Clustering

Sonal Ghanshani

Difference

Supervised learning	Unsupervised learning
Mathematical relationship Single answer More insights, inference, Quantify impact Used in predictive analytics	No mathematical relationship No single correct answer Multiple interpretations Limited information, insight Can't quantify implications Used in EDA - exploratory data analysis

Unsupervised learning

- Clustering: Hierarchical clustering
- Association: Market basket analysis

What is the basis of clustering

- Similarity
- How much similarity? Distances.
- How many clusters:
- Max no. of clusters = no. of datapoints = each datapoint is one cluster
- Min clusters = 1 = 1 all datapoints together