FACULTY OF COMPUTER SCIENCE AND ENGINEERING
Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi

Lab Duration:3 hr.	CS351L Intro to AI Lab
	Marks: 10 Lab No: 8

Instructor: Memoona Saleem

◆ **Task Statement:**

1.1 Objective:

Utilize clustering algorithms to organize music preferences and construct a personalized playlist recommendation system.

1.2 Dataset:

Use the "Million Song Dataset" containing a diverse array of music attributes like artist, genre, tempo, key, duration, etc.

Data Acquisition and Preprocessing:

Obtain the Million Song Dataset,

Preprocess the dataset to extract relevant features for clustering (e.g., artist, genre, tempo).

Clustering Music Preferences:

Apply various clustering techniques (e.g., K-Means, Hierarchical Clustering) to group similar music preferences together.

Explore different representations of music features to identify meaningful clusters.

Cluster Analysis and Visualization:

Analyze formed clusters to understand common music characteristics within each cluster.

Personalized Playlist Recommendation System:

Develop a personalized playlist recommendation mechanism based on the clusters created.

User-Specific Playlist Recommendation:

✓ For a given user, determine the cluster label that represents their music preferences. Retrieve songs or artists that are popular or highly liked within that

cluster. Recommend songs or create a playlist for the user comprising tracks that align with the preferences of their cluster.

✓ Creating Cluster-Centric Playlists: Generate playlists based on each cluster's preferences.

✓ Aggregate the most popular or frequently listened-to songs within each cluster. Create cluster-centric playlists that represent the collective music tastes of users within a cluster.

Task 2 : Lab Activity: Clustering Analysis

1. Load a dataset suitable for clustering analysis.
2. Explore the dataset through summary statistics and visualizations.
3. Preprocess the data by handling missing values and scaling features.
4. Implement the K-means clustering algorithm.
5. Experiment with different values of K (number of clusters) and evaluate the results.
6. Visualize the clustering results using appropriate plots.
7. Explore other clustering algorithms (e.g., hierarchical clustering, DBSCAN).
8. Compare the performance and characteristics of different clustering algorithms.