

Assignment #1 – K-Means Clustering Visualization Algorithm  
Course: Software Engineering for Distributed Systems

Definition of K-Mean Clustering: K-Means clustering is an unsupervised machine learning algorithm used to group similar data points into clusters.

Programming Language: C++ Library

Used: OpenCV

Description: This project implements the K-Means clustering algorithm in C++ and visualizes its execution step-by-step using OpenCV. The goal is to help users understand how data points are assigned to clusters and how centroids move iteratively until convergence. The animation clearly illustrates cluster formation and centroid adjustments over time.

Input:

- Number of clusters- Selecting number of clusters to divide data points.
- Initial centroid coordinates (entered via console).
- Animation delay (in milliseconds) for controlling speed.

Output: A graphical OpenCV window that shows:

- Colored data points grouped by cluster Centroids visualized as X-marks.
- Each iteration step is visualized sequentially.
- Final result displayed after convergence.

How k-mean algorithm works:

1. Initialise – Initialise k centroids as per number of clusters required.
2. Updates - Updating data points and making clusters as per distance from centroid.
3. Repeat- Iterate until the algorithm converges.