

Assignment 8

Page : _____

Date : / /

Unsupervised learning Kmeans

Aim: Implement Kmean algorithm for clustering to create a cluster on dataset (Using python).

Objective: To understand the concept of clustering.

To implement Kmeans clustering algorithm.

Theory: Kmeans clustering aims to partition n observation into K clusters in which each observation belongs to the cluster with the nearest mean, serving as a program.

Working: The Kmeans clustering algorithm attempts to split a given anonymous data set.

Initially K number of so called centroids are chosen randomly and are unique.

Then, clusters are formed Based on these centroids, and the new centroid are selected. This process is iteratively repeated until centroids stop changing.

- Advantages

- Simple to implement
- Scale to large dataset
- Adaptive

- Disadvantages

- Sensitive to outliers
- Depends on initial values.

• Conclusion.

Thus, we have understood the concept of K-means and successfully implemented it to cluster our dataset.