# Assignment 5 Instructions

Two data set files dataset1.npy and dataset2.npy have been provided. These files contain semi-processed LiDAR data. Share.py file provides the framework to work with LiDAR data. Follow the instructions in the document to complete task 1-3. The source and the result should be committed to github to your personal folder. Try to separate the tasks with different version of the commit.

Link to discussion: https://emaintenancelab-my.sharepoint.com/:v:/p/jaya\_kumari/ER4mFOUrtklDjlBYY0Yn\_xoBOXxXTcPwIhX0M46rVQ3VGg?e=11RI6n&isSPOFile=1

**Outcome**

* To learn processing of raw LiDAR data using Python
* To learn application of ML methods and their tuning as per the data set
* Documentation of the project on Github

## Task 1 (3)

* Find the best value for the ground level
* One way to do it is using a histogram np.histogram
* Update the function get\_ground\_level() with your changes

### Submission

For both the datasets

* Report the ground level in the readme file in your github project
* Add the histogram plots to your project Readme

## Task 2 (4)

* Find an optimized value for eps.
* Plot the elbow and extract the optimal value from the plot
* Apply DBSCAN again with the new eps value and confirm visually that clusters are proper

Use the following links to understand DBSCAN and calculation of optimal epsilon value

https://www.analyticsvidhya.com/blog/2020/09/how-dbscan-clustering-works/

https://machinelearningknowledge.ai/tutorial-for-dbscan-clustering-in-python-sklearn/

### Submission

For both the datasets

* Report the optimal value of eps in the Readme to your github project
* Add the elbow plots to your github project Readme
* Add the cluster plots to your github project Readme

## Task 3 (5)

* Find the largest cluster, since that should be the catenary, beware of the noise cluster.
* Use the x,y span for the clusters to find the largest cluster

### Submission

For both the datasets

* Report min(x), min(y), max(x), max(y) for the catenary cluster in the Readme of your github project
* Add the plot of the catenary cluster to the Readme