



भारतीय सूचना प्रौद्योगिकी संस्थान गुवाहाटी
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY GUWAHATI

CS 360: Machine Learning Lab
Practice Assignment 6

Instructions: This is only for practice. Complete it by 12:00 PM today. Your completion will be reviewed by the Teaching Assistants.

1. Download [The Mall customer dataset](#).

Write a program to do the followings:

- (a) Read the dataset.
 - (b) Cluster the dataset into different categories of customers using DBSCAN Clustering method(**without using in-built Python library/ packages**).
Perform hyper-parameter tuning (using grid-search method) for the following hyper parameters considering the given values:
 - minpts: {4 - 10}
 - epsilon: {2, 2.5, 3, 3.5, 4}
 - (c) Determine the optimal value of epsilon using [k-distance graph](#) , w.r.t. the optimal minpts-value from (b), and then apply DBSCAN clustering. Consider the optimal minpts-value as same as obtained in 1(b).
 - (d) Report and visualize the obtained final clusters for both 1(b) and 1(c), and write the result analysis.
2. (a) Cluster the above [Mall customer dataset](#) into different categories of customers using Hierarchical clustering method, with single-linkage, complete-linkage and centroid-linkage strategies separately(**without using in-built Python library/ packages**).
 - (b) Report and visualize the obtained final clusters for all the three strategies, and write the result analysis.