **Yushan Jiang, Yongxin Liu,Dahai Liu,Houbing Song.** 2020. 2020 IEEE Intl Conf on Dependable,Intl Conf on Cloud and Big Data Computing. p. 8.
5. *FLIGHT DELAY PREDICTION USING MACHINE LEARNING.* **Sarah Ajmeria, Srushti V,Prof. Kavitha S Patil.** Bangalore, India : IJIREEICE, 2022. DOI: 10.17148/IJIREEICE.2022.10584. p. 5.
6. *An explainable machine learning approach to improve take-off time.* **Ramon Dalmau, Franck Ballerini,Herbert Naessens,Seddik Belkoura.** 2021, Journal of Air Transport Management, p. 12.
7. *Flight Delay Prediction Based on Aviation Big Data.* **Guan Gui, Senior Member, IEEE, Fan Liu, Student Member, IEEE, Jinlong Sun, Member, IEEE,.** 2020, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. XX, NO. XX, XXX 2015, p. 11.

2.3 Problem Statement Definition

Analysis of flight delay and causal factors is crucial in maintaining airspace efficiency and safety. However, delay samples are not independent since they always show a certain aggregation pattern. Therefore, this study develops a novel spatial analysis approach to explore the delay and causal factors which is able to take dependence and the possible problem involved including error correlation and variable lag effect of causal factors on delay into account using a Machine Learning Model.

Problem Statement – Flight Delay Prediction

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	**Passenger**	**Go on time**	**Flight is Delay**	**Inclement weather,such as thunderstorm,hurricane,or blizzard**	**Frustrated**
PS-2	Businessman	Check the status of my flight	I was unable to accurately track my information	There is no reliable prediction software or tool.	Disappointed

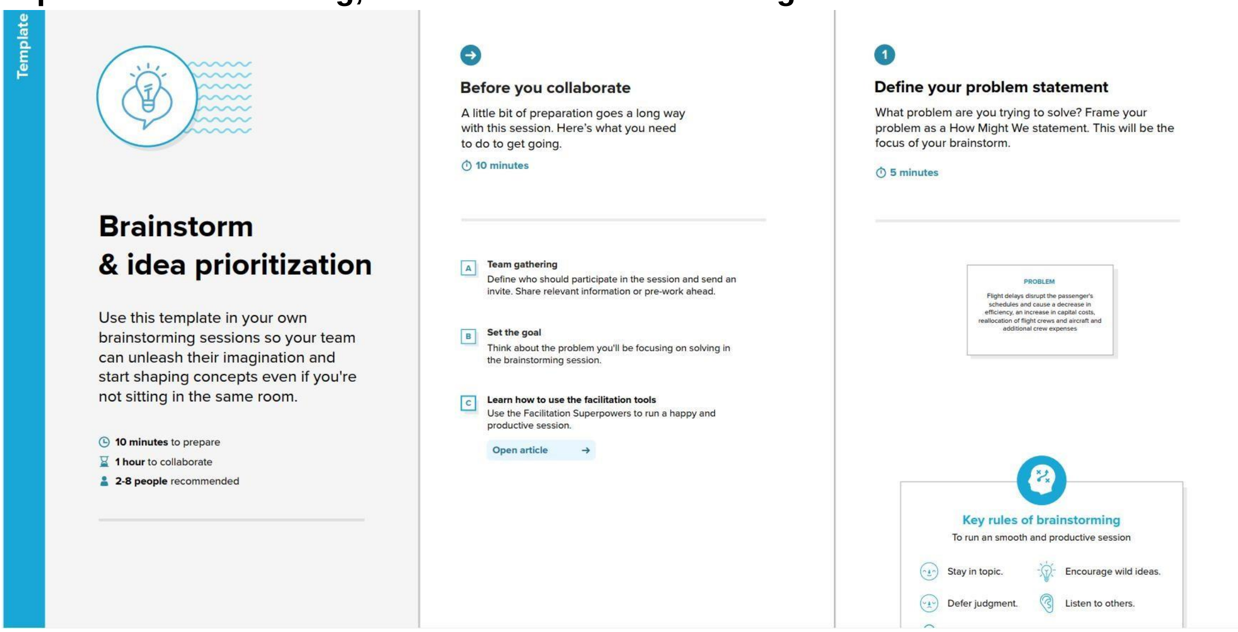
CHAPTER 3 IDEATION & PROPOSED SOLUTION

3.1.EMPATHY MAP CANVAS

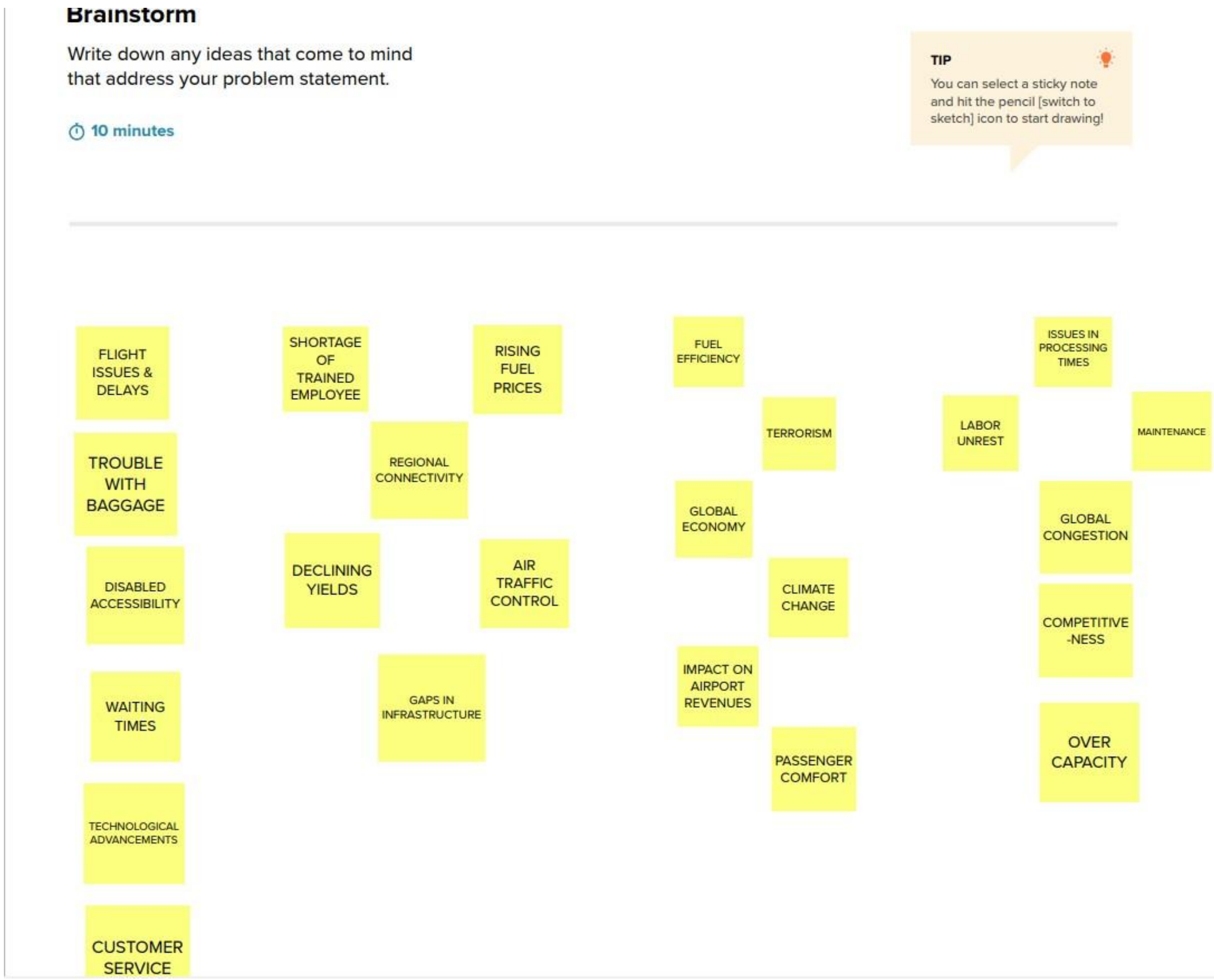
Figure 3.1. Empathy Map

3.2. IDEATION & BRAINSTORMING

Step 1 - Team Gathering, Collaboration and Selecting the Problem Statement



Step 2 - Brainstorm, Idea Listing and Grouping



3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes

reduce congestion by auctioning taking off & landing slots

Charge customers for services to raise ancillary revenue.

Conducting regular inspections and maintenance tasks will help in ensuring aircrafts stay serviceable for the time they fly.

TIP

Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mural.

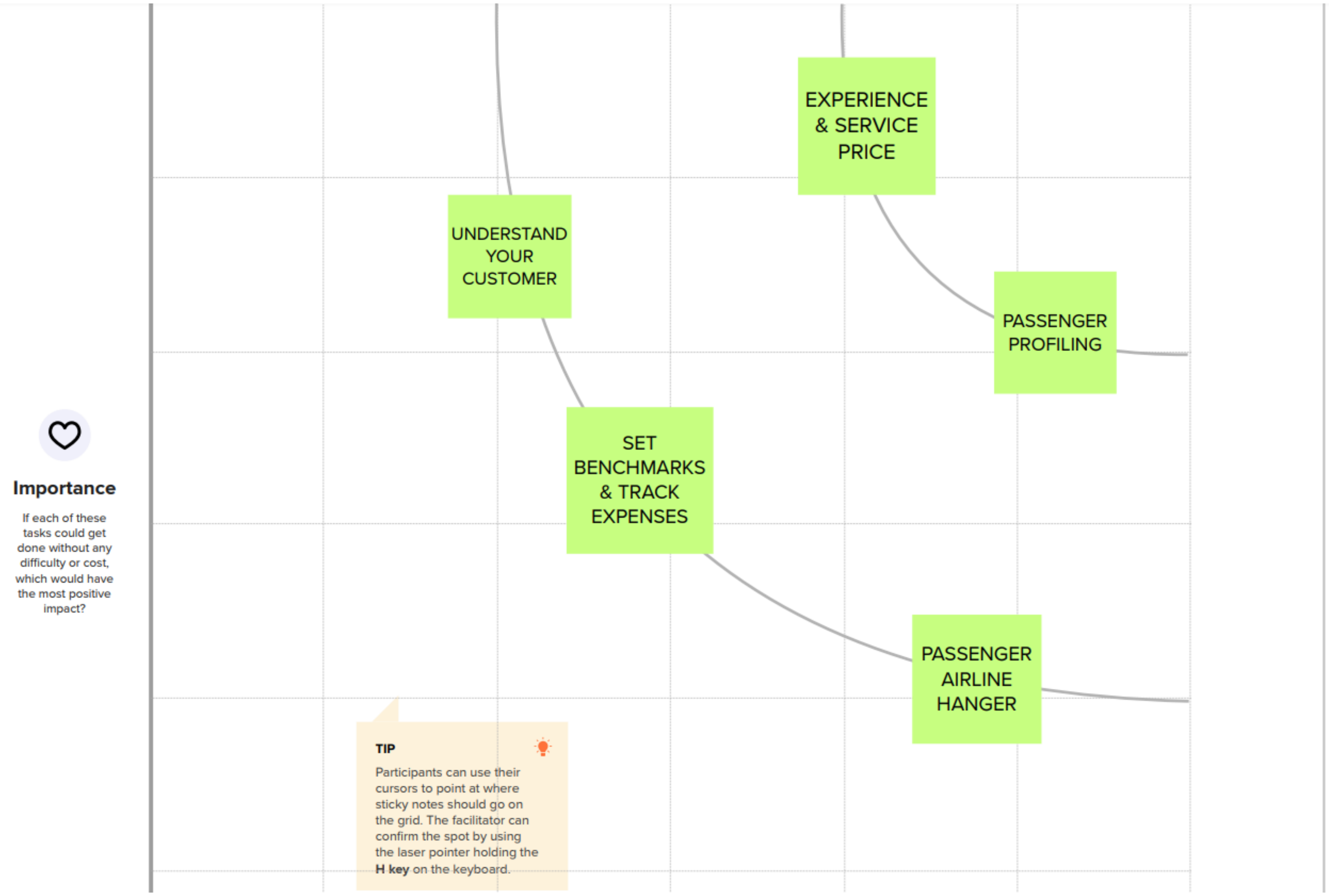
make it more aerodynamically efficient and help reduce the aircraft's weight and fuel use.

Know What Your Customers Think

Improvement of Available Choices

Airlines can save money through fuel hedging contracts that secure a fixed price for fuel over a certain period.

Step 3 - Idea Prioritization



3.3. PROPOSED SOLUTION

S.No.	Parameter	Description
1	Problem Statement (Problem to be solved)	Flight arrival delays can be predicted using a machine learning algorithm. Rows of feature vectors, such as departure date, delay, travel time between the two airports, and scheduled arrival time, provide the input to our algorithm. The decision tree classifier is then used to determine whether or not the flight arrival will be delayed. When there is more than a 15-minute gap between the scheduled and actual arrival timings, a flight is deemed to be delayed. Additionally, for various figures of merit, we contrast the decision tree classifier with logistic regression and a straightforward neural new
2	Idea / Solution description	Using ML Algorithms to predict the delay in flight arrival, informing them to the customers using a Mobile Application or a Web Application. We are developing a software that will allow passengers who use airplanes to foresee flight delays. They may effectively plan their travel using this application, which will help them save time. The tool will have an intuitive user interface. To estimate delays and execute the most effective and efficient methods in the tool, we will use a variety of machine learning algorithms.

3	Novelty / Uniqueness	1. Building a full-fledged application in which the customers can track whether the flights will be delayed or not. 2. Combining the results of one or more ML models using the techniques of ensembling
4	Social Impact / Customer Satisfaction	Flight delays not only anger and disturb air travelers' plans, but they also reduce productivity, raise capital costs, reallocate flight crews and aircraft, and add to crew costs. Higher operating costs for airline firms are unavoidable as flight delays necessitate the consumption of more labour, capital, and other necessary inputs. Flight delays could make the transportation system less effective and have a negative impact on how an airport is planned. Delayed flights subject airlines to penalties, fines, and additional expenses.

5	Business Model (Revenue Model)	The cost of airline tickets and flight delays are now uncertain. Even for the same airplane and seat class, ticket costs are dynamic and frequently change. To increase their revenue, airline firms use a variety of algorithms to adjust the prices dynamically. These models are not accessible to the general public due to the intense competition among airline operators. Additionally, the flight is delayed due to a number of micro and macro causes. The air route status, the prior flight's delay, airplane capacity, air traffic management, airline properties, etc. are the main elements that have an impact on airlines. To save "Time and Money," it is necessary to forecast airline flight delays and ticket costs.
6	Scalability of the Solution	The proposed system can be scaled up to take actions – book another flight for passengers or if a particular flight is getting delayed often, the same can be examined by memorizing the outputs of this system. This can be scaled up to predict the delay of flights in every airport.

3.4. PROBLEM SOLUTION FIT

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

All the passengers who are taking

the flight and the flight are delayed

due to some reasons

6. CUSTOMER CONSTRAINTS CC

-No refunds will be given to the passengers

-Cannot pay or book an alternative flight
-Not satisfied with the benefits

5. AVAILABLE SOLUTIONS AS

- The delay of flights are informed earlier

- Airline benefits are given
- Book for an alternate flight
- Enjoys the benefits from the airline
- Go to different places they are at

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

2. JOBS-TO-BE-DONE / PROBLEMS J&P

The problem that is addressed to the customer is the delay of flights

9. PROBLEM ROOT CAUSE RC

- Mechanical issue
- Unpredictable weather condition
- Consecutive delay of previous flights.
- Air traffic due to weather

7.BEHAVIOUR BE

- Get information from the airlines in prior
- Try to book another flight if emergency
- Reach the airport early
- Book a nearby hotel if the delay of flight is prolonged

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS TR

Many may respond to the problem differently but the common response will be tension, anger or maybe even relaxed.

4. EMOTIONS: BEFORE / AFTER EM

BEFORE:
Perturbed, discouraged, bored not knowing what to do, stressed out and full of rage

AFTER:
Relaxed, and content
Gets benefit from the airlines

10. YOUR SOLUTION SL

The solution to the delay of flight is by developing a flight delay prediction model by using machine learning to predict and declare the delay of flights.

8. CHANNELS of BEHAVIOUR CH

8.1 **ONLINE**
- Checks the airline application to know aboutthe delay
- Checks the nearby hotel with accommodations

8.2 **OFFLINE**
- Checks with the attendees about alternative flight and about how long the delay of the flight will be for.

Reaches the airport soon

Identify strong TR & EM

CHAPTER 4 REQUIREMENT ANALYSIS

4.1. FUNCTIONAL REQUIREMENT

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User Login	Login using credentials
FR-4	User Verification	To check if a user is authorized or not
FR-5	Search Flights	The system should allow users to search for their flight details .
FR-6	Flights Status Notification	Passengers can view the status of their flight anytime.

4.2. NON-FUNCTIONAL REQUIREMENTS

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Ease of use Ease of access
NFR-2	Security	Information about the users and their flight details is kept private. Provides assurance to users by informing them of possible flight delay
NFR-3	Reliability	Should provide accurate predictions
NFR-4	Performance	Should provide an uninterrupted connection. High-speed performance
NFR-5	Availability	The system should be available at all times.
NFR-6	Scalability	Can handle multiple users at the same time Accessible even in remote areas

CHAPTER 5

PROJECT DESIGN

5.1. DATA FLOW DIAGRAMS

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

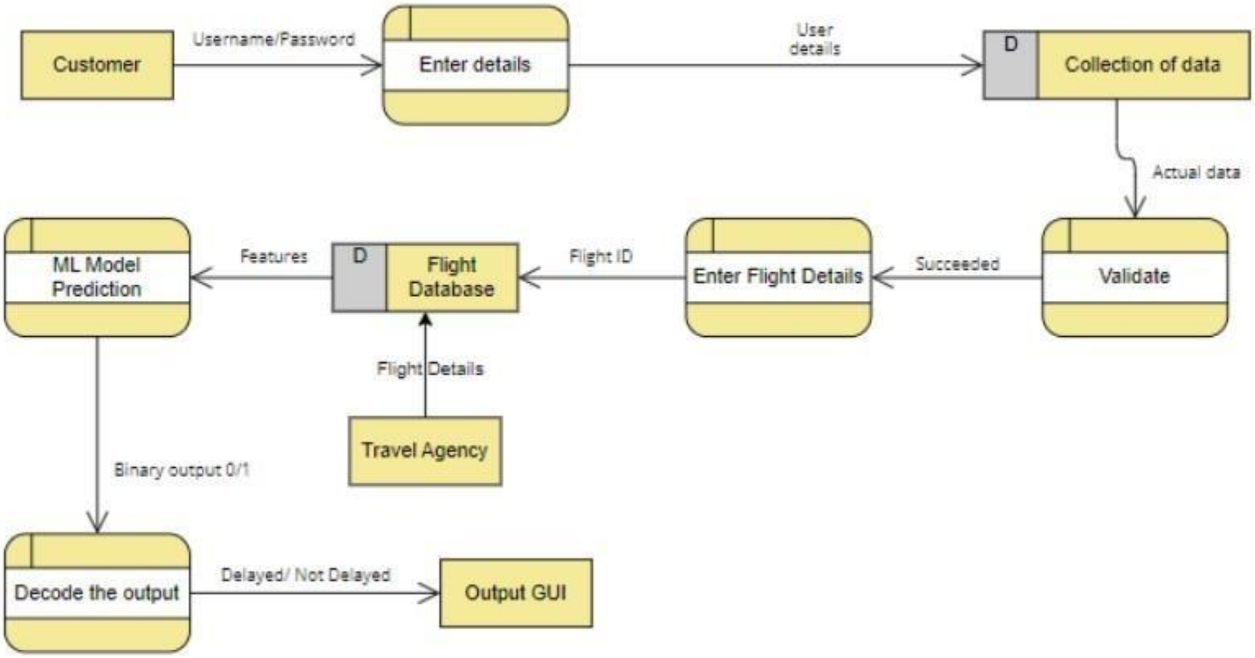


Figure 5.1. Data flow diagram

5.2. SOLUTION & TECHNICAL ARCHITECTURE

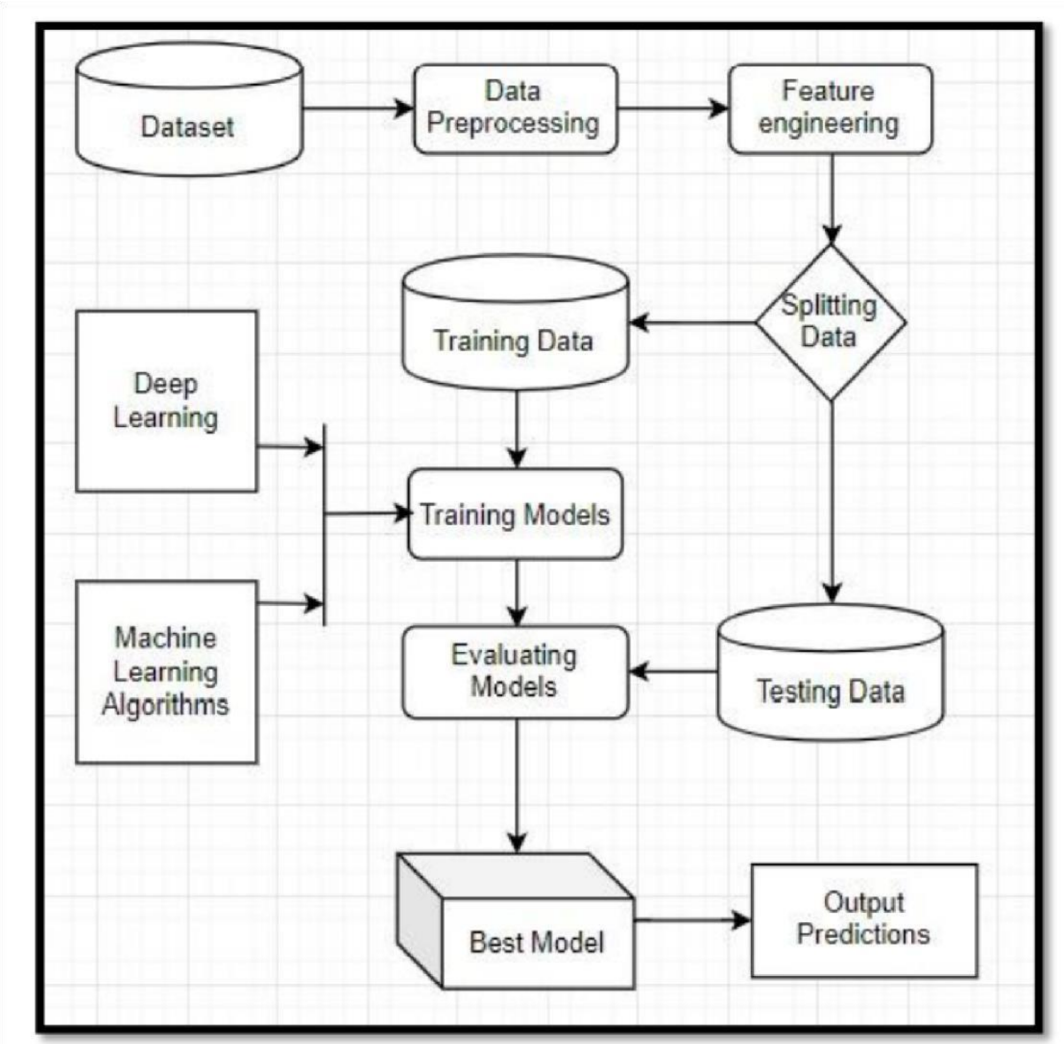


Figure 5.2. Solution Architecture

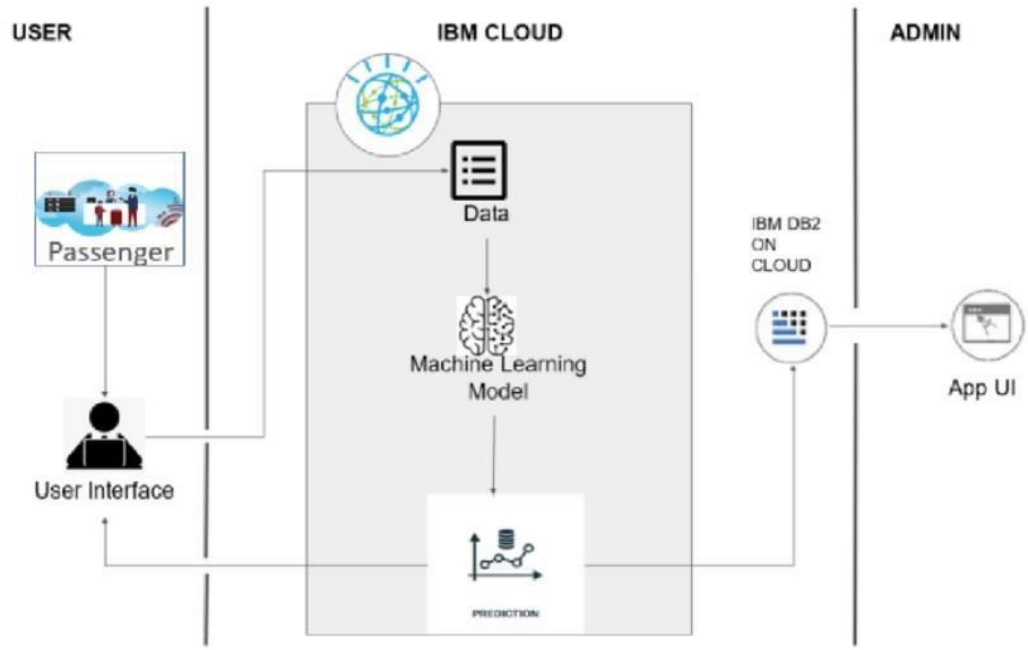


Figure 5.3. Technology Stack

Components & Technologies

S.No	Component	Description	Technology
1.	User Interface	Web Application to interact with the user.	Flask
2.	Login/Sign up	Login/ Sign up – The user can enter the details and get them validated	Python
3.	Database	The Database to store the login details of the user	MySQL
4.	Cloud Database	The database to keep track of the flight details from the travel agency, input to the Machine Learning Model	Firebase
5.	Machine Learning Model	To Predict whether the flight will get delayed or not.	SVM, KNN Classifier, Logistic Regression, Decision Trees
6.	Deep Learning Model	To Predict whether the flight will get delayed or not	Fully Connected Neural Networks
7.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	IBM Cloud

Application Characteristics

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Web application – Flask ML – Sklearn, Tensorflow, Keras API	Deep Learning, Python
2.	Security Implementations	The data is secured that it is encrypted in IBM cloud	AES (256-bit)
3.	Scalable Architecture	Can be scaled upto many airports, many users with more training	Firebase
4.	Availability	The status will be updated frequently	IBM Cloud
5.	Performance	Can make as many number of requests per second to get the prediction	IBM Cloud

5.3. User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail	I can register & access the dashboard with Gmail Login	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password	I can login & access the dashboard	High	Sprint-1
	Core	USN-6	As a user, I can enter my flight details	I can feed the inputs to the system	High	Sprint-2
		USN-7	As a user, I can look at the flight details	I can see whether my flight is getting delayed or not	High	Sprint-3

CHAPTER 6 PROJECT PLANNING & SCHEDULING

6.1. SPRINT PLANNING & ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration and Login	USN-1	As a new user, I can register for the application by entering my email and my password.	2	High	Shaik Suleman
Sprint-2	Confirmation email	USN-2	As a user, I will receive confirmation email once I have registered for the application	2	Medium	Pallavarapu Chandu
Sprint-1	User login	USN-3	As a user, I can login into the application by entering the registered email-id and password	2	High	Puvvada Uday
Sprint-2	Admin Panel	USN-4	As an admin, I can authenticate the registration and login credentials of the passengers	2	High	Pasupuleti Midhilesh
Sprint-3	Arrival and Departure time of flights	USN-5	As a user, I can find all the details of a specific flight with its number or name	2	High	Pallavarapu Chandu
Sprint-3		USN-6	As a user, I can find exactly how long the flight will be delayed	2	High	Pasupuleti Midhilesh
Sprint-4	Helpdesk	USN-7	As a customer care executive, I can provide the contact details of the airlines	1	Medium	Puvvada Uday
Sprint-4		USN-8	As a passenger, I can find alternative flights to the destination that are available	1	High	Pasupuleti Midhilesh Shaik Suleman Puvvada Uday Pallavarapu Chandu
Sprint-4	Feedback	USN-9	As a user, I can provide my suggestions and feedback for the improvement of the application	2	Medium	Shaik Suleman

6.2. SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	4	6 Days	24 Oct 2022	29 Oct 2022	4	29 Oct 2022
Sprint-2	4	6 Days	31 Oct 2022	05 Nov 2022	4	05 Nov 2022
Sprint-3	4	6 Days	07 Nov 2022	12 Nov 2022	4	12 Nov 2022
Sprint-4	4	6 Days	14 Nov 2022	19 Nov 2022	4	19 Nov 2022

Velocity:

We have a 24-day sprint duration, and the velocity of the team is 4 (points per sprint). Thus the team's average velocity (AV) per iteration unit (story points per day) is as follows

AV = sprint duration / velocity

= 24/16

= 1.5

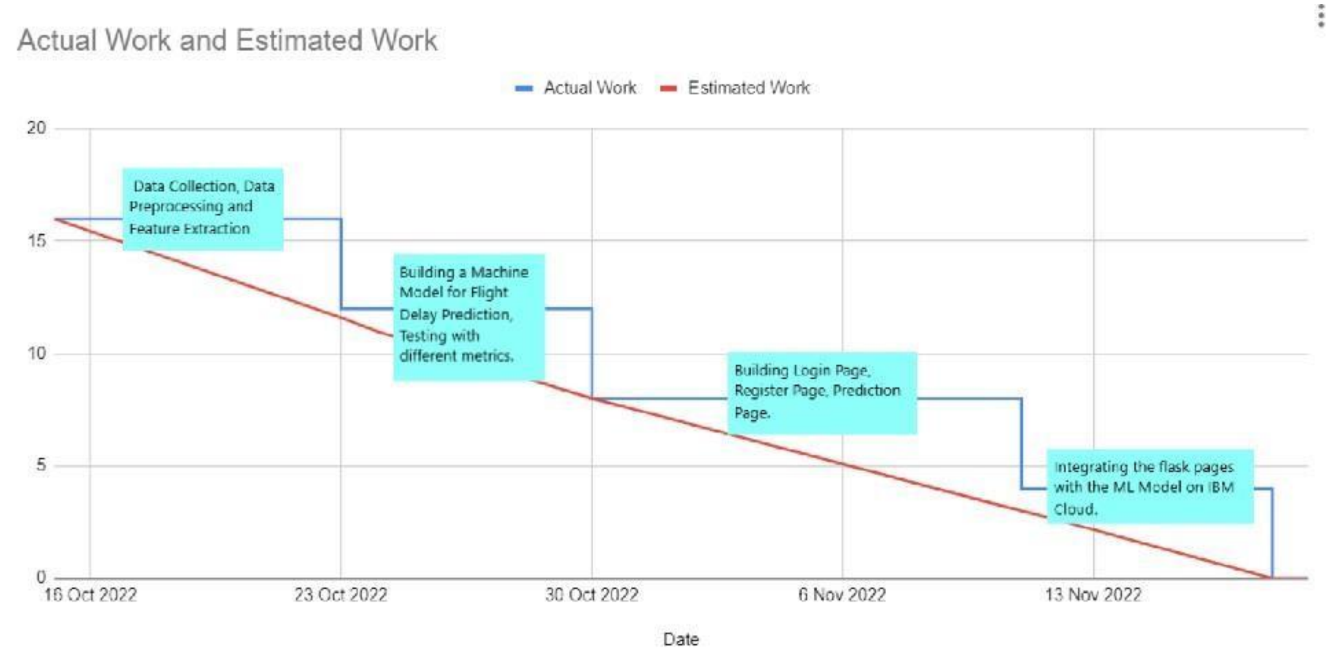


Figure 6.1 - Burndown Chart

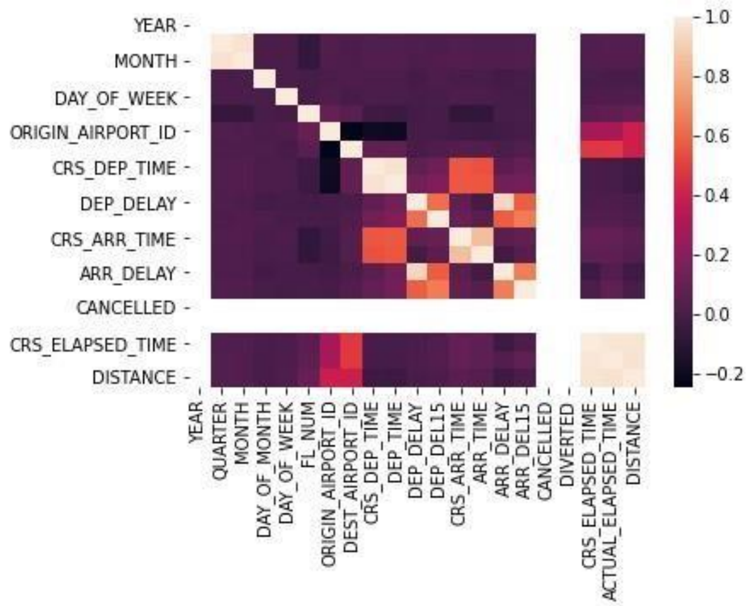
CHAPTER 7

CODING AND SOLUTIONING

7.1. FEATURE 1 - CORRELATION BETWEEN THE VARIABLES IN THE DATASET

```
In [19]: sns.heatmap(data.corr())

Out[19]: <AxesSubplot:>
```



This will help us to find out the correlation between the variables in the dataset which would help us to find out the columns that are unnecessary and hence to be dropped.

7.2. FEATURE 2 - ONE HOT ENCODING

```
In [39]: data=pd.get_dummies(data,columns=['ORIGIN','DEST'])

In [40]: data['ARR_DEL15'].value_counts()

Out[40]: 0.0    9668
         1.0    1375
         Name: ARR_DEL15, dtype: int64

In [41]: data.tail()

Out[41]:
```

	FL_NUM	MONTH	DAY_OF_MONTH	DAY_OF_WEEK	CRS_ARR_TIME	DEP_DEL15	ARR_DEL15	ORIGIN_0	ORIGIN_1	ORIGIN_2	ORIGIN_3	ORIGIN_4
11226	1715	12	30	5	12	0.0	0.0	0	1	0	0	0
11227	1770	12	30	5	20	1.0	0.0	0	0	0	0	1
11228	1823	12	30	5	22	0.0	0.0	0	1	0	0	0
11229	1901	12	30	5	18	0.0	0.0	1	0	0	0	0
11230	2005	12	30	5	9	0.0	0.0	1	0	0	0	0

The cities in both Origin and Destination are one-hot encoded using the above code.

7.3. FEATURE 3 - SAVING THE MODEL WEIGHTS FOR DEPLOYMENT

SAVING THE MODEL

```
In [63]: pickle.dump(classifier,open('flight_new.pkl','wb'))

In [64]: from sklearn.metrics import confusion_matrix
         confusion_matrix(predicted, y_test)

Out[64]: array([[1825,  129],
                [ 138,  117]], dtype=int64)

In [66]: from sklearn.metrics import classification_report
         print(classification_report(predicted, y_test, labels=[1, 2]))

                       precision    recall  f1-score   support

                   1       0.48      0.46      0.47       255
                   2       0.00      0.00      0.00         0

           micro avg       0.48      0.46      0.47       255
           macro avg       0.24      0.23      0.23       255
        weighted avg       0.48      0.46      0.47       255
```

The above code will save the model weights for further deployment in IBM Cloud and also measure the performance metrics.

7.4. FEATURE 4 - FLASK INTERFACE - UI

from flask import Flask, request, render_template

import numpy as np
import pandas as
pd import pickle
import os import
joblib
model = joblib.load(open('flight.pkl', 'rb'))
app = Flask(_name_)

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

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

def predict():
    name = request.form['name']   month =
request.form['month']   dayofmonth =
request.form['dayofmonth']   dayofweek =

```
request.form['dayofweek']     origin =
request.form['origin']     if (origin == "msp"):
origin1, origin2, origin3, origin4, origin5 = 0, 0,
0, 0, 1     if (origin == "dtw"):         origin1, origin2,
origin3, origin4, origin5 = 1, 0,
0, 0, 0     if (origin
== "jfk"):
        origin1, origin2, origin3, origin4, origin5 = 0, 0,
1, 0, 0     if (origin == "sea"):         origin1, origin2,
origin3, origin4, origin5 = 0, 1,
0, 0, 0     if (origin == "alt"):         origin1, origin2,
origin3, origin4, origin5 = 0, 0,
0, 1, 0     destination =
request.form['destination']     if (destination
== "msp"):         destination1, destination2,
destination3, destination4, destination5 = 0,
0, 0, 0, 1     if (destination == "dtw"):
destination1, destination2, destination3,
destination4, destination5 = 1, 0, 0, 0, 0     if
(destination == "jfk"):
        destination1, destination2,
destination3, destination4, destination5 = 0,
0, 1, 0, 0     if (destination == "sea"):
destination1, destination2, destination3,
destination4, destination5 = 0, 1, 0, 0, 0     if
(destination == "atl"):         destination1,
destination2, destination3, destination4,
destination5 = 0, 0, 0, 1, 0     dept =
request.form['dept']     arrtime =
request.form['arrtime']     actdept =
request.form['actdept']     dept15 = int(dept) -
int(actdept)
    total = [[name, month, dayofmonth, dayofweek,
origin1, origin2, origin3, origin4, origin5,
destination1, destination2, destination3,
destination4, destination5, dept,
arrtime]]     #y_pred =
model.predict(total)     y_pred =
model.predict(total)     print(y_pred)     if
(y_pred == [0.]):         ans = "The Flight
will be on time"     else:
        ans = "The Flight will be Delayed"
    return render_template("predict.html",
showcase=ans)
```

```
if _name_ == '_main_':
app.run(debug=True)
```

Explanation:

The above code will be able to get the details of the flight from the user in the respective text fields created using the HTML, scale the inputs and give the inputs to the model which has been developed already. The predictions are shown in another HTML page.

7.5. FEATURE 5 – Index.html:

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

<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Flight Prediction Details</title>



</head>

<body>
    <img src="flightpic.jpg" alt="">
    <h2 style="color: black;"> Flight Delay Prediction using Machine Learning</h2>
<h2>Team ID : PNT2022TMID18451</h2>

    <form action="{{url_for('predict') }}" method="post">
        <div class="form-control">
<label for="name">
Flight Number :            </label>
            <input type="text" placeholder="Enter your Flight Number" name="name"
required />
        </div>
        <div class="form-
control">            <label
for="month">                Month
:            </label>
            <input type="text" placeholder="Enter your Month" name="month" required />
        </div>
```

```
        <div class="form-control">
<label for="dayodmonth">
Day Of Month :
            </label>
            <input type="text" id="dom" placeholder="Enter your Day of Month"
name="dayofmonth" required />
        </div>
        <div class="form-control">
<label for="dayofweek">
Day Of Week :
            </label>
            <input type="text" id="dow" placeholder="Enter your Day of week"
name="dayofweek" required />
        </div>
        <div class="form-control">
<label for="origin">
Origin :
            </label>
            <!-- <input type="text"
id="ori"
    placeholder="Enter your Origin" name="ori" required /> -->


            <select name="origin" id="ori">
                <option value="msp">msp</option>
                <option value="dtw">dtw</option>
                <option value="jfk">jkf</option>
                <option value="sea">sea</option>
                <option value="atl">atl</option>
            </select>
        </div>
        <div class="form-control">
<label for="destination">
Destination :
            </label>
            <select name="destination" id="dest">
<option value="msp">msp</option>
                <option value="dtw">dtw</option>
                <option value="jfk">jfk</option>
                <option value="sea">sea</option>
                <option value="atl">atl</option>
            </select>
        </div>
```

```
        <div class="form-control">
            <label for="dept">
                Scheduled Departure Time :
            </label>
            <input type="text" id="dept" placeholder="Enter your scheduled departure time"
name="dept" required />
        </div>
        <div class="form-control">
<label for="arrtime">
                Scheduled Arrival Time :
            </label>
            <input type="text" id="arrtime" placeholder="Enter your scheduled arrival time"
name="arrtime" required />
        </div>
        <div class="form-control">
<label for="actdept">
Actual Departure Time :
            </label>

            <input type="text" id="actdept" placeholder="Enter your actual arrival time"
name="actdept" required />
        </div>
        <button>
            <div class="svg-wrapper-1">
                <div class="svg-wrapper">
                    <svg xmlns="http://www.w3.org/2000/svg" viewBox="0 0 24 24" width="24"
height="24">
                        <path fill="none" d="M0 0h24v24H0z"></path>
                        <path fill="currentColor" d="M1.946 9.315c-.522-.174-.527-
.455.01.634l19.087-6.362c.529-.176.832.12.684.638l-5.454 19.086c-.15.529-.455.547-
.679.045L12 14l6-8-8 6-8.054-2.685z"></path>
                    </svg>
                </div>
            </div>
            <span>Predict</span>
        </button>
    </form>
    <b>{{showcase}}</b>
</body>

</html>
```


Flight Delay Prediction using Machine Learning
Team ID : PNT2022TMID18451
Flight Number :
Enter your Flight Number
Month :
Enter your Month
Day Of Month :
Enter your Day of Month
Day Of Week :
Enter your Day of week
Origin :
msp
Destination :
msp
Scheduled Departure Time :
Enter your scheduled departure time
Scheduled Arrival Time :
Enter your scheduled arrival time
Actual Departure Time :
Enter your actual arrival time
Predict

Predict.html:

```
<!DOCTYPE html>
<html lang="en">
   <head>
      <meta charset="UTF-8">
      <meta http-equiv="X-UA-Compatible" content="IE=edge">
      <meta name="viewport" content="width=device-width, initial-scale=1.0">
      <title>Flight Delay Prediction</title>
      <!-- <link rel="stylesheet" href="predict.css" -->

   </head>

   <body>
      <h1>Predicted Analysis of your Flight</h1>
      <h2>{{ showcase }}</h2>
      <a href="/">Go Back</a>
   </body>
</html>
```

Style.css:

```
body {     background-image:
url('preflight.jpg');     background-
repeat: no-repeat;     background-
attachment: fixed;     background-size:
cover;     text-align: center;
    /* background-blend-mode:overlay; */
}

h1 {     margin-top:
10px;     font-size:
55px;     margin-bottom:
20px;     text-align:
center; letter-spacing:
2px;     font-weight:
300px;     color:
rgb(255, 255, 255);
```

```
    opacity: .9;
}

h2 {     margin-top: 10px;
font-size: 30px;     margin-
bottom: 5px;     text-align:
center;     font-weight: 30px;
color: rgb(255, 222, 181);
    opacity: .9;
}

form {
    /* background-color: rgba(255, 255, 255, 0.804); */
background-color: #2d3c0f;     max-width: 500px;
margin: 50px auto;     padding: 30px 20px;     width:
30rem;     height: 1050px;     border-radius: 30px;
    box-shadow: 2px 5px 10px rgba(0, 0, 0, 0.5);
}

/* Styling form-control Class */
.form-control {     text-
align: left;
    margin-bottom: 25px;
}

/* Styling form-control Label */
.form-control label {
display: block;     margin-
bottom: 10px;     font-size:
25px;     color: #19096b;
    font-weight: bolder;
    font-style: oblique;
}

/* Styling form-control input,
select, textarea */ .form-control
input {
    /* border: 1px solid #777; */
/* border-radius: 2px; */
padding: 10px;     display: block;
width: 95%;     width: 80%;     /*
margin-bottom: 8px; */     height:
500%;     /* margin-top: 30%; */
```

```
margin: 5% auto;     /* border:
none; */     font-size: 14px;
outline: none;     background-
color: #c4c4c462;     border-
radius: 10px;
}

.form-control select {     padding:
10px;     display: block;     width:
95%;     width: 80%;     /*
margin-bottom: 8px; */     height:
500%;     /* margin-top: 30%; */
margin: 5% auto;     /* border:
none; */     font-size: 14px;
outline: none;     background-
color: #c4c4c462;
   border-radius: 10px;
}

button {
   font-size: 20px;
   background: rgb(21, 65, 196);
   color: white;     padding:
0.7em 1em;     height:
50px;     width: 165px;     /*
align-content: center; */     /*
padding-left: 0.em; */
display: flex;     border-
radius: 16px;     overflow:
hidden;     transition: all
0.2s;     align-items: center;
margin-left: 30%;
}

button span {     display: block;
margin-left: 0.9em;     transition:
all 0.3s ease-in-out;
}

button svg {     display: block;
transform-origin: center center;
transition: transform 0.3s ease-in-out;
}
```

```
button:hover .svg-wrapper {
   animation: fly-1 0.6s ease-in-out infinite alternate;
}

button:hover svg {
   transform: translateX(1.2em) rotate(45deg) scale(1.1);
}

button:hover span {
   transform: translateX(5em);
}

button:active {
   transform: scale(0.95);
}

@keyframes fly-1 {
from {
      transform: translateY(0.1em);
   }
to {
      transform: translateY(-0.1em);
   }
}

.reg {
   margin-left: 100px;
}
```

Predict.css:

```
@import
url('https://fonts.googleapis.com/css2?family=Poppins:ital,wght@0,400;0,500;0,600;0,7
00;0,800;0,900;1,300&display=swap');
 *{
   margin: 0%;
padding: 0%;
}

body {
      background-image: url('preflight.jpg');
background-repeat: no-repeat;
```

```
background-attachment: fixed;
background-size: cover;     text-align:
center;
    /* background-blend-mode:overlay; */
}

h1 {     margin-top: 10px;
font-size: 55px;
margin-bottom: 20px;
text-align: center;
letter-spacing: 2px;
font-weight: 300px;
color: rgb(255, 255, 255);
opacity: .9;
}

h2 {     margin-top: 10px;
font-size: 120px;
margin-bottom: 5px;
text-align: center;
font-weight: 30px;
color: rgb(2, 48, 94);
opacity: .9;
}

a {     font-size:
35px;
}
```

CHAPTER 8

TESTING

8.1. TEST

User No	Flight No	Month	Day of month	Day of week	Origin	Destination	Scheduled Departure Time	Scheduled Arrival Time	Actual Departure Time	Actual Inputs
1	1232	1	1	1	ATL	MSP	1905	2305	1945	Delayed
2	1399	1	1	1	ATL	SEA	1805	2410	1855	Delayed
3	2351	1	2	3	ATL	DTW	1305	2305	1305	Not Delayed
4	2637	2	1	3	DTW	ATL	1500	2410	1505	Not Delayed

8.2. USER ACCEPTANCE TESTING

This report shows the number of test cases that have passed and failed

User No	Flight No	Month	Day Of Month	Day Of Week	Origin	Destination	Scheduled Departure Time	Scheduled Arrival Time	Actual Departure Time	Actual Output	Predict - ed Output	Correct-ne ss
1	1232	1	1	1	ATL	MSP	1905	2305	1945	Delayed	Delayed	Correct
2	1399	1	1	1	ATL	SEA	1805	2410	1855	Delayed	Delayed	Correct
3	2351	1	2	3	ATL	DTW	1305	2305	1305	Not Delayed	Not Delayed	Correct
4	2637	2	1	3	DTW	ATL	1500	2410	1505	Not Delayed	Not Delayed	Correct

CHAPTER 9

RESULTS

9.1. PERFORMANCE METRICS

Training Accuracy

MODEL EVALUATION

```
acc=accuracy_score(predicted,y_test)
```

```
acc
```

```
0.8791308284291535
```

```
from sklearn.metrics import classification_report
print(classification_report(predicted, y_test, labels=[1, 2, 3]))
```

	precision	recall	f1-score	support
1	0.48	0.46	0.47	255
2	0.00	0.00	0.00	0
3	0.00	0.00	0.00	0
micro avg	0.48	0.46	0.47	255
macro avg	0.16	0.15	0.16	255
weighted avg	0.48	0.46	0.47	255

Confusion Matrix

```
from sklearn.metrics import confusion_matrix
confusion_matrix(predicted, y_test)
```

```
array([[1825,  129],
       [ 138,  117]], dtype=int64)
```

CHAPTER 10

ADVANTAGES AND DISADVANTAGES

Advantages

- Customers are happy.
- With this model, we can easily simplify the extensive traffic at the airport and can prevent the major confusions over flight delays.
- This can enable customer satisfaction and incomes of major airlines.
- Accuracy is measured with the previous models and we have analyzed that this model is much more effective in every way.
- The delay prediction model can make the concerned authorities be well prepared for any possible problem.
- The model can easily be understood by a layman: the model is simple and effective.

Disadvantages

- Wrong prediction due to noise of input data
- If the prediction is wrong, then there will be extra expenses for the agencies, passengers and airport
- Passengers with medical emergencies gets affected
- This model needs to be more compact and flexible. The interoperability feature should be more enhanced.
- The model can be automated instead of manually entering data from the user. Manually entering data is hectic work for the user.

CHAPTER 11

CONCLUSION

In this project, we use flight data, weather, and demand data to predict flight departure delay. In the end, our model correctly predicts the delayed and non-delayed flights correctly. As a result, there can be additional features related to the causes of flight delay that are not yet discovered using our existing data sources.

CHAPTER 12

FUTURE SCOPE

Based on data analysis from the year 2008, this project. There is a sizable dataset accessible from 1987 to 2008, but managing a larger dataset necessitates extensive preprocessing and purification of the data Therefore, adding a larger dataset is a part of this project's future effort. Preprocessing a bigger dataset can be done in a variety of methods, such as establishing a Spark cluster on a computer or using cloud services like AWS and Azure. Now that deep learning has advanced, we can employ neural networks algorithms to analyze aviation and meteorological data. Neural networks employ a form of pattern matching.

The project's focus is primarily on flight and weather data for India, but we can also include data from other nations like China, the United States, and Russia. We can broaden the project's scope by including flight information from international flights rather than just domestic flights.

CHAPTER 13 APPENDIX

GITHUB & PROJECT DEMO LINK

Github link: https://github.com/IBM-EPBL/IBM-Project-27409-1660055754

Project Demo link:

https://drive.google.com/file/d/18VSWNOSLTIru8mEt1vWB0x_ROEoBFxw/view?usp=drivesdk