

Devanshu Surana

PC-23, 1032210755

Panel C, Batch C1

## DEC Lab Assignment 8

Problem Statement: Consider a suitable dataset and apply different clustering techniques.

Objectives :- To build the cluster using different cluster techniques

- To implement the K-means and Hierarchical clustering
- To check the performance of clustering algorithm

### Conclusion :

In conclusion the clustering is a powerful technique in machine learning that helps uncover hidden patterns and structures within datasets. Demonstrated their implementation using Python programming language and the sci-kit learn library.

### FAQ's

1) Differentiate bet<sup>n</sup> unsupervised and supervised learning.

→ In supervised learning the algo is trained on a labeled dataset, where the input data is paired with corresponding output labels. The goal is to learn a mapping from inputs to outputs and the model makes predictions based on this learned mapping.

In unsupervised learning the algorithm is given unlabelled data and must find patterns or relationships within the data without explicit guidance. The goal is often to discover the underlying structure or distribution of the data.

Q2) What is the purpose of using cluster analysis in data science?

→ Cluster analysis in data science is used to group similar data points together on their inherent characteristics. The purpose is to uncover patterns, identify natural groupings and gain insights into the structure of the data, facilitating tasks like segmentation, anomaly detection and pattern recognition.

Q3) What are the different types of clustering algorithms available?

→ There are several types of clustering algorithms, including

- 1) K-means clustering
- 2) Hierarchical clustering
- 3) DBSCAN (Density based Spatial Clustering of Applications with noise)
- 4) Gaussian Mixture Models
- 5) Agglomerative clustering
- 6) Affinity propagation
- 7) Mean shift clustering
- 8) Self-organizing Maps (SOM)