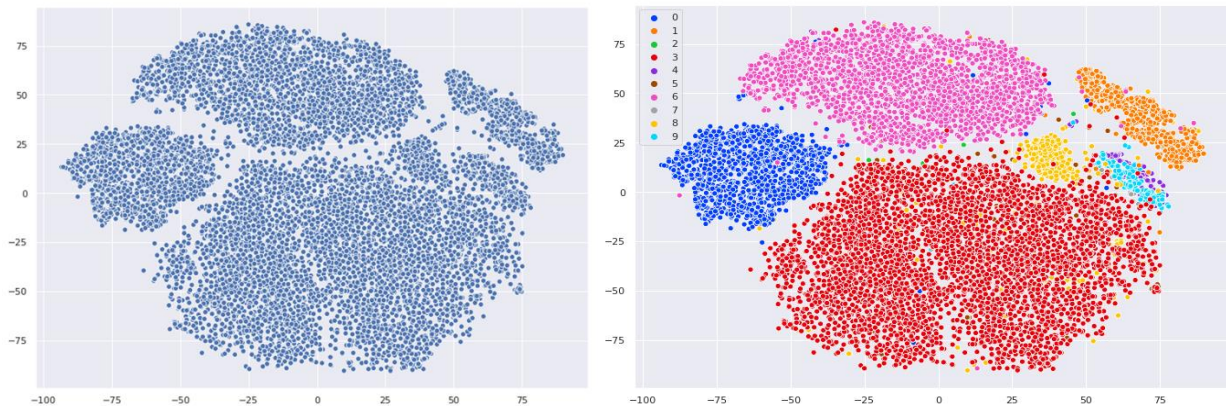


Assignment 9

Huh, wonder what's out there? – **Active Learning Based Rare Class Discovery**



In this assignment you write code “from scratch” to implement an active learning based rare class discovery algorithm, and test it on both the MNIST_C derived dataset and the MNIST-skewed dataset. Unlike the previous assignment, the goal is not to build a classifier with high accuracy; the goal is to discover all the classes with as few queries to the oracle as possible, using the information gained along the way to inform your choice about which point to query next.

Specifically, you will:

- Get the provided datasets from D2L. Then for each dataset:
- Visualize the data w/ labels using 2 or 3-D tSNE.
- Write your own version of an active learning rare class discovery algorithm.
- Run your code on the dataset and keep track of the number of classes discovered vs. number of queries.
- Plot that (# classes discovered vs. # queries).
- Rerun the same experiment using a random query strategy.
- Plot the results from the random algorithm on the same plot.

You will turn in both the code and a short report, discussing what you've implemented, how well it worked, what you've learned, etc.

Remember, you must understand what you turn in – you may be asked to explain it to me and/or the class.