Tik-61.261 Principles of Neural Computing

Raivio, Venna

Exercise 3

1. To which of the two paradigms, learning with a teacher and learning without a teacher, do the following algorithms belong? Justify your answers.
   (a) nearest neighbor rule
   (b) k-nearest neighbor rule
   (c) Hebbian learning
   (d) error-correction learning

2. Consider the difficulties that a learning machine faces in assigning credit for the outcome (win, loss, or draw) of a game of chess. Discuss the notations of temporal credit assignment and structural credit assignment in the context of this game.

3. A supervised learning task may be viewed as a reinforcement learning task by using as the reinforcement signal some measure of the closeness of the actual response of the system to the desired response. Discuss this relationship between supervised learning and reinforcement learning.

4. Heteroassociative memory $\mathbf{M}$, a matrix of size $c \times d$, is a solution to the following group of equation systems:
$$\mathbf{M}\mathbf{x}_j = \mathbf{y}_j, \quad j = 1, \ldots, N,$$
where $\mathbf{x}_j$ is the jth input vector of size $d \times 1$ and $\mathbf{y}_j$ is the corresponding desired output vector of size $c \times 1$. The ith equation of the jth equation system can be written as follows:
$$\mathbf{m}_i^T \mathbf{x}_j = y_{ij},$$
where $\mathbf{m}_i^T = [m_{i1}, m_{i2}, \ldots, m_{id}]$. Derive a gradient method which minimizes the following sum of squared errors:
$$\sum_{j=1}^{N} (\mathbf{m}_i^T \mathbf{x}_j - y_{ij})^2.$$
How it is related to the LMS-algorithm (Widrow-Hoff rule)?

5. Show that $\mathbf{M} = \mathbf{Y}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is a solution to the following group of equation systems:
$$\mathbf{M}\mathbf{x}_j = \mathbf{y}_j, \quad j = 1, \ldots, N.$$
Vectors $\mathbf{x}_j$ and $\mathbf{y}_j$ are the jth columns of matrixes $\mathbf{X}$ and $\mathbf{Y}$, respectively.