Name: .....

Student ID#: .....

Statistical Pattern Recognition (CE-725)
Department of Computer Engineering
Mini Exam #1 (SVM & Kernels) - Spring 2011

1. a) (20 extra points) Prove that if we remove non-support vector data from the data set the solution of the SVM will remain unchanged.

b) (20 points) Consider the standard two-class SVM with the hinge loss. Argue that under a given value of C,

$$\text{Leave\_One\_Out error} \leq \frac{\#\text{SVs}}{l},$$

where l is the size of training data and #SVs is the number of support vectors obtained by training SVM on the entire set of training data. Hint: use part a.

2. (30 points) Assume that the embedding spaces of the kernels k_1 and k_2 are represented as φ_1 and φ_2, respectively. Find the embedding spaces of these kernels:

a) $k = k_1 + k_2$

b) $k = k_1 . k_2$

3. (50 points) Suppose we have six training points from two classes. We have four points from class 1: (0.2, 0.4), (0.4, 0.8), (0.4, 0.2), (0.8, 0.4) and two points from class 2: (0.4, 0.4), (0.8, 0.8).

Unfortunately, the points cannot be separated by a linear classifier. The kernel trick is to find a mapping of x to some feature vector $\phi(x)$ such that there is a function K called kernel which satisfies $K(x,x') = \phi(x)^T \phi(x')$. And we expect the points of $\phi(x)$ to be linearly separable in the feature space. Here, we consider the following normalized kernel:

$$k(x,x') = \frac{x^T x'}{\|x^T\| \|x'\|}$$

a) What is the feature vector $\phi(x)$ corresponding to this kernel? Draw $\phi(x)$ for each training point x, and specify from which point it is mapped.

b) Now the feature vectors are linearly separable in the feature space. The maximum-margin decision boundary in the feature space will be a line in $\mathbb{R}^2$, which can be written as $w_1x + w_2y + c = 0$. What are the values of the coefficients w_1 and w_2? (Hint: you don't need to compute them).

c) Specify the support vectors.

d) Draw the decision boundary in the original input space resulting from the normalized linear kernel. Briefly explain your answer.

Good Luck!