SLIIT UNI
THE KNOWLEDGE UNIVERSITY

IT2011 - Artificial Intelligence and Machine Learning

Department of Information Technology, Faculty of Computing

Year 2 semester 1 (2025)

Workshop 03

Unsupervised Learning Workshop

Part 1 – K Means Clustering

1. How do you import the necessary Python libraries for K-Means clustering?
2. How can you load a dataset using pandas and display its first five rows?
3. How do you select the relevant features for clustering?
4. How can you visualize the dataset using a scatter plot before applying clustering?
5. How do you create and fit a K-Means clustering model with a chosen number of clusters ?
6. How can you retrieve the cluster labels assigned to each data point?
7. How do you add these cluster labels as a new column to the dataset?
8. How can you visualize the clusters and their centroids using a scatter plot?
9. How do you calculate the Within-Cluster Sum of Squares (WCSS) for different numbers of clusters?
10. How can you use the Elbow Method to determine the optimal number of clusters?
11. How do you re-train the K-Means model with the optimal number of clusters and plot the final clusters?

Part 2 – Hierarchical Clustering

12. How do you import the required Python libraries for hierarchical clustering and visualization?
13. How can you load the dataset using pandas and display its first few rows?
14. How do you select the relevant features for clustering?
15. How can you plot a dendrogram to visualize the hierarchical clustering process?

16. How do you perform Agglomerative Hierarchical Clustering with a chosen number of clusters?

17. How can you retrieve the cluster labels assigned to each data point?

18. How do you add the hierarchical cluster labels as a new column to the dataset?

19. How can you visualize the resulting clusters with different colors on a scatter plot?

20. How do you compare the clusters obtained from hierarchical clustering with those from K-Means?