Guided Projects Artificial Intelligence & Machine Learning

Guided Projects: Unsupervised Learning

Spherical K-Means: Pattern Discovery in Textures

The **classical k-means** method of clustering minimizes the sum of squared distances between **cluster centres and cluster members**. The intuition is that the **radial distance** from the Cluster- Centre should be similar for all elements of that cluster. **The spherical k-means** algorithm, however, is equivalent to the k-means algorithm with cosine similarity, a popular method for clustering high-dimensional data. The idea is to set the centre of each cluster such that it makes the angle between components both **uniform and minimal**.

Question:

Generate a dummy dataset using **Scikit-Learn having high dimensionality (number of features >10) and total 4 classes**. For this dataset, first implement **K-Means clustering** and then use the clusters for **classification purpose**. Now using the same dataset, implement spherical clustering and then check accuracy for classification. Notice the change in accuracy. You may also plot the obtained clusters from both the methods using **t-SNE plots** or by projecting data into **two dimensions using PCA.**