CSE 5526 - Autumn 2014

Introduction to Neural Networks

Homework #3

Due Tuesday, October 28

Grader: Yuzhou Liu
Email: liu.2376@osu.edu
Office: Dreese Labs 578
Office hours: Monday & Wednesday 3 - 4pm

Problem 1. Given the following linearly separable training patterns:

$$\boldsymbol{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad d_1 = +1$$
$$\boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad d_2 = -1$$
$$\boldsymbol{x}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad d_3 = -1$$

Find $\boldsymbol{w}$, b, and $\boldsymbol{a}$ for the maximum margin hyperplane separating the two classes by optimizing the Lagrangian function

$$L(\boldsymbol{w}, b, \boldsymbol{a}) = \frac{1}{2}\|\boldsymbol{w}\|^2 - \sum_p a_p(d_p(\boldsymbol{w}^T\boldsymbol{x}_p + b) - 1).$$

Write down the discriminant function, $y(\boldsymbol{x})$, using these values for $\boldsymbol{w}$ and b and specify which of the input patterns are support vectors.

Problem 2. Prove that the $N \times N$ symmetric kernel matrix, K, formed using an inner-product kernel function on N data points, $\{\boldsymbol{x}_p\}_{p=1}^N$, such that

$$K_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \phi^T(\boldsymbol{x}_i)\phi(\boldsymbol{x}_j)$$

is positive semidefinite, i.e.,

$$a^T K a \geq 0 \text{ for all } a \in \mathbb{R}^N$$