

SMAI Mid Evals Report

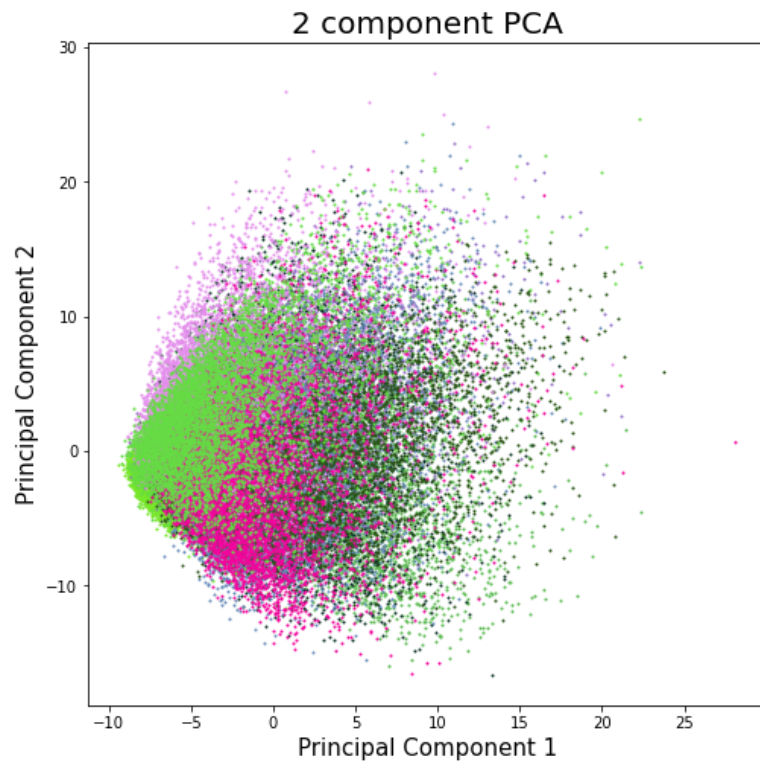
Progress

- Cleaned all 4 datasets and loaded into our jupyter notebook
 - Loaded MNIST using the python-mnist python library
 - Loaded Oliveretti from sklearn datasets
 - Downloaded COIL-20, restructured and classified
 - Requested and downloaded Animals-10, loaded with appropriate classifications
- Tested dimensionality reduction using PCA
 - Reduced all the datasets to 2 components using PCA and plotted on graphs
 - Reduced to 30 dimensions for testing t-SNE in the future as done in the original paper
- Started implementing t-SNE taking the base as SNE
 - Implemented functions to calculate:
 - Euclidean distances
 - Exponential Average
 - Perplexity Calculation
 - Binary Search for sigmas
 - Q joint probability
 - P joint probability

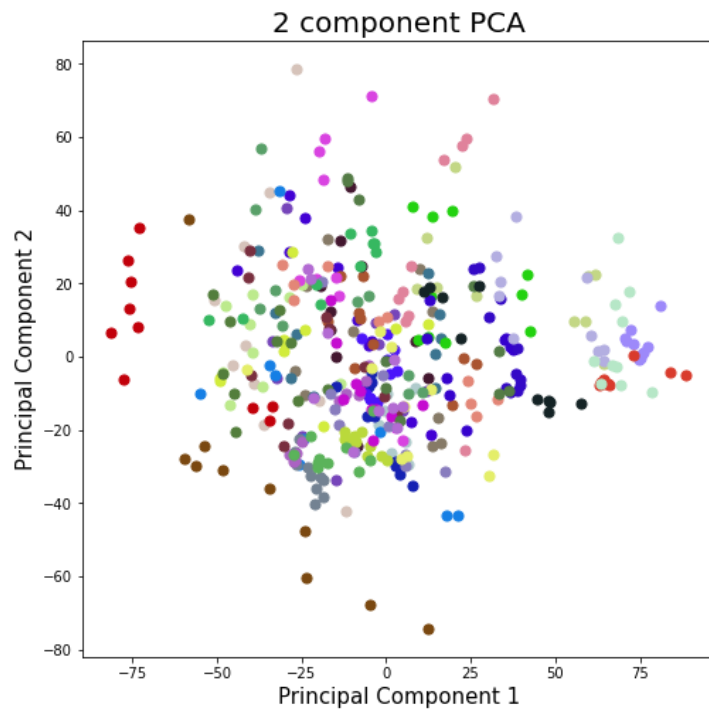
Link to code and datasets: <https://github.com/kjain1810/t-SNE>

Datasets

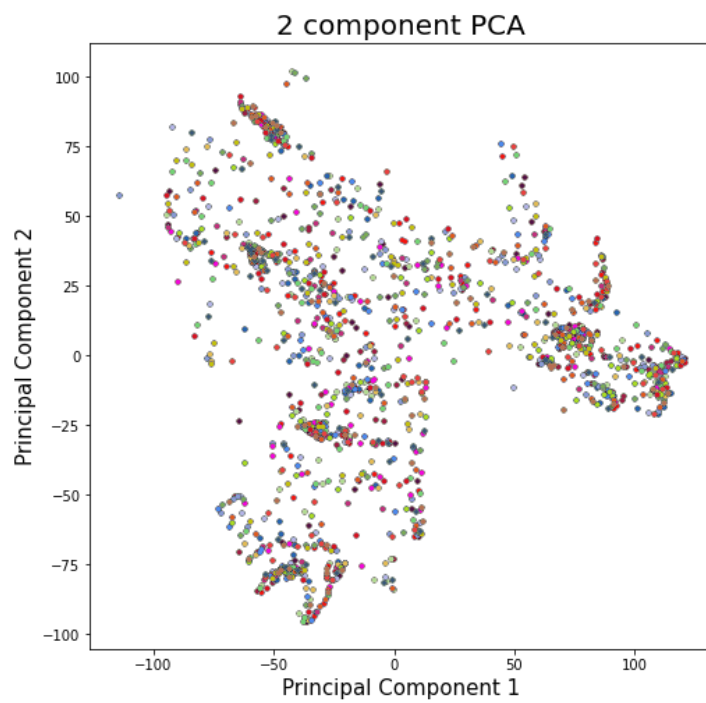
MNIST



Oliveretti



COILS-20



Animals - 10

