# AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING



Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai NAAC Accredited Institution "Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.
ANNA UNIVERSITY COUNSELLING CODE: 1101



## BONAFIDE CERTIFICATE

Certified that this project titled "EXPLORATORY ANALYSIS OF RAINFALL DATA IN INDIA FOR AGRICULTURE" is the bonafide work MOHAMMED ISMAIL KAFIL(110119205027), MOHAMMED ISMAIL KM(110119205026), MOHAMMED FASEHIULLAH(110119205025), SYED IJAZ AHAMED(11011920503" who carried out the project work under my supervision

## PROJECT REPORT SUBMISSION

## TABLE OF CONTENT

## 1. INTRODUCTION

- a. Project Overview
- b. Purpose

## 2. LITERATURE SURVEY

- a. Existing problem
- b. References
- c. Problem Statement Definition

## 3. IDEATION & PROPOSED SOLUTION

- a. Empathy Map Canvas
- b. Ideation & Brainstorming
- c. Proposed Solution
- d. Problem Solution fit

## 4. REQUIREMENT ANALYSIS

- a. Functional requirement
- b. Non-Functional requirements

## 5. PROJECT DESIGN

- a. Data Flow Diagrams
- b. Solution & Technical Architecture
- c. User Stories

## 6. PROJECT PLANNING & SCHEDULING

a. Sprint Planning & Estimation

- b. Sprint Delivery Schedule
- c. Reports from JIRA
- 7. CODING & SOLUTIONING (Explain the features added in the project along with code)

#### 8. RESULTS

- a. Performance Metrics
- 9. ADVANTAGES & DISADVANTAGES
- 10. CONCLUSION

#### 11.APPENDIX

Source Code

GitHub & Project Demo Link

# **Exploratory Analysis Of RainFall Data In India For Agriculture**

## 1. INTRODUCTION

a. Project Overview

Agricultureis the backbone of the Indian economy. For agriculture, the most important thing is water source, i.e. rainfall. The prediction of the amount of rainfall gives alertness to farmers by knowing early they can protect their crops from rain. So, it is important to predict the rainfall accurately as much as possible. Exploration and analysis of data on

rainfall over various regions of India and especially the regions where agricultural works have been done persistently in a wide range. With the help of analysis and the resultant data, future rainfall prediction for those regions using various machine learning techniques such as XGBoost classifier, SVM classifiers, Decision tree, Naive bayes classifier, Logistic regression etc.

## b. Purpose

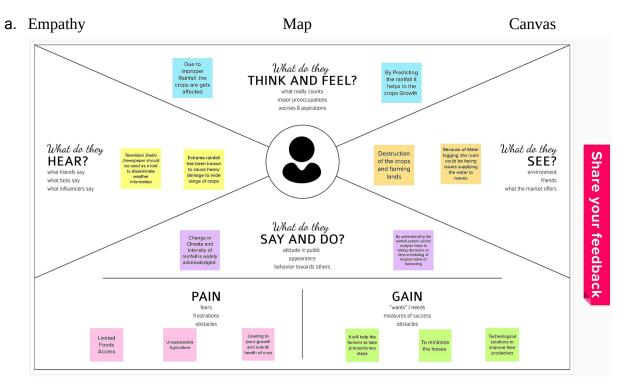
Rainfall has been a major concern these days. Weather conditions have been changing for time being. Rainfall forecasting is important otherwise, it may lead to many disasters. Irregular heavy rainfall may lead to the destruction of crops, heavy floods that can cause harm to human life. It is important to exactly determine the rainfall for effective use of water resources, crop productivity, and pre-planning of water structures.

#### 2. LITERATURE SURVEY

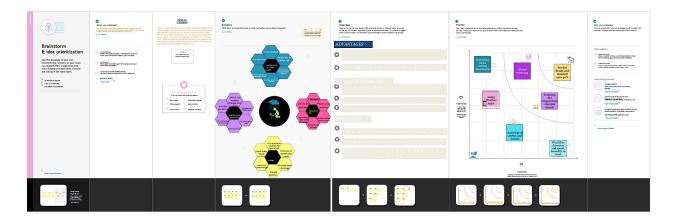
- a. Existing problem
- b. References
- They can serve as model to investigate the metabolic processes and behavior/reaction of a
  cell under chemical or photic stimuli, or determine the dynamic of natural populations in
  response to variations of environmental conditions. In fact, algae can occasionally bloom in
  enormous amount becoming a serious public health and environmental problem (Artiola,
  Pepper, & Brusseau, 2004; Wiersma, 2004).
  - c. Problem Statement Definition
  - Climate is a important aspect of human life. So, the Prediction should accurate as much as possible. In this paper we try to deal with the prediction of the rainfall which is also a major aspect of human life and which provide the major resource of human life which is Fresh Water.
     Fresh water is always a crucial resource of human survival – not only for the drinking purposes but also for farming,
  - Making a good prediction of climate is always a major task now a day because of the climate change.

 A bad rainfall prediction can affect the agriculture mostly framers as their whole crop is depend on the rainfall and agriculture is always an important part of every economy. So, making an accurate prediction of the rainfall somewhat good.

#### 3. IDEATION & PROPOSED SOLUTION



b. Ideation & Brainstorming



## c. Proposed Solution

Problem Statement (Problem to be solved):

- Climate is a important aspect of human life. So, the Prediction should accurate as much as possible. In this paper we try to deal with the prediction of the rainfall which is also a major aspect of human life and which provide the major resource of human life which is Fresh Water.
- Now climate change is the biggest issue all over the world. Peoples are working on to detect the patterns in climate change as it affects the economy in production to infrastructure

## Idea / Solution description:

• In rainfall also making prediction of rainfall is a challenging task with a good accuracy rate. Making prediction on rainfall cannot be done by the traditional way, so scientist is using machine learning and deep learning to find out the pattern for rainfall prediction. • Provides extra support to maintain the agriculture.

## Novelty / Uniqueness:

- This application is useful for the beginners in agriculture.
- Seed maturity selection features are available.

## Social Impact / Customer Satisfaction:

• Different types of crops can be planted for good health.

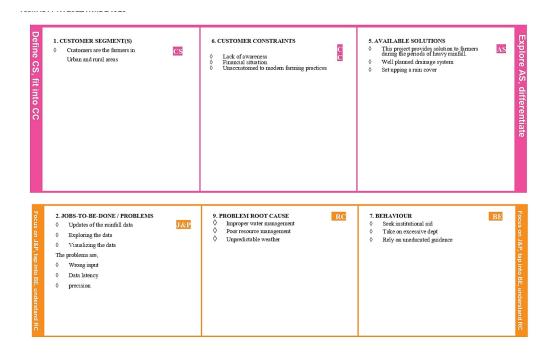
· Helps in producing healthy crops and good fields

Business Model (Revenue Model):

 This comparative study is conducted concentrating on the following aspects: modeling inputs, Visualizing the data, modeling methods, and pre-processing techniques. The results provide a comparison of various evaluation metrics of these machine learning techniques and their reliability to predict rainfall by analyzing the weather data. We will be using classification algorithms such as Decision tree, Random forest, KNN, and xgboost.

## Scalability of the Solution:

- When we predict rainfall correctly, it helps growth of crop and yielding will be better.
- d. Problem Solution fit





# 4. REQUIREMENT ANALYSIS

## a. Functional requirement

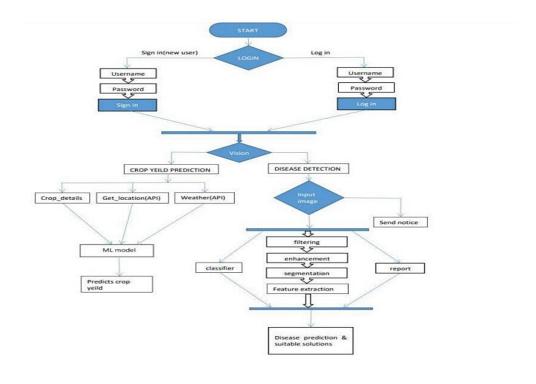
	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR No.		
FR-1	Import necessary packages	Importing packages like NumPy, pandas, seaborn, etc
FR-2	Download and load dataset	Download the dataset Load the Appropriate dataset
FR-3	Pre-processing of data	Making data suitable for building a good model
FR-4	Building Machine learning mode C	Choose the best algorithm. Check for the best optimised result
FR-5	Train the data	Train the model using training data.

# b. Non-Functional requirements

FR No	Non-Functional Requirement	Description
NFR-	Usability	The system should be easy to use
NFR- 2	Security	Security is given over the model, so the user can use this with full trust. The system should protect the data and information related to the farms
NFR- 3	Reliability	Good connectivity and a supporting device. The system should be reliable and not crash when using it
NFR-	Performance	The system should output results of different inputs in a reasonable time.
NFR- 5	Availability	Any person can use this and this is an open-source model
NFR-	Scalability	Farmers, Vegetable sellers, citizens can use this, prediction of data is accurate

# 5. PROJECT DESIGN

a. Data Flow Diagrams



# b. Solution & Technical Architecture

S.No	Component	Description	Technology
1.	User Interface	The user interacts with the	HTML, CSS,python,
		application through a webUI	Flask
		and a chatbot	
2.	Application Logic-1	Logic for registration Registration	Python
3.	Application Logic-2	Logic for login to the application	Python
			•
4.	Application Logic-3	Integrating machine	Flask
		learning model and	
		thewebpage	

5.	Database	Numeric data	MySQL
6.	File Storage	Tostore files suchas prediction report	Local Filesystem
7.	External API	Allows developers access to critical forecasts, alerts, and observations, alongwith other weatherdata	IBM WeatherAPI, etc.
8.	Machine Learning Model	Predictive modeling is a statistical technique using machine learning and data mining to predict and forecast likely future outcomes with the aid ofhistorical and existing data	Predictive modeling
9.	Infrastructure (Server )	Application Deployment on Local System Local Server Configuration: built-in flask web server	Flask web server

## c. User Stories

User Type	Functional Requireme nt (Epic)	User Story Numb er	User Story <i>l</i> Task	Acceptance criteria	Priority	Relea se
Customer (Mobileuse r)	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 confirmati on email oncel have registered for the application As a user, I can register for the application through Facebook	I can receive confirmation email &click confirm  I can register & access the dashboard with Facebook Login	High	Sprint-1
	USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
Dashboard	USN-6	As a user, I can view the details about the page and navigate throughthe entire pages	I can navigate through the pages.	Medi um	Sprint-1
Prediction	USN-7	User can searchfor the area/ place wherethe user wantsto know the prediction	Searchingf or the regionwith in INDIA onlybe accepted	High	Sprint-1

			of rainfall			
		USN-8	The prediction or analysis for the desired region for the future or past events		High	Sprint-1
		USN-9	respectively User can see the visualization of the rainfall data for the specific region in INDIA for a specified timeperiod		High	Sprint-1
	News	USN-10	User can view the latest news articles related to agriculture.	I can view the news articles.	Medium	Sprint-2
Customer (Webuser)	Support	USN-11	User can ask queries aboutthe system.	I can rectifymy doubts	High	Sprint-3

Customer	USN-12	The team	High	Sprint-3
Care		must		
Executive		analyse all		
		the queries		
		and debug		
		it in the		
		next update		

# 6. PROJECT PLANNING & SCHEDULING

# a. Sprint Planning & Estimation

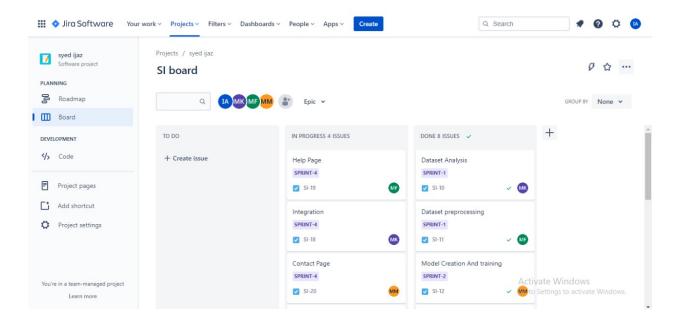
Spri nt	Functional Requirement (Epic)	User Story Numb	User Story / Task	Story Poin ts	Priority	Team Membe rs
Sprin t-1	Dataset Analysis	USN-1	Download theweather dataset and analyze it.	1	High	Ismail Kafil
Sprin t-1	Dataset preprocessi ng	USN-2	Examine the dataset and perform preprocessi ng steps	1	Medium	Fasehiullah
Sprin t-2	Model Creation and Training	USN-3	Create a modelfrom the training data	2	Low	Ismail K M
Sprin t-2	Registration	USN-4	As a user,I can register for the application.	2	Low	ljaz z
Sprin t-2	Login	USN-5	As a user, I can log into the application by entering email& password	1	High	Ismail Kafil
Sprin t-3	Dashboard	USN-6	As a user, once I log in, I can view the Rainfall Prediction page	1	High	Fasehiullah
Sprin t-3	Predictor	USN-7	As a user, I can specify all the values for prediction andget accurate results	1	High	Ismail K M

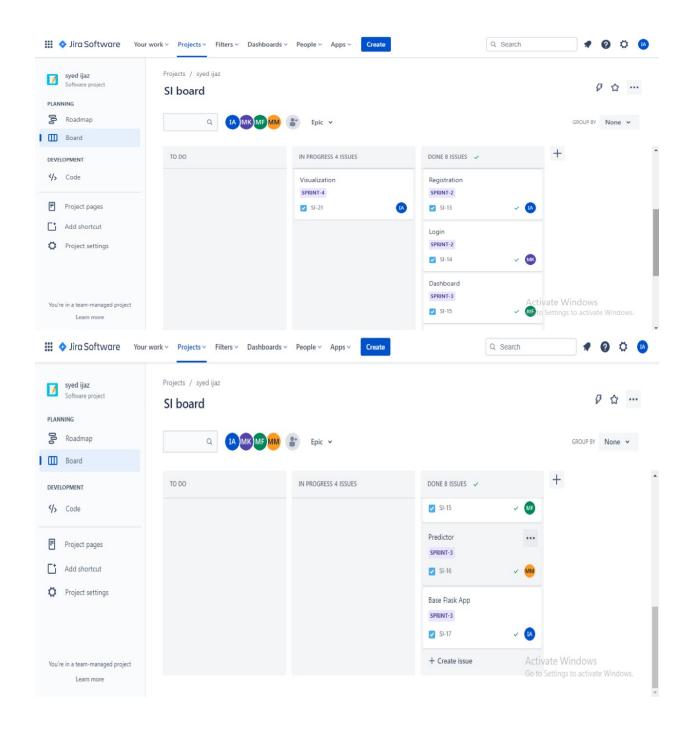
Sprin	Base Flask App	USN-8	Integrate Flask and the	2	High	ljaz z
t-3			builtmodel			
Sprin	Integration	USN-9	Integrate the app on	2	High	Ismail Kafil
t-4			IBMcloud			
Sprin	Help page	USN-10	As a user, I can get	1	Medi	Fasehiullah
t-4			directions on how to		um	
			use the predictor to			
			gain valuable insights			

# b. Sprint Delivery Schedule

Sprint	Functional	User	User Story /	Story	Priority	Team Members
	Requirement	Story	Task	Points		
	(Epic)	Number				
Sprint-	Contact page	USN-11	As a user, I	1	Medium	Ismail K M
4			can get my			
			queries			
			clarified by			
			the admin			
Sprint-	Visualization	USN-12	As a user,I can	2	Medium	ljaz z
4			visualize the			
			datausing			
			various plots			

# c.Reports from JIRA





				Т	
v 🖸	SI-1 Sprint-1		DONE		
	✓ SI-10 Dataset Analysis	DONE I	монамм		
	✓ SI-11 Dataset preproce	ssing DONE	монам		
	SI-2 Sprint-2				
	✓ SI-12 Model Creation A	nd tra DONE	монам		
	✓ SI-13 Registration	DONE	IJAZ AHA		
	☑ <del>SI-14</del> Login	DONE	монамм		
v <b>5</b>	SI-3 Sprint-3				
	✓ SI-15 Dashboard	DONE	монам		
	SI-16 Predictor	DONE	монам		
	SI-17 Base Flask App	DONE	IJAZ AHA		
· [7	SI-4 Sprint-4	i.			
	☑ SI-19 Help Page	IN PROGRESS	монам	1	
	☑ SI-18 Integration	IN PROGRESS	монамм		
	SI-20 Contact Page	IN PROGRESS	монам		
	SI-21 Visualization	IN PROGRESS	IJAZ AHA		

## 7. **RESULTS**

- a. Performance Metrics
- b. USER ACCEPTANCE TESTING

# **Purpose of Document**

The purpose of this documentis to briefly explain the test coverageand openissues of the Exploratory Analysis of Rainfall Data in india for Agriculture project at the time of the release to User Acceptance Testing (UAT).

# **Defect Analysis**

This report showsthenumber of resolvedor closed bugs at each severity level, and how they were resolved

Resoluti on	Severi ty1	Severi ty2	Severi ty3	Severi ty4	Subtot al
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduc ed	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	24	14	13	26	7 7

## 8. ADVANTAGES & DISADVANTAGES

## **ADVANTAGES:**

- -Preserves the sequence of dry/wet days , It accounts for different correction different time windows.
  - -It allows for distinct corrections between mean and variance.
- -The frequency of preciption is corrected, no theoretical distribution is assumed.

## **DISADVANTAGES:**

- -It only corrects the mean precipitation.
- -Does not account for changes in the length of dry/wet spells.
- -Requires large computation time and data preparation.

#### 11.CONCLUSION

Rainfall Prediction is the application area of data science and machine learning to predict the state of the atmosphere. It is important to predict the rainfall intensity for effective use of water resources and crop production to reduce mortality due to flood and any disease caused by rain.

#### 12.FUTURE SCOPE

Predicting weather accurately doesn't just help our daily lives but has deeperimpact for food security and disaster management. Good news for monsoon-dependent Indiais that we are gettingbetter at predicting. New technologies, such as Internetof Things (IoT) and Artificial Intelligence (AI) are helpingmeteorological experts to give better information to predict agricultural output and natural disasters.

#### 9. APPENDIX

Source Code python code

import pickle import re import sqlite3 as sql import time import ibm\_db

import joblib import matplotlib import matplotlib.pyplot as plt import numpy as np import pandas

```
import requests
from flask import Flask, render_template, request
API_KEY = "qJSD5R0m29i2iJGncQiHjCmDMgTx_563xkzVMRZ3Wvw3"
token_response = requests.post('https://iam.cloud.ibm.com/identity/token',
data={"apikey":
API_KEY, "grant_type": 'urn:ibm:params:oauth:grant-type:apikey'})
mltoken = token_response.json()["access_token"]
header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}
conn = ibm_db.connect("DATABASE=bludb;HOSTNAME=125f9f61-9715-46f9-9399-
c8177b21803b.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:31048;PORT=31498
;SECURITY=SSL;SSLServerCertificate=DigiCertGlobalRootCA.crt;UID=wcg74792;PWD=N
ogApdfJJmMbemby",",")
app = Flask(__name__)
model=pickle.load(open('rainfall.pkl','rb'))
scale=pickle.load(open('scale.pkl','rb'))
@app.route('/signin',methods =['GET', 'POST'])
def signin():
  global userid
  msg = "
  if request.method == 'POST':
    un = request.form['username']
    pd = request.form['password']
    sql = "SELECT * FROM register WHERE username =? AND password=?"
    stmt = ibm_db.prepare(conn, sql)
    ibm_db.bind_param(stmt,1,un)
    ibm_db.bind_param(stmt,2,pd)
    ibm_db.execute(stmt)
    account = ibm_db.fetch_assoc(stmt)
    print (account)
    if account:
      session['loggedin'] = True
      session['id'] = account['USERNAME']
      userid= account['USERNAME']
      session['username'] = account['USERNAME']
      msg = 'Logged in successfully!'
```

```
return render_template('dashboard.html', msg = msg)
    else:
      msg = 'Incorrect username / password !'
  return render_template('signin.html', msg = msg)
@app.route('/signup', methods=['POST','GET'])
def signup():
  msg="
  if request.method == "POST":
    username=request.form['username']
    email=request.form['email']
    pw=request.form['password']
    sgl='SELECT * FROM register WHERE email =?'
    stmt = ibm_db.prepare(conn, sql)
    ibm_db.bind_param(stmt,1,email)
    ibm_db.execute(stmt)
    acnt=ibm_db.fetch_assoc(stmt)
    print(acnt)
    if acnt:
      msg='Account already exits!!'
    elif not re.match(r'[^@]+@[^@]+\.[^@]+', email):
      msg='Please enter the avalid email address'
    elif not re.match(r'[A-Za-z0-9]+', username):
      msg='name must contain only character and number'
    else:
      insert_sql='INSERT INTO register VALUES (?,?,?)'
      pstmt=ibm_db.prepare(conn, insert_sql)
      ibm_db.bind_param(pstmt,1,username)
      ibm_db.bind_param(pstmt,2,email)
      ibm_db.bind_param(pstmt,3,pw)
      ibm_db.execute(pstmt)
      msg='You have successfully registered click signin!!'
      return render_template("signin.html")
```

```
elif request.method == 'POST':
    msg="fill out the form first!"
  return render_template("signup.html",msg=msg)
@app.route('/')
def home():
  return render_template('index.html')
@app.route('/predict',methods=["POST","GET"])
def predict():
  input_feature=[x for x in request.form.values()]
  feature_values=[np.array(input_feature)]
names=[['Location','MinTemp','MaxTemp','Rainfall','WindGustDir','WindGustSpeed','WindDir
9am', 'WindDir3pm', 'WindSpeed9am', 'WindSpeed3pm', 'Humidity9am', 'Humidity3pm', 'Pressu
re9am','Pressure3pm','Cloud9am','Cloud3pm','Temp9am','Temp3pm','RainToday']]
  response_scoring = requests.post('https://us-
south.ml.cloud.ibm.com/ml/v4/deployments/cae560db-c5dc-4eba-b799-
99e910c25da3/predictions?version=2022-11-04', json=names,
  headers={'Authorization': 'Bearer ' + mltoken})
  print("Scoring response")
  print(response_scoring.json())
  data=pandas.DataFrame(feature_values,columns=names)
  data=scale.fit_transform(data)
  data=pandas.DataFrame(data,columns=names)
  prediction =model.predict(data)
  pred_prob=model.predict_proba(data)
  print(prediction)
  if prediction == "Yes":
    return render_template("chance.html")
  else:
    return render_template("nochance.html")
if __name__ == "__main__":
 app.run(debug=True)
HTML CODE
<!DOCTYPE html>
<html>
```

```
<head>
  <meta charset="UTF-8">
  <title>Rainfall prediction</title>
</head>
<body>
  <div class="login">
    <center>
      <h1>Rainfall Prediction</h1>
    </center>
    <h1>Please enter the following details<h1>
        <form action="{{url_for('predict')}}" method="post">
          <label for="Location">Location:</label>
          <select id="Location" name="Location">
            <option value=albury>albury</option>
            <option value=4>badgery</option>
            <option value=10>cobar</option>
            <option value=11>coff</option>
            <option value=21>moree
            <option value=24>newcastle</option>
            <option value=26>norah
            <option value=27>pen</option>
            <option value=30>a</option>
            <option value=34>a</option>
            <option value=37>a</option>
            <option value=38>a</option>
            <option value=42>a</option>
            <option value=45>a</option>
            <option value=47>b</option>
            <option value=9>c</option>
            <option value=40>v</option>
            <option value=23>v</option>
            <option value=5>n</option>
            <option value=6>m</option>
            <option value=35>m</option>
```

```
<option value=19>m</option>
            <option value=18>m</option>
            <option value=20>m</option>
            <option value=25>m</option>
            <option value=33>m</option>
            <option value=44>m</option>
            <option value=5>m</option>
            <option value=5>m</option>
          </select>&nbsp; &nbsp;
          <label for="MinTemp"></label>
          <input type="float" id="MinTemp" name="MinTemp" placeholder="MinTemp">
          <label for="MaxTemp"></label>
          <input type="float" id="MaxTemp" name="MaxTemp"
placeholder="MaxTemp">
          <label for="Rainfall"></label>
          <input type="float" id="Rainfall" name="Rainfall" placeholder="Rainfall">
          <label for="Windspeed9am"></label>
          <input type="float" id="Windspeed9am" name="Windspeed9am"
placeholder="Windspeed9am">
          <label for="Windgustspeed"></label>
          <input type="float" id="Windgustspeed" name="Windgustspeed"
placeholder="Windgustspeed">
          <label for="Windspeed3pm"></label>
          <input type="float" id="Windspeed3pm" name="Windspeed3pm"
placeholder="Windspeed3pm">
          <label for="Humidity9am"></label>
          <input type="float" id="Humidity9am" name="Humidity9am"
placeholder="Humidity9am">
          <label for="Humidity3pm"></label>
          <input type="float" id="Humidity3pm" name="Humidity3pm"
placeholder="Humidity3pm">
          <label for="Pressure9am"></label>
          <input type="float" id="Pressure9am" name="Pressure9am"
placeholder="Pressure9am">
          <label for="Pressure3pm"></label>
          <input type="float" id="Pressure9pm" name="Pressure3pm"</pre>
placeholder="Pressure3pm">
```

```
<label for="Temp9am"></label>
          <label for="Cloud9am"></label>
          <input type="float" id="Cloud9am" name="Cloud9am"
placeholder="Cloud9am">
          <label for="Cloud3am"></label>
          <input type="float" id="Cloud3am" name="Cloud3am"</pre>
placeholder="Cloud3am">
          <input type="float" id="Temp9am" name="Temp9am"</pre>
placeholder="Temp9am">
          <label for="Temp3pm"></label>
          <input type="float" id="Temp3pm" name="Temp3pm"</pre>
placeholder="Temp3pm">
          <label for="RainToday">RainToday:</label>
          <select id="Rain" name="RainToday">
            <option value=2>YES</option>
            <option value=4>NO</option>
          </select>
          <label for="Windgustdir">Windgustdir:</label>
          <select id="Windgustdir" name="Windgustdir">
            <option value=2>W</option>
          </select>
          <label for="Winddir9am">Winddir9am:</label>
          <select id="Winddir9am" name="Winddir9am">
            <option value=2>W</option>
          </select>
          <label for="Winddir3pm">Winddir3pm:</label>
          <select id="Winddir3pm" name="Winddir3pm">
            <option value=2>W</option>
          </select>
```

```
<br/>
```

GitHub & Project Demo Link

GitHub link: https://github.com/IBM-EPBL/IBM-Project-42195-1660655698.git