

Friday April 12 - with Dr.Diggans (mentor)

Last week: reorganized the code of class-by-class approach. Tried the grid approach suggested by Dr.Diggans.

Problem: the new approach we came up with last week in order to figure out a clearer correlation between the entropy value and test accuracy is not working well. Even if we combined all the most representative subsets from each class, the overall entropy value of the combined subsample is not outstanding. As a result, the test accuracy of the trained CNN using this combined subsample did not show a prominent performance in classification accuracy.

Next week: consider other approaches like comparing the eigenvectors between a subset and the whole training set (maybe using cosine similarity) to see if this subset is representative of the training set so that we can employ it to do data efficient training on CNN models.