UNSUPERVISED LEARNING

→ Unsupervised learning is where we have only input data (X) and no corresponding output variable.

→ The goal of unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.

→ These are called unsupervised learning because unlike supervised learning there is no correct answer. Algorithm are left on their own devises to discover and present the interesting structure in the data.

→ Unsupervised learning model can be further grouped into clustering and association problems.

① Clustering → A clustering problem is where we want to discover the inherit groupings in the data, such as grouping customers by purchasing customers.

② Association → An association rule mining problem is where we want to discover rules that describe large portions of data, such as people that buy X also tend to buy Y.

Different type of Clustering →

Exclusive (Partioning) – In this clustering method, every data is clustered grouped in such a way that one data can belong to one cluster only. Example – K Means.

Agglomerative – In this clustering technique, every data is cluster. The iterative unions between the two nearest cluster reduce the number of clusters. Example – Hierarchical clustering.

→ Clustering analysis learns to group, or segment, datasets with shared attributes in order to extrapolate algorithmic relationships.

→ Association rule mining can be of two types – 1) Apriori
2) Eclat

→ Different type of Association rule mining –

1) Apriori → Apriori algorithm uses frequent itemsets to generate association rules and it is designed to work on the datasets that contain transactions. With the help of association rule, it determines how strongly or how weak two objects are connected. It is the iterative process for finding frequent itemset from the dataset.

2) Eclat – Full form is Equivalence Class Clustering & bottom up lattice Traversal. More efficient and scalable version of Apriori algo. Apriori works on horizontal sense (Breadth first Search, BFS), Eclat works on DFS (Depth first search) vertical approach.