T-61.5030 Advanced course in neural computing Karhunen, Peltonen

Exercise 1 Sept. 21, 2006

Please note that although Exercise 1 is relatively straightforward, the goal in the exercises is not always to find the "correct" answer, that is, the same answer the assistant will propose. Like in research, work, and real life in general, some problems may have multiple correct solutions, and all problems do not need to have clear-cut answers or answers at all. In such cases the task is to find a satisfactory solution whenever possible. (A proof that there exists no answer is a satisfactory solution!) If the problem setting is misleading it may even be a good idea to find a better formulation to the problem and solve the revised version.

The solutions must, however, be justified and mathematically correct when applicable.

The exercises:

1. Recapitulation. The network structures considered in the Basic Course were Perceptrons, Multilayer Perceptrons, Radial-Basis Function Networks, and Self-Organizing Maps.

   Classify these network architectures along the following dimensions:

   (a) Learning type: error-correction learning, memory-based learning, Hebbian learning, or competitive learning.

   (b) Category of architecture: feedforward network, recurrent network, competitive network.

   (c) Type of task: supervised, unsupervised, reinforcement learning.

   (d) Functions of neurons: projections or distances from points in space.

   (e) Suitable for these tasks: pattern association, pattern recognition, function approximation, control, filtering, density estimation, visualization, summarization, other

2. (Easy!) Prove that if the data comes from the regressive model

   $$D = f(\mathbf{X}) + \epsilon$$

   then the average of the mean-square prediction error is equal to

   $$\mathcal{E}(\mathbf{x}) = E_\epsilon\{(d - F(\mathbf{x}, \mathcal{T}))^2 | \mathbf{x}, \mathcal{T}\} = (f(\mathbf{x}) - F(\mathbf{x}, \mathcal{T}))^2 + E_\epsilon\{\epsilon^2\},$$

   where $\mathcal{T}$ denotes the training set.

3. Derive the bias/variance decomposition

   $$E_\mathcal{T}\{(f(\mathbf{x}) - F(\mathbf{x}, \mathcal{T}))^2\} = (f(\mathbf{x}) - E_\mathcal{T}\{F(\mathbf{x}, \mathcal{T})\})^2 + E_\mathcal{T}\{(F(\mathbf{x}, \mathcal{T}) - E_\mathcal{T}\{F(\mathbf{x}, \mathcal{T})\})^2\}$$

   Cf. problem 2.22 in Haykin.