

# Rainfall Prediction System using Machine Learning

**Submitted By:**

Richa Maurya

Sameeksha Yadav

**Submitted To:**

Ms.Alka Singh  
Assistant Professor  
CSE Dept,BBDEC

# Agenda / Topics

- Abstract
- Introduction
- Problem Statement
- Proposed System
- Objective
- Hardware Requirements
- Software Requirements
- Technologies Required
- Methodology
- Use case diagram
- Work Flow diagram
- Modules
- Screenshots
- Database table
- Advantages
- Limitations
- Experimental Results
- Conclusions and Future Scope
- References

# Abstract

- In this project, we were asked to experiment with a real world dataset, and to explore how machine learning algorithms can be used to find the patterns in data.
- We were expected to gain experience using a common data-analysis and machine learning library, and were expected to submit a report about the dataset and the algorithms used. After performing the required tasks on a dataset of my choice, herein lies my final report.
- This project have gone through experiments which involve the use of machine learning techniques to build models to predict whether it is going to rain tomorrow or not based on weather data for that particular day in major cities.
- The results provide a comparison of various evaluation metrics of these machine learning techniques and their reliability to predict the rainfall by analyzing the weather data.

# Introduction

- Rainfall prediction remains a serious concern and has attracted the attention of governments, industries, risk management entities, as well as the scientific community.
- To solve this uncertainty, we have developed a software system using machine learning techniques.
- The project “*Rainfall Prediction using Machine Learning*” is a software system which is capable of predicting rainfall for various locations .
- The software has the ability to browse dataset from the system, clear the selected dataset, preprocess to visualise the dataset and then prediction on the basis of dataset.
- It is an application which gives a helping hand to the organisation to manage the data, resources, performance and also manages all the internal tasks and communication for fruitful result.

# Problem Statement

- Rainfall prediction can broadly be classified into two categories.
- One by analysing the different physical laws that governs rainfall in a particular region.
- Though, the approach sound fair enough , it has been found that the number of such constraints that governs rainfall is both spatial and temporal. Thus, the prediction involves too much calculations and hence it is not computationally feasible.
- The second approach involves expert systems to be involved and discovering hidden patterns of how different features that affects rainfall are actually related with physical rainfall.
- The later approach has been found to be more suitable.

# Proposed System

- So, by taking the consideration of modern advancement in research of Machine Learning we have developed a “*software application(Rainfall Prediction System)*” based on ML algorithms which are exceptionally accurate and robust in solving real life problems.
- The weather dataset contains 142,193 daily weather observations from 49 weather stations across Australia over the period November 2007 to June 2017.
- The system is based on logistic regression supervised machine learning algorithm.
- A software system which is user friendly, simple, fast and cost-effective.
- The application in the training phase can work as an optimising problem where an error function is minimised.

# Objective

- The objective is to implement rainfall prediction with machine learning techniques such as logistic regression or random forest or decision tree.
- To minimise the mathematical calculations for prediction.
- The ultimate goal of prediction is to enable the forecaster to increase the accuracy of forecasts made routinely.
- The objective is to improve organizational performance—more revenue, more profit, increased customer satisfaction through this application.
- Helps in determining future climate expectations.

# Hardware Requirements

	Minimum System Requirements	Recommended System Requirements
Processor	7 <sup>th</sup> Gen Intel Core i7	9 <sup>th</sup> Gen Intel Core i7 or better
RAM	4GB	8GB or more
Storage	256GB SSD	512GB or more
Display	14-inch FHD (1920 X 1080)	15.6-inch FHD IPS(1920 X 1080)
Graphics	4GB NVIDIA GeForce GTX 1060	8GB NVIDIA GeForce GTX 2070
Battery	Up to 2 hours	Up to 5 hours

# Software Requirements

## Tools Used:

- NumPy
- Pandas
- Matplotlib
- Seaborn
- Scikit Learn
- PyQt5 Tools

## Back End Tools / IDE:

- PyCharm
- Jupyter Notebook
- MySQL

## Front End Tools:

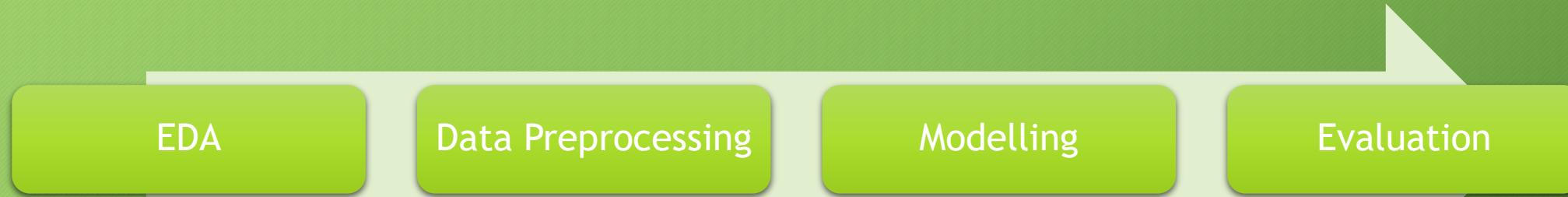
- Qt Designer

# Technologies Required

- **Python-** Python is a dynamic, high level, free open source and interpreted programming language. It supports object-oriented programming as well as procedural oriented programming. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. We can use python for developing desktop GUI applications, websites and web applications. The simple rules of python allow you to express concepts without writing additional code.
- **Machine Learning-** Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

# Methodology

- The System is developed in python using ML and the overall architecture include four major components:  
Data Exploration and Analysis, Data Pre-processing, Model Implementation, and Model Evaluation.



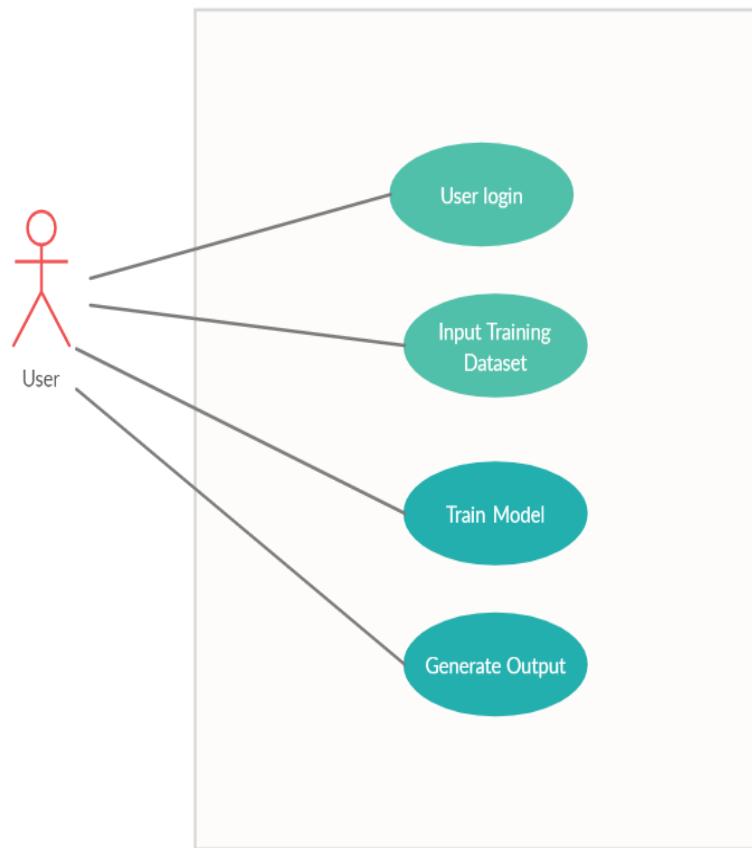
- Data Exploration
- Visualisation

- Missing Values
- Feature Expansion
- Categorical values
- Feature Scaling
- Feature Selection
- Over Sampling
- Under Sampling

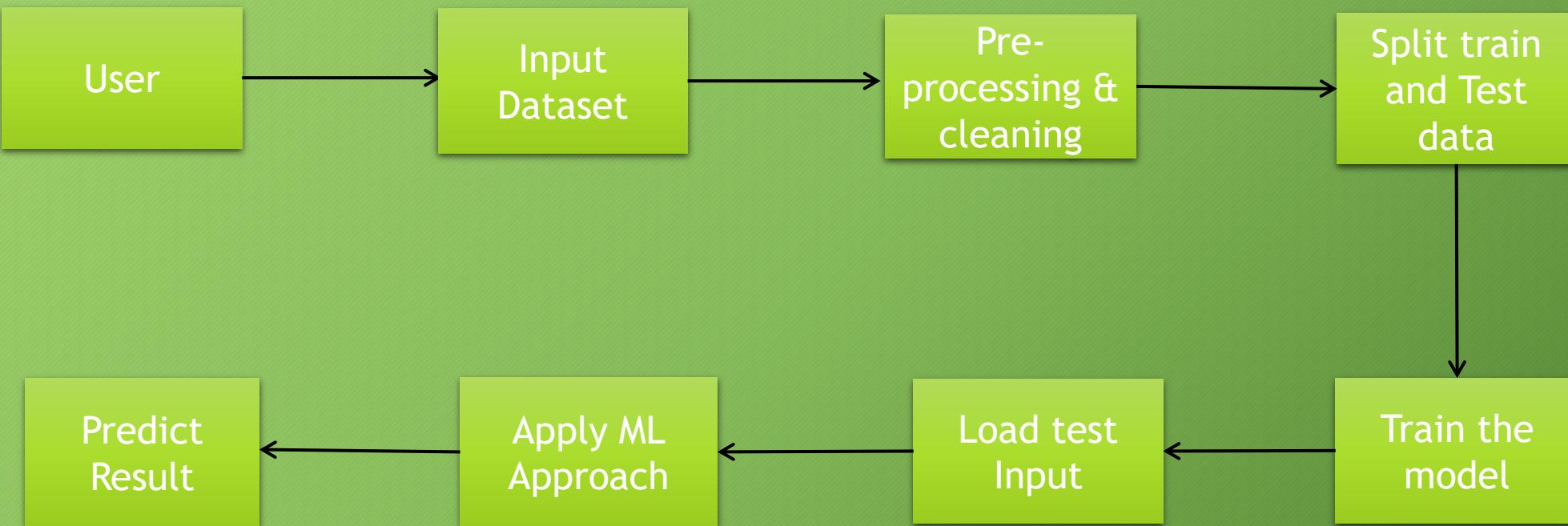
- Logistic Regression
- Decision Tree
- Random Forest

- Accuracy
- Precision
- Recall
- AUC/ROC
- F Score

# Use Case Diagram



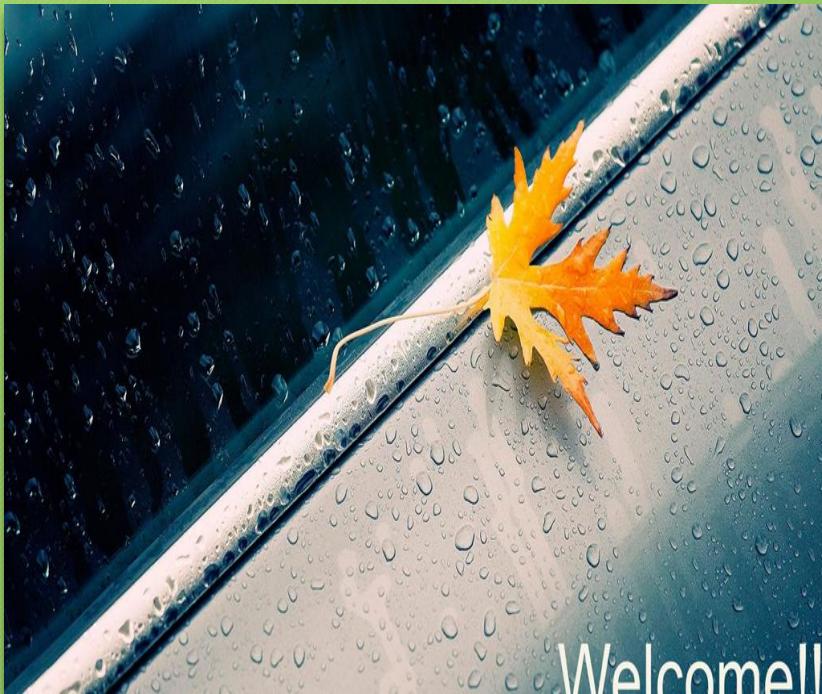
# Workflow Diagram



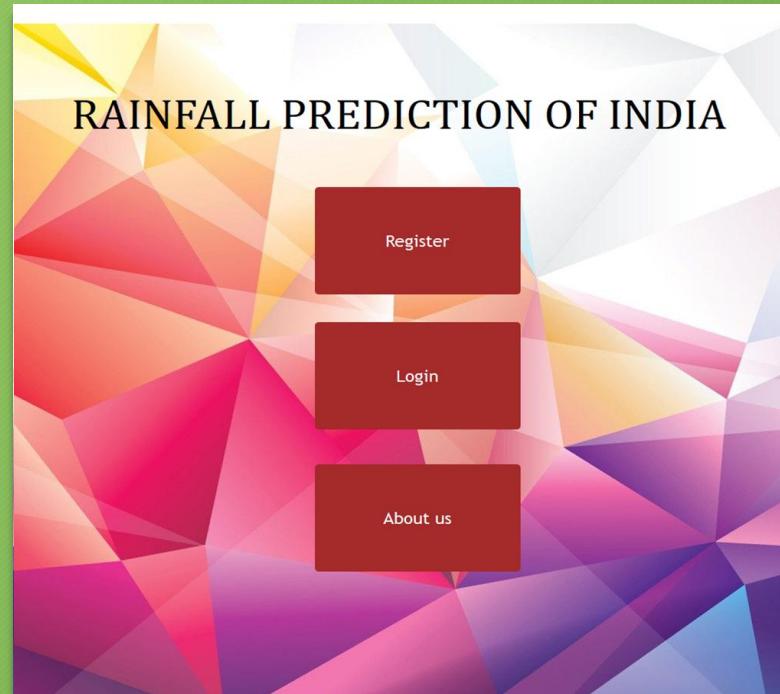
# Modules

- Welcome Frame : First screen of our application developed using pyttsx3 lib. for text to speech.
- Main Frame : User have the option to Login, Register or know About Us.
  - ❖ Login Frame : Only the authorized user can access the login successfully.
  - ❖ Register Frame : User can give their details to register themselves for the application.
  - ❖ About Us: User can get the information about the developers .
- Menu Frame : After successful login user have the menu page with different options like-
  - ❖ Browse File:- Here, user can upload the dataset and preprocess the data ,also have option to deselect the dataset.
  - ❖ Prediction:- User can get a predicted value on the basis of uploaded dataset.
  - ❖ Weather Forecasting :- It's add on functionality of this application in which user can get the current weather information of any country or city from all over the world.
  - ❖ Performance graph:- It gives accuracy/precision time of ML algorithms.
  - ❖ Logout Frame:- It's the last screen of our application through which user can exit from the frame.

# Screenshots of Application



Welcome Frame



Main Frame

A registration form titled "REGISTER". It includes five input fields with labels: "UserID", "Password", "Contact", "Email", and "Name", each accompanied by a grey input bar. Below the input fields is a large red button labeled "Register".

Register Frame



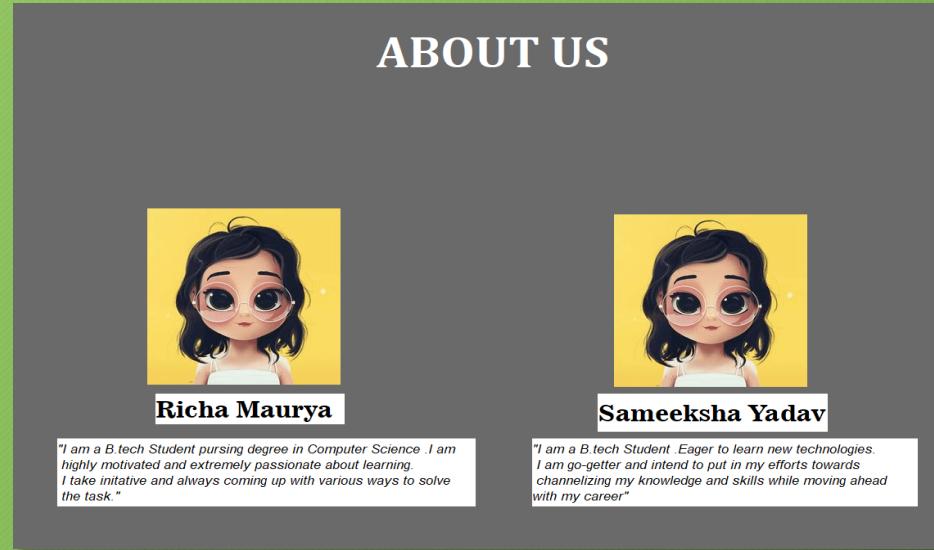
LOGIN HERE

Email

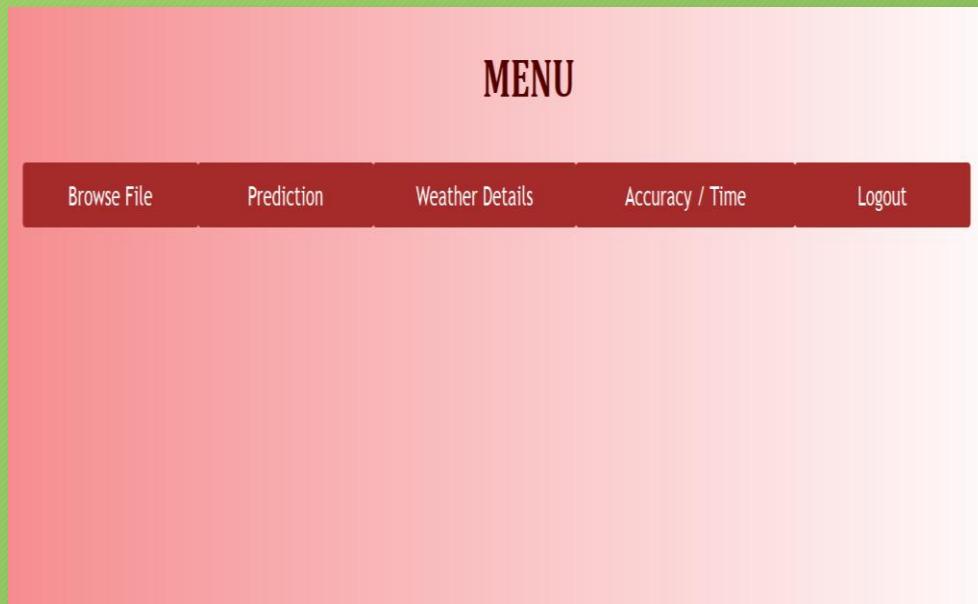
Password

**Login**

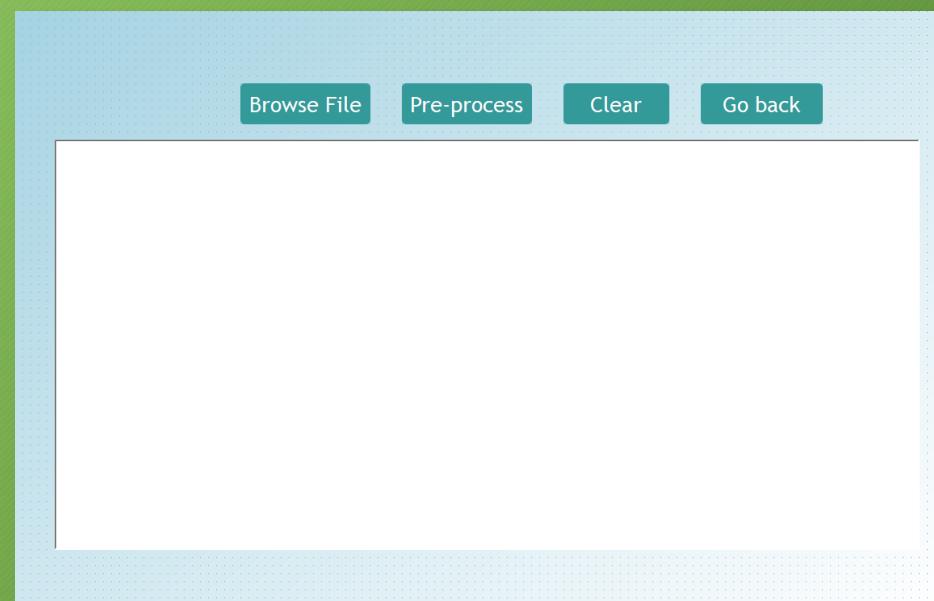
Login Frame



About Us Frame



Menu Frame



Browse Frame

### Rainfall Predictor

MinTemp	<input type="text"/>	MaxTemp	<input type="text"/>
Rainfall	<input type="text"/>	Evaporation	<input type="text"/>
Sunshine	<input type="text"/>	WindGustSpeed	<input type="text"/>
WindSpeed9am	<input type="text"/>	WindSpeed3pm	<input type="text"/>
Humidity9am	<input type="text"/>	Humidity3pm	<input type="text"/>
Pressure9am	<input type="text"/>	Pressure3pm	<input type="text"/>
Cloud9am	<input type="text"/>	Cloud3pm	<input type="text"/>
Temp9am	<input type="text"/>	Temp3pm	<input type="text"/>
RainToday	<input type="text"/>		
RainTomorrow:	<input type="button" value="Predict"/>		

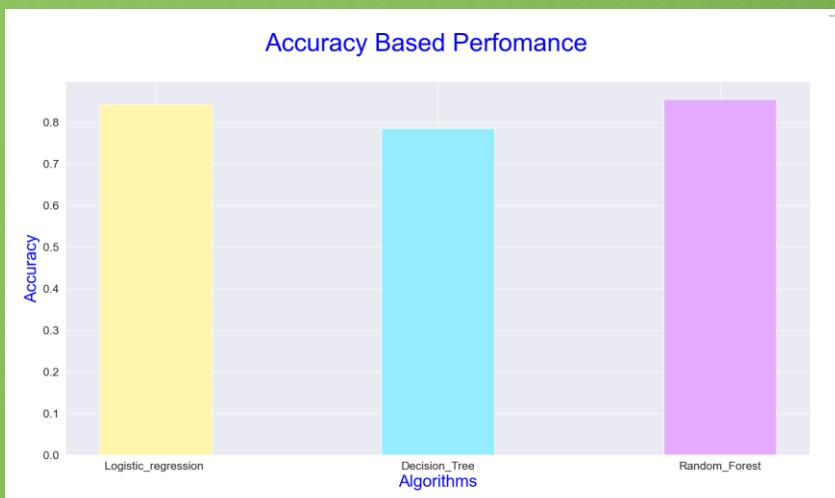
Prediction Frame

### WEATHER DETAILS

Enter City	<input type="text"/>	<input type="button" value="Click here"/>
Temperature	<input type="text"/>	
Description	<input type="text"/>	
Country	<input type="text"/>	



Weather Frame



Performance Frame

# Table

## Registration Table

FIELD NAME	DATATYPE	CONSTRAINT
UserID	INTEGER	PRIMARY KEY & AUTO INCREMENT
Password	VARCHAR(45)	NOT NULL
Contact	INTEGER	NOT NULL
Email	VARCHAR(45)	NOT NULL
Name	VARCHAR(45)	NOT NULL

# Advantages

- Give good accuracy for the considered dataset.
- Allows comparison between different supervised machine learning model.
- Can easily calculate mean absolute errors for each model and can calculate mean standard deviation for each dataset accuracy.
- Can be used for weather forecast.

# Experimental Result

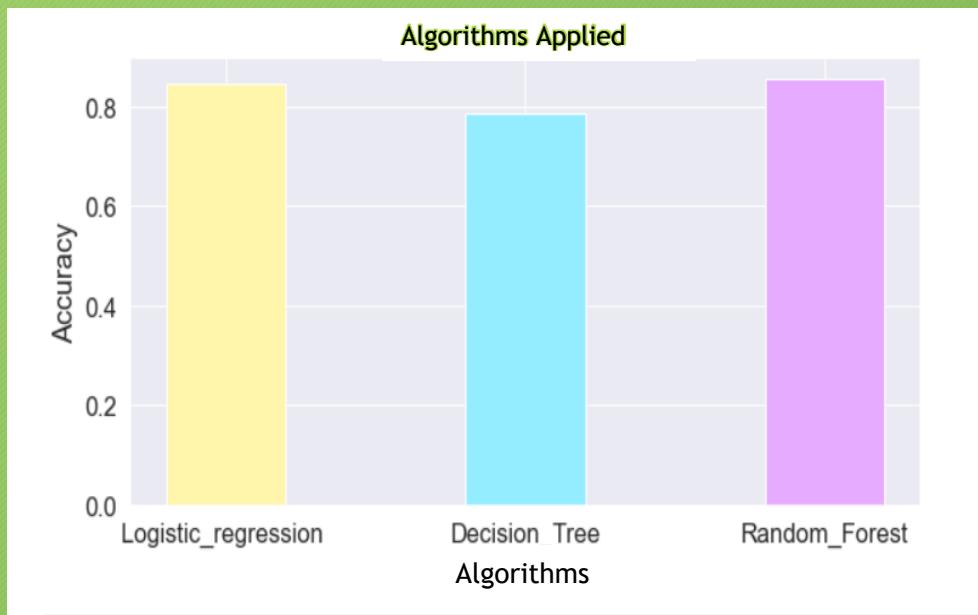
- The result obtained from the selected dataset on our project model:

	Accuracy	Time Taken Score
Logistic Regression	0.8377931713492036	1.413895606994629
Decision Tree	0.7842047891979325	1.3723666667938232
Random Forest	0.8564647139491544	24.639346837997437

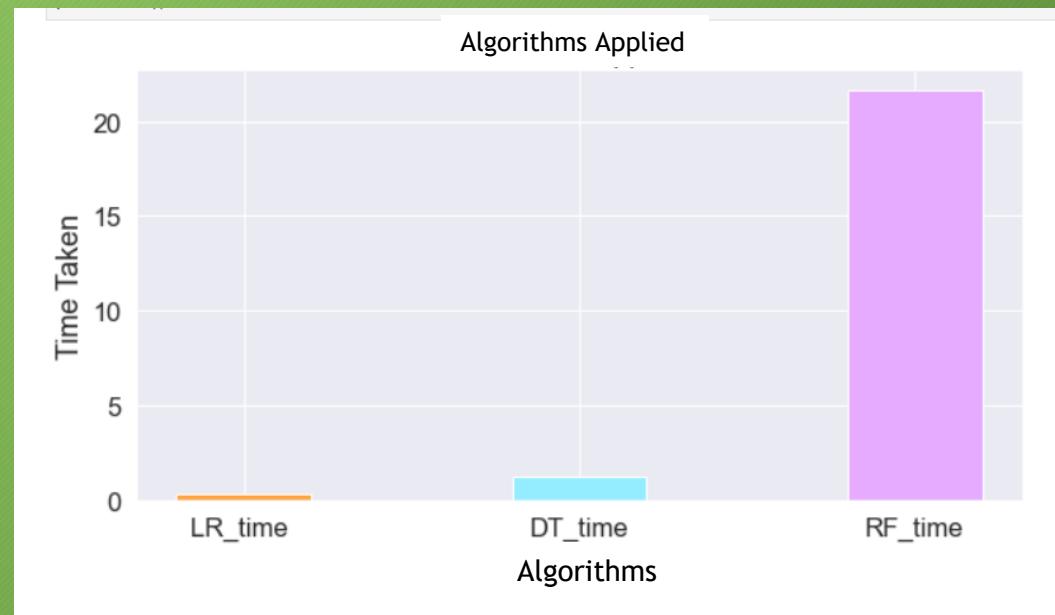
- For each regression model we have calculated accuracy and time taken score to get the best algorithm for deployment of our model.

# Comparison among applied model

Comparison of accuracy



Comparison of Time taken score



# Limitations

- This prediction is for short term.
- The prediction may change considerably by very small changes in the output parameters.
- It has been revealed that the standard algorithms may be unable to approximate the exact pattern of the data.

# Conclusion & Future Work

- Rainfall takes a vital role in deciding the weather condition and also a deciding factor of natural disasters such as flood, drought, etc. Agriculture sectors can avail the benefit of knowing weather condition in advance and take precautionary steps accordingly.
- This directly helps in improvement of national economy as well. In the necessity of an accurate and robust model to predict rainfall in the present work on Machine Learning based prediction has been proposed.
- A two step method coupled with a pre training phase of feature selection has been employed on a dataset collected dataset from Kaggle. Proposed model can efficiently predict rainfall status with an accuracy of 89.54% and quantitative rainfall prediction than other existing model.

# References

- APTRON Training Institute, guidance towards completion of project.
- Python.org documentation for the concepts of python libraries while developing the project.
- Kaggle website for the collection of dataset.
- Javatpoint tutorials for the concept of the machine learning algorithms .
- Various channels of ‘YouTube’ to get updated about the different predictions models and graphs.
- Creatly.com website for the Usecase Diagram and workflow diagram.
- Help from Research Paper by Nikhil Oswal from university of Ottawa, Canada (EECS).



# Thank You