

PRACTICAL 8

Aim :- Performing Clustering using Scikit Learn for a given problem statement.

Theory :- Un-supervised Machine Learning.

It is a branch of ML that deals with unlabeled data. The algorithms are tasked with finding patterns and relationships within the data without any prior knowledge of data's meaning.

★ Clustering.

- Clustering is an unsupervised machine learning algorithm which groups the unlabeled dataset into different clusters.
- It assigns data points to one of the K clusters depending on their distance from center of the clusters.
- The goal of clustering is to divide the set of data points into no. of groups so that these points are comparable to one another in its group & different from data points of others group.
- The advantages of clustering are that hidden patterns are found & can be used to explore new data. Clustering is widely used in customer segmentation, Fraud Detection, Recommendation Systems & Natural Language Processing (NLP).