Pattern Recognition and Machine Learning

James L. Crowley

ENSIMAG 3 Final Exam - Jan 2017

Conditions: You have the right to use any notes or written material. You may answer questions in English or in French. When appropriate, illustrate your answer with mathematics. Your written answers must be clear and legible. Illegible text will not be graded. Duration: 3 hours.

1 (4 points) Two classes, C_1 and C_2 are to be recognized from a feature, X, with $N=4$ possible values. You provided with the following histograms computed from $M=40$ samples, and asked to construct a detector for class C_1 using a ratio of histograms.

x	1	2	3	4
$h_1(x)$	2	4	6	8
$h_2(x)$	8	6	4	2

a) Plot the ROC curve for this detector.
b) Give the Precision and Recall for this detector.

2) (4 points) You are provided with a Viola-Jones style face detector composed of a cascade of committees. Each committee has been trained to accept 10% false positives and 1% false negatives. Is it possible to determine the probability of error for windows that pass each committee? If yes, provide the formula for cumulative probability of error for the first and second committee. If no, explain why.

3) (6 point) You are provided with a data-base of 1024 gray-scale images of handwritten text of size 1032 by 1032 pixels. You are asked to use Principal Components Analysis to learn the weights for the $D^{(2)}$ hidden units of a convolutional neural network that serves as a first hidden layer in a neural network. Present and describe the calculations involved in learning the weights. What is the training data? How many training samples do you have? What is the maximum number of possible hidden units? How would you select the subset to use in your network? How would you compute the weights for each unit?

4) (6 points) The goal of this exercise is to use back-propagation to train a 3-layer network to compute the Exclusive OR function (XOR). Your network will have 2 inputs, 1 output and a single hidden layer composed of 3 hidden units and will use a sigmoid activation function. What is the input training data? What is the loss function? How do you initialize the network parameters? Show the equations for a feed-forward pass followed by error back-propagation.