Week 9 Practical

Dr Simon Denman
CAB420: Machine Learning

This weeks practical will focus on clustering, and will compare the four models we have used for a face clustering application.

Problem 1. Select one of the previous face datasets that we have used (YaleB, ORL or Yale) and the method of your choice for extracting a face representation (i.e. PCA, LDA or a Siamese Network). Using these:

1. Transform the input faces from the dataset of your choice, using the method of your choice;
2. Cluster the resultant embeddings using:
   (a) K-Means
   (b) A GMM
   (c) HAC
   (d) DBScan
3. For each method, choose appropriate hyper-parameters. Use prior knowledge (i.e. the number of true faces in the dataset, a typical distance between two faces) to aid your parameter select.
4. Compute the purity, completeness and v-measure for each clustering method.