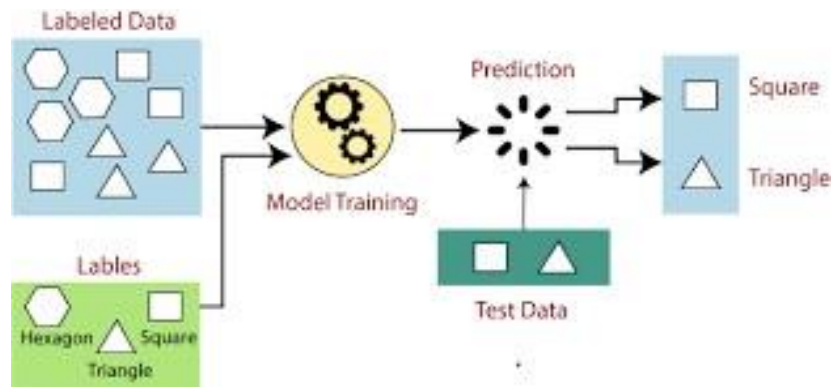# Prior Knowledge

| Date	17 November 2022
Team ID	PNT2022TMID48329
Project Name	Project – Early Detection of Chronic Kidney Disease using Machine Learning

Supervised Learning:

- Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence.
- It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately.
- Supervised learning uses a training set to teach models to yield the desired output. This training dataset includes inputs and correct outputs, which allow the model to learn over time. The algorithm measures its accuracy through the loss function, adjusting until the error has been sufficiently minimized.
- Algorithms,
  1. Naïve Bayes
  2. Logistic Regression
  3. K Nearest Neighbour
  4. Support Vector Machine

Unsupervised Learning:

Raw Data   Output
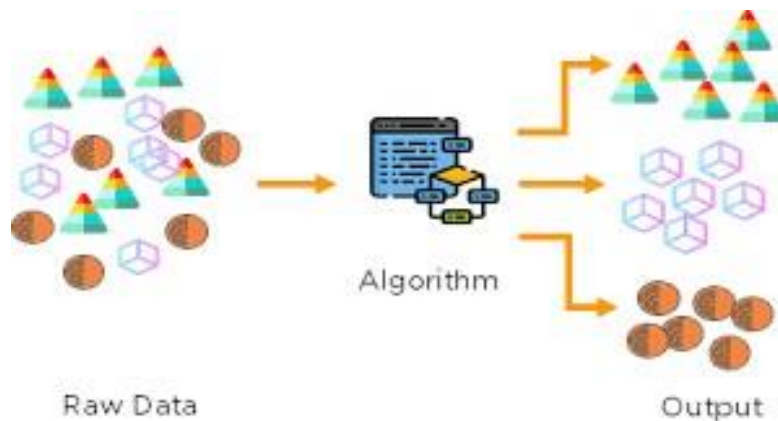
- Unsupervised learning is a type of machine learning in which models are trained using unlabeled dataset and are allowed to act on that data without any supervision.
- Unsupervised learning cannot be directly applied to a regression or classification problem because unlike supervised learning, we have the input data but no corresponding output data. The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format.
- Algorithms,
  1. Hierarchal clustering
  2. Anomaly detection
  3. Neural Networks
 4. Principle Component Analysis

Regression:
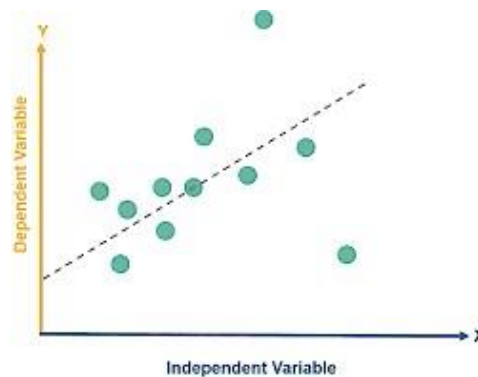
- Regression is used to understand the relationship between dependent and independent variables. It is commonly used to make projections, such as for sales revenue for a given business.

Clustering:
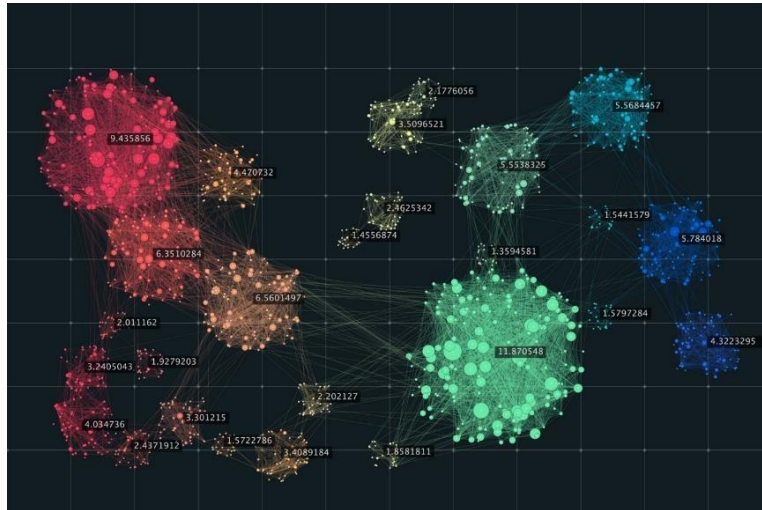
- Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. It is basically a collection of objects on the basis of similarity and dissimilarity between them.

Classification:
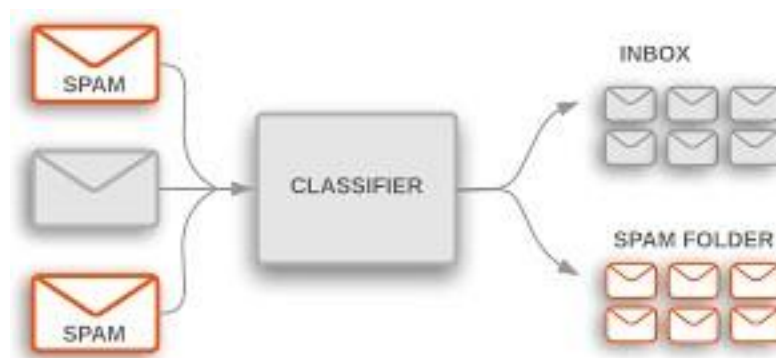
- Classification uses an algorithm to accurately assign test data into specific categories. It recognizes specific entities within the dataset and attempts to draw some conclusions on how those entities should be labeled or defined.

Flask:

Logistic Regression: