Exploratory Analysis of Rainfall Data in India for Agriculture

Machine Learning

Submitted

by

ROHITH T

YUVA PRAKASH M

RITHIK M R

SANJAY SELVAN R

TEAM ID - PNT2022TMID00722

Department of Computer Science and Engineering ,

PANIMALAR ENGINEERING COLLEGE

Faculty Mentor: VIJAYALAKSHMI

Industrial Mentor: Lalitha Gayathri

1. INTRODUCTION:

1.1. Project Overview:

With the changes in the climatic conditions and irregular pattern of weather conditions, predicting their occurrence for preventing lifeloss to humankind and environment is an utmost societal needed problem of the society.Drastic changes in climate have occur ed over the past years and with change in revolution proper preventive measures are needed.Heavy rainfa l can lead to floods.Flash floods are catastrophic.Climate change is increasing the frequency, intensity and magnitude of disasters, leading to a higher number of deaths and injuries, as we l as increased property and economic losses. In the past 20 years, 90% of major disasters have been caused by weather-related events such as heatwaves, storms, floods and droughts, according to the UN Of icefor Disaster Risk Reduction (UNISDR).Natural disasters are increasing in strength and frequency. Shifting weather pat erns make predictions and emergency planning difficult. Hence , we focus on the ef ective prediction of the probability of the flood occuring in a particular region and recommending an evacuation area nearby by performing an exploratory analysis of the data co lected.

1.2. Purpose

To design a disaster management system by forecasting a flood event to control flood risk by recommending an evacuation area from flood hazard areas which ultimately helps to manage theenvironment and water resource system.This also serves a purpose of the Early warning system by training a model and selecting the best prediction algorithm among the classifiers.The occurrence of flash floods can cause catastrophic damage to the society. They first mainly affect the people livingnear to the riverbeds.Evacuating them from the hazard areas and providing them the shelter they needed.With the irregular change in climate patterns, it's been difficult to predict the occur ence of floods using traditional methods leading to massive destruction. Thus to cope with flash floods and to handle critical situations new methodologies are invented

to overcome such difficulties.

Technology has to be more aware to reduce the loss that a flash flood would make. In the modernizing era, it's made even easier to predict the occurrence of floods and recommend nearby evacuation areas. Hazard areas that are prone to destruction and devastating loss are monitored regularly and the rainfa l readings are co lected, integrated from multiple resources, curated, mined, analyzed and prediction is done over pat erns. With the prediction, recommendation areas are listedfor the society.Early warning systems are climate change adaptation measures that use integrated

communication systems to help communities prepare for dangerous climate-related events. An earlywarning system ' s success saves lives and jobs, land and infrastructure, and supports long-term sustainability.

2. LITERATURE SURVEY:

2.1. Existing Solutions

This paper deals with the idea of predicting floods using the algorithm Artificial neural networks(ANN) and with the support of the Internet of Things. This system looks after the humidity, temperature, pressure, rainfall, and river water level periodica ly to the temporal cor elative information for flood prediction analysis. Flood data is dynamic and non-linear in nature. The sensors read the data and inform the system. With those values, the prediction is doneand the decision is taken on the occur ence of a flood.[1]

Precipitation in any form such as snow, rain or hail can af ect the routine of the society.Therefore predicting the occur ence of rainfa l beforehand and warning the society about the day ' s condition can be helpful in a lot more ways.Providing accurate results for forecasting rainfall has been a major issue with the drastic change in climatic conditions.Using a fusion of machine learning techniques can help in providing much more accurate results about the occurrenceof rainfa l.Four supervised learning algorithms has been used to get out the accurate results for prediction.The four effective algorithms that results in accurate prediction are decision tree, Naïve Bayes, K-nearest neighbors, and support vector machines.The effectiveness of the algorithm is checked by incorporating the technology known as fuz y logic.A twelve year historical weather data of city lahore is considered for training ,validating and for testing .In such a way that this fusion model outperformed other existing models.[2]

The drastic change in climatic conditions has caused severe impact on the society and environment. A country's economic and financial condition is mainly dependent on the country ' sagriculture. Farming and agriculture are considered to be India ' s backbone of economic

conditions. In such a way any climate change af ects the agricultural development which in directlyaf ects the economic andfinancial conditions of the country. Therefore predicting the occurrence ofrainfa l is one of the most important aspect for the safety of the society as well as the country withits economic conditions. Loss in agriculture could lead to famine and create a huge economic crisis. Prediction made should be to the point. The traditional methods of predicting rainfa l have

gone out of control with the drastic change in climatic conditions and development of the country.With the rise in global warming conditions, rough humidity and change in the oceans predicting rainfall with any modest technologies that results in the precise results is an utmost needof the society.Applying machine learning classification algorithms to predict the accurate results ofrainfall has been implemented.UCI repository dataset has been considered for training,validating and testing.[3]

With nature being unpredictable the intensity of the rainfa l varies according to the climaticconditions and the pressure of the wind. Under such conditions, urban floods can be a great disasterfor society. This paper deals with a classification-based real-time flood prediction model with the support of a numerical analysis model based on hydraulic theory and the required machine learningmodels. The Flood database has been created beforehand with the help of the Environmental Protection Agency-Storm Water Management model and from a two-dimensional inundation model. Using the Latin hypercube sampling and probabilistic neural network are used for categorizing the flood depth data into five categories. This machine learning model is constructed to identify the respective cumulative volume if the observed rainfall data is entered. Therefore a system that's capable of generating a real-time flood map by cumulative volume of each grid to the cumulative volume using linear regression and nonlinear regression. The developed system can predict the rainfa l-induced flooding potential in such a way that reduces the risk due to disaster and minimizes damage to health and properties. Therefore a useful disaster management system has been developed for preventing huge losses due to disasters.[4]

On a high note, research has been continuously carried out on achieving ef icient and accurate prediction technology or systems. With the help of machine learning techniques and algorithms, prediction can be made easy to obtain accurate and earlier results such as making the required ar angements and evacuating people from the hazard areas. Over the two decades, neuralnetworks have shown an extraordinary outcome in predicting the occur ence of floods with the given rainfa l

data providing bet er results and cost-ef ective solutions. This paper is novel in the way of analyzing databases by Multi-layer perceptron classifier to read data such as dynamic identification, deficit treatment, data validation, and data cleaning to be carried across the database.Advancements in every note can provide better results based on the preprocessing of data.[5]

2.2. References

[1] Swapnil Bande, Virendra V. Shete, "Smart flood disaster prediction system using IoT& neuralnetworks", 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)

[2] Atta-ur Rahman,Sagheer Abbas, Mohammed Gollapa li, Rashad Ahmed, Shabib Aftab, MunirAhmad,Muhammad Adnan Khan, Amir Mosav, "Rainfall Prediction System Using Machine Learning Fusion for Smart Cities", 2022 May National Library of Medicine

[3] Vikas Kumar, Vishal Kumar Yadav, Er. Sandeep Dubey, "Rainfa l Prediction using MachineLearning", Ijraset Journal For Research in Applied Science and Engineering Technology, 2022.

[4] Ho Jun Keum, Kun Yeun Han & Hyun Il Kim, "Real-Time Flood Disaster PredictionSystem by Applying Machine Learning Technique", KSCE Journal of Civil Engineering 24, 2835-2848(2020)

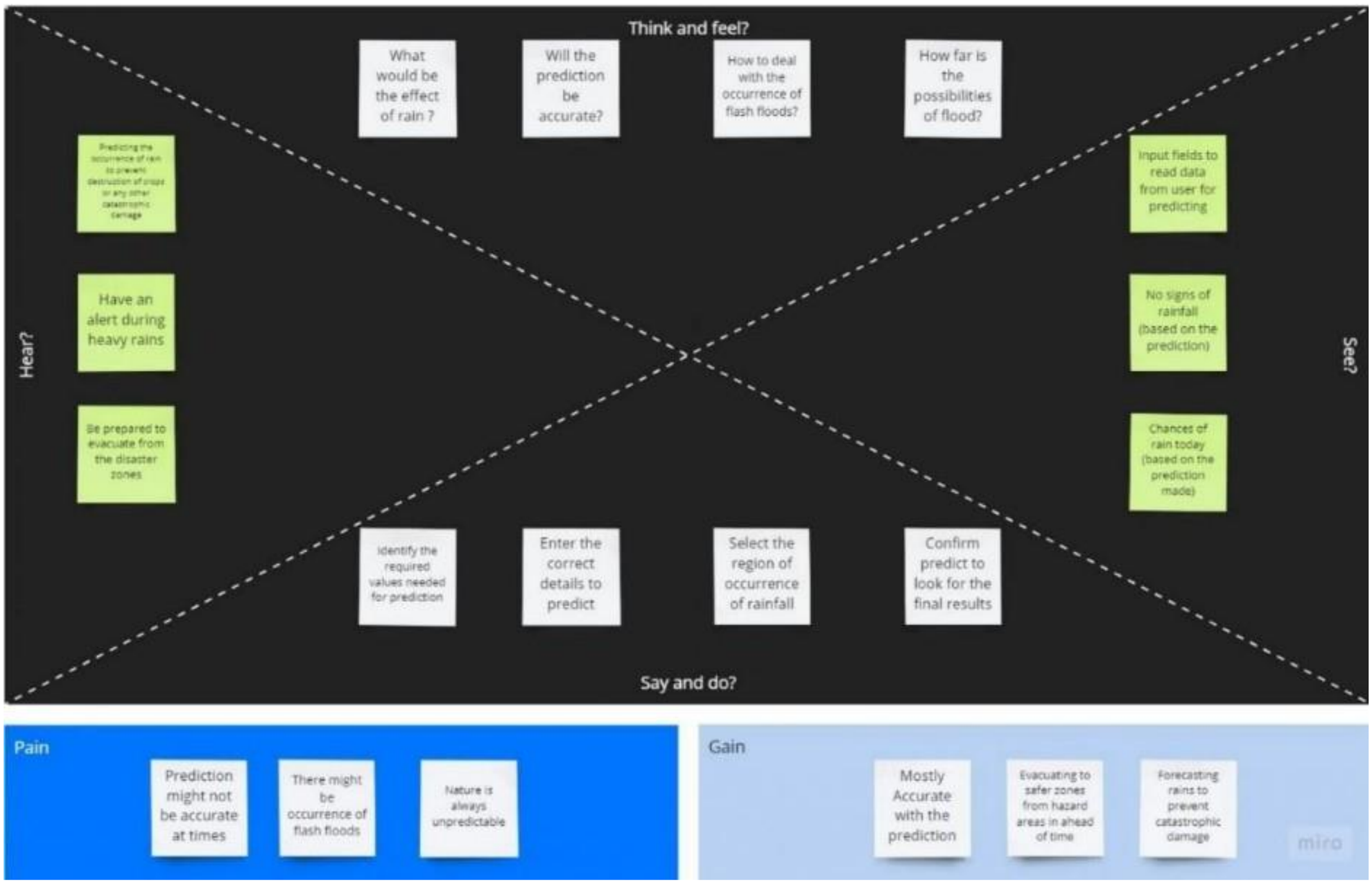
[5] Thegeshwar Sivamoorthy, Asif Mohammed Ansari, Dr. B. Sivakumar, V. Nallarasan, "Flood Prediction Using ML Classification Methods on Rainfa l Data", IJRASET Journal For Research inApplied Science and Engineering Technology

2.3. Problem Statement Definition

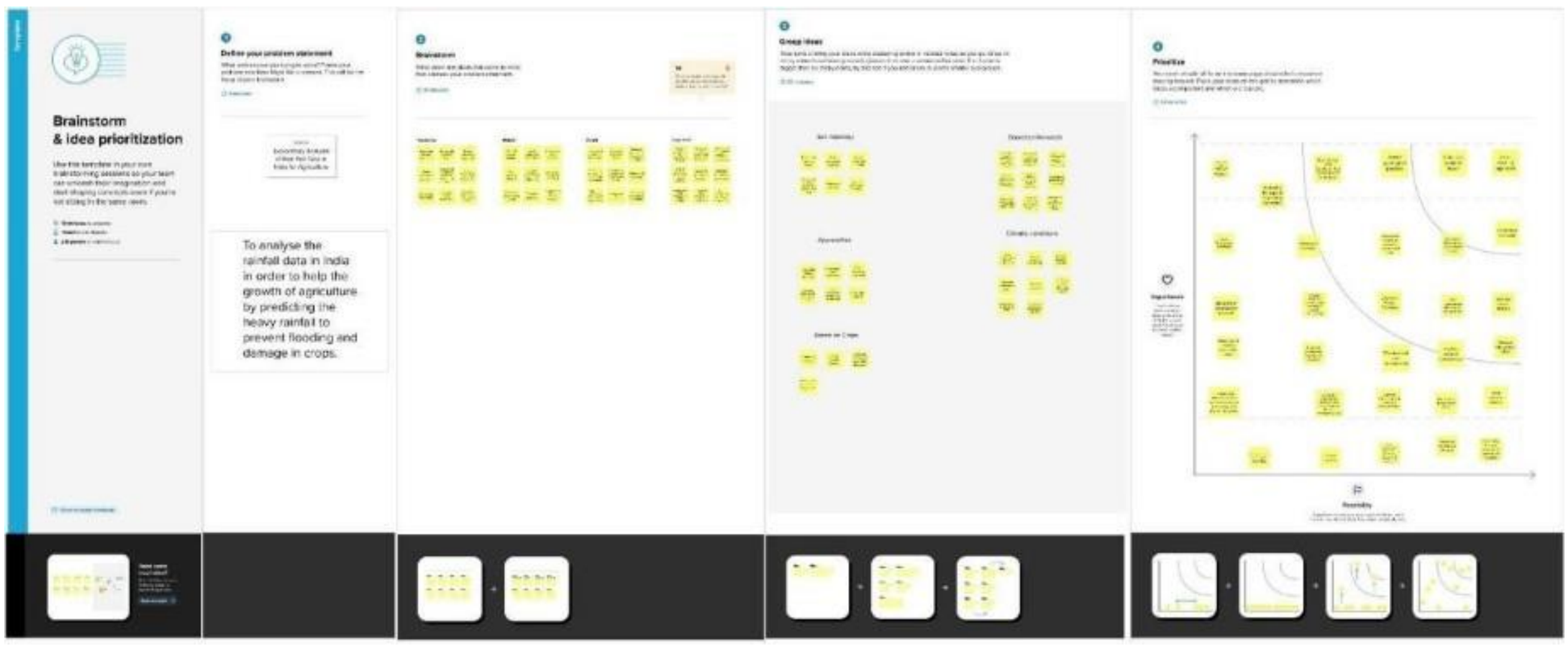
Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	Farmer	Predict the heavy rainfall to take precautionary measures to protect the crops from destruction	Can't predict the heavy rainfall	The climate changes made difficult to make the season as usual	Sad and anxious
PS-2	Department agencies	Make a continuous and good supply of the crops	The crops are destructed and are in shortage.	The heavy rainfall can't be predicted beforehand	Frustrated

3. IDEATION AND PROPOSED SOLUTION:

3.1. Empathy Map Canvas:



3.2. Ideation and Brainstorming:



3.3. Proposed Solution:

S.No.	Parameter	Description
1.	Problem Statement	To analyze the rainfa l data in India in order tohelp the growth of agriculture by predicting theheavy rainfall to prevent flooding and damage in crops.
2.	Idea / Solution description	In our analysis we are trying to understand thebehavior of rainfa l in India over the years, by months and different subdivisions. The trend analysis of Annual rainfall considering India as whole, show that it is important to study subdivision for better forecasting.
3.	Novelty / Uniqueness	Rainfa l prediction is important as heavy rainfallcan lead to many disasters. The prediction helps people to take preventive measures and moreoverthe prediction should be accurate. So, it is important to analyse rainfa l in order to prevent damage in crops.

4.	Social Impact / CustomerSatisfaction	India is an agricultural country and secondary agro based market will be steady with a good monsoon. Thus, the prediction of amount of rainfall can play a major role in agricultural field.
5.	Business Model (RevenueModel)	• To create customer value by satisfying needs of a farmer (i. e.) predicting when the rainfa l is highand providing early warning. • Marketing the product among farmers and farmers associations.
6.	Scalability of the Solution	The Scalability of the project is ensuring that model ' s high-performance computing can support agriculture by delivering more accuratepredictions, using higher resolution and more complex mode ling, greater use of ensembles and vastly increased volumes of data of a l forms.

3.4. Solution Fit

Project Title: Exploratory Analysis of Rainfall data in India for Agriculture **Project Design Phase-I - Solution Fit** **Team ID:** PNT2022TMID21508

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

(1) Farmers and farming industries

(2) Government departments and agencies

(3) Scientific journals.

6. CUSTOMER CONSTRAINTS CC

- Budget
- Network connection in rural areas
- Basic statistical knowledge

5. AVAILABLE SOLUTIONS AS

An Exploratory Study on Occurrence and Impact of Climate Change on Agriculture in Tamil Nadu, India - examine the occurrence of climate change in Tamil- -Nadu, and its impact on rainfall pattern which is a primary constraint for agricultural production

Flood forecasting using Internet of things and artificial neural networks - India is one of the worst flood-affected countries in the world based on the annual rainfall. They use number of IOT and based techniques but the challenge is that no one has attempted the possibility of occurrence of flood rainfall intensity.

Explore AS, differentiate

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

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

- To create customer value by satisfying needs of a farmer (i.e.) predicting when the rainfall is high and providing early warning.
- Marketing the product among farmers and farmers associations.

9. PROBLEM ROOT CAUSE RC

An important aspect to be understood regarding the relationship between rainfall and agriculture is that rainfall is the major factor in the growth and production of food crops both at the germination and fruit development stage. But with a change in the world's climate, temperatures will rise and rainfall will increase in some places. In other places, rainfall will decrease.As a result of global warming, the world's climate is changing and its effect is being felt the world over. And one of the most important parameters of climate is rainfall. So inorder to find an effective solution for finding the right time for the cultivation of crops ,an algorithm is needed to predict the rainfall rate and derive an useful model out of it.

7. BEHAVIOUR BE

The model's high-performance computing can support agriculture by delivering more accurate predictions, using higher resolution and more complex modelling, greater use of ensembles and vastly increased volumes of data of all forms.

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

3. TRIGGERS TR

By monitoring the data and metrics mentioned above, farmers find a wealth of benefits, including higher production quality and quantity. Other benefits include: Save costs: smart farming leaders to lower costs on labor, water, and nutrients for crops. Save water resources: knowing the exact rainfall for each crop can help optimize watering, thus preventing overwatering, which can impact not only crop health, but the environment.

4. EMOTIONS: BEFORE / AFTER EM

Anxious > Clear and calm

(i.e) the farmers can be pre-prepared to face the heavy rainfall and can reduce the destruction of crops.

10. YOUR SOLUTION SL

In our analysis we are trying to understand the behavior of rainfall in India over the years, by months and different subdivisions. The trend analysis of Annual rainfall considering India as whole, shows that it is important to study subdivision for better forecasting. We can also extend the scope of the project by predicting the probability of flood.

8. CHANNELS of BEHAVIOUR CH

8.1 **ONLINE**

The prediction is done online through a server-client model.

8.2 **OFFLINE**

The predicted result can be downloaded and made offline.

4. REQUIREMENT ANALYSIS:

4.1. Functional Requirements

S.No	Component	Description	Technology
1.	User Interface	The user interacts with the application through a webUI and a chatbot	HTML, CSS, python, Flask
2.	Application Logic-1	Logic for registration Registration	Python
3.	Application Logic-2	Logic for login to theapplication	Python
4.	Application Logic-3	Integrating machine learning model and thewebpage	Flask
5.	Database	Numeric data	MySQL
6.	File Storage	To store files such aspredictionreport	Local Filesystem
7.	External API	Allows developers access to critical forecasts, alerts, and observations, along with other weather data.	IBM Weather API

8.	Machine LearningModel	Predictive modeling is a statistical technique using machine learning and datamining to predict and forecast likely future outcomes with the aid of historical and existing data	Predictive modeling
9.	Infrastructure(Server)	Application Deployment on LocalSystem Local Server Configuration: built-in flask webserver	Flask web server

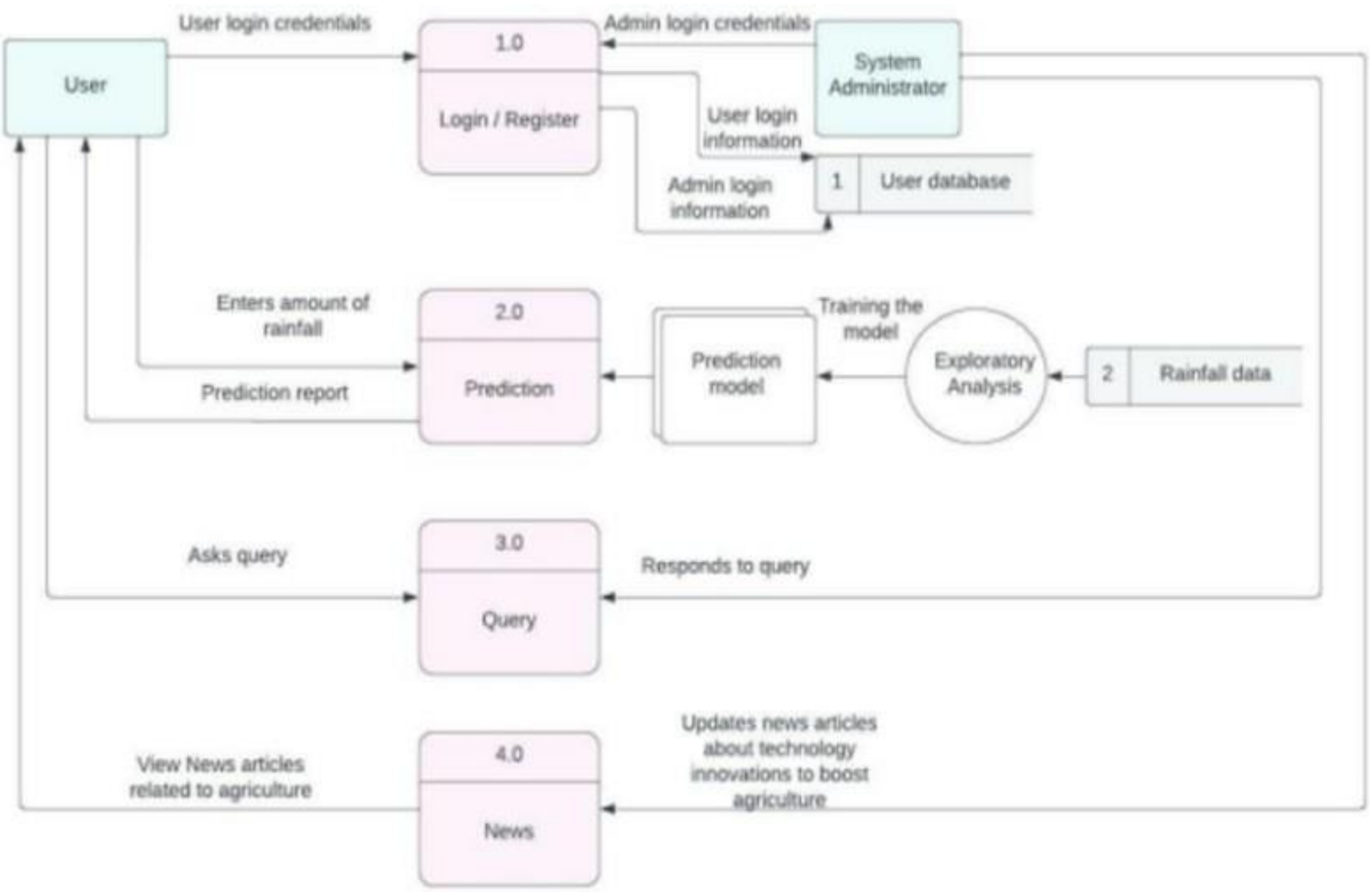
4.2. Non Functional Requirements

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Flask	Micro web framework written in Python
2.	Security Implementations	Basic HTTP authentication, Session based authentication, User Registration, LoginTracking	Flask Security

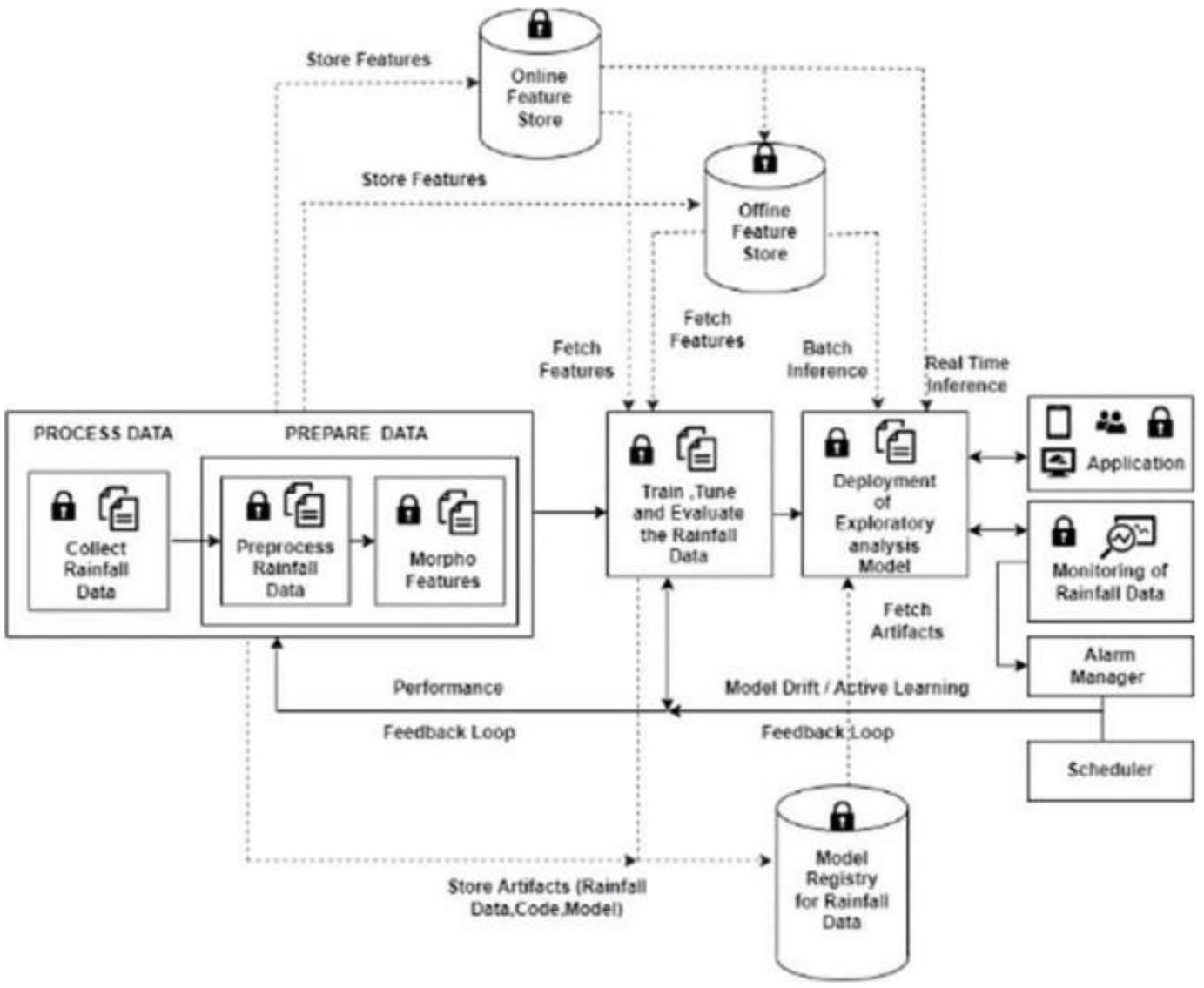
3.	Scalable Architecture	Size is everything, and Flask' s status as a microframework means that you can use it to grow a tech project such as a web app incredibly quickly. Its simplicity of use and few dependencies enable it to run smoothly even as it scales up and up.	Flask
4.	Availability	Higher compatibility with latest technologies and allowscustomization	Flask
5.	Performance	• Integrated support for unit testing. • RESTful request dispatching. • Uses Jinja templating. • Supportfor secure cookies (client side sessions) • 100% WSGI 1.0 compliant.	Flask

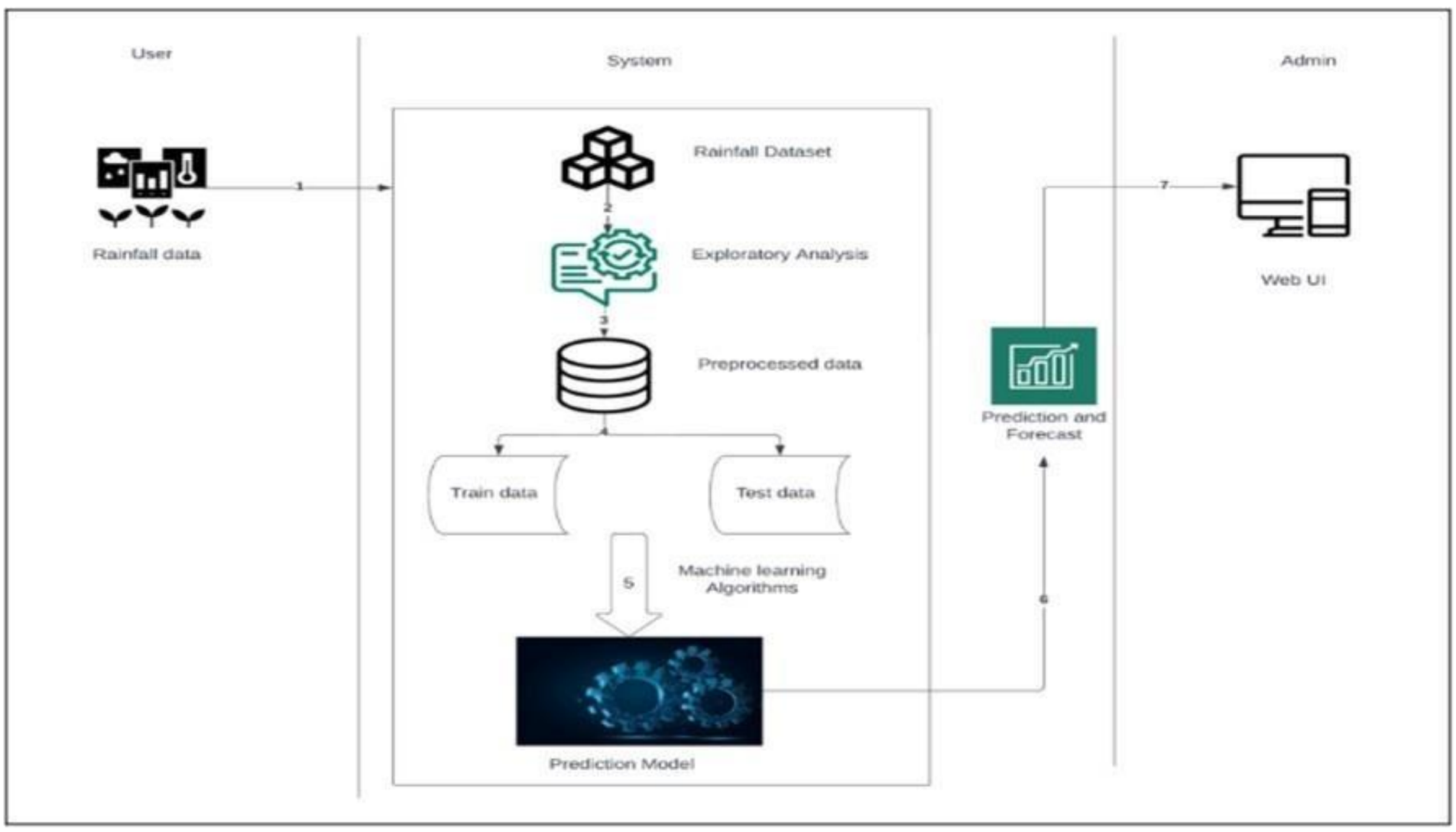
5. PROJECT DESIGN

5.1. Data Flow Diagram:



5.2. Solution And Technical Architecture:





5.3. User Stories:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (web 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
	Registration	USN-2	As a user,I will receive confirmation email once I have registeredfor the application	I can receive confirmation email & click confirm	Medium	Sprint-2
	Login	USN-3	As a user,I can log into the application by entering	I can accessthe system	High	Sprint-1

			email & passwor d			
	Dashboard	USN-4	As a user, I can view the details about the system and can navigate	I can navig ate throu g h pages	High	Sprint-3

			through thepages.			
	Prediction	USN-5	As a user, I can enter therainfall amount and get the prediction results	I can get the predic tion result	High	Sprint-4
	News	USN-6	As a user, I can view latest news article s relate d to agricul ture	I can viewthe articles	Medium	Sprint-3
Cust omer Care Exec utive	Contact	USN-7	As a user,I can ask queries regardin gthe system	I can clarifymy doubts	High	Sprint-3
	Chat bot	USN-8	As a user, Ican interact with chatbot to ask queries	I can getmy queries clear instantl y	Low	Sprint-4
Admini strator	Login	USN-9	As a user, I can regist er for the application by enterin g my email,	I can view and update the system	High	Sprint-1

			password, and confirming my password.			
	Prediction	USN-10	As a user, I can see the prediction result from the model trained by the	I can train the predicti on model	High	Sprint-3

			system administrator			

6. PROJECT PLANNING & SCHEDULING:

6.1. Planning & Estimation:

Sprint	Total Story Points	Duration	Sprint StartDate	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint ReleaseDate (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

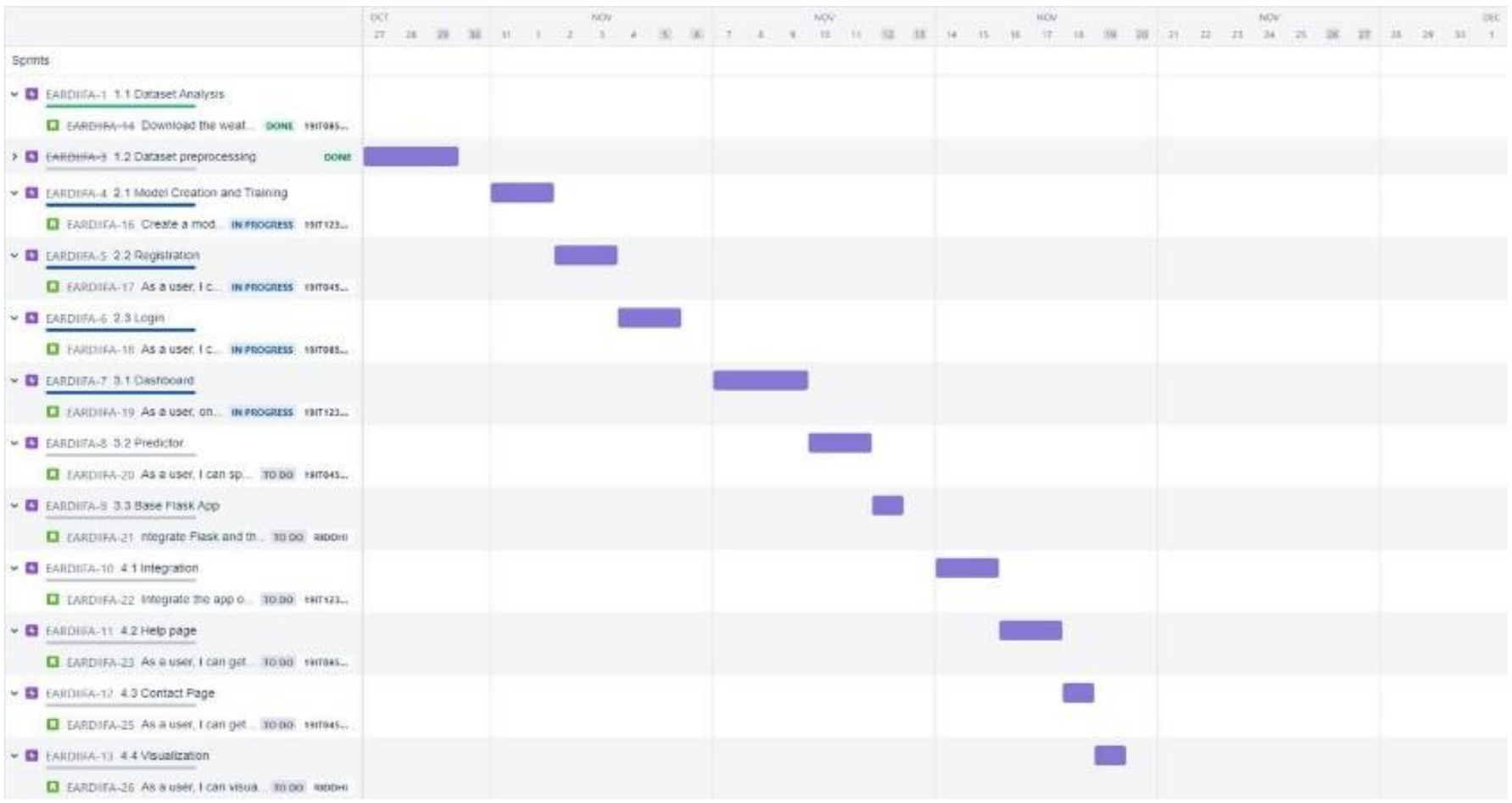
6.2. Sprint Delivery Schedule:

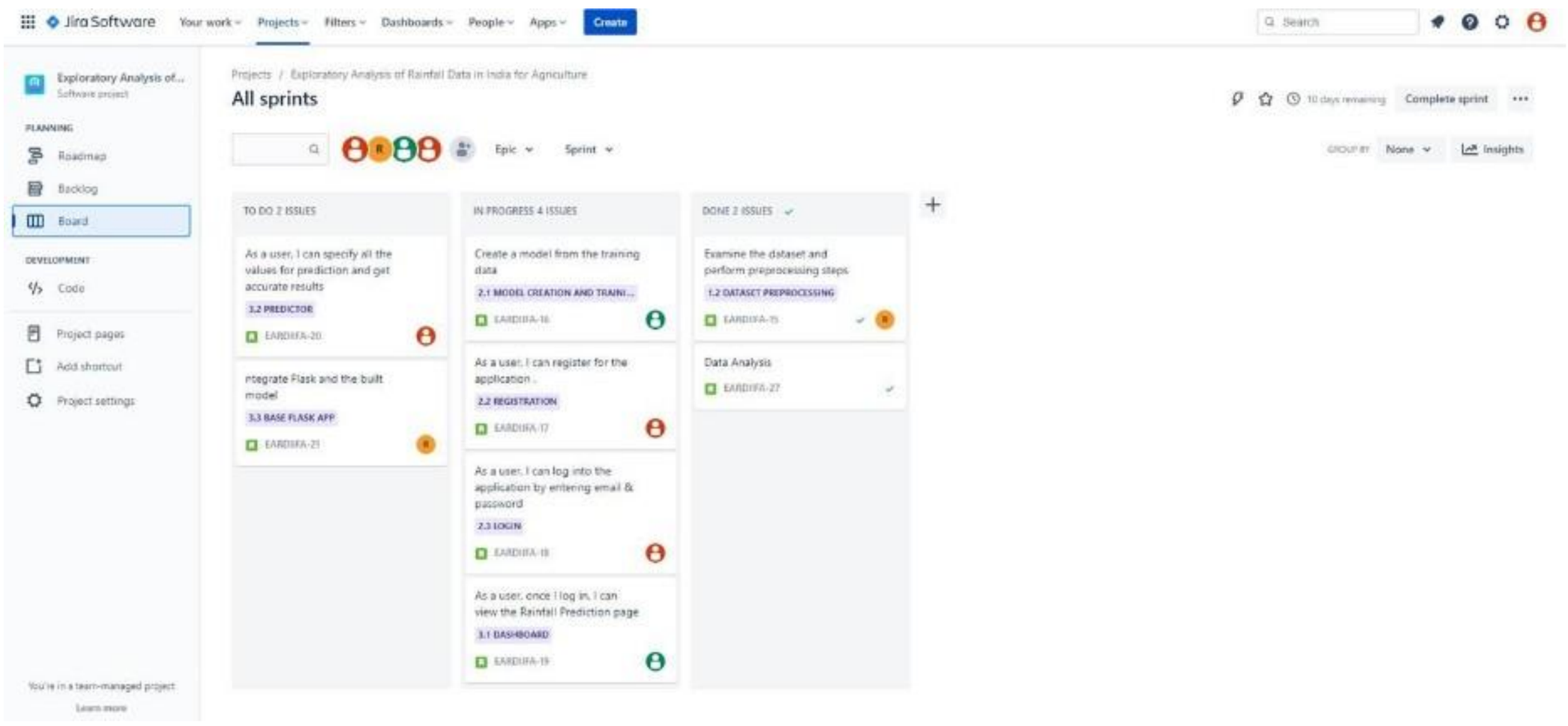
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Dataset Analysis	USN-1	Download the weather datasetand analyze it.	1	High	ROHITH T
	Dataset preprocessing	USN-2	Examine the dataset and performpreprocessing steps	1	High	YUVA PRAKASH M
Sprint-2	Model Creation and Training	USN-3	Create a model from the trainingdata	2	High	RITHIK M R
	Registration	USN-4	As a user, I can register for theapplication .	2	Low	SANJAY SELVAN R
	Login	USN-5	As a user, I can log into the application by entering email & password	1	Low	ROHITH T
	Dashboard	USN-6	As a user, once I log in, I can view the Rainfa l Prediction page	1	High	YUVA PRAKASH M

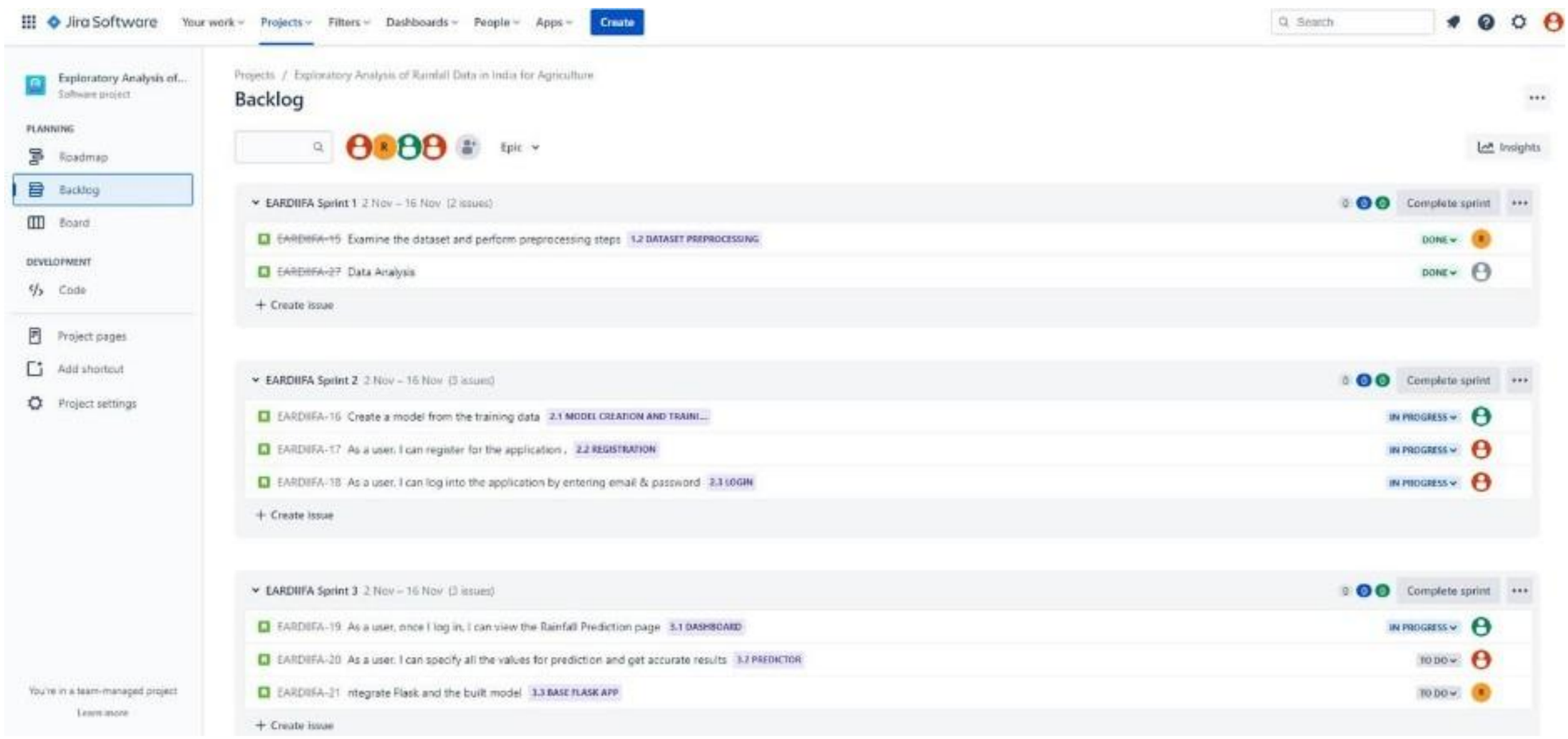
Sprint-3	Predictor	USN-7	As a user, I can specify all the values for prediction and getaccurate results	1	High	RITHICK M R	
	Base Flask App	USN-8	Integrate Flask and the built model	2	High	SANJAY SELVAN R	
	Integration	USN-9	Integrate the app on IBM cloud	2	High	ROHITH T	

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members	
Sprint - 4	Help page	USN-10	As a user, I can get directions onhow to use the predictor to gain valuable insights	1	Medium	ROHITH T	
	Contact page	USN-1	As a user, I can get my queriesclarified by the admin	1	Medium	YUVA PRAKASH M	
	Visualization	USN-12	As a user, I can visualize the datausing various plots	2	Medium	RITHIK M R	

6.3. Reports From Jira







7. CODE AND SOLUTIONING:

7.1. Feature 1: To retrieve information from IBM cloud account using API Key

```
import
requests
import json
API_KEY = "PQBr9MBF7mFuSh2VVLfOE-lilA04VH-h5VEk8EfjFluw"

token_response = requests.post('ht ps://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"]

print("ML Token",mltoken)

header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}
```

7.2. Feature 2: To get predictions when the user enters the values and connecting to thedeployed ML model using scoring end point

```
def predict():
   res = " "
    # If a form is submitted
   if request.method == "POST":
      Location =
      request.form.get('Location')
      MinTemp =
      request.form['MinTemp'] MaxTemp
      = request.form['MaxTemp'] Rainfa l
      = request.form["Rainfall"]
      WindGustSpeed =
      request.form['WindGustSpeed']
      WindSpeed9am =
      request.form['WindSpeed9am']
      WindSpeed3pm =
      request.form['WindSpeed3pm']Humidity9am
      = request.form['Humidity9am'] Humidity3pm
      = request.form['Humidity3pm'] Pressure9am
      = request.form['Pressure9am'] Pressure3pm
      = request.form['Pressure3pm'] Temp9am =
      request.form['Temp9am']
      Temp3pm = request.form['Temp3pm']
      RainToday = request.form.get('RainToday')
      WindGustDir =
      request.form.get('WindGustDir')
```

```
WindDir9am =

request.form.get('WindDir9am')

WindDir3pm =

request.form.get('WindDir3pm')

 new_row =
{'Location':Location,'MinTemp':MinTemp,'MaxTemp':MaxTemp,'Rainfall':Rainfa l,'WindG
ustSp
eed':WindGustSpeed,'WindSpeed9am':WindSpeed9am,'WindSpeed3pm':WindSpeed
3pm,'Hu
midity9am':Humidity9am,'Humidity3pm':Humidity3pm,'Pressure9am':Pressure9am,'Pre
ssure
```

```
3pm':Pressure3pm,'Temp9am':Temp9am,'Temp3pm':Temp3pm,'RainToday':RainToday,'WindG
ustDir':WindGustDir,'WindDir9am':WindDir9am,'WindDir3pm':WindDir3pm}

print(new_row

)new_df =
pd.DataFrame(columns=['Location','MinTemp','MaxTemp','Rainfa l','WindGustSpeed','WindSpe
ed9am','WindSpeed3pm','Humidity9am','Humidity3pm','Pressure9am','Pressure3pm','Temp9a
m','Temp3pm','RainToday','WindGustDir','WindDir9am','WindDir3pm'])

new_df =

new_df.append(new_row,ignore_index=True)

labeled =
new_df[['Location','MinTemp','MaxTemp','Rainfa l','WindGustSpeed','WindSpeed9am',
'WindSp
eed3pm','Humidity9am','Humidity3pm','Pres ure9am','Pres ure3pm','Temp9am','Temp3
pm','R ainToday','WindGustDir','WindDir9am','WindDir3pm']]

X =

labeled.values

print(X)

payload_scoring = {"input_data": [{"field":
[['Location','MinTemp','MaxTemp','Rainfa l','WindGustSpeed','WindSpeed9am','WindSpeed3pm
','Humidity9am','Humidity3pm','Pressure9am','Pressure3pm','Temp9am','Temp3pm','RainyTod
ay','WindGustDir','WindDir9am','WindDir3pm']], "values": X.tolist()}]}

response_scoring =
requests.post('ht ps://us-south.ml.cloud.ibm.com/ml/v4/deployments/73230b85-51ea-
45d b-baa7-e86b5d528fbe/predictions?version=2022-11-14',
json=payload_scoring,headers={'Authorization': 'Bearer ' + mltoken})

print("Scoring response")

predictions =

response_scoring.json()

print(predictions)

output =

predictions['predictions'][0]['values'][0][0]

print(output)

else:

output =

""if output ==
```

```
1:
    return
redirect(url_for('chance'))elif
output == 0:
```

```
        return redirect(url_for('nochance')
    return render_template("index.html", output = res)
```

7.3. Feature 3 : To navigate between pages

```
<div class="navbar">
    <ul>
        <div class="nav"><a href="">HOME</a></div>
        <div class="nav"><a href="{{ url_for('predict') }}">PREDICTOR</a></div>
        <div class="nav"><a href="{{ url_for('help') }}">HELP</a></div>
        <div class="nav"><a href="{{ url_for('contact') }}">CONTACT</a></div>
    </ul>
</div>
```

8. TESTING:

8.1. Test Cases

Test case ID	Feature Type	Component	Test Scenario	Prerequisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Comments	TC for Automation(Y/N)	BUG ID
HomePage_TC_001	UI	Home Page	Verify all the UI elements in Home page rendered properly	HTML	1. Enter URL and click go 2. Verify all the UI elements displayed or not		All the UI elements rendered properly	Working as expected	Pass		N	
HomePage_TC_002	Functional	Home page	Verify the Data Entry page can be reachable.	HTML, CSS	1. click the predict tab in navigation bar. 2. Verify all the UI elements displayed or not.		User should navigate to Predictor page	Working as expected	Pass		N	

Predict_Page_TC_003	UI	Predict Page	Verify all the UI elements in Predict page rendered properly	HTML,CSS	1. Enter URL and click go 2. Verify all the UI elements displayed or not		All the UI elements rendered properly	Working as expected	Pass		N	
PredictPage_TC_004	Functional	Predict Page	Enter all the values and verify the prediction	Flask	1. Enter URL and click go 2. Enter the values for 17 attributes 3. Click Predict	NewCastle 13.4 22.6 0.6 44 21 24 70 78 1007.7 1007.1 34 32 Yes WSW NNW ESE	Redirect to corresponding html page (chance/no chance)	Working as expected	Pass		N	
OutputPage_TC_005	Functional	Chance Page	Verify whether it is redirected to chance page		1. Enter URL and click go 2. Enter the values and click predict button 3. If prediction equals one, chance page is displayed.	Prediction = 1	Redirect to chance page	Working as expected	Pass		N	
OutputPage_TC_006	Functional	No chance Page	Verify whether it is redirected to no chance page		1. Enter URL and click go 2. Enter the values and click predict button 3. If prediction equals zero, no chance page is displayed.	Prediction = 0	Redirect to no chance page	Working as expected	Pass		N	

S.NO	Test Scenerios
1	Verify all the UI elements in Home page rendered properly.
2	Verify the Data Entry page can be reachable.
3	Verify all the UI elements in Predict page rendered properly
4	Enter all the values and verify the prediction
5	Verify whether it is redirected to chance page
6	Verify whether it is redirected to no chance page

8.2. User Acceptance Testing

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of theProject Exploratory Analysis of Rainfall data in India for Agriculture at the time of the release to User Acceptance Testing (UAT).

2. Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and howthey were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	0	0	0	0	0
Duplicate	0	0	0	0	0
External	0	0	0	0	0
Fixed	0	0	0	0	0
Not Reproduced	0	0	0	0	0
Skipped	0	0	0	0	0
Won't Fix	0	0	0	0	0
Totals	0	0	0	0	0

3. Test Case Analysis

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

Section	Total Cases	Not Tested	Fail	Pass
Home Page	2	0	0	2
Predict Page	4	0	0	4

9. RESULTS:

9.1. Performance Metrics

<table>
<tr><th>S.No.</th><th>Parameter</th><th>Values</th><th>Screenshot</th></tr>
<tr><td>1.</td><td>Metrics</td><td>Classification Model: Random Forest

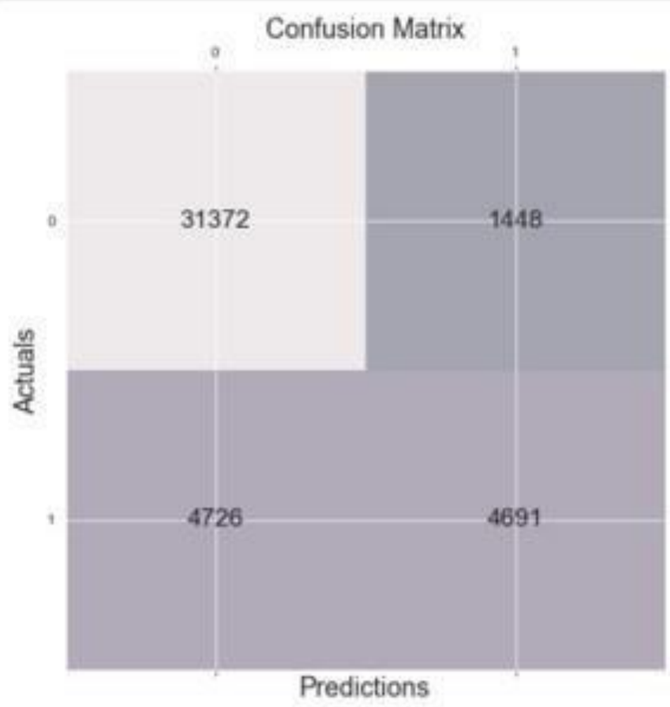
Confusion Matrix –
[[31372 1448]
[4726 4691]]

Accuracy Score-
0.8538248455145963

Classification Report –
Accuracy:
0.8538248455145963
Precision:
0.7641309659553673
Recall:
0.49814165870234683
F1-score:
0.6031113396760092</td><td>Random forest Confusion matrix

conf_matrix = metrics.confusion_matrix(y_test,t1)

fig,ax = plt.subplots(figsize=(7.5,7.5))
ax.matshow(conf_matrix,alpha=0.3)
for i in range(conf_matrix.shape[0]):
for j in range(conf_matrix.shape[1]):
ax.text(x=j, y=i, s=conf_matrix[i,j], va ='center', ha='center',size='xx-larg
plt.xlabel('Predictions',fontsize=18)
plt.ylabel('Actuals',fontsize=18)
plt.title('Confusion Matrix',fontsize=18)
plt.show()

t1 = Rand_forest.predict(X_test_scaled)

print("Rand_forest:",metrics.accuracy_score(y_test,t1))

Rand_forest: 0.8538248455145963</td></tr>
<tr><td></td><td></td><td></td><td>print("*"*10, "Classification Report", "*"*10)

print("-"*30)
print(classification_report(y_test, t1))
print("-"*30)

********** Classification Report **********
----
precision recall f1-score support

0 0.87 0.96 0.91 32820
1 0.76 0.50 0.60 9417

accuracy 0.85 42237
macro avg 0.82 0.73 0.76 42237
weighted avg 0.85 0.85 0.84 42237

------------------------------</td></tr>
</table>

2.	Tune the Model	Hyperparameter Tuning & Validation Method - RandomizedSearchCV	**Hyperparameter Tuning**
			`from sklearn.ensemble import RandomForestRegressor`
			`rf = RandomForestRegressor(random_state = 42)`
			`from pprint import pprint`
			`# Look at parameters used by our current forest`
			`print('Parameters currently in use:\n')`
			`pprint(rf.get_params())`
			Parameters currently in use:
			{'bootstrap': True,
			'ccp_alpha': 0.0,
			'criterion': 'mse',
			'max_depth': None,
			'max_features': 'auto',
			'max_leaf_nodes': None,
			'max_samples': None,
			'min_impurity_decrease': 0.0,
			'min_impurity_split': None,
			'min_samples_leaf': 1,
			'min_samples_split': 2,
			'min_weight_fraction_leaf': 0.0,
			'n_estimators': 100,
			'n_jobs': None,
			'oob_score': False,
			'random_state': 42,
			'verbose': 0,
			'warm_start': False}

```
n_estimators = [10,20,30,50]
max_features = ['auto', 'sqrt']
max_depth = [int(x) for x in np.linspace(10, 50, num = 8)]
min_samples_split = [4, 8, 10]
min_samples_leaf = [2, 4, 6]
bootstrap = [True, False]
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'bootstrap': bootstrap}
```

```
from sklearn.model_selection import RandomizedSearchCV
rf = RandomForestRegressor()
rf_random = RandomizedSearchCV(estimator = rf,param_distributions = random_grid,r
```

```
rf_random.fit(X_train_scaled, y_train)
```

```
Fitting 5 folds for each of 100 candidates, totalling 500 fits

RandomizedSearchCV(cv=5, estimator=RandomForestRegressor(), n_iter=100,
                   n_jobs=-1,
                   param_distributions={'bootstrap': [True, False],
                                        'max_depth': [10, 15, 21, 27, 32, 38,
                                                      44, 50],
                                        'max_features': ['auto', 'sqrt'],
                                        'min_samples_leaf': [2, 4, 6],
                                        'min_samples_split': [4, 8, 10],
                                        'n_estimators': [10, 20, 30, 50]},
                   random_state=35, verbose=2)
```

```
best_params = rf_random.best_params_
```

```
print ('Best Parameters is', best_params)
```

```
Best Parameters is {'n_estimators': 50, 'min_samples_split': 10, 'min_samples_l
eaf': 6, 'max_features': 'sqrt', 'max_depth': 21, 'bootstrap': False}
```

```
print(f'Accuracy =: {round(rf_random.score(X_train_scaled, y_train) * 100, 2)}%')
```

```
Accuracy =: 75.07%
```

10. ADVANTAGES & DISADVANTAGES:

10.1. Advantages:

- As Weather conditions have been changing for the time being this helps people to knowabout the rainfall prediction
- To avoid unnecessary floods by opening dams with the help of rainfa l prediction
- Farmers and fisherman will get the most advantage of these rainfa l details so that we theycan plan accordingly
- During the monsoon days it helps the government to find the evacuation areas to avoid lossof human life and costly things

10.2. DisAdvantages:

- As the data was co lected from limited places so it helps only for the people who located inthose areas.
- In case the data was co lected being wrong the algorithm wi l produce the wrong prediction
- As of now have collecting only a limited number of data set, In feature, we wi l make thealgorithm to work worldwide

11. CONCLUSION:

Floods are the most common natural disasters and have widespread ef ect flood forecasting is hence an important research area and various possible solutions have been presented in literature to this end the input data were selected based on a cor elation and uncertainty analysis of the rainfa l and flood data and a classification based real-time flood prediction model was developed heavy rainfall that may occur in urban areas was analyzed in advance and the expected range of an urban flood was predicted in real time using the proposed model

12. FUTURE SCOPE:

With the change in climatic conditions and rainfa l patterns this can lead to flash floods causing catastrophic damage to the environment.The system can be further enhanced with a flood prediction system along with rainfall prediction. Evacuation areas can be included along with the flood prediction system in such a way that the system recommends the user as we l as to the community if there might

be an occur ence of flood. A recommendation system integrated with theprediction system sha l sound good for society.

GITHUB LINK : https://github.com/IBM-EPBL/IBM-Project-5571-1658808571