T-61.5030 Advanced course in neural computing Karhunen, Peltonen

Exercise 3, Oct. 5, 2006

1. Consider a committee machine consisting of K experts. The input-output function of the kth expert is denoted by $F_k(\mathbf{x})$ where $\mathbf{x}$ is the input vector and $k = 1, \ldots, K$. The individual outputs of the experts are linearly combined to form the overall output y, defined by

$$y = \sum_{k=1}^{K} w_k F_k(\mathbf{x}) \ ,$$

   where w_k is a linear weight assigned to $F_k(\mathbf{x})$. The requirement is to evaluate w_k so that y provides a least-squares estimate of the desired response d corresponding to $\mathbf{x}$. Given a set of training data $\{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$, determine the required values of the w_k's to solve this parameter estimation problem.

2. What can you say about the bias and variance of a committee in which the output is computed with ensemble averaging? Hint 1: Any answer is better than none at all. Hint 2: the output y is obtained like in the previous problem but with equal weights for each expert, w_k=1/K.

3. Discuss ensemble averaging from the Bayesian point of view.

4. How much data is required to train a committee machine consisting of three experts, when the training is boosted by filtering? Hint: make some simplifying assumptions. Use the error rates of the experts in your estimate.