Unsupervised Learning and Clustering

1. Introduction:

Unsupervised learning discovers hidden patterns in data without labels. Its mainly used for clustering and dimensionality reduction.

2. K-Means Clustering:

- Divides data into K clusters
- Iteratively assigns points to closest centroid
- Requires pre-defined K value

3. Hierarchical Clustering:

- Builds a tree of clusters (dendrogram)
- Agglomerative or divisive
- Doesnt require K

4. DBSCAN:

- Density-based clustering
- Detects arbitrary-shaped clusters and outliers

5. PCA - Principal Component Analysis:

- Reduces dimensionality
- Projects data onto lower-dimensional space
- Captures maximum variance

6. Applications:

- Market segmentation
- Anomaly detection
- Image compression

7. Summary:

Unsupervised learning is powerful for data exploration. Choosing the right technique depends on data size, noise, and structure.