Homework 1

Machine Learning
1st Semester (FEUP-PDEEC)

1 October, 2013
15-782PP CMU-Portugal

Submit by 15 October, 2013, 23h59
by email to `jaime.cardoso@fe.up.pt`
and `vikramg@andrew.cmu.edu`

Problem 1

Prove that $E[L] \leq E[L|L > n]$ for any discrete random variable L, where the inequality is strict if $Prob(L \leq n) > 0$

Problem 2

Load the MATLAB variables `trainingset.mat`. The training set consist of three sets of observations, X_1, X_2 and X_3, where $X_i \in N(\mu_i, \sigma_i)$, $i = 1,,3$. There is also 100 observations in the variable X_x which we wish to classify. Each row in the matrices is an observation. The goal of the exercise is to use the training data to estimate mean and (co)variance of each distribution, and then use this information to classify the observations in X_x.

1. Display the training set using plot. Use a different color for each of the variables.
2. Based on the values in the training set, estimate the parameters μ and σ of each distribution using the MATLAB commands `mean` and `cov`.
3. We would like to create a discriminant function for each of the three distributions. We assume equal prior probabilities and equal losses. Armed with pen and paper,
   (a) find the analytical expressions for the class boundaries u12, u13 and u23
   (b) create the necessary heuristics to make the proper decisions between R1, R2 and R3
4. Having a set of discriminant functions, we are now ready to make a classification given a new observation. Write the Matlab code to classify X_x and plot the result.
5. Describe how we statistically could calculate the probability of a selected class given an observation. What do we call this expression?

Homework 1

MACHINE LEARNING — 1 October, 2013
1st Semester (FEUP-PDEEC) — 15-782PP CMU-Portugal

Problem 3

Consider two classes A and B with probabilities $P(C_A) = 2/5$ and $P(C_B) = 3/5$; assume that the corresponding probability density functions are :
$p(x|C_A) = e^{-x}, x \geq 0$
$p(x|C_B) = \frac{1}{\sqrt{2\pi}} e^{-(x-2)^2}$

1. Compute the regions of decision of the Bayes classifier. Draw the plot with the decision functions and the probability density functions.

2. For the observation $x = 1$ state the predicted class and the probability of error.

3. Assume now the error cost are not uniform. The penalty term for deciding class C_B although the pattern belongs to C_A is 1.2; the penalty term for deciding class C_A although the pattern belongs to C_B is 0.8; the cost of a correct decision is zero. What would now be the regions of decision minimizing the expected value of the cost?