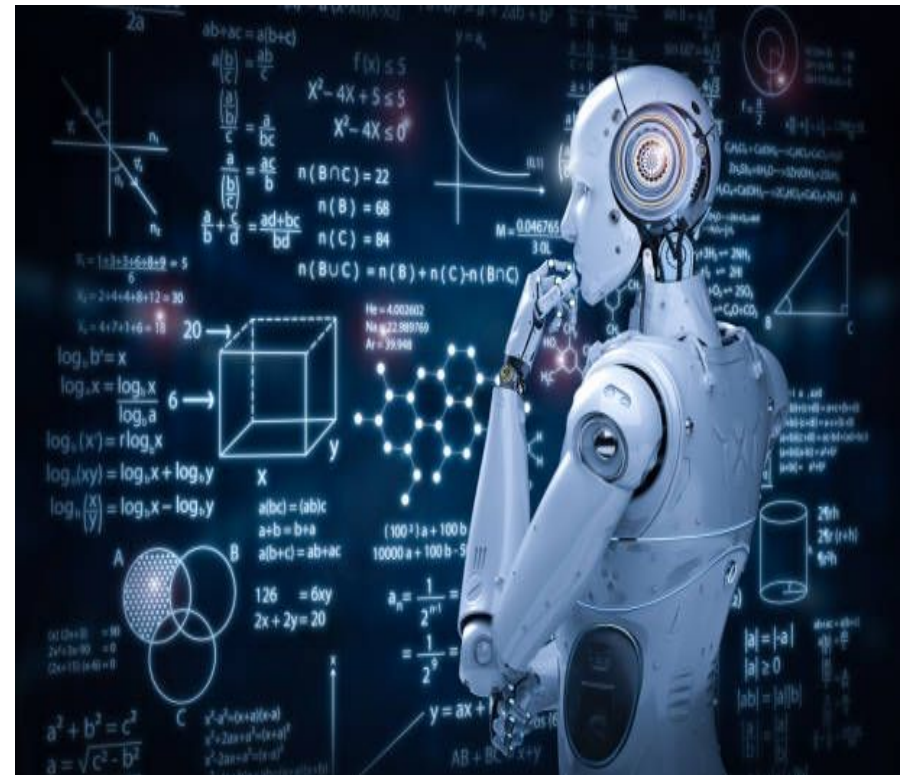# Machine Learning Activities
Prof. Rajesh Pradhan

1. Collecting Data :-

- As you know, machines initially learn from the **data** that you give them.

- It is of the utmost importance to collect reliable data so that your machine learning model can find the correct patterns.

- The quality of the data that you feed to the machine will determine how accurate your model is.

1. Collecting Data :-   Cont...

- If you have incorrect or outdated data, you will have wrong outcomes or predictions which are not relevant.

- Make sure you use data from a reliable source, as it will directly affect the outcome of your model.

- Good data is relevant, contains very few missing and repeated values, and has a good representation of the various subcategories/classes present.

2. Preparing the Data :-

After you have your data, you have to prepare it. You can do this by :-

- Putting together all the data you have and randomizing it.
- This helps make sure that data is evenly distributed, and the ordering does not affect the learning process.
- Cleaning the data to remove unwanted data, missing values, rows, and columns, duplicate values, data type conversion, etc.

2. Preparing the Data :-   Cont….

- You might even have to restructure the dataset and change the rows and columns or index of rows and columns.

- Visualize the data to understand how it is structured and understand the relationship between various variables and classes .

- Splitting the cleaned data into two sets - a training set and a testing set.

- The training set is the set your model learns from. A testing set is used to check the accuracy of your model after training.


Prof. Rajesh Pradhan   **GPG, SURAT**

3. Choosing a Model :-

- A machine learning model determines the output you get after running a machine learning algorithm on the collected data.

- It is important to choose a model which is relevant to the task at hand.

- Over the years, scientists and engineers developed various models suited for different tasks like speech recognition, image recognition, prediction, etc.
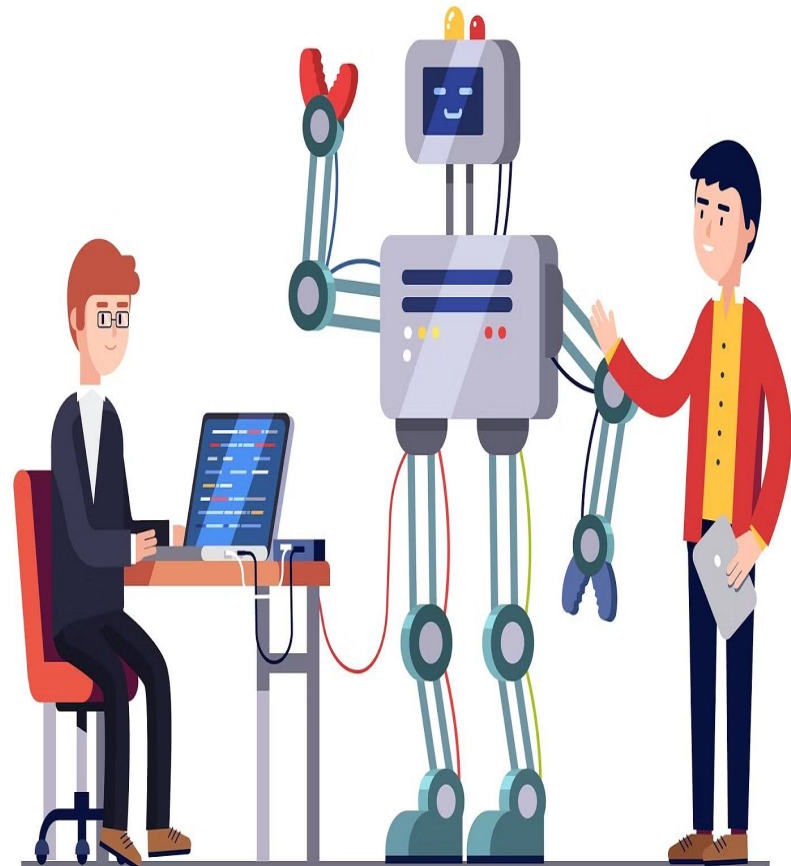
4. Training the Model :-

- Training is the most important step in machine learning.
- In training, you pass the prepared data to your machine learning model to find patterns and make predictions.
- It results in the model learning from the data so that it can accomplish the task set.
- Over time, with training, the model gets better at predicting.

Prof. Rajesh Pradhan   **GPG, SURAT**
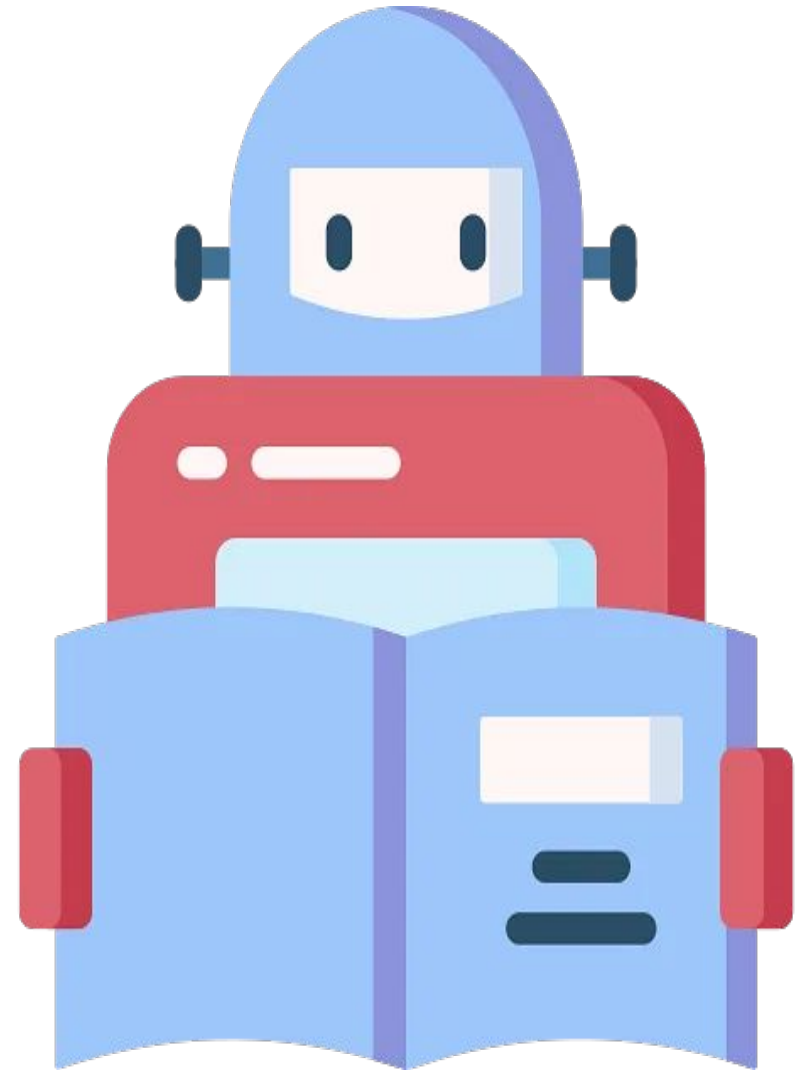
5. Evaluating the Model :-

- After training your model, you have to check to see how it's performing.
- This is done by testing the performance of the model on previously unseen data.
- If testing was done on the same data which is used for training, you will not get an accurate measure, as the model is already used to the data, and finds the same patterns in it, as it previously did.
- This will give you disproportionately high accuracy.

6. Parameter Tuning :-

- Once you have created and evaluated your model, see if its accuracy can be improved in any way.

- This is done by tuning the parameters present in your model.

- Parameters are the variables in the model that the programmer generally decides.

- At a particular value of your parameter, the accuracy will be the maximum.

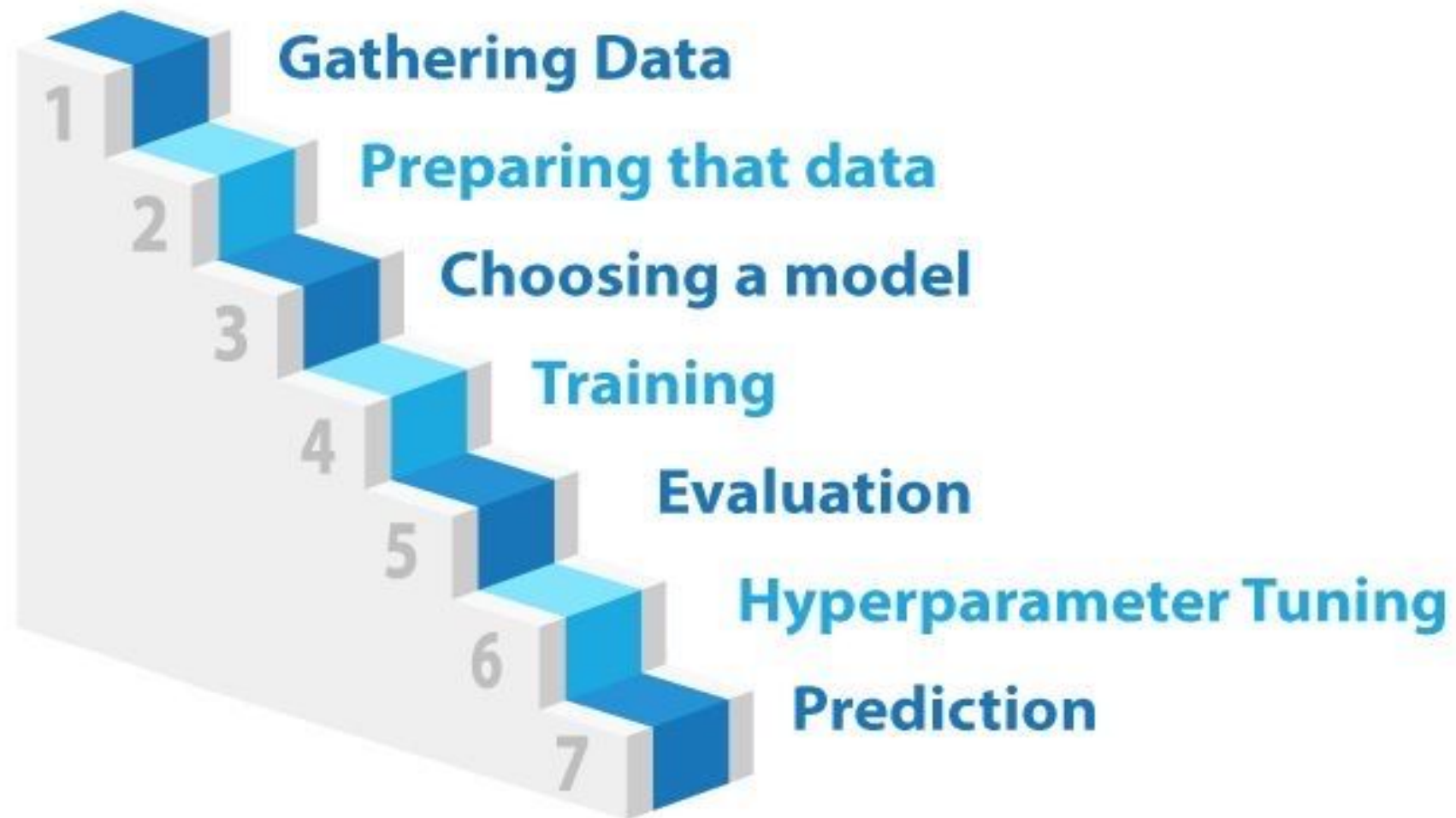- Parameter tuning refers to finding these values.
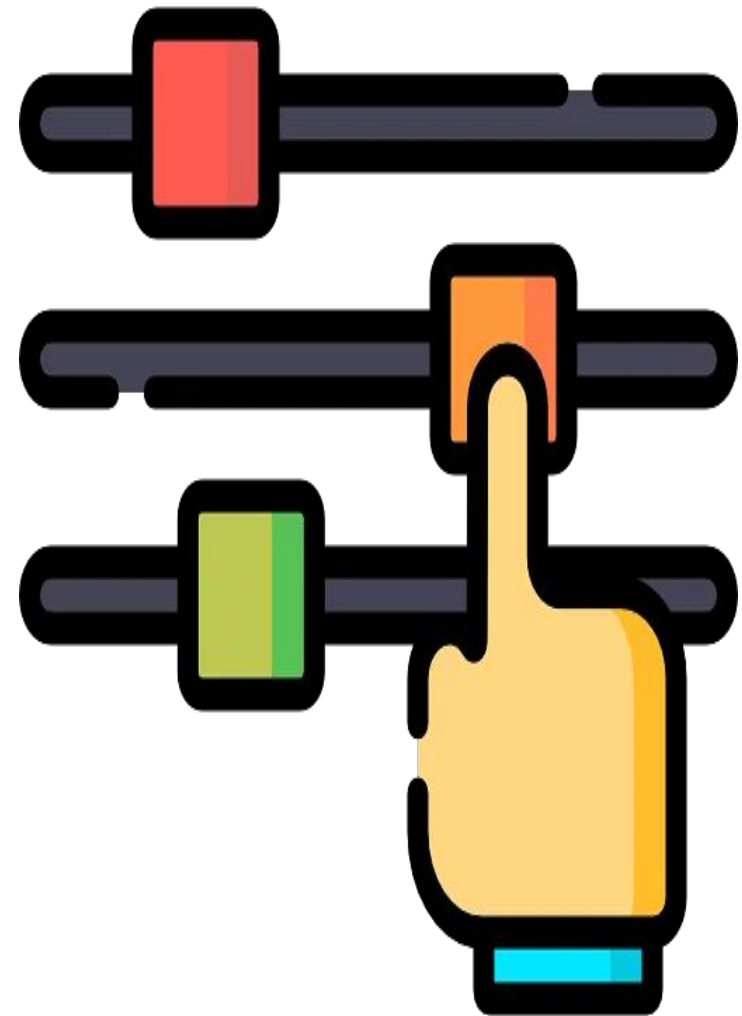
7. Making Predictions :-

- **In the end, you can use your model on unseen data to make predictions accurately .**

7 steps of Machine Learning

1. **Gathering Data**
2. **Preparing that data**
3. **Choosing a model**
4. **Training**
5. **Evaluation**
6. **Hyperparameter Tuning**
7. **Prediction**

References :-

https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-steps

https://www.youtube.com/watch?v=nKW8Ndu7Mjw

Prof. Rajesh Pradhan   **GPG, SURAT**