

**DATA.STAT.770 Dimensionality Reduction and  
Visualization, Spring 2021, Exercise set 10**

Md. Abdullah-Al Mamun

**Precision and Recall**

**Part F: Metric Learning**

**Problem F1: Metric learning with Neighborhood Components Analysis**

**Problem F2: Other approaches in metric learning**

Dong, Yanni, et al. "Maximum margin metric learning based target detection for hyperspectral images." ISPRS Journal of Photogrammetry and Remote Sensing 108 (2015): 138-150.

Differences between Maximum Margin Metric Learning (MMML) and Neighborhood Components Analysis (NCA) are as follows:

- MMML improves the efficiency in calculations.
- MMML aims at maximum the margins between cases in different classes. I think MMML is good for support vector machines and NCA good for KNN classifiers.
- MMML can well separate the target sample even when the target sample number is very small.
- MMML is capable of handling high-dimensional data while NCA cannot.

Du M, Ding S, Jia H. Study on density peaks clustering based on k-nearest neighbors and principal component analysis. *Knowledge-Based Syst.* 2016;99:135-145. doi: 10.1016/j.knosys.2016.02.001.

Differences between density peaks clustering based on k nearest neighbors (DPC-KNN) and Neighborhood Components Analysis PCA (DPC-KNN-PCA).

- DPC does not perform well when data sets have relatively high dimension. PCA works well on high dimensional data.
- DPC generates wrong number of clusters of real-world data sets.
- DPC follows KNN principles and perform density peaks clustering.
- It is used for local density computation
- To overcome several problem s, PCA into the DPC-KNN and works for high dimensional data preprocessing.
- k-means algorithm and spectral clustering (SC) algorithms are used to find out accuracy those are highly effective.