

Friday April 27- with Dr.Diggans (mentor)

Last week: think about other possible approaches to attain optimal subset to train ML models.

1. To resolve the embedding function problem before (need to use data to train it before it can be used to extract feature space), we can try to flatten the raw pixels of the images directly to form a matrix and do SVD on it directly.
2. To create a biased subset, we can also try to oversample from one particular class and subsample from the remaining classes.
3. Improvement on the Boltzmann Shannon Interaction Entropy: try to consider the direction of the eigenvector as well besides comparing the entropy value between different subsets.
4. Artificial Noise Injection and Data Augmentation to change the original pixel values of the images so that we might be able to attain some distinct discrepancy in terms of entropy values of the subsets.

Next week: start to implement these approaches.