T-61.5030 Advanced course in neural computing Karhunen, Peltonen

Exercise 2 Sept. 28, 2006

1. Consider a function that resembles functions implemented by decision trees: $F(\mathbf{x}, \mathbf{w})$ is one when $\mathbf{x}$ is within a rectangular region aligned with one of the axes in a plane, and zero elsewhere. (Stated in another way, any such rectangle can be formed by changing the value of $\mathbf{w}$.) Show that the VC dimension is four.

2. Consider a linear binary pattern classifier whose input vector $\mathbf{x}$ has dimension m. The first element of the vector $\mathbf{x}$ is constant and set to unity so that the corresponding weight of the classifier introduces a bias. What is the VC-dimension of the classifier?

3. The inequality (2.97) in Haykin's book defines a bound on the rate of uniform convergence, which is basic to the principle of empirical risk minimization.

   - Justify the validity of Eq. (2.98), assuming that the inequality (2.97) holds.
   - Derive Eq. (2.99) that defines the confidence interval ϵ_1.

4. Consider the pupils of a large Finnish school. Take a random sample of 100 teenagers who will participate in an experiment in which they have to choose between strawberry and chocolate cakes. You notice that 90% of the girls prefer the strawberry cake and 90% boys prefer the chocolate cake.

   (a) What is the VC-dimension of the classifier which assumes that girls prefer strawberry cake and boys prefer chocolate cake?

   (b) The experiment is repeated for the rest of the teenagers of the school immediatly after the first one. Compute the 99,9% confidence interval on the classification error rate of the classifier estimated in the first experiment.

   (c) In the first experiment, a reward was hidden in the chocolate cakes for girls and strawberry cakes for boys. Does this change the situation?