ComS 573: Machine Learning
Spring 2008

Homework 4
Due Monday, March 31, 2008 in class

Note: Please do not hesitate to contact the instructor or TA if you have difficulty understanding or getting started with solving any of the problems.

1. (10 pts.) By making use of kernel substitution, formulate the (1-)nearest-neighbor classifier (that is based on the Euclidean distance metric) for a general kernel.

2. (20 pts.) Given valid kernels $K_1(\vec{x}, \vec{x}')$ and $K_2(\vec{x}, \vec{x}')$, prove that the following will also be valid kernels:

   (a) $K(\vec{x}, \vec{x}') = cK_1(\vec{x}, \vec{x}')$ where $c > 0$ is a constant

   (b) $K(\vec{x}, \vec{x}') = f(x)K_1(\vec{x}, \vec{x}')f(x')$

   (c) $K(\vec{x}, \vec{x}') = K_1(\vec{x}, \vec{x}') + K_2(\vec{x}, \vec{x}')$

   (d) $K(\vec{x}, \vec{x}') = K_1(\vec{x}, \vec{x}')K_2(\vec{x}, \vec{x}')$

   (e) $K(\vec{x}, \vec{x}') = \vec{x}^t A \vec{x}'$ where A is a symmetric positive semidefinite matrix.

3. (10 pts.) The entropy of a probability distribution $P = (p_1, \ldots, p_n)$ where $\sum_i p_i = 1$ is defined as

$$H(P) = -\sum_i p_i \log p_i$$

   Find the distribution with maximum entropy.

4. (20 pts.) Minimize

$$f(x, y) = (x - 1)^2 + (y - 3)^2$$

   subject to

$$x + y \leq 1; \quad x - y \leq 0$$

5. (20 pts.) Consider the maximum margin classifier. Let α_i be the solution of the dual problem. Show that the margin for the maximum margin hyperplane is given by

$$\frac{1}{\sqrt{\sum_{i \in sv} \alpha_i}}$$

   where the summation is over indices of support vectors.

6. (20 pts.) The Soft-Margin classifier may use an alternative objective function and the primal problem is given by

$$\begin{array}{ll} \text{Minimize} & \frac{1}{2}\vec{w}^t\vec{w} + \frac{C}{2}\sum_i \xi_i^2 \\ \text{subject to} & t_i(\vec{w}^t\Phi(\vec{x}_i) + b) \geq 1 - \xi_i, \quad i = 1, \ldots, n \end{array}$$

Derive the dual problem, and the KKT conditions.