

Project 2

Following is a task of detecting anomaly using auto-encoder. Consider we have unknown dataset `X_train` and `X_test`, both of them are from the same dataset. In the dataset, most of the points obey a certain distribution, while the rest obey a different distribution and is far from most of the points, which are called anomalous points. To complete the task, please:

- Write an auto-encoder in PyTorch to do the forward propogation.
- Implement a train function and set your favorite parameters to train the network with the input as `X_train.npy` (located in the folder `Project2`) and using mean square error loss function.
- Using the test data `X_test.npy` (located in the folder `Project2`) to display the distribution of Mean Square Error(MSE) with your trained auto-encoder.
- Assuming anomalous points in the testset are well separated in the distribution of MSE, give the number of anomalous points in the testset by using a cut on the value of MSE.