IBM–NALAIYATHIRANPROJECT

DEVELOPINGAFLIGHTDELAYPREDICTIONMODELUSING MACHINELEARNING

INDUSTRY MENTOR :Lalitha Gayathri

FACULTYMENTOR :V RAMESH

TEAMID	**:**	PNT2022TMID44521
TEAMLEAD	**:**	JEEVA E
TEAMMEMBER	**:**	CHANDRU R.S
TEAMMEMBER	**:**	SELVAN P
TEAMMEMBER	**:**	YOGESH M
TEAMMEMBER	**:**	SRIGURUPRASATH

ABSTRACT

Flight delays in air transportation are a major concern that has adverse effects on the economy, the passengers, and the aviation industry. This matter critically requires an accurate estimation for future flight delays that can be implemented to improve airport operations and customer satisfaction. Having said that, a massive volume of data and an extreme number of parameters have restricted the way to build an accurate model.

Many existing flightdelay prediction methods are based on small samples and/or are complex to interpret with little or no opportunity for machine learning deployment. This paper develops a prediction model by analysing the data of domestic flights within the United States of America (USA). The proposed model gains insight into factors causing flight delays, cancellations and the relationship between departure and arrival delay using exploratory data analysis.

In addition, Random Forest (RF) algorithm is used to train and test the big dataset to help the model development. A web application has also been developed to implement the model and the testing results are presented with the limitation discussed.

TABLEOFCONTENT

CHAPTER	CONTENTS	PAGENO
1	**INTRODUCTION** 1. PROJECTOVERVIEW 2. PURPOSE	06 07
2	**LITERATURESURVEY** 1. EXISTING PROBLEM 2. REFERENCES 3. PROBLEM STATEMENT DEFINITION	08
3	**IDEATION&PROPOSEDSOLUTION** 1. EMPATHYMAPCANVAS 2. IDEATION & BRAINSTROMING 3. PROPOSED SOLUTION 4. PROBLEM SOLUTIONFIT	13 14 15 23
4	**REQUIREMENTANALYSIS** 1. FUNCTIONAL REQUIREMENT 2. NON-FUNCTIONAL REQUIREMENTS	24 25
5	**PROJECTDESIGN** 1. DATAFLOW DIAGRAMS 2. SOLUTION & TECHNICAL ARCHITECTURE 3. USERSTORIES	26 27 27

6	**PROJECTPLANNING&SCHEDULING** 1. SPRINT PLANNING & ESTIMATION 2. SPRINT DELIVERY SCHEDULE	**30** **31**

7	**CODING & SOLUTIONING** 1. FEATURE1 2. FEATURE2 3. FEATURE3 4. FEATURE4	 32 33 34 34
8	**TESTING** 1. TEST CASES 2. USER ACCEPTANCE TESTING	 42 43
9	**RESULTS** 9.1PERFORMANCE METRICS	**44**
10	**ADVANTAGES&DISADVANTAGES**	**46**
11	**CONCLUSION**	**47**
12	**FUTURESCOPE**	**48**
13	**APPENDIX**	**49**

CHAPTER 1

INTRODUCTION

Travelers have begun to favor 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

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

2.LITERATURESURVEY

EXISTINGPROBLEM:

[1] H. Khaksar and A. Sheikholeslami, “Airline delay prediction by machine learning algorithms”, Scientia Iranica, Transactions A: Civil Engineering 26 (2019) 2689-2702.

Proposed work:

This paper proposes a flight delay prediction model through different methods which includes Bayesian modeling, decision tree, cluster classification, random forest, and hybrid methods. These methods were applied to estimate the occurrences and magnitude of delay in a network.

[2] Miguel Lambelho, Mihaela Mitici, Simon Pickup, Alan Marsden,”Assessing strategic flight schedules at an airport using machine learning-based flight delay and cancellation predictions”, Journal of Air Transport Management, Volume 82, 2020, 101737, ISSN 0969-6997.

[3] Navoneel Chakrabarty, ”A Data Mining Approach to Flight Arrival Delay Prediction for
American Airlines“, The 9th Annual Information Technology, Electromechanical and Microelectronics Conference (IEMECON 2019).

Proposed work:

This paper aims at analyzing flight information of US domestic flights operated by American Airlines, covering top 5 busiest airports of the US and predicting possible arrival delay of the flight using Data Mining and Machine Learning Approaches.

[4] Kaiquan Cai, Yue Li, Yiping Fang, Yanbo Zhu," A Deep Learning Approach for Flight
Delay PredictionthroughTime-EvolvingGraphs".IEEE Transactions on Intelligent Transportation Systems,IEEE,In press,pp.1-11.ff10.1109/TITS.2021.3103502ff. ffhal-03428046f.

Proposed work:

This paper is about the flight delay prediction problem is investigated from a network perspective (i.e., multi-airport scenario). To model the time-evolving and periodic graph-structured information in the airport network, a flight delay prediction approach based on the graph convolutional neural network (GCN) is developed in this paper.

[5] Yi Ding," Predicting flight delay based on multiple linear regression" ,2017 IOP Conf.
Ser.: Earth Environ. Sci. 81 012198

Proposed work: This paper proposes a method to model the arriving flights and a multiple linear regression algorithm to predict delay, comparing with Naive-Bayes and

C4.5 approach.

[6] Qu, J., Zhao, T., Ye, M. et al. “Flight Delay Prediction Using Deep Convolutional Neural Network Based on Fusion of Meteorological Data.”, Neural Process Lett 52, 1461–1484 (2020).

Proposed work:

This paper provides two flight delay prediction models using deep convolutional neural networks based on fusion of meteorological data. The first model is DCNN (Dual- channel Convolutional Neural Network), which refers to the ResNet network structure. The second model is SE- DenseNet (Squeeze and ExcitationDensely Connected Convolutional Network).

[7] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou and D. Zhao, "Flight Delay Prediction Based

on Aviation Big Data and Machine Learning," in IEEE Transactions on Vehicular Technology, vol. 69, no. 1, pp. 140-150, Jan. 2020, doi: 10.1109/TVT.2019.2954094.

Proposed work:

This paper 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. To build a dataset for the proposed scheme, automatic dependent surveillance-broadcast (ADS-B) messages are received, pre- processed, and integrated with other information such as weather condition, flight schedule, and airport information.

[8] Yu, Bin; Guo, Zhen; Asian, Sobhan; Wang, Huaizhu; Chen, Gang (2019),”Flight

delay prediction for commercial air transport: A deep learning approach." Transportation Research Part E: Logistics and Transportation Review.

Proposed work:

This paper analyzes high-dimensional data from Beijing International Airport and presents a practical flight delay prediction model. Following a multifactor approach, a novel deep belief network method is employed to mine the inner patterns of flight delays. Support vector regression is embedded in the developed model to perform a supervised fine-tuning within the presented predictive architecture

9] Esmaeilzadeh, Ehsan; Mokhtarimousavi, Seyedmirsajad (2020)."Machine Learning
Approach for Flight Departure Delay Prediction and Analysis". Transportation Research Record: Journal of the Transportation Research Board.

Proposed work:

This paper employs a support vector machine (SVM) model to explore the non- linear relationship between flight delay outcomes. Individual flight data were gathered from 20 days in 2018 to investigate causes and patterns of air traffic delay at three major New York City airports

[10] Etani, Noriko (2019),"Development of a predictive model for on-time arrival flight ofairliners by discovering correlation between flight and weather data.", Journal of

Big
Data,2019.

Proposed work:

This paper aims to discover the correlation between fight data and weather data. A predictive model of on-time arrival flight is proposed using flight data and weather data. The feasibility of the predictive model is evaluated by developing a tool of on-time arrival fight prediction

CHAPTER 3

IDEATION & PROPOSED SOLUTION

3.1. EMPATHY MAP CANVAS

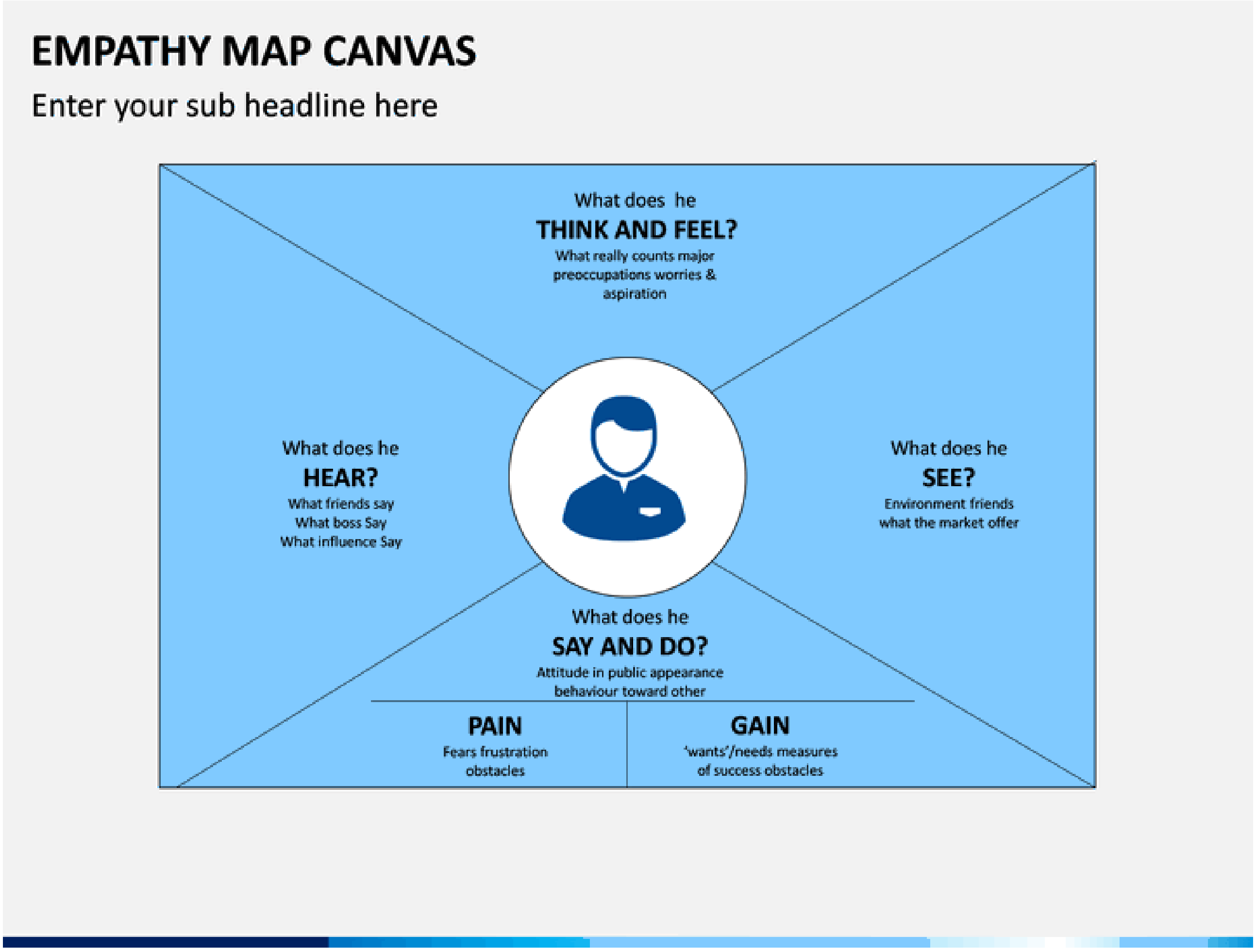
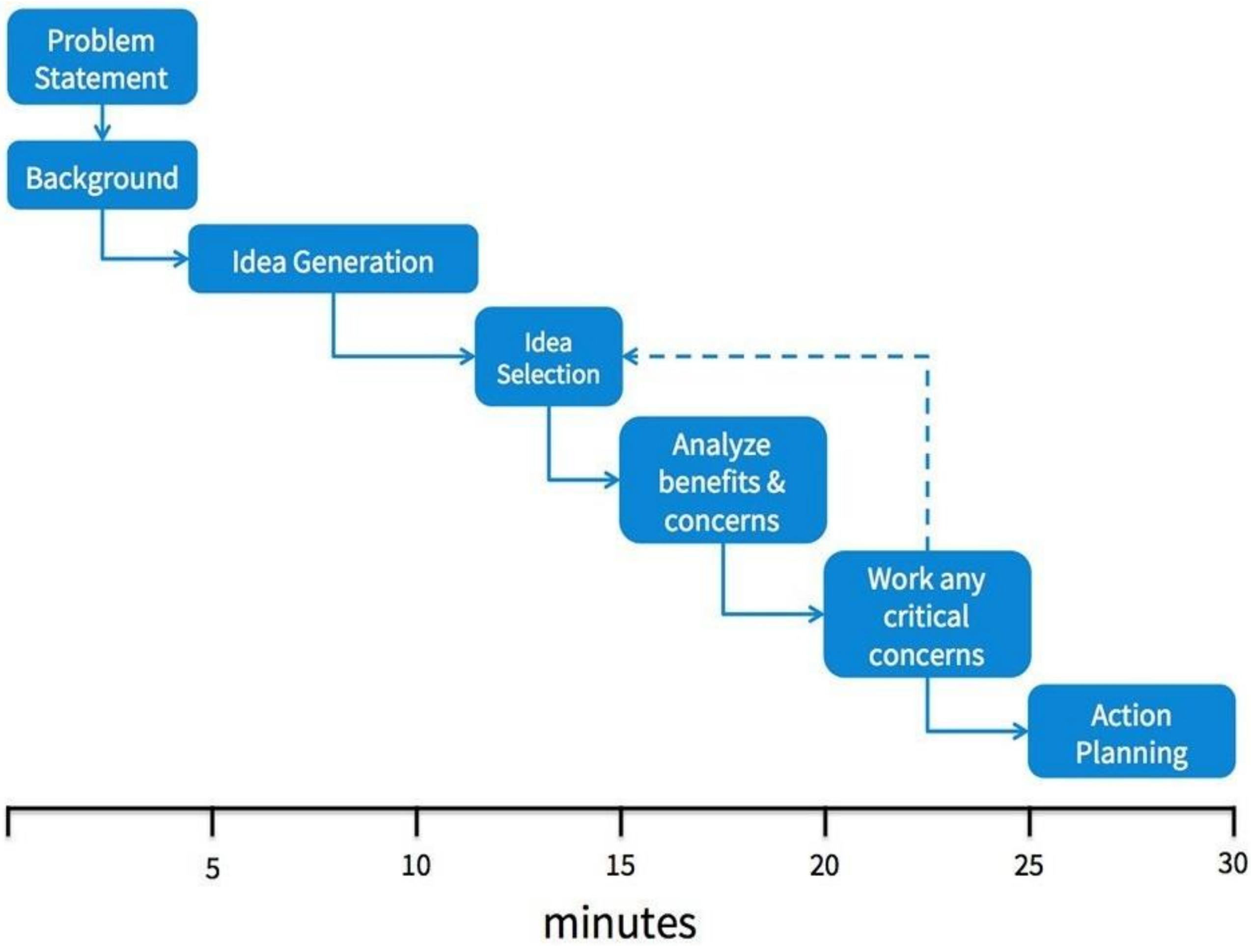


Figure 3.1. Empathy MaP

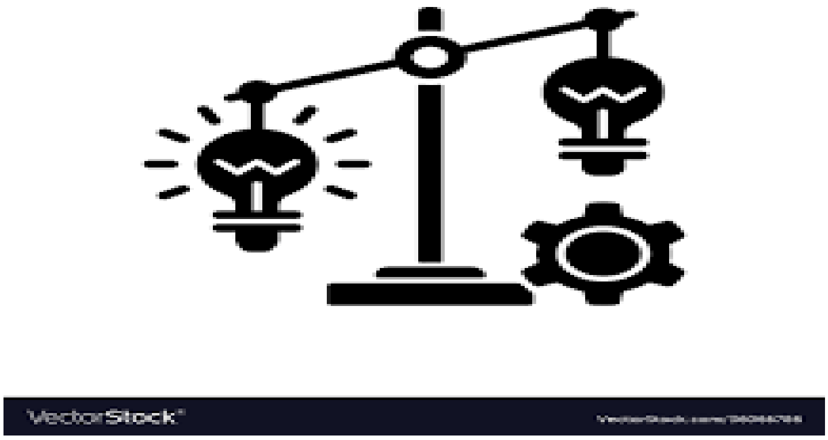
3.2. IDEATION & BRAINSTORMING

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



Step 2 - Brainstorm, Idea Listing and Grouping

Step 3 - Idea Prioritization



3.PROPOSED\SOLUTION

S.No.	Parameter	Description
	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 network.
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.
	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

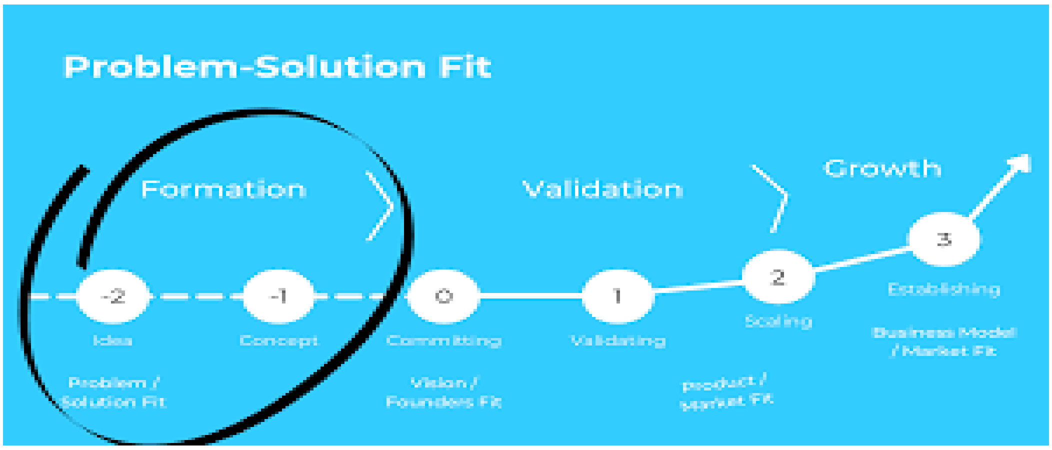
		more ML models using the techniques of ensembling
4	Social Impact / Customer Satisfaction	Flight delays not only anger and disturb air travelers' plans, but they also reduce productivity, raise capital costs, reallocate flight crews and aircraft, and add to crew costs. Higher operating costs for airline firms are unavoidable as flight delays necessitate the consumption of more labor, 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

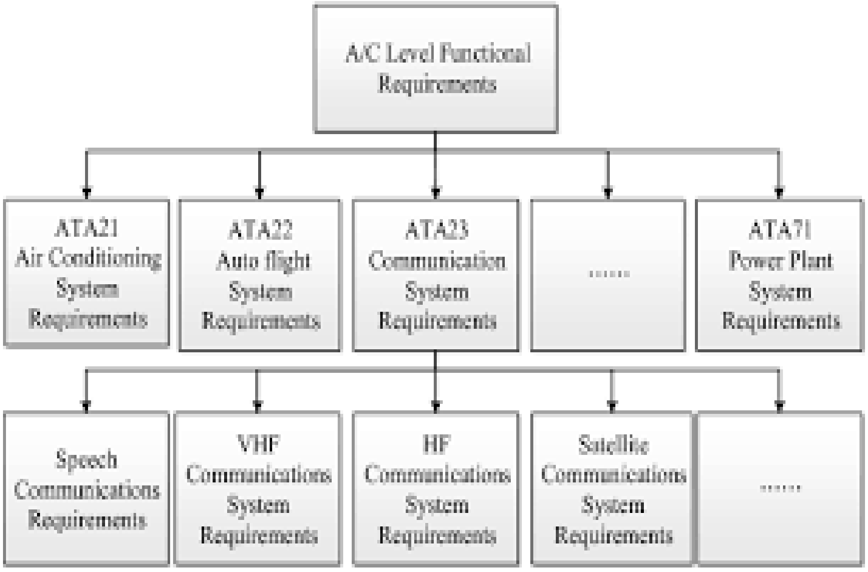
3.4. PROBLEM SOLUTION FIT

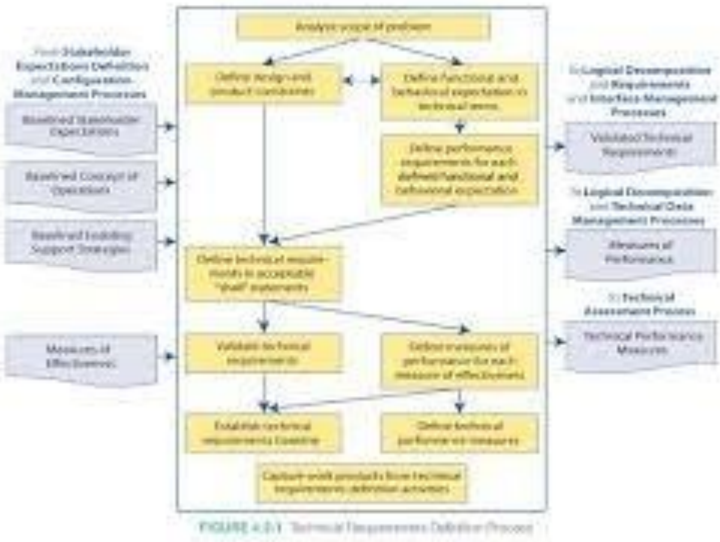


CHAPTER 4

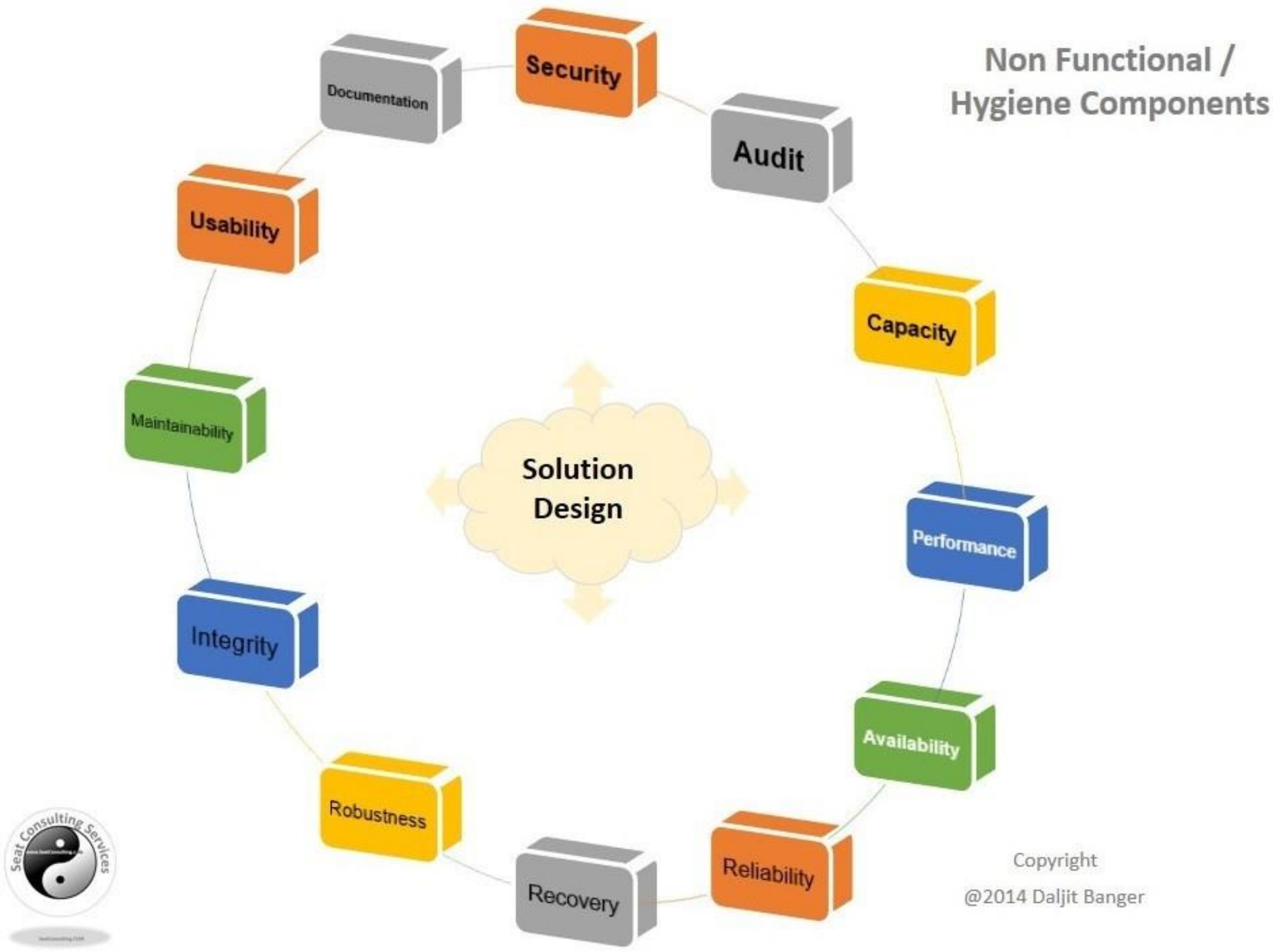
REQUIREMENT ANALYSIS

4.1. FUNCTIONAL REQUIREMENT





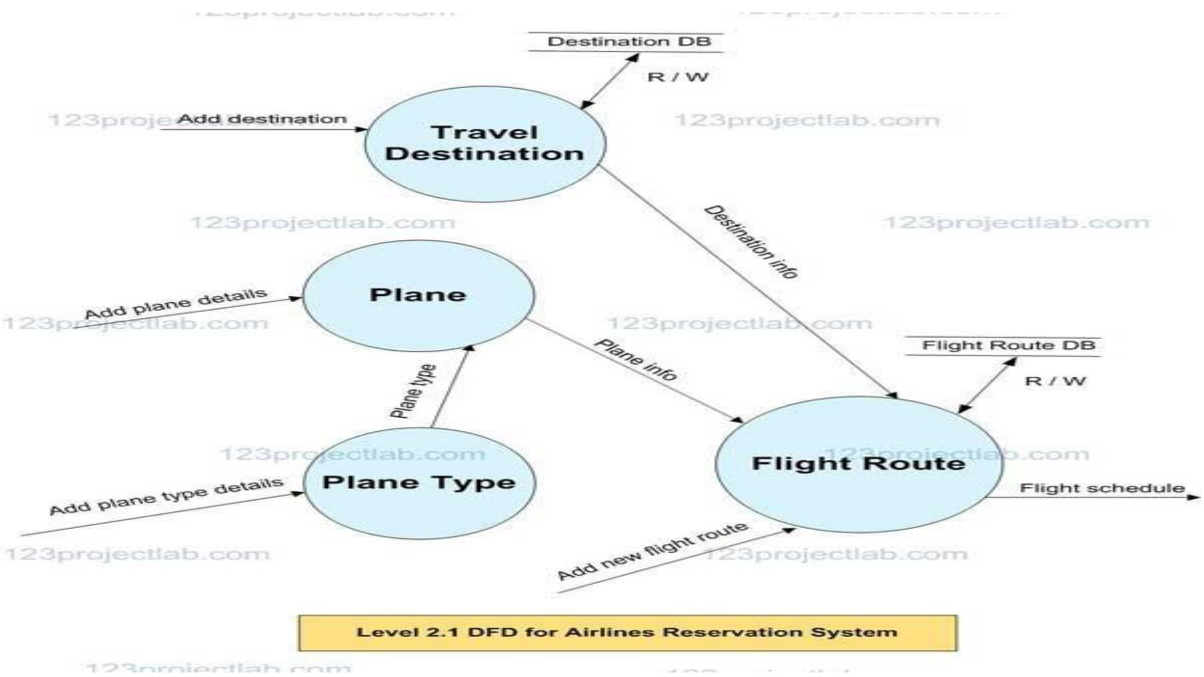
4.2. NON-FUNCTIONAL REQUIREMENTS



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.



Level 2.1 DFD for Airlines Reservation System

FIG 5.1.DATA FLOW DIAGRAMS

5.2. SOLUTION & TECHNICAL ARCHITECTURE

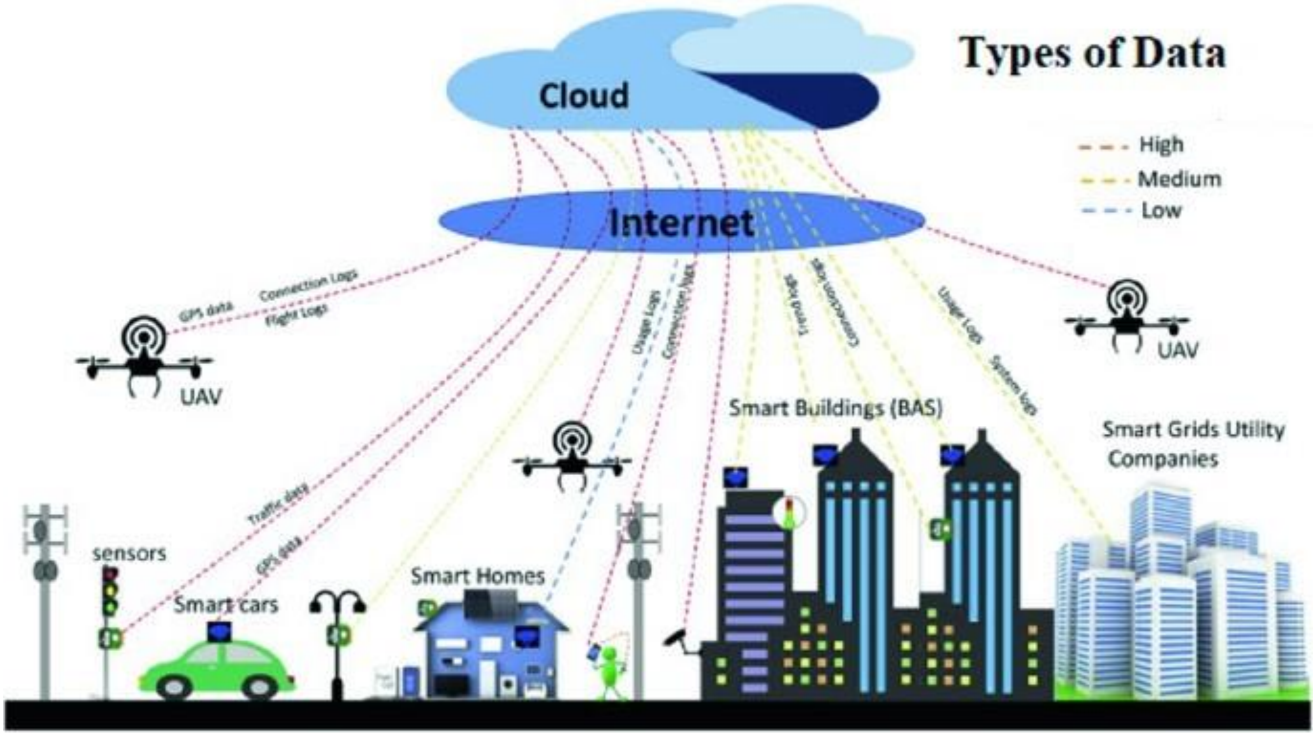
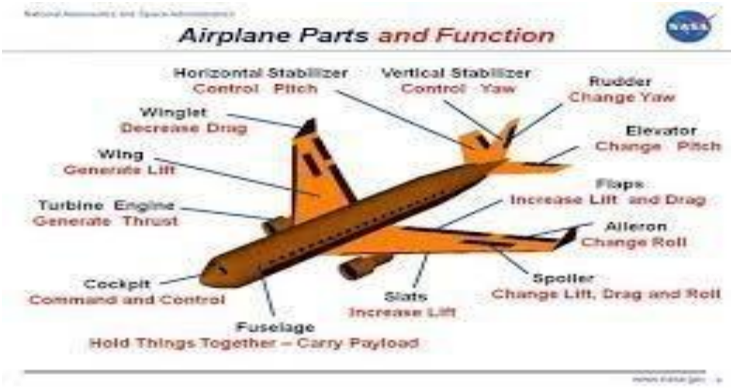


FIG 5.2. SOLUTION & TECHNICAL ARCHITECTURE

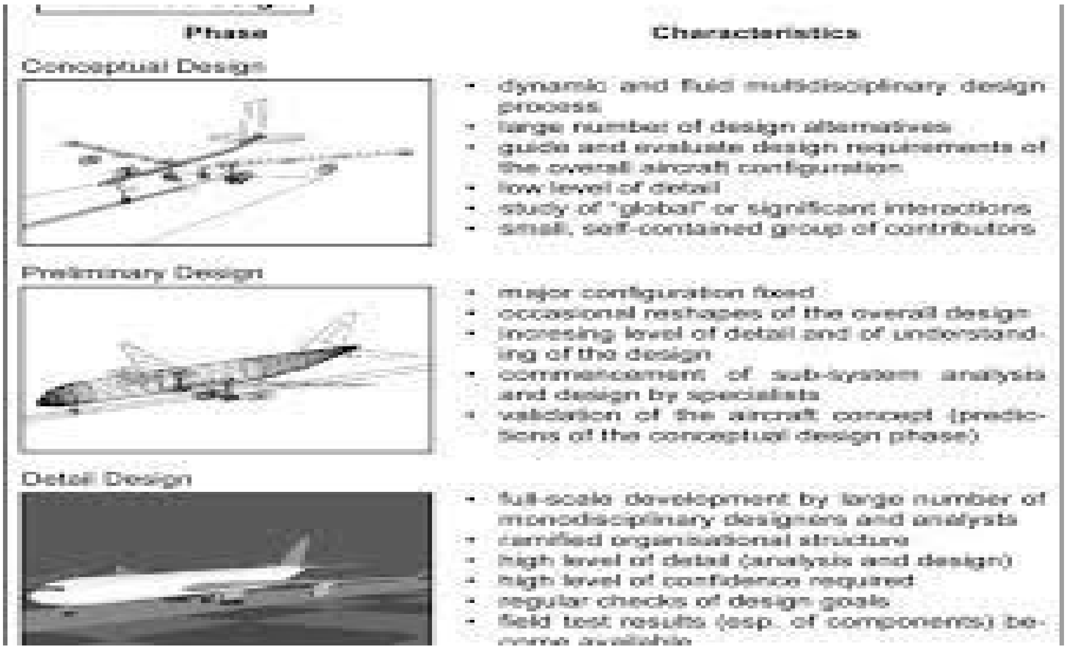
5.3. Technology Stack

Figure 5.3. Technology Stack

Components & Technologie

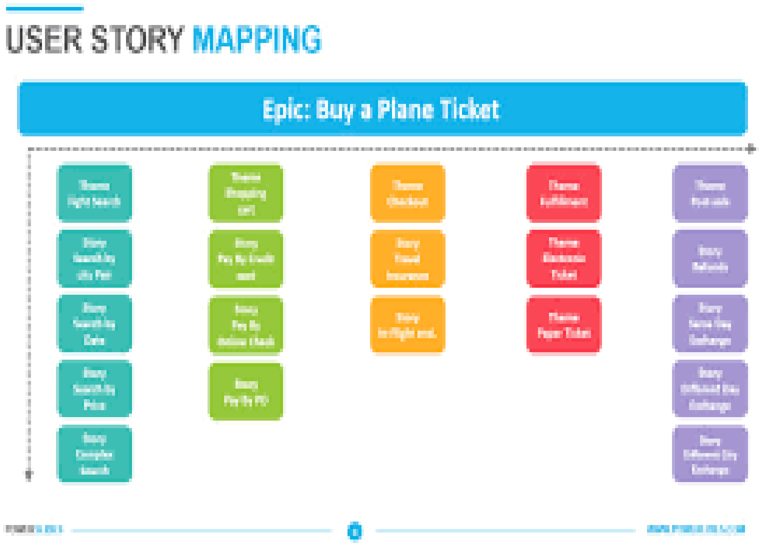


Application Characteristics



Phase	Characteristics
Conceptual Design	• dynamic and fluid multidisciplinary design process • large number of design alternatives • guide and evaluate design requirements of the overall aircraft configuration • low level of detail • study of "global" or significant interactions • small, self-contained group of contributors
Preliminary Design	• major configuration fixed • occasional reshapes of the overall design • increasing level of detail and of understanding of the design • commencement of sub-system analysis and design by specialists • validation of the aircraft concept (predictions of the conceptual design phase)
Detail Design	• full-scale development by large number of monodisciplinary designers and analysts • ramified organisational structure • high level of detail (analysis and design) • high level of confidence required • regular checks of design goals • field test results (esp. of components) be-

5.3. User Stories



CHAPTER 6

PROJECT PLANNING & SCHEDULING

6.1. SPRINT PLANNING & ESTIMATION

fig 6.1 SPRINT PLANNING

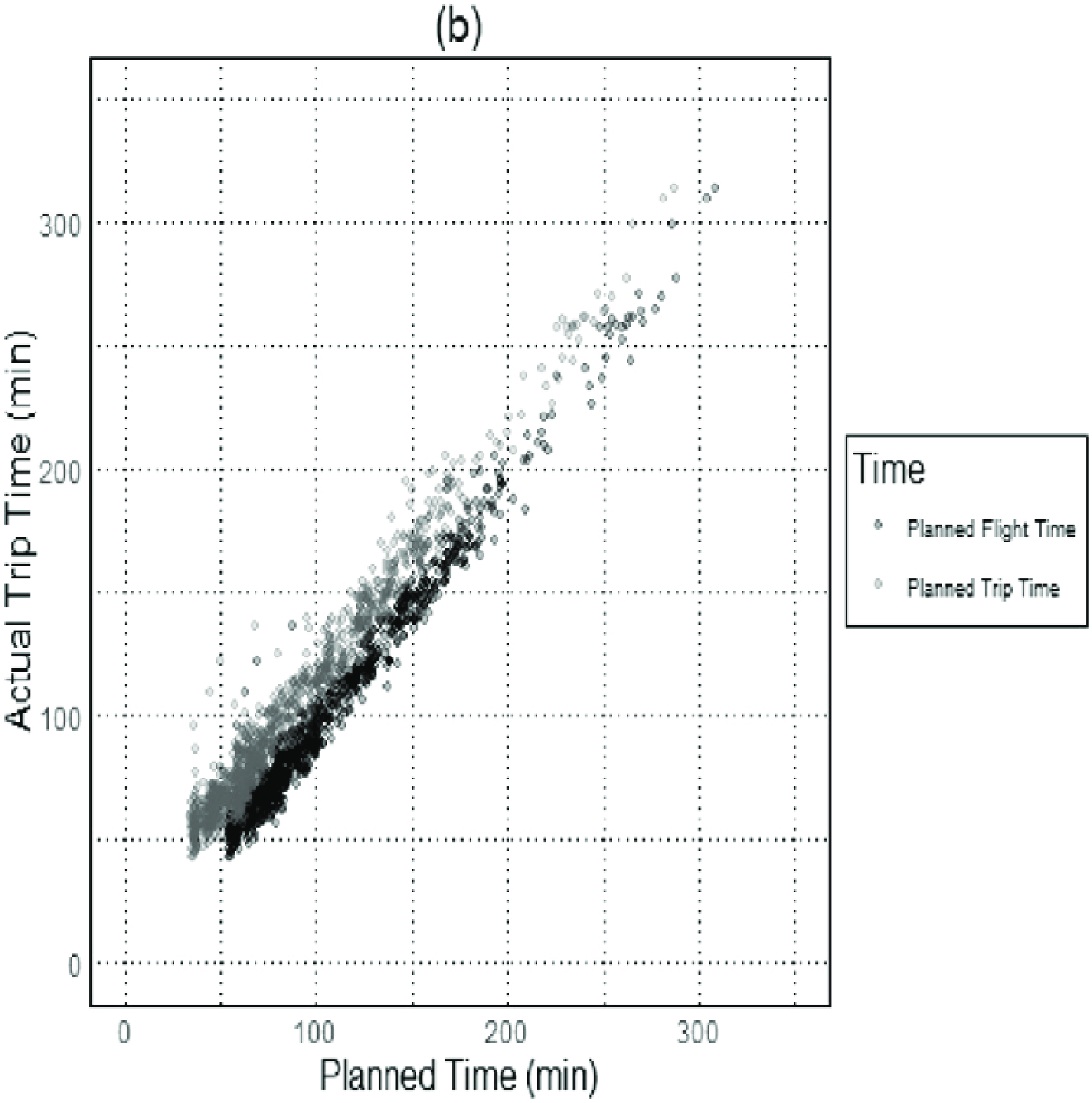
6.2. SPRINT DELIVERY SCHEDULE

fig 6.2. SPRINT DELIVERY SCHEDULE

CHAPTER 7

CODING AND SOLUTIONING

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



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 [271]:

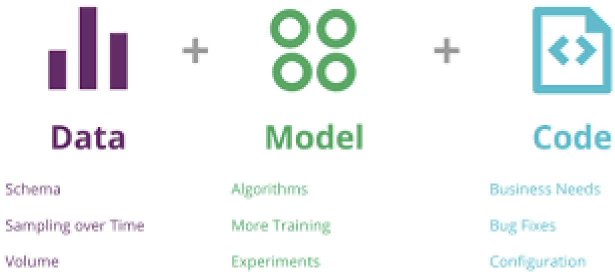
```
df = pd.DataFrame([['male', 82, 'graduate'],
                   ['female', 54, 'postgraduate'],
                   ['female', 42, 'highschool'],
                   ['female', 28, 'highschool']])
df.columns = ['gender', 'weight', 'degree']
df.head()
```

Out[271]:

	gender	weight	degree
0	male	82	graduate
1	female	54	postgraduate
2	female	42	highschool
3	female	28	highschool

fig 7.2. FEATURE 2 - ONE HOT ENCODING

7.3. FEATURE 3 - SAVING THE MODEL WEIGHTS FOR DEPLOYMENT



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
model = pickle.load(open('flight_new.pk1','rb'))
app = Flask(__name__)
```

```
@app.route('/')
def home():
return render_template("mainpage.html")
@app.route('/prediction',methods=['GET','POST'])
def predict():
name = request.form['fname']
month = request.form['month']
dayofmonth = request.form['daymonth']
dayofweek = request.form['dayweek']
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 == "atl"):
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['sdeparttime']
arrtime = request.form['sarrivaltime']
actdept = request.form['adeparttime']
dept15 = int(dept)-int(actdept)
```

```
total =
[[name,month,dayofmonth,dayofweek,arrtime,dept15,origin1,origin2,orig
in3,origin4,origi
n5,destination1,destination2,destination3,destination4,destination5]]
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("index.html",data = ans)
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 - HTML PAGES FOR FRONTEND DESIGN:

```
<html>
<div align="center" class="logbg">
<head>
<meta charset="UTF-8">
<center>
<table>
<tr>
<td><h1><br>Prediction of Flight Delay<br><br></h1></td>
</tr>
</table>
</center>
</head>
<body background='C:\Users\Public\project\templates\flight_4.jpg'>
<form action="http://localhost:5000/prediction" method="POST" >
<center>
<table>
<tr>
<td>Enter the flight number:</td>
```

```
<td><input type="number" name="fname"><br></td>
</tr>
<tr>
<td>Month:</td>
<td><input type="number" name="month"><br></td>
</tr>
<tr>
<td>Day of Month:</td>
<td><input type="number" name="daymonth"><br></td>
</tr>
<tr>
<td>Day of Week:</td>
<td><input type="number" name="dayweek"><br></td>
</tr>
<tr>
<td>Origin:</td>
<td><select name="origin">
<option value="atl">ATL</option>
<option value="dtw">DTW</option>
<option value="sea">SEA</option>
```

```
<option  value="msp">MSP</option>
<option  value="jfk">JFK</option>
</select></td>
<tr>
<tr>
<td>Destination:</td>
<td><select name="destination">
<option  value="atl">ATL</option>
<option value="dtw">DTW</option>
<option  value="sea">SEA</option>
<option  value="msp">MSP</option>
<option  value="jfk">JFK</option>
</select></td>
<tr>
<tr>
<td>Scheduled Departure Time:</td>
<td><input  type="number"  name="sdeparttime"><br></td>
</tr>
<tr>
<td>Scheduled Arrival Time:</td>
```

```
<td><input type="number" name="sarrivaltime"><br></td>
</tr>
<tr>
<td>Actual Departure Time:</td>
<td><input type="number" name="adeparttime"><br></td>
</tr>
<tr>
<td><br><input type="submit" class="btn" value="SUBMIT"></br>
</tr>
</table>
</center>
</form>
</body>
</div>
</html>
```

CHAPTER 8

TESTING

8.1. TEST

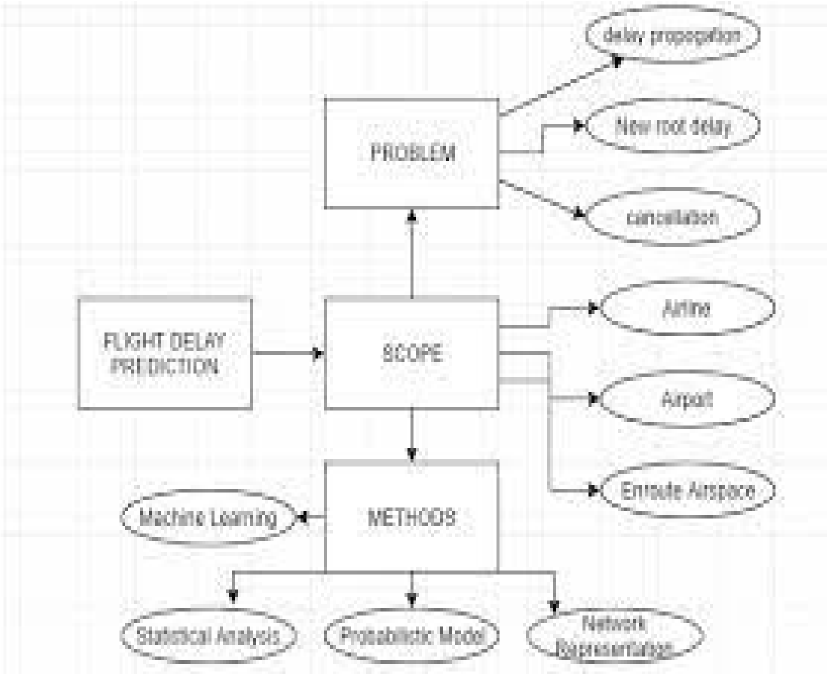


fig 8.1. TEST

8.2. USER ACCEPTANCE TESTING

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

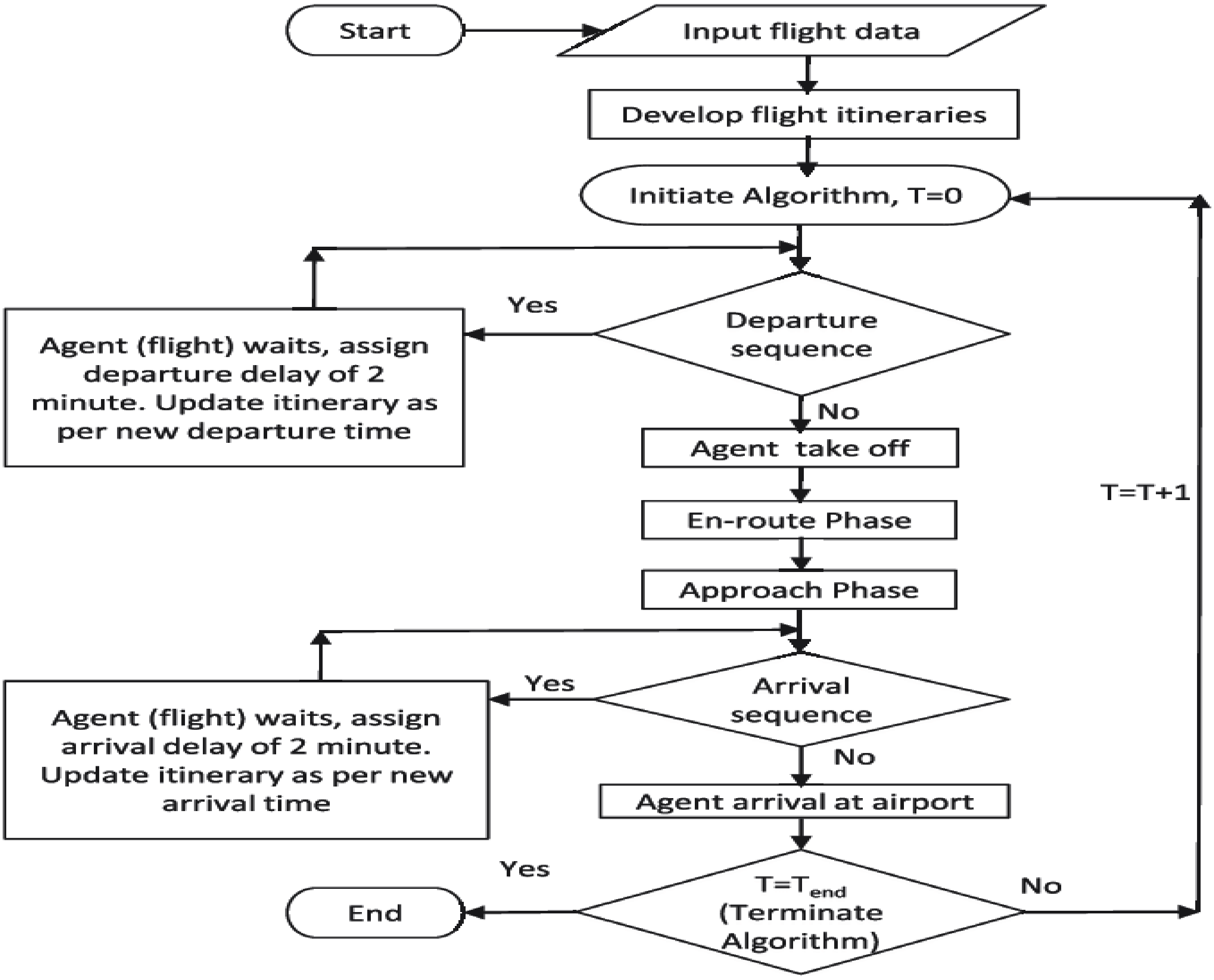


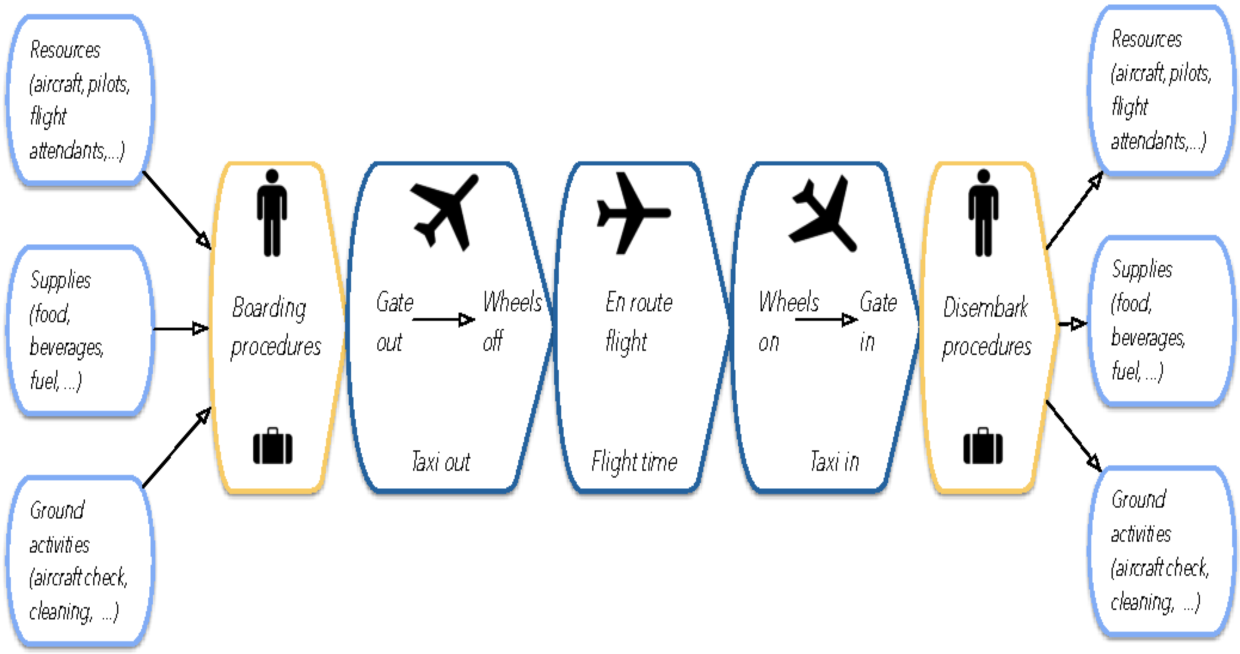
fig8.2. USER ACCEPTANCE TESTING

CHAPTER 9

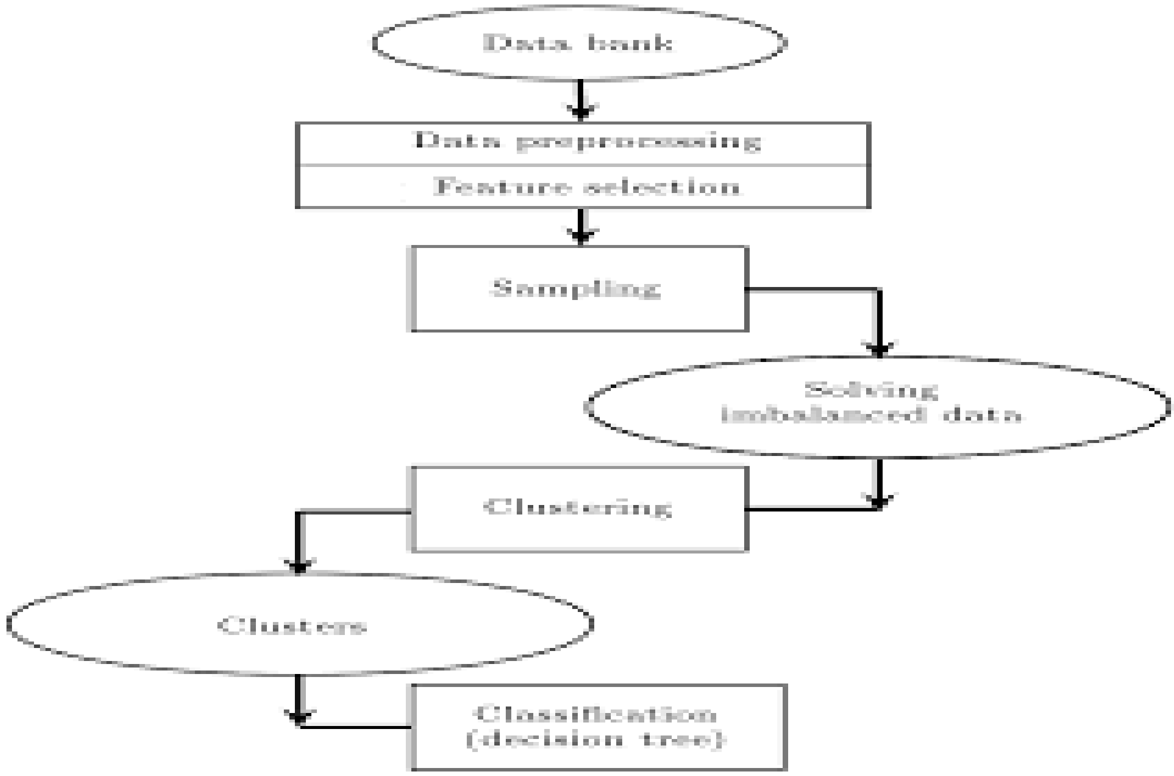
RESULTS

9.1. PERFORMANCE METRICS

Training Accuracy



Classification Model



CHAPTER 10

ADVANTAGES AND DISADVANTAGES

Advantages

- Customers are happy
- The available flights are easily identified
- Prior information will be sent if in case the flight is delayed
- The current status of the flight can be tracked.

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

CHAPTER 11

CONCLUSION

In this project, we use flight data, weather, and demand data to predict flight departure delay. Our result shows that the Random Forest method yields the best performance compared to the SVM model. Somehow the SVM model is very time consuming and does not necessarily produce better results.

In the end, our model correctly predicts 91% of the non-delayed flights. However, the delayed flights are only correctly predicted 41% of time. 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.In the second part of the project, we can see that it is possible to predict flight delaypatterns from just the volume of concurrently published tweets, and their sentiment and objectivity. This is not unreasonable; people tend to post about airport delays on Twitter; it stands to reason that these posts would become more frequent, and more profoundly emotional, as the delays get worse.

Without more data, we cannot make a robust model and find out the role of related factors and chance on these results. However, as a proof of concept, there is potential for these results. It may be possible to routinely use tweets to

ascertain an understanding of concurrent airline delays and traffic patterns, which could be useful in a variety of circumstances.

CHAPTER 12

FUTURE SCOPE

This project is based on data analysis from year 2008. A large dataset is available from 1987-2008 but handling a bigger dataset requires a great amount of preprocessing and cleaning of the data. Therefore, 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. Neural Network works on the pattern matching methodology. It is divided into three basic parts for data modelling that includes feed forward networks, feedback networks, and self- organization network. Feed-forward and feedback networks are generally used in the areas of prediction, pattern recognition, associative memory, and optimization calculation, whereas self-organization networks are generally used in cluster analysis. Neural Network offers distributed computer architecture with important learning abilities to represent nonlinear relationships.

CHAPTER 13

APPENDIX

13.1. SOURCE CODE

13.1.1. Train the ML Model

```
# # SPRINT-2
#   **TRAIN-TEST-SPLIT**
# In[45]:
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=
0)
#  In[46]:
x_test.shape
# In[47]:
x_train.shape
# In[48]:
y_test.shape
# In[49]:
y_train.shape
#   **Scaling**
# In[50]:
sc =  StandardScaler()
```

```
# In[51]:
x_train=sc.fit_transform(x_train)


# In[52]:
x_test=sc.fit_transform(x_test)
# **Model Building**
# In[53]:
classifier = DecisionTreeClassifier(random_state=0)
# In[54]:
classifier.fit(x_train,y_train)
# In[55]:
predicted = classifier.predict(x_test)
# In[56]:
predicted
# In[57]:
y_test
# **MODEL EVALUATION**
# In[58]:
acc=accuracy_score(predicted,y_test)
# In[59]:
acc
```

```
# In[  ]:
#  In[60]:
data[data['ARR_DEL15']>0].iloc[33].values
# In[61]:
sample=[[1.187e+03, 1.000e+00, 1.500e+01, 5.000e+00, 1.900e+01,
1.000e+00,
0.000e+00, 0.000e+00,  0.000e+00,  1.000e+00,  0.000e+00,
0.000e+00, 0.000e+00, 0.000e+00,  0.000e+00,  1.000e+00]]
# In[62]:
classifier.predict(sample)
#  **SAVING THE  MODEL**
# In[63]:
pickle.dump(classifier,open('flight_new.pk1','wb'))
# In[64]:
from sklearn.metrics import confusion_matrix
confusion_matrix(predicted, y_test)
# In[66]:
from sklearn.metrics import classification_report
print(classification_report(predicted, y_test, labels=[1,  2]))
# In[ ]:
```

13.1.2. Mainpage – HTML Code

```
<html>
<div align="center" class="logbg">
<head>
<meta charset="UTF-8">
<center>
<table>
<tr>
<td><h1><br>Prediction of Flight Delay<br><br></h1></td>
</tr>
</table>
</center>
</head>
<body background='C:\Users\Public\project\templates\flight_4.jpg'>
<form action="http://localhost:5000/prediction" method="POST" >
<center>
<table>
<tr>
<td>Enter the flight number:</td>
<td><input type="number" name="fname"><br></td>
```

```
</tr>
<tr>
<td>Month:</td>
<td><input type="number" name="month"><br></td>
</tr>
<tr>
<td>Day of Month:</td>
<td><input type="number" name="daymonth"><br></td>
</tr>
<tr>
<td>Day of Week:</td>
<td><input type="number" name="dayweek"><br></td>
</tr>
<tr>
<td>Origin:</td>
<td><select name="origin">
<option value="atl">ATL</option>
<option value="dtw">DTW</option>
<option value="sea">SEA</option>
<option value="msp">MSP</option>
```

```
<option  value="jfk">JFK</option>
</select></td>
<tr>
<tr>
<td>Destination:</td>
<td><select name="destination">
<option  value="atl">ATL</option>
<option value="dtw">DTW</option>
<option  value="sea">SEA</option>
<option  value="msp">MSP</option>
<option  value="jfk">JFK</option>
</select></td>
<tr>
<tr>
<td>Scheduled Departure Time:</td>
<td><input type="number" name="sdeparttime"><br></td>
</tr>
<tr>
<td>Scheduled Arrival Time:</td>
<td><input type="number" name="sarrivaltime"><br></td>
</tr>
```

```
<tr>
<td>Actual Departure Time:</td>
<td><input type="number" name="adeparttime"><br></td>
</tr>
<tr>
<td><br><input type="submit" class="btn" value="SUBMIT"></br>
</tr>
</table>
</center>
</form>
</body>
</div>
</html>
```

13.1.3 Prediction Page - HTML Code

```
<!doctype html>
<html>
<body background="C:\Users\Public\project\templates\flight_2.jpg">
<center>
<h1><strong>Thanks for asking</strong></h1>
<h2>{{data}}</h2>
<a href='/'>Go back to home page</a>
</center>
```

```
</body>
</html>
```

13.1.4. Flask Application

```
from flask import Flask, request, render_template
import numpy as np
import pandas as pd
import pickle
import os
model = pickle.load(open('flight_new.pk1','rb'))
app = Flask(__name__)

@app.route('/')
def home():
return render_template("mainpage.html")
@app.route('/prediction',methods=['GET','POST'])
def predict():
name = request.form['fname']
month = request.form['month']
dayofmonth = request.form['daymonth']
dayofweek = request.form['dayweek']
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
f(origin ==  "atl"):
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['sdeparttime']
arrtime =  request.form['sarrivaltime']
```

```
actdept = request.form['adeparttime']
dept15 = int(dept)-int(actdept)
total =
[[name,month,dayofmonth,dayofweek,arrtime,dept15,origin1,origin2,origin3,o
rigin4,origi
n5,destination1,destination2,destination3,destination4,destination5]]
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("index.html",data = ans)
app.run(debug=True)
```

13.2. GITHUB & PROJECT DEMO LINK

Github link

https://github.com/IBM-EPBL/IBM-Project-47661-1660800914

Project Demo link

https://www.youtube.com/watch?v=WH1ItFUmakE