Name: ____

Problem	Points	Score
1	40	
2(a)	10	
2(b)	10	
2(c)	10	
3(a)	10	
3(b)	10	
3(c)	10	
Total	100	

Notes:

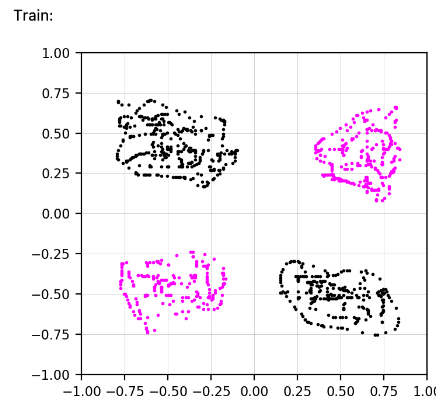
(1) The exam is closed books and notes.

(2) Please indicate clearly your answer to the problem.

(3) Note that ungrammatical sentences, incoherent statements, or general illegible scratches will get zero credit.

(4) If I can’t read or follow your solution, it is wrong, and no partial credit will be awarded.

(40 pts) Problem No. 1: Given the data shown to the right, discuss the differences between classifying the data using class-independent PCA, class-dependent PCA and, most importantly, using EM to estimate the parameters of a two-mixture Gaussian distribution per class.



The black class (upper left, lower right) is class 0, and the blue class (upper right, lower left) is class 1. Within these classes, you know which of the component distributions the data belongs to. For example, class 1 could represent "pets": "dogs" (upper left) and "cats" (bottom right). Class 2 could represent "farm animals": "horses" (lower left) and "cows" (upper right). Your classifier should decide between the two classes ("pets" and farm animals").

For EM, use a two-mixture Gaussian distribution per class. Estimate the means and covariances for each component from the data assuming you know the class labels for each data point, and the sub-class labels as well. Discuss the challenges associated with this and assume a best case scenario where training goes well (this is purposefully a bit vague).

Be very thorough in explaining your answers. Plan your answer before you start writing.

Problem No. 2: Consider a two-class discrete distribution problem:

$$\omega_1: \{[0,0], [2,0], [2,2], [0,2]\}$$
$$\omega_2: \{[1,1], [3,1], [1,3], [3,3]\}$$

(10 pts) (a) Compute the minimum achievable error rate by a linear machine (hint: draw a picture of the data). Assume the classes are equiprobable.

(10 pts) (b) Assume the priors for each class are: $P(\omega_1) = \alpha$ and $P(\omega_2) = 1 - \alpha$. Sketch $P(E)$ as a function of α for a maximum likelihood classifier based on the assumption that each class is drawn from a multivariate Gaussian distribution. Compare and contrast your answer with your answer to (a). Be very specific in your sketch and label all critical points. Unlabeled plots will receive no partial credit.

(10 pts) (c) Assume you are not constrained to a linear machine. What is the minimum achievable error rate that can be achieved for this data? Is this value different than (a)? If so, why? How might you achieve such a solution? Compare and contrast this solution to (a).

Problem No. 3: Let $p(x|\omega_i) = N(\mu_i, \sigma^2)$ for a two-category one-dimensional problem with $P(\omega_1) = P(\omega_2) = 1/2$.

(10 pts) (a) Create a sketch that demonstrates how you would calculate the minimum probability of error.

(10 pts) (b) Show that the minimum probability of error is given by: $P_e = \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-u^2/2}\, du$, where $a = |\mu_2 - \mu_1|/2\sigma$.

(10 pts) (c) Use the inequality $P_e = \frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-u^2/2}\, du \leq \frac{1}{\sqrt{2\pi}} e^{-a^2/2}$ to show that P_e goes to zero as $|\mu_2 - \mu_1|/\sigma$ goes to infinity.