

Note: UG: 100+10, G: 100

Problem 0 (Self-Study) Use the dataset provided in HW2, calculate the projection direction using FLD and PCA. Plot on the same figure (1) the vector $(\mathbf{m}_1 - \mathbf{m}_2)$ where \mathbf{m}_1 and \mathbf{m}_2 are the means from each category, (2) the projection vector using FLD, and (3) the major axis from PCA. Comment on the difference.

Problem 1 (40/40) On Minkowski distance. Refer to Chapter 4.6.

- (a) (20/20) Illustrate/Plot, in a **2-D** space, the points that have equal Minkowski distance to the origin, when the order (k) equals to 1, 2, 3 and infinity.
- (b) (20/20) Search for the definition of “Hamming Distance” and show that it is or isn’t a distance metric

Problem 2 (60+10/60) On classifier fusion. Assume there are two classifiers (L1, L2), performing classification tasks on three objects ($\omega_1, \omega_2, \omega_3$). Assume there are 30 samples in the training data for each category. Following is the confusion matrices generated from the training data.

| L1 | ω_1 | ω_2 | ω_3 |
|------------|------------|------------|------------|
| ω_1 | 20 | 5 | 5 |
| ω_2 | 3 | 24 | 3 |
| ω_3 | 0 | 9 | 21 |

| L2 | ω_1 | ω_2 | ω_3 |
|------------|------------|------------|------------|
| ω_1 | 25 | 2 | 3 |
| ω_2 | 3 | 22 | 5 |
| ω_3 | 5 | 6 | 19 |

In the confusion matrix, each row represents the ground truth label and each column represents the label given by the classifier.

- (a) (10/10) Suppose for a certain testing sample, classifier 1 assigns it to ω_1 , and classifier 2 assigns it to ω_2 , what is the assigned label after fusing both classifiers using Naïve Bayes rule? (Note that by dividing each matrix element by the sum of all elements in the same row, you can arrive at a probability value that indicates the likelihood. Naïve Bayes assumes the classifiers are independent from each other)
- (b) (30/20) Derive a lookup table that includes the fused result from all possible combinations of labels from the two classifiers

| L_1, L_2 | Fused label |
|------------|-------------|
| 1, 1 | |
| 1, 2 | |
| 1, 3 | |
| 2, 1 | |
| 2, 2 | |
| 2, 3 | |
| 3, 1 | |
| 3, 2 | |
| 3, 3 | |

- (c) (Only 571) (+10/10) Provide an analytical study on how Naïve Bayes-based fusion approaches find the fused label according to the maximum posterior probability.