1. Given $P(\omega_1) = P(\omega_2) = 0.5$, $\mu_1 = [0, 1]$, $\mu_2 = [0, -1]$ and $\Sigma_1 = \Sigma_2 = I$, derive the analytic Bayes form and sketch the decision boundary.

2. Let k be number of misclassified points out of n total points.

   (a) What is the distribution of k, $P(k)$

   (b) Show that Maximum Likelihood Estimate of $Err = k/n$ and estimate its value. (Hint: equals the partial derivative of $P(k)$ with respect to probability of error to 0)

3. Consider a binary classifier and finite training set. Your task it to find the number of optimal features k, select these features, train the classifier and estimate its performance. Briefly describe the process. You can use logical notation, pseudo-code or using your own words.

4.

   (a) When does the Bayesian estimation equal Maximum Likelihood Estimation? Why?

   (b) What is the difference between Maximum A Posteriori and Maximum Likelihood?

   (c) Why is "kernel trick" useful?