EXPLORATORY ANALYSIS OF RAINFALL DATA IN INDIA FOR AGRICULTURE

TEAM ID: PNT2022TMID12583

TEAM LEADER: Vikneshwar A

TEAM MEMBER 1: Srivathssan V V

TEAM MEMBER 2: Vignesh M

TEAM MEMBER 3: Vignesh A

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 PROJECT OVERVIEW

The ability to get rain is crucial for agriculture. India's economic development depends heavily on agriculture. Due to its enduring applications in crisis management, agriculture, and the measurement of pollutant concentrations, rainfall forecasting or prediction has gained research relevance. But because of the extreme fluctuation in climatic changes, it is now exceedingly difficult to anticipate the amount of rain that will fall. Therefore, a more accurate forecasting model based on classification models like SVM, XGBoost, KNN, and Decision Tree is required.

1.2 PURPOSE

After too much rain damaged the winter crops, farmers were faced with the arduous chore of harvesting their harvest and transporting the produce to markets. A new phase of intervention is anticipated to be introduced to the impacted sectors confronted with the adverse propensities of rainfall extremes by accurate and timely rainfall forecast.The purpose of this project is to help farmers and to effectively use the water resources, it is necessary to have a programme that can predict rainfall more accurately. Heavy rainfall might have negative effects, such as crop damage or destruction.

CHAPTER 2
LITERATURE SURVEY

2.1 EXISTING PROBLEM

Huge rainfall cause disasterous effect like cyclones and floods. It also adversely affects the agriculture. This becomes a threat to the economy of a country like India whose economy is hugely dependent on agriculture.In the existing scenario, we get information regarding weather foreasting from media like television, news channels et c.But forecasting from these media are time constrained and it is difficult to obtain forecasting during a desired time. Also, the accuracy of the existing forecasting media is much low.

2.2 REFERENCES

1. Rainfall forecasting model using machine learning models: Case study Terengganu, Malaysia

a. Authors: Wannie M. Ridwan, Michelle Sapitang, Awatif Aziz, Khairul Faizal Kushiar, Ali Najah Ahmed, Ahmed El-Shafie

2. A study on rainfall prediction techniques

a. Authors: Dr. C K. Gomathy, Annapareddy bala narashimha reddy ,Aravapallii pavan kumar, Aile Lokesh.

3. Predicting Rainfall for agriculture in India using Regression

a. Authors: Sandeep Dhyani

2.2.1.RAINFALL FORECASTING MODEL USING MACHINE LEARNING MODELS

Wannie M. Ridwan, Michelle Sapitang, Awatif Aziz, Khairul Faizal Kushiar, Ali Najah Ahmed, Ahmed El-Shafie

In order to forecast the rainfall this study makes use of machine learning algorithms by evaluating different scenarios like humid season, winter season etc, and time horizon using two different methods. This paper takes Tasik Kenyir in Terengganu of Malaysia as case study.

They have used multiple machine learning algorithms like Bayesian Linear Regression (BLR), Boosted Decision Tree Regression (BDTR), Decision Forest Regression (DFR) and Neural Network Regression (NNR).

(i) Bayesian Linear Regression (BLR) :

a. In Bayesian Linear Regression, the linear regression is formulated using probability distributions rather than point estimates. Thus the dependent variable is assumed to be drawn from a probability distribution which is most probably normal distribution.

b. The BLR model estimates the response using the formula,

Y~N(B^TX,(sigma)^2*I)

c. The aim of BLR is not to find the single “best” value of model parameters but to determine the posterior distribution for model parameters

(ii) BOOSTED DECISION TREE REGRESSION (BDTR) :

a. Boosting is one of the methods to create ensemble models. Gradient boosting is a machine learning algorithm that repetitively build models fine tuning it in subsequent steps.

b. It creates subsequent models in a step-wise fashion, using a predefined loss function to measure the error, and correcting it in each successive step

c. The BDTR involves a decision tree which has its model parameters getting finetuned at each step and finally comes up with the best version and the most efficient model.

(iii) DECISION FOREST REGRESSION (DFR) :

a. Decision Forest Regression is an ensemble learning algorithm that uses multiple models trained over the same data, to find a more powerful predictive result.

b. This Model contains many decision tress internally and each decision tree differs from one another with respect to the value of hyper parameters like number of nodes, number of leaf nodes etc.

c. Each of these estimators (Decision Trees) makes independent predictions and the correct prediction is made by averaging the results

(iv) NEURAL NETWORK REGRESSION (NNR) :

a. Neural Networks are huge networks of neurons/ perceptrons that are connected with one another and associated with different weights according to their significance.
b. These networks are used to predict with higher accuracy

2.2.2. A STUDY ON RAINFALL PREDICTION TECHNIQUES

Dr. C K. Gomathy, Annapareddy bala narashimha reddy ,Aravapallii pavan kumar, Aile Lokesh.

This study focuses in reviewing various prediction techniques available out there and the challenges that are associated with predicting rainfall
This study intents to offer non-experts with easy access to techniques, approaches utilized in the sector of precipitation prediction and provide a comparison between various machine learning techniques
This paper deals with different Regression analysis techniques available out there. Regression analysis deals with the dependence of one variable called as the dependent variable on one or more other variables called as the independent variables.
This helps in testing the strong relationship between one dependent variable and many independent variables.
It also allows researchers to control the extraneous factors and helps assessing the cumulative effect of multiple factors.

(i) PROPOSED METHOD :

a. Conversion of the dataset into the appropriate format to perform good analysis of data and observe variation in the patterns of rainfall.

b. Then use the dataset to train machine learning models like MLR, SVR and test them to monitor their accuracies.

(ii) SUPPORT VECTOR REGRESSION :

a. Support vector regression is a bit different from SVM since it tries to find as many instances as possible on the street while limiting margin violations instead of trying to fit the largest possible street between two classes while limiting margin violations

b. Hyper planes in SVM which are separation lines between the data classes are used to define the equation of a straight line that will help us predict a continuous value

c. SVR performs regression in a higher dimensional space and the equations of hyperplanes are as follows: Wx+b=e and WX+b=-e. And the equation that satisfies the SVR is -e<= y -Wx-b <= +e

(iii) LASSO REGRESSION :

a. Lasso is Least Absolute Shrinkage and Selection Operator Lasso regression works by introducing a bias term but instead of squaring the slope, the absolute value of the slope is added as a penalty term.

On the whole this study concludes that the SVR is a valuable and adaptable strategy which helps the clients to manage the impediments related to distributional properties of fundamental factors, geometry of the information and normal issue if model over fitting.

This study finds SVR better than MLR since MLR can’t catch the non-linearity in a dataset and SVR winds up such circumstances. Also this study suggests that MAE (Mean Absolute Error) for both MLR and SVR models to assess the execution of the models.

2.2.3.PREDICTING RAINFALL AGRICULTURE IN INDIA USING REGRESSION

Sandeep Dhyani

This research mainly focuses to use machine learning regression models that can predict the rainfall in the most effective way on the given data.

This study makes use of five regression machine learning algorithms name decision trees, KNN (Kth Nearest Neighbor), Multiple Linear Regression (MLR), RFE (Random Forest Regressor) and Support Vector Machine.

This study aims to develop the ML models and then verify the outcomes of it which can predict the rainfall keeping the agriculture sector in mind. This study also compares the accuracies of various regression models along with R-Squared score and RMSE package.

This study concludes that RFE (Random Forest Regressor) is the best suited model with much accuracy among all the models for the given dataset.

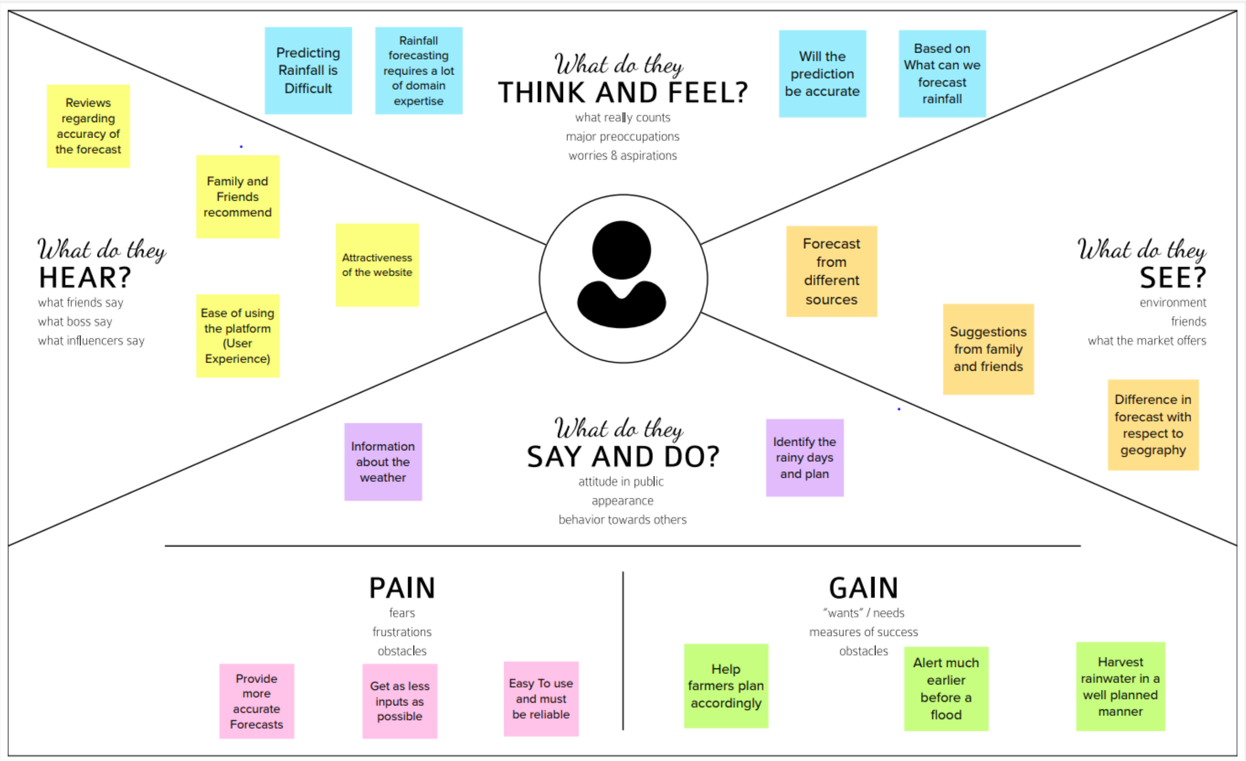
2.3 PROBLEM STATEMENT DEFINITION

The foundation of the Indian economy is agriculture. The most crucial factor for agriculture is the water source, or rainfall. Farmers are alerted by the rainfall prediction because they can safeguard their crops from the rain early on. Therefore, it is critical to make as accurate of a rainfall prediction as possible. Investigation and analysis of rainfall data across different locations in India, with a focus on those where agricultural activity has been done consistently across a wide range. Future rainfall prediction for certain regions utilising various machine learning approaches, such as XGBoost classifer, SVM classifers, Decision tree, Naive Bayes classifer, Logistic regression, etc., with the aid of the analysis of the resulting data.

CHAPTER 3

IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP



Predicting Rainfall is Difficult
Rainfall forecasting requires a lot of domain expertise
What do they THINK AND FEEL?
what really counts
major preoccupations
worries & aspirations
Will the prediction be accurate
Based on What can we forecast rainfall
Reviews regarding accuracy of the forecast
Family and Friends recommend
Attractiveness of the website
What do they HEAR?
what friends say
what boss say
what influencers say
Ease of using the platform (User Experience)
Forecast from different sources
What do they SEE?
environment
friends
what the market offers
Suggestions from family and friends
Difference in forecast with respect to geography
Information about the weather
What do they SAY AND DO?
attitude in public
appearance
behavior towards others
Identify the rainy days and plan
PAIN
fears
frustrations
obstacles
Provide more accurate Forecasts
Get as less inputs as possible
Easy To use and must be reliable
GAIN
"wants" / needs
measures of success
obstacles
Help farmers plan accordingly
Alert much earlier before a flood
Harvest rainwater in a well planned manner

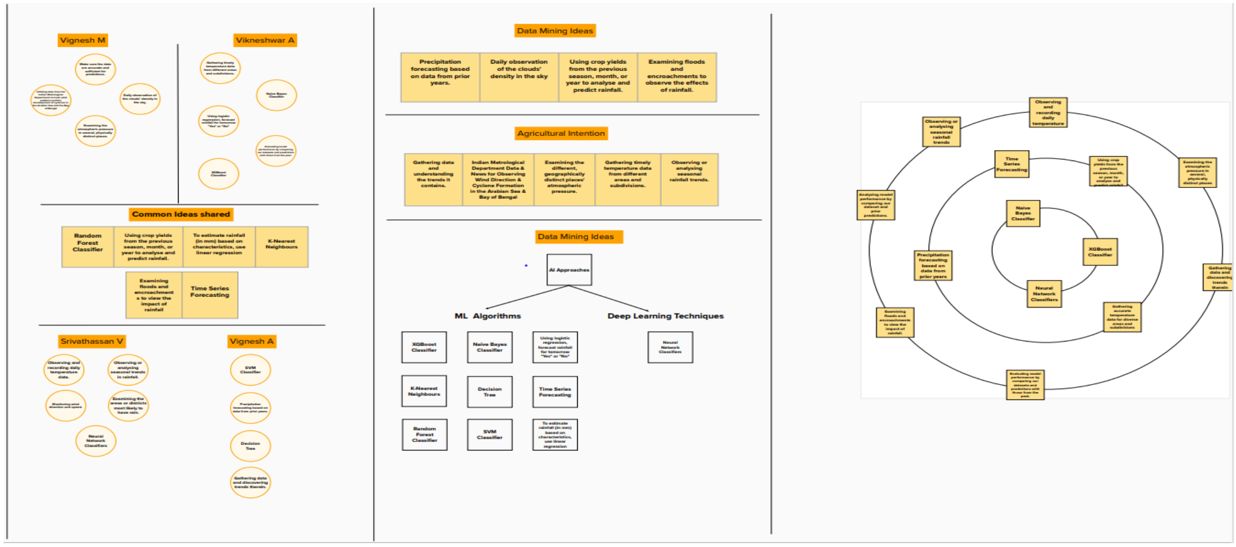
3.2 IDEATION & BRAINSTORMING

In the ideation phase, we examine the atmospheric pressure in a number of geographically distinct regions, observe wind patterns and cyclone formation in the Arabian Sea and the Bay of Bengal using data and news from the Indian Meteorological Department, and compare our model's performance to that of earlier predictions and datasets.

The most frequently mentioned concepts include analyzing and predicting rainfall based on seasonal, monthly, or yearly agricultural yields, as well as looking at floods and encroachments to see the consequences of rainfall.

The data mining ideas include keeping track of daily temperatures and comparing our dataset and past predictions to assess the success of the model.

Based on information from the previous year, the agricultural intentions forecast the months that will have rain. Daily observation of the cloud density by gazing up into the skyData gathering and learning are included in pattern recognition and the Trends found there and examining the atmospheric pressure in various geographically distinct regions.


Vignesh M
Vikneshwar A
Naive Bayes Classifier
Common Ideas shared
Random Forest Classifier
Using crop yields from the previous season, month, or year to analyse and predict rainfall.
To estimate rainfall (in mm) based on characteristics, use linear regression
K-Nearest Neighbours
Examining floods and encroachments to view the impact of rainfall
Time Series Forecasting
Srivathassan V
Vignesh A
Observing and recording daily temperature data.
Observing or analysing seasonal trends in rainfall.
Examining the areas or districts most likely to have rain.
SVM Classifier
Data Mining Ideas
Precipitation forecasting based on data from prior years.
Daily observation of the clouds' density in the sky
Using crop yields from the previous season, month, or year to analyse and predict rainfall.
Examining floods and encroachments to observe the effects of rainfall.
Agricultural Intention
Gathering data and understanding the trends it contains.
Indian Metrological Department Data & News for Observing Wind Direction & Cyclone Formation in the Arabian Sea & Bay of Bengal
Examining the different, geographically distinct places' atmospheric pressure.
Gathering timely temperature data from different areas and subdivisions.
Observing or analysing seasonal rainfall trends.
Data Mining Ideas
AI Approaches
ML Algorithms
Deep Learning Techniques
XGBoost Classifier
Naive Bayes Classifier
Using logistic regression, forecast rainfall for tomorrow "Yes" or "No"
Neural Network Classifiers
K-Nearest Neighbours
Decision Tree
Time Series Forecasting
Random Forest Classifier
SVM Classifier
To estimate rainfall (in mm) based on characteristics, use linear regression
Time Series Forecasting
Naive Bayes Classifier
XGBoost Classifier
Neural Network Classifiers

3.3 PROPOSED SOLUTION

S. NO	Parameter	Description
1	Problem Statement	Heavy rains can cause huge effects as leading to disasters like floods, Landslides etc and can have adverse effect effects on agriculture. This could severely affect countries like India wherein Agriculture serves as a predominant occupation. So, if we can predict the rainfall, we could help farmers plan accordingly thereby reducing the damage of crops.
2.	Ideation/ Solution Description	We propose a web application that takes in the weather conditions as input like the temperature, wind direction etc and using Machine Learning algorithms to forecast the rainfall for a given date thereby helping farmers to plan accordingly. This application gives you a precise information as the wind direction, the intensity of the rain expected like whether it is High, Low, Moderate etc. It also generates a report which provides guidance to famers as in how much water resources do they need to prepare, what crops could they cultivate etc

3	Novelty/ Uniqueness	Providing a report on the rain forecast for a given time period and some recommended measures that can be taken Display the annual and monthly rainfall received by states in Tamil Nadu as a pie chart If the farmers are not sure about the input, they can simply specify the region
4.	Social Impact/ Customer Satisfaction	Prevents crop damage due to heavy rainfall Improves crop yield Prior planning can be done using rain forecast
5.	Business / Revenue Model	We can provide pop-up ads, overlay ads, and other advertising services from third-party advertisers A pro version without any ads and with added benefits.
6.	Scalability of the Solution	Inbuilt Machine Learning Model as a pickle file so training or testing isn’t required and prediction doesn’t take long time Provide enhanced customer support .

3.4 PROBLEM SOLUTION FIT

1. CUSTOMER SEGMENTS

- Farmers
- Those Who are engaged in agricultural activities
- Government departments which would like to take some preventive

2. JOBS-TO-BE-DONE / PROBLEMS

- Find the appropriate dataset
- Preprocess the dataset and find the appropriate machine learning algorithm
- Train and Test till higher accuracy is achieved
- Recommend Measures

3. TRIGGERS

- Current losses and debts
- Crop damage and financial loss
- Market competition

4. EMOTIONS: BEFORE / AFTER

BEFORE: Incur huge financial loss, crop damage and poor disaster management and recovery

AFTER: Increase in crop production and well organized and pre planned disaster management

5. AVAILABLE SOLUTIONS

- Weather forecasting channels in television and internet
- Weather forecasting apps
- Rainfall related information printed on news papers

6. CUSTOMER CONSTRAINTS

- Predict the rainfall with 100% accuracy faster
- Recommend best practices depending on the rainfall
- Tolerate Unstable Network connection

7. BEHAVIOUR

- Consult authorities and those who are experienced in this domain
- Take decisions by observing the previous data

8. CHANNELS of BEHAVIOUR

8.1 ONLINE

- Receive notifications via SMS or mobile devices

8.2 OFFLINE

- Gather forums or meetings where farmers gather to discuss the problems

9. PROBLEM ROOT CAUSE

- Irregular rainfall causes a lot of destruction to crops
- Adverse effect on agriculture
- Disastrous effects like flood, landslides etc

10. OUR SOLUTION

- Building a low-cost ML based solution for predicting rainfall and taking appropriate measures
- Region based analysis of rainfall in India to get the annual rainfall and seasonal crop yield statewise

CHAPTER 4

REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

FRNo.	**Functional Requirement(Epic)**	**Sub Requirement (Story / Sub-Task)**
FR-1	Input Area	Provide Input and Get Result
FR-2	Dashboard	Viewing the visualisations, pages navigation
FR-3	Generate Report	Download the report as .pdf, .docx files
FR-4	Model Update	Update the model periodically on the updated dataset and replace with the updated model
FR-5	Prediction	Predicting the future rainfall from the dataset using ML based model
FR-6	Suggestion forum	Suggest measures to be taken and recommended practises based on the rainfall condition

FR - 7	Automatic field fill	Fill in the details of the input fields if region is specified and is unknown to the user
FR - 8	Feedback	Collecting feedback against the accuracy of the prediction for further improvement

4.2 NON FUNCTIONAL REQUIREMENT

FRNo.	**Non-Functional Requirement**	**Description**
NFR-1	Usability	● quality attribute that assesses how easy ● user interfaces are to use. ● the system doesn't expect any technical ● pre-requisite from the user
NFR-2	Security	● User details should be secured ● The confirmation of valid user is required
NFR-3	Reliability	● The extent to which an experiment, test, or measuring procedure yields the same ● Results on repeated trials.
NFR-4	Performance	● The degree to which the result of a measurement, calculation, or specification ● Can be depended on to be accurate.

NFR-5	Availability	● The version of application should be available even at the time of maintanance and updation ● The page would not take a lot of time to ● Load the content and display them
NFR-6	Scalability	● The application should in the way of adding ● New functionalities without affecting the existing functionalities ● Less prone to errors

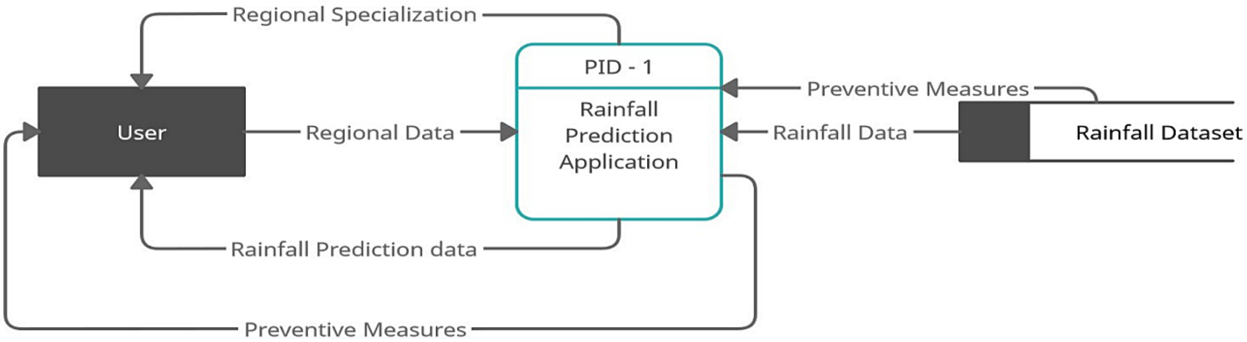
CHAPTER 5

PROJECT DESIGN

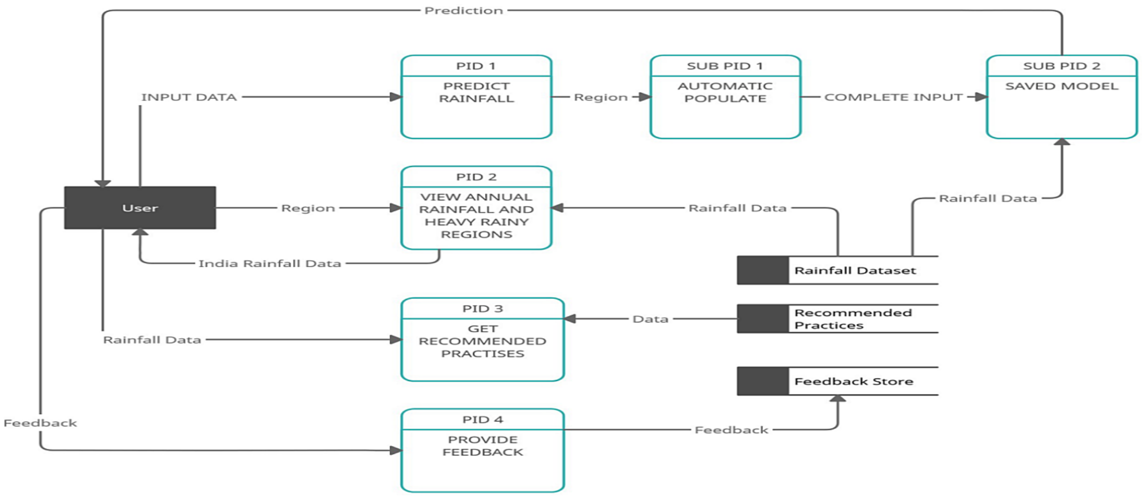
5.1 DATA FLOW DIAGRAM

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.

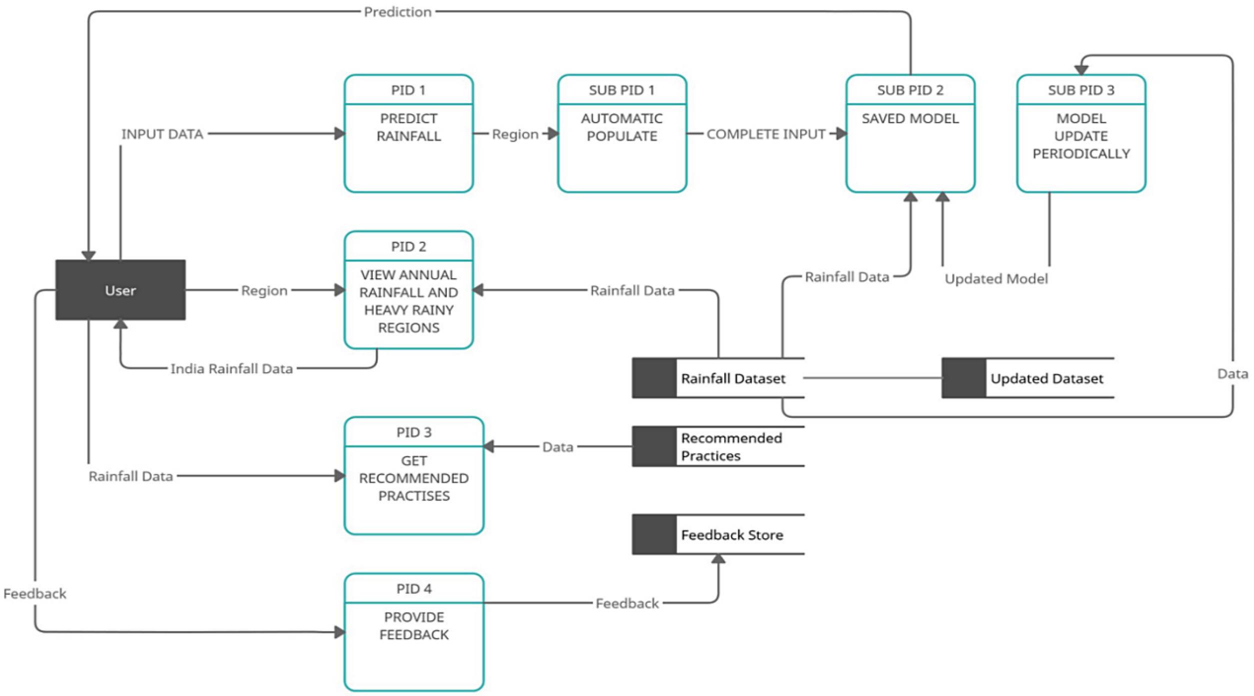
0- LEVEL DATA FLOW DIAGRAM:



1- LEVEL DATA FLOW DIAGRAM



2 - LEVEL DATA FLOW DIAGRAM



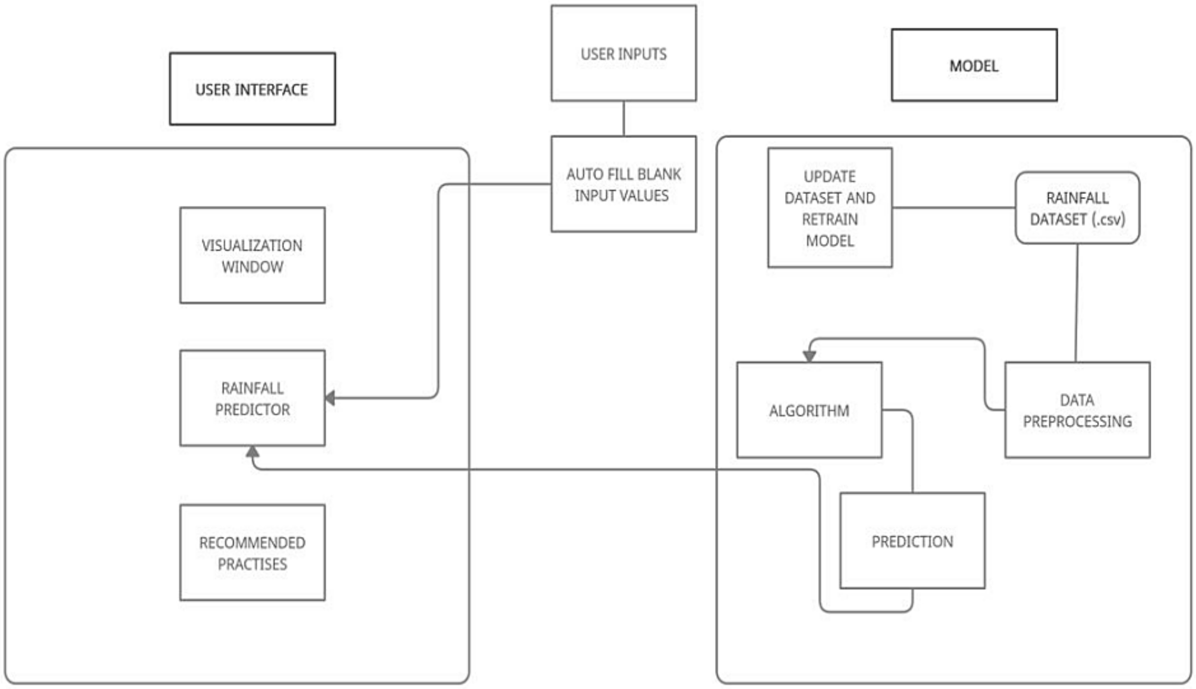
5.2 SOLUTION & TECHNICAL ARCHITECTURE

5.2.1 Solution Architecture:

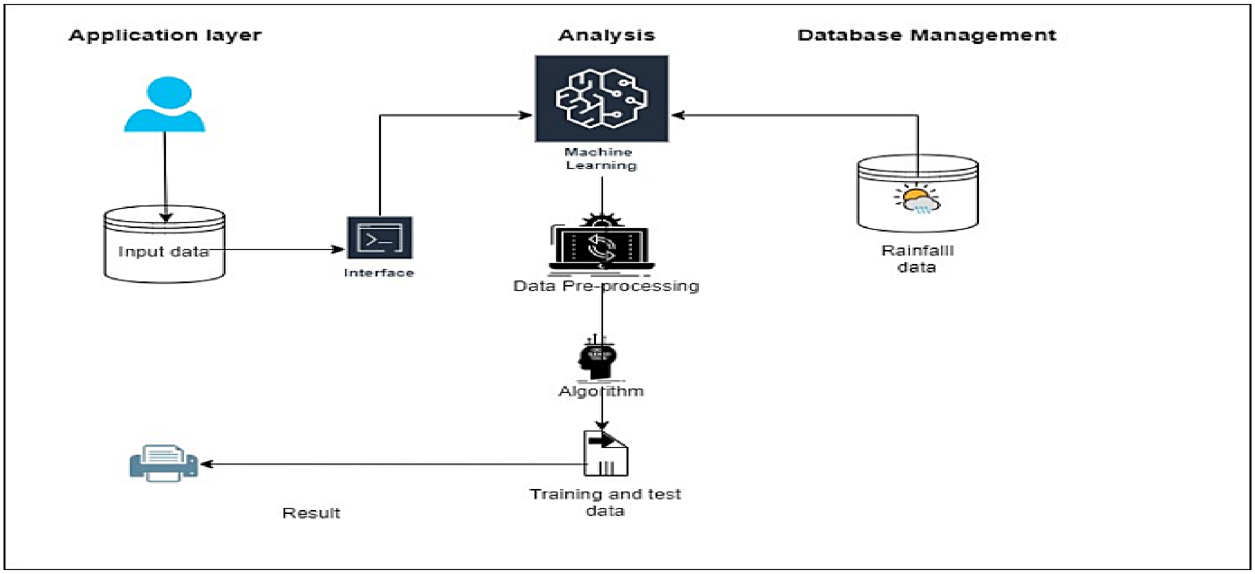
Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.

5.2.2 Solution Architecture Diagram:



5.2.3 Technical Architecture:



5.2.4 Components and technologies:

S. No	Component	Description	Technology
1.	User Interface	User interacts with the application using the web UI .	HTML, CSS, Python and flask
2.	Registration	Registration of the user to the application	Python
3.	Login	Login of user to the application	Python
4.	Integration	Integrating machine learning model and the web page	Flask
5.	Database	Numeric data stored in the form of tables	MySQL

6.	Cloud Storage	Database service on cloud	IBM, DB2, IBM cloudant etc
7.	Data Pre-processing	Data is processed to remove outliers and missing values, so that the data can be used for training the model	Pandas, Numpy, Keras, Matplotlib modules of python
8.	Machine learning model	Random forest, KNN, CNN are used to improve the accuracy of the model.	Sklearn, Seaborn
9.	Result	This application shows the predicted rainfall data with the crops’s suggestions.	Python Flask

10 .	Crops	This shows the list of crops and details about it.	HTML, CSS. Flask

5.2.5 Application characteristics:

S.No	Characteristics	Description	Technology
1.	Open Source Frameworks	Flask	Micro web framework in python
2.	Security Implementations	Basic HTTP authentication, Session based authentication, User Registration, Login, Tracking	Flask Security
3.	Scalable Architecture	It can grow and adapt with ease. It is designed for scalability and flexibility	Python, Flask
4.	Availability	Infrastructure of the system provides recoverability and protection from system failure	Flask

5.	Performance	Integrated support for unit testing RESTful request dispatching Uses Jinja templating Support for secure cookies	Flask

5.3 USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
User in Website	Basic Form Template	USN-1	Develop a simple html file with a basic form where users can provide their inputs like they can select their region of interest, wind direction etc.	User should be able to select any region in India and enter the values	High	Sprint-1
	Auto Fill Inputs	USN-2	The input fields must be auto filled If the user is not sure about the values given the region	The inputs must get auto filled once the region is selected	High	Sprint-1

	Basic Flask Server	USN-3	A basic flask server that has all the necessary routes and integrated with the website A basic flask server that has all the necessary routes and integrated with the website	The corresponding functions must execute correctly once route is hit and checked by tools like postman	Medium	Sprint-1

Backend Team	Model Building and Testing	USN-4	Construct Machine Learning models trained on the pre-processed dataset and observe their accuracy. Save the Models as .pkl file	Queries can be raised in situation of doubts	High	Sprint-2
Frontend Team	Design the website and add all required pages	USN-5	Develop Website with all the required pages	All necessary pages must be developed and appropriate content is filled	High	Sprint-2
	Add Recommended Features	USN-6	Recommend best practises based on the rainfall	The user will get all their doubt clarified	High	Sprint-2

	Visualization	USN-7	Add Visualisation for rainfall data across all the states in India	The visualizations has to be dynamically loaded and charts must be responsive	High	Sprint-3
Developer team	Auto Update Model	USN - 8	Update the ML model with the updated dataset	Model stays up to date	High	Sprint - 3

6.PROJECT PLANNING AND SCHEDULING

6.1 Sprint Planning and Estimation and Sprint Delivery and Schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Basic Form Template	USN-1	Develop a simple html file with a basic form where users can provide their inputs like they can select their region of interest, wind direction etc.	5	High	Vigneshwar, Srivathssan
Sprint-1	Auto Fill Inputs	USN-2	The input fields must be auto filled If the user is not sure about the values given the region	6	High	Vignesh A, Vignesh M

Sprint-1	Basic Flask Server	USN-3	A basic flask server that has all the necessary routes and integrated with the website	5	Medium	Vigneshwar, Srivathssan
Sprint- 2	Model Building and Testing	USN-4	Construct Machine Learning models trained on the pre-processed dataset and observe their accuracy. Save the Models as .pkl file	5	High	Vikneshwar, Srivathssan
Sprint- 2	Design the website and add all required pages	USN-5	Develop Website with all the required pages	6	High	Vignesh A, Vignesh M
Sprint - 2	Add Recommended Features	USN - 6	Recommend best practises based on the rainfall	5	Medium	Vikneshwar, Srivathssan

Sprint – 3	Visualisation	USN – 7	Add Visualisation for rainfall data across all the states in India	8	High	Vignesh A, Vignesh M
Sprint - 3	Auto Update Model	USN -8	Update the ML model with the updated dataset	8		Vikneshwar, Srivathssan

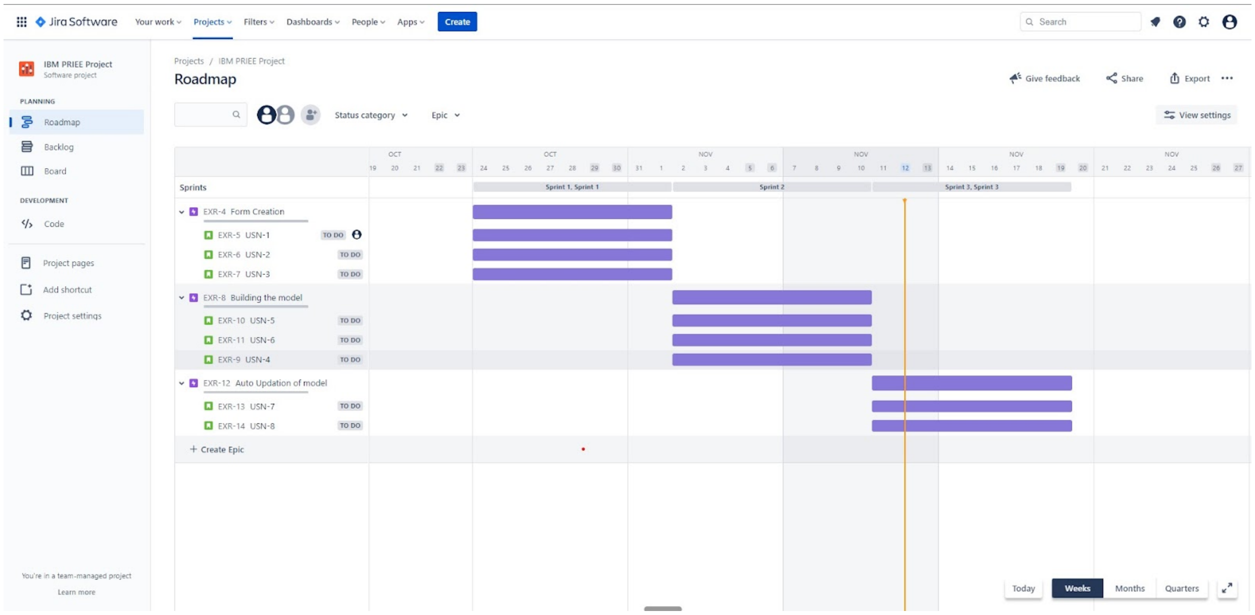
6.2 Sprint Delivery Schedule

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	16	8 Days	24 Oct 2022	01 Nov 2022	15	03 Nov 2022
Sprint-2	16	8 Days	02 Nov 2022	10 Nov 2022	15	10 Nov 2022
Sprint-3	16	8 Days	11 Nov 2022	19 Nov 2022	15	19 Nov 2022

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let’s calculate the team’s average velocity (AV) per iteration unit (story points per day)

6.3 Reports from JIRA



Projects / IBM PRIEE Project

Roadmap

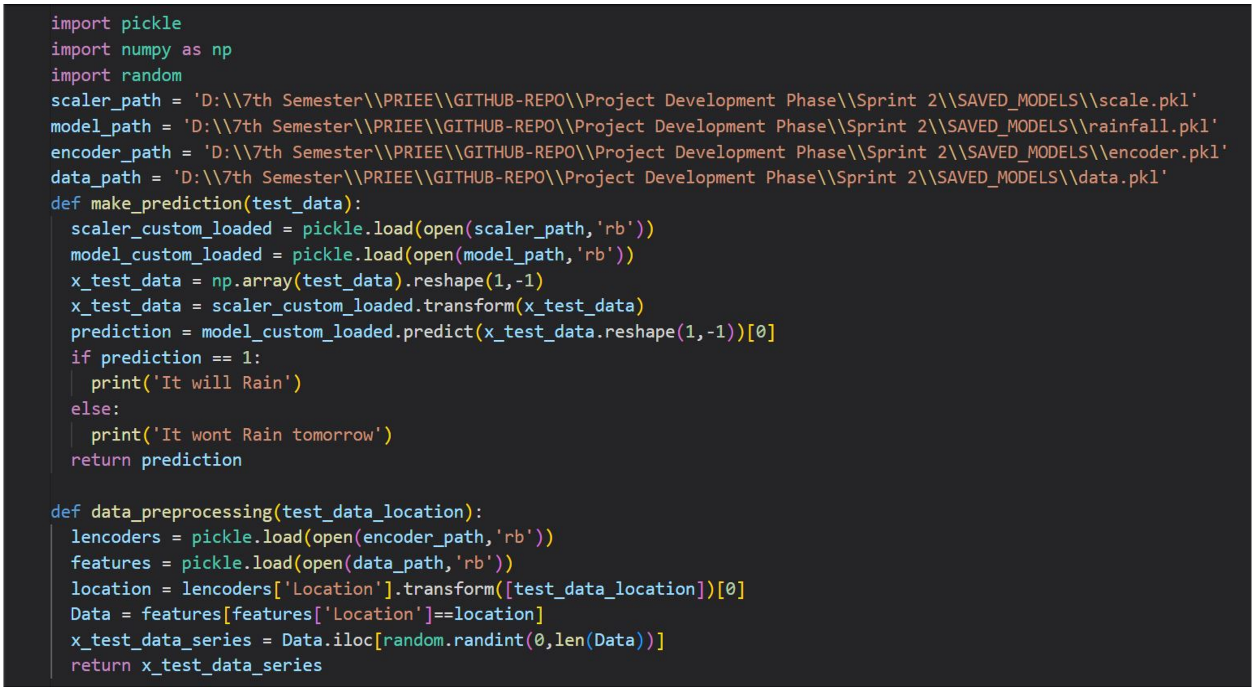
Sprints: Sprint 1, Sprint 1 | Sprint 2 | Sprint 3, Sprint 3

- EXR-4 Form Creation
  - EXR-5 USN-1 TO DO
  - EXR-6 USN-2 TO DO
  - EXR-7 USN-3 TO DO
- EXR-8 Building the model
  - EXR-10 USN-5 TO DO
  - EXR-11 USN-6 TO DO
  - EXR-9 USN-4 TO DO
- EXR-12 Auto Updation of model
  - EXR-13 USN-7 TO DO
  - EXR-14 USN-8 TO DO

+ Create Epic

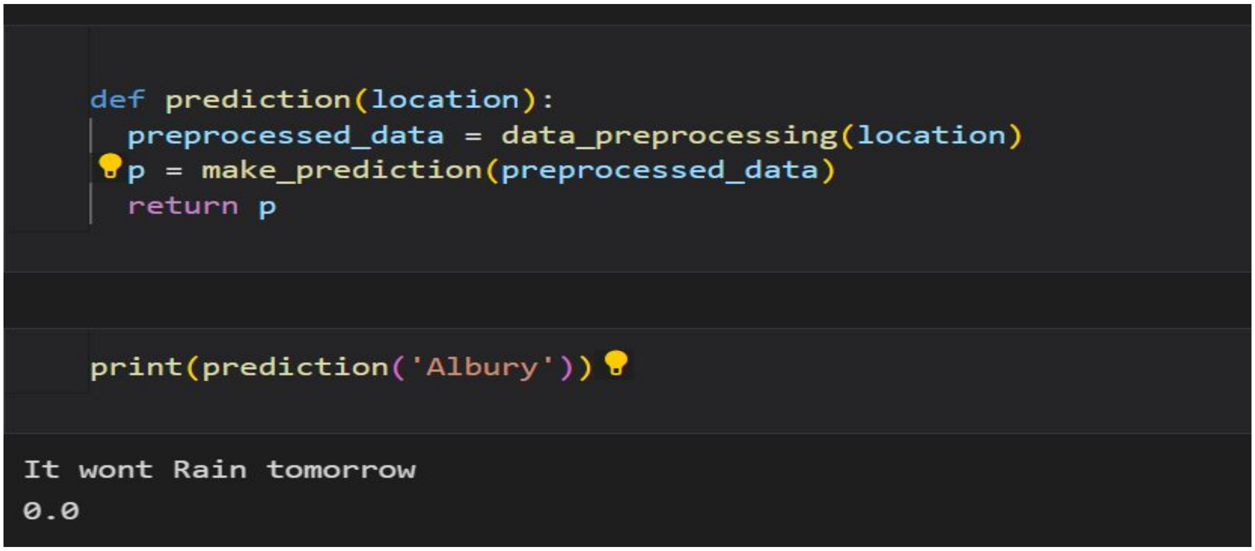
7. CODING & SOLUTIONING

7.1 DATA PRE-PROCESSING



```
import pickle
import numpy as np
import random
scaler_path = 'D:\\7th Semester\\PRIEE\\GITHUB-REPO\\Project Development Phase\\Sprint 2\\SAVED_MODELS\\scale.pkl'
model_path = 'D:\\7th Semester\\PRIEE\\GITHUB-REPO\\Project Development Phase\\Sprint 2\\SAVED_MODELS\\rainfall.pkl'
encoder_path = 'D:\\7th Semester\\PRIEE\\GITHUB-REPO\\Project Development Phase\\Sprint 2\\SAVED_MODELS\\encoder.pkl'
data_path = 'D:\\7th Semester\\PRIEE\\GITHUB-REPO\\Project Development Phase\\Sprint 2\\SAVED_MODELS\\data.pkl'
def make_prediction(test_data):
  scaler_custom_loaded = pickle.load(open(scaler_path,'rb'))
  model_custom_loaded = pickle.load(open(model_path,'rb'))
  x_test_data = np.array(test_data).reshape(1,-1)
  x_test_data = scaler_custom_loaded.transform(x_test_data)
  prediction = model_custom_loaded.predict(x_test_data.reshape(1,-1))[0]
  if prediction == 1:
    print('It will Rain')
  else:
    print('It wont Rain tomorrow')
  return prediction

def data_preprocessing(test_data_location):
  lencoders = pickle.load(open(encoder_path,'rb'))
  features = pickle.load(open(data_path,'rb'))
  location = lencoders['Location'].transform([test_data_location])[0]
  Data = features[features['Location']==location]
  x_test_data_series = Data.iloc[random.randint(0,len(Data))]
  return x_test_data_series
```

7.2 MODEL PREDICTION

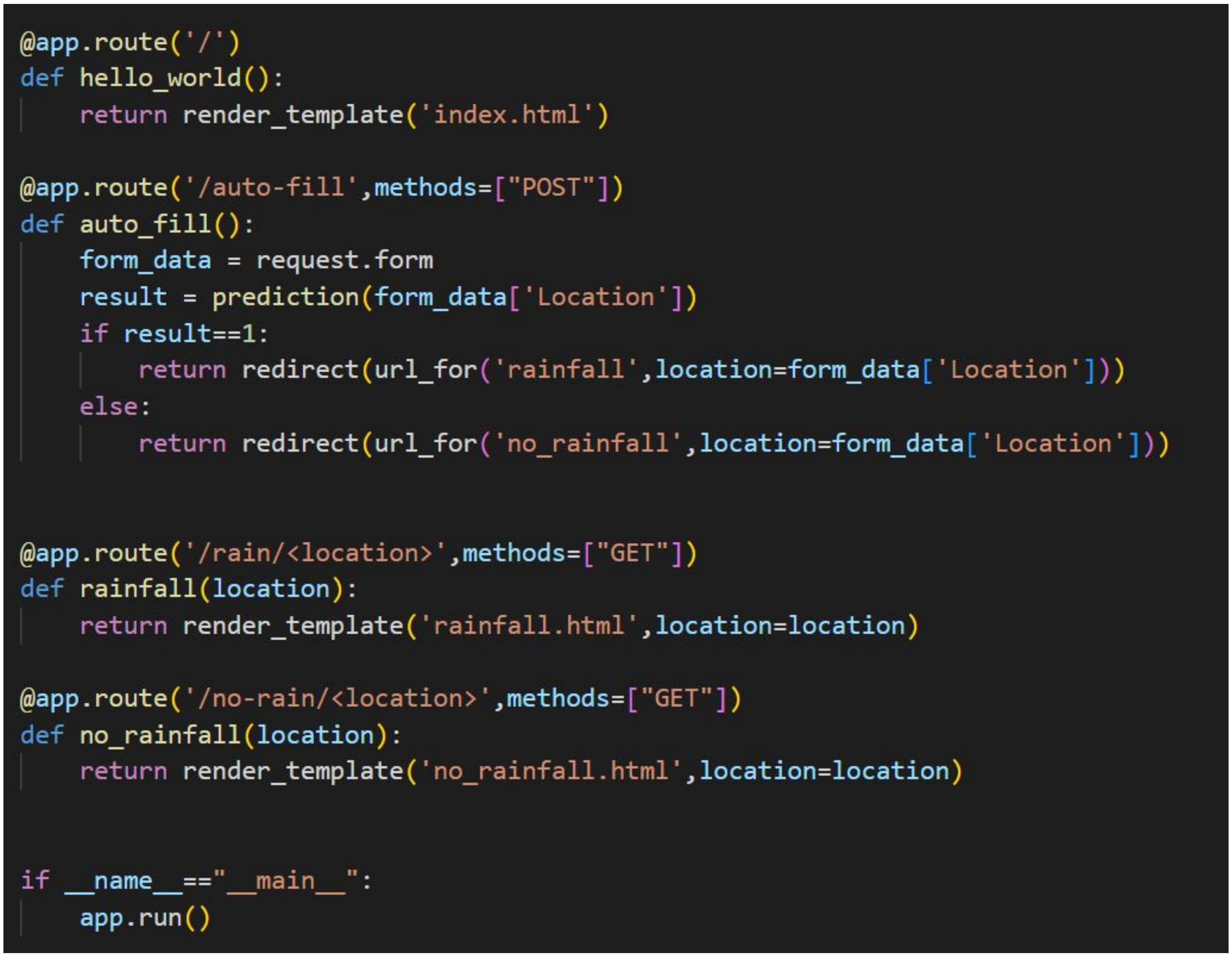


```
def prediction(location):
  preprocessed_data = data_preprocessing(location)
  p = make_prediction(preprocessed_data)
  return p
```

```
print(prediction('Albury'))
```

```
It wont Rain tomorrow
0.0
```

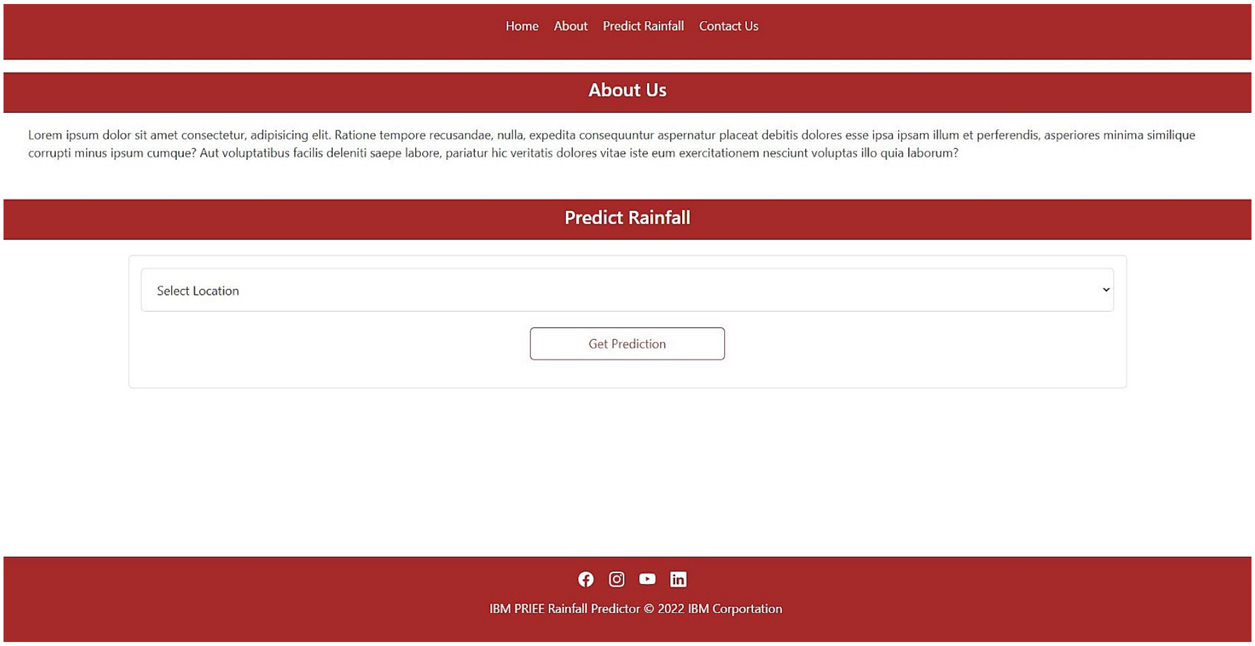
7.3 FLASK SERVER



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

@app.route('/auto-fill',methods=["POST"])
def auto_fill():
    form_data = request.form
    result = prediction(form_data['Location'])
    if result==1:
        return redirect(url_for('rainfall',location=form_data['Location']))
    else:
        return redirect(url_for('no_rainfall',location=form_data['Location']))


@app.route('/rain/<location>',methods=["GET"])
def rainfall(location):
    return render_template('rainfall.html',location=location)

@app.route('/no-rain/<location>',methods=["GET"])
def no_rainfall(location):
    return render_template('no_rainfall.html',location=location)


if __name__=="__main__":
    app.run()
```

8. TESTING

8.1 HOME SCREEN



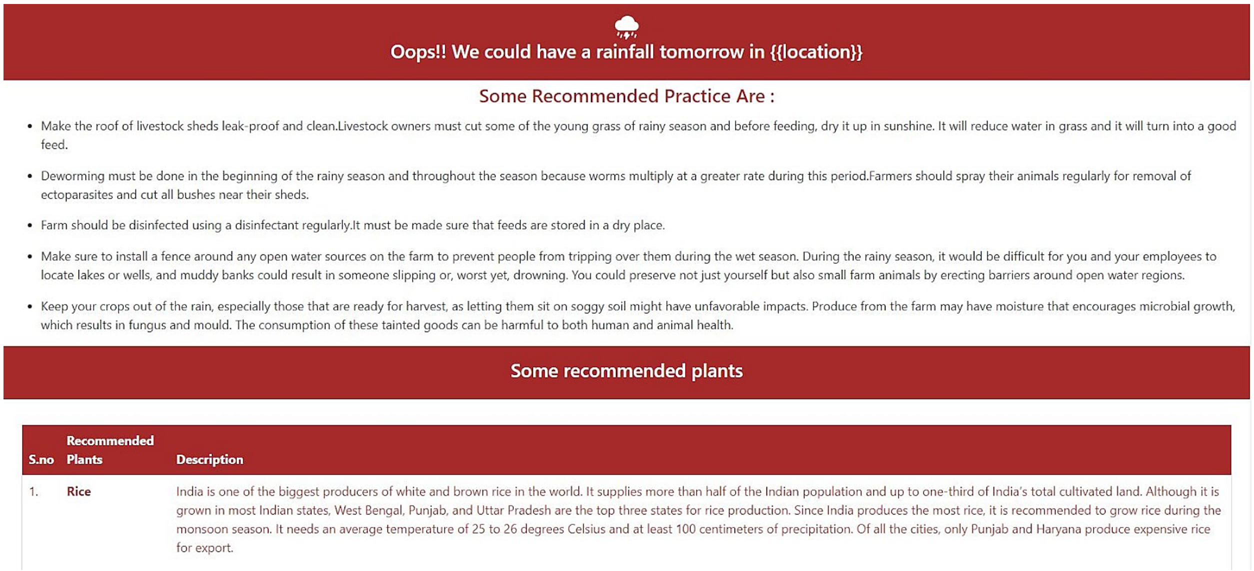
About Us

Lorem ipsum dolor sit amet consectetur, adipisicing elit. Ratione tempore recusandae, nulla, expedita consequuntur aspernatur placeat debitis dolores esse ipsa ipsam illum et perferendis, asperiores minima similique corrupti minus ipsum cumque? Aut voluptatibus facilis deleniti saepe labore, pariatur hic veritatis dolores vitae iste eum exercitationem nesciunt voluptas illo quia laborum?

Predict Rainfall

IBM PRIEE Rainfall Predictor © 2022 IBM Corportation

8.2 RESULT PAGE - RAINFALL TOMORROW



Oops!! We could have a rainfall tomorrow in {{location}}

Some Recommended Practice Are :

- Make the roof of livestock sheds leak-proof and clean.Livestock owners must cut some of the young grass of rainy season and before feeding, dry it up in sunshine. It will reduce water in grass and it will turn into a good feed.
- Deworming must be done in the beginning of the rainy season and throughout the season because worms multiply at a greater rate during this period.Farmers should spray their animals regularly for removal of ectoparasites and cut all bushes near their sheds.
- Farm should be disinfected using a disinfectant regularly.It must be made sure that feeds are stored in a dry place.
- Make sure to install a fence around any open water sources on the farm to prevent people from tripping over them during the wet season. During the rainy season, it would be difficult for you and your employees to locate lakes or wells, and muddy banks could result in someone slipping or, worst yet, drowning. You could preserve not just yourself but also small farm animals by erecting barriers around open water regions.
- Keep your crops out of the rain, especially those that are ready for harvest, as letting them sit on soggy soil might have unfavorable impacts. Produce from the farm may have moisture that encourages microbial growth, which results in fungus and mould. The consumption of these tainted goods can be harmful to both human and animal health.

Some recommended plants

S.no	Recommended Plants	Description
1.	**Rice**	India is one of the biggest producers of white and brown rice in the world. It supplies more than half of the Indian population and up to one-third of India's total cultivated land. Although it is grown in most Indian states, West Bengal, Punjab, and Uttar Pradesh are the top three states for rice production. Since India produces the most rice, it is recommended to grow rice during the monsoon season. It needs an average temperature of 25 to 26 degrees Celsius and at least 100 centimeters of precipitation. Of all the cities, only Punjab and Haryana produce expensive rice for export.

9. RESULTS

On observing the application developed, it successfully loads the saved model pickle file and makes the prediction.

The website makes it easier to provide input for the user as he only needs to choose the location by selecting it under a drop-down box.

The latency involved is much lesser as the model is saved with highest accuracy of about 93.5%.

The application predicts rainfall and according to the prediction it suggests some recommended practises and suitable .

This helps farmers as it suggests some crops and best practises.

10.Advantages and Disadvantages

PROS:

Making sound planning is one of the most obvious advantages of weather forecasting. It might be beneficial for a farmer's business decisions in the agriculture industry. They may better prepare for the numerous daily decisions by using forecasts. These choices include when to fertilize, when to irrigate crops, and which days are best for fieldwork. Farmers' choices will determine whether or not their crop is lucrative.

CONS:

Many years worth of climatic data must be included into certain forecast models.

11.Conclusion

Predicting rainfall has always been a potential field of research, since it has lot of benefits. Huge rains may have disasterous effects like cyclones, flood etc. This can have adverse effect on the economy of countries like India , whose economy is solely based on agriculture.

With advent of Machine Learning and subsequent developments in it, predicting rainfall has become much easier. It helps people with making prediction anywhere and anytime they wish to.

12.Future Scope

This application can further be developed to work efficiently in poor connectivity areas.

The model can be made more accurate by using advanced artificial technologies like Deep Neural Networks etc.

The feature of automatic location detection can be integrated with the application.

13.Appendix

Github Link:

The source code for the application is provided in the link