PESIT Department of Computer Science and Engineering

Course: Data Mining
Semester: 2016 Spring (January – May)
Instructor: BNR (Dr. B. Narsing Rao)

Assignment: 10
Topic: Image Clustering
Due by: Midnight on **Tuesday, March 22, 2016**
Method: Send program and output by email using standard naming conventions.

Use a modified version of the DBSCAN algorithm (that takes into account color distance) to identify clusters in the accompanying image file **dm-a10.gif.**

The output of your program should be the following:

1. Number of points (pixels) in the original image
2. Number of clusters and noise points
3. A table that shows for each cluster the following: cluster number, cluster size, average cluster color
4. A regeneration of the image, using different colors, based on the clusters found by the program

Answer the following questions:

1. What criterion did you use to decide whether a pixel is a core object?
2. What values of epsilon (in pixel space and in color space) and MinPts were used and why? (Note that you may want to try out several sets of values before deciding on one set)
3. What logic was used to determine whether a pixel was in the neighborhood of another pixel?

Submit the following using standard naming conventions:

- Program source
- Program output
- Answers to questions