

# Exploratory Analysis of RainFall Data in India for Agriculture.

Team ID : PNT2022TMID38676

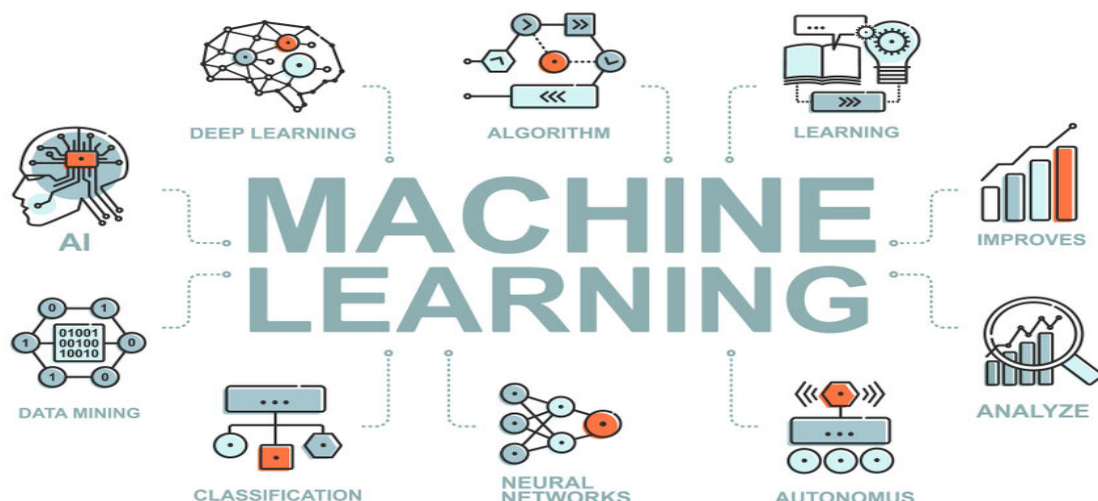
## Prior Knowledge :

**Prior knowledge** is the information and educational context a learner already has before they learn new information. A learner's understanding of educational material can be improved by taking advantage of their prior knowledge before dealing with the new material. General understanding about the way students learn states that the success of learning is determined by how much the learner already knows about a given topic or related topics.

**Prior knowledge** refers to the information, no matter how limited, a learner has at the start of learning a new topic. This knowledge will likely have been gathered over time in a variety of ways. If the prior knowledge is correct and consistent with the new information being taught, the effect on learning is positive. However, if prior knowledge conflicts with new information, the effect on new learning can be negative.

## What is Machine Learning?

Machine learning is a growing technology which enables computers to learn automatically from past data. Machine learning uses various algorithms for building mathematical models and making predictions using historical data or information. Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system, and many more. This machine learning tutorial gives you an introduction to machine learning along with the wide range of machine learning techniques such as Supervised, Unsupervised, and Reinforcement learning. You will learn about regression and classification models, clustering methods, hidden Markov models, and various sequential models.



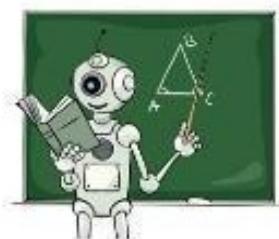
## Supervised and unsupervised learning:

In Supervised Learning, a machine is trained using 'labelled data. Datasets are said to be labelled when they contain both input and output parameters. In other words, the data has already been tagged with the correct answer. So, the technique mimics a classroom environment where a student learns in the presence of a supervisor or teacher. On the other hand, unsupervised learning algorithms let the models discover information and learn on their own data items in their experience.

Supervised machine learning is immensely helpful in solving real-world computational problems. The algorithm predicts outcomes for unforeseen data by learning from labeled training data. Therefore, it takes highly- skilled data scientists to build and deploy such models. Over time, data scientists also use their technical expertise to rebuild the models to maintain the integrity of the insights given.

It is neither based on supervised learning or unsupervised learning. Moreover, here the algorithms learn to react to an environment on their own. It is rapidly growing and moreover producing a variety of learning algorithms. These algorithms are useful in the field of Robotics, Gaming etc.

For a learning agent, there is always a start state and an end state. However, to reach the end state, there might be a different path. In the Reinforcement Learning Problem an agent tries to manipulate the environment. The agent travels from one state to another. The agent gets the reward(appreciation) on success but will not receive any reward or appreciation on failure. In this way, the agent learns from the environment.



**Supervised Learning**

VS



**Unsupervised Learning**

VS



**Reinforcement Learning**

## TYPES OF ML ALGORITHM PROBLEMS

- Classification
- Regression
- Clustering

### CLUSTERING, CLASSIFICATION, REGRESSION:

In the field of machine learning we all know the type of problems are different, sometimes we predict the value of the previous set of data – Where data learns from available dataset, or sometimes grouping them into some cluster. So today we are going to see what these terms are – **Clustering, Classification and Regression** means in the Data science field. Let's dive into this concept.

Generally machine learning algorithms are categorized on the basis of output type and type of problem that need to be addressed. So these algorithm are divided into three categories –

1. Classification
2. Regression
3. Clustering

### Classification

Classification is the type of supervised machine learning. For any given input, the classification algorithm helps in the prediction of the class of the output variables. There can be multiple types of classification – binary classification, multi-class classification.

#### Types of classification

- K – Nearest Neighbour
- Logistic regression
- Decision tree
- Random forest
- Naive Bayes
- SVM (Support vector machine)

### Regression

Regression is the type of supervised machine learning. When the output is continuous like age, height etc. one of the very popular regression algorithms is **Linear Regression**.

## **Clustering:**

Clustering is an unsupervised machine learning algorithm, it is used to group data points having similar characteristics as clusters.

Clustering is divided into two groups

1. Hard clustering – In hard clustering, the data point is assigned to one of the clusters only.
2. Soft clustering – It provides a probability of a data point to be in each of the clusters.

## **DATA VISUALIZATION**

In today's world, a lot of data is being generated on a daily basis. And sometimes to analyse this data for certain trends, patterns may become difficult if the data is in its raw format. To overcome this data visualization comes into play. Data visualization provides a good, organized pictorial representation of the data which makes it easier to understand, observe, analyse.

Python provides various libraries that come with different features for visualizing data. All these libraries come with different features and can support various types of graphs. In this tutorial, we will be discussing four such libraries.

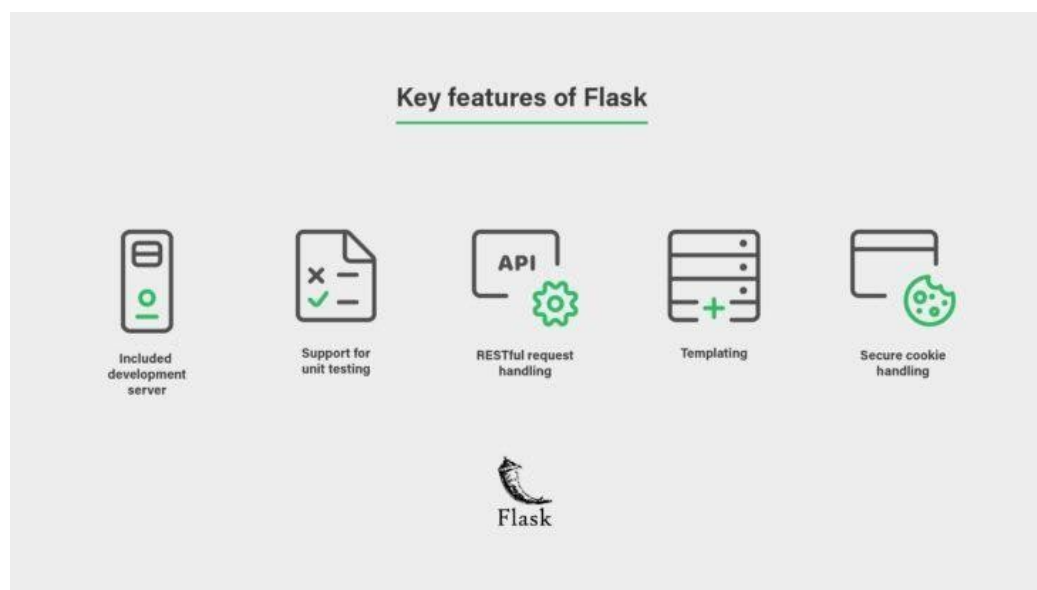
- Matplotlib
- Seaborn
- Bokeh
- Plotly



## Python Flask :

Flask API provides the basic and advanced concepts of the Python Flask framework. Our Flask tutorial is designed for beginners and professionals.

Flask is a web framework that provides libraries to build lightweight web applications in python. It is developed by **Armin Ronacher** who leads an international group of python enthusiasts (POCCO).



## What is Flask?

Flask is a web framework that provides libraries to build lightweight web applications in python. It is developed by **Armin Ronacher** who leads an international group of python enthusiasts (POCCO). It is based on the WSGI toolkit and jinja2 template engine. Flask is considered as a micro framework.

Industries do use Flask but more complex web apps are built on Django. Django's modularity makes it easier for multiple developers to work on different apps in one project. Beginners prefer Flask over Django though, because of its minimalistic design.

For example, the development team at popular companies like **Netflix, Lyft, and Reddit** all use Flask as backend development for their mobile and web applications.

## WSGI

Web Server Gateway Interface (WSGI) has been adopted as a standard for Python web application development. WSGI is a specification for a universal interface between the web server and the web applications.

## Werkzeug

It is a WSGI toolkit, which implements requests, response objects, and other utility functions. This enables building a web framework on top of it. The Flask framework uses Werkzeug as one of its bases.

## Jinja2

Jinja2 is a popular templating engine for Python. A web templating system combines a template with a certain data source to render dynamic web pages.

Flask is often referred to as a micro framework. It aims to keep the core of an application simple yet extensible. Flask does not have a built-in abstraction layer for database handling, nor does it have form validation support. Instead, Flask supports the extensions to add such functionality to the application.