

ECE/CS 559 - Fall 2022 - HW #6

Due: 11/04/2022, 11:00pm Chicago time.

Q1. (100 pts) Download hw6.py from Piazza under “General Resources” folder. This script trains a denoising autoencoder.

- (a) Explain how the denoising autoencoder works.
- (b) The model involves batch normalization. Learn what batch normalization is and briefly explain here.
- (c) Write a script (to be appended to the end of the script - see the file) that generates 9 random images of digits arranged in a 3×3 matrix. Labels are not important. Include the images with your report. Comment on the results.
- (d) Write a script that clusters the images in the dataset (to be appended at the end of the script - see the file). To do this, first obtain all encoder outputs for all elements of the training set and cluster these outputs using the k -means algorithm. You will obtain a sequence of cluster assignments, like (1,1,2,2,2) under the assumption of 5 dataset elements and 3 clusters (you will have 48000 dataset elements and 10 clusters). You also have true labels from the dataset itself, such as (2,2,1,1,1). Now find the difference between the assignments found by k -means and the true labels to obtain the accuracy. You should find a good reassignment of indices such that the accuracy is maximized, because k -means is agnostic to labels. For example, the above example should yield zero errors, because we can assign 1s to 2s and 2s to 1s achieving perfect accuracy between the k -means labels and the true labels. Report the accuracy you obtain over the training set, and describe your index reassignment algorithm.
- (e) Update only the new code you have written with the file name 06-IDNumber-LastName.py to Box. Also include your code on your report.