T-61.5130 Machine Learning and Neural Networks
Examination 12th March 2009/Karhunen

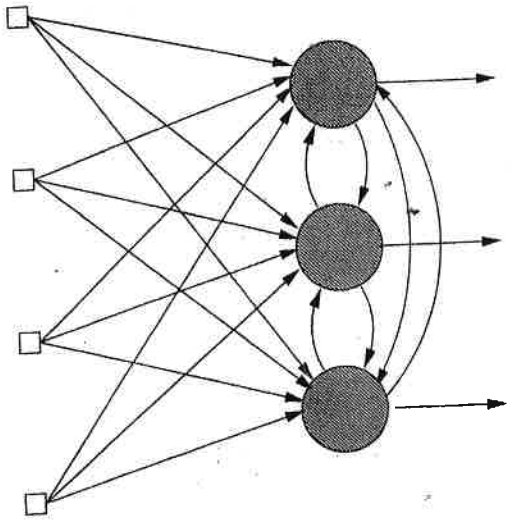
You are allowed to have in the examination a collection of mathematical formulas, but not any of the teaching material. (Voit vastata tenttiin myös suomeksi.)

1. Answer briefly (using a few lines) to the following questions:
   (a) What for is Oja's rule used in neural computing?
   (b) How does Newton's optimization method differ from standard steepest descent method?
   (c) Explain briefly the bias-variance decomposition.
   (d) Name the methods that you know for selecting the centers of radial-basis function networks.
   (e) What is a Voronoi cell?
   (f) Present the schematic diagram of focused neuronal filter.

2. Assume that the relationship between the input vector $\mathbf{x}$ and the desired response (output) vector $\mathbf{d}$ is of the form
$$\mathbf{d} = \mathbf{h}(\mathbf{x}) + \mathbf{e}$$
where $\mathbf{h}(\mathbf{x})$ is the true mapping between $\mathbf{x}$ and $\mathbf{d}$ and $\mathbf{e}$ is the error or noise vector. Consider modeling the unknown true mapping $\mathbf{h}(\mathbf{x})$ by the output $\mathbf{y}(\mathbf{x}, \mathbf{w})$ of a neural network, where the vector $\mathbf{w}$ contains all the adjustable weights of the neural network. Assume that you have at your disposal N training pairs $(\mathbf{x}_i, \mathbf{d}_i)$ of the mapping. Show that if the training pairs are independent, and the noise vector $\mathbf{e}$ is Gaussian with zero mean and covariance matrix $\sigma^2\mathbf{I}$, the standard least-squares method and maximum likelihood method provide the same results.

3. Explain what is Extreme Learning Machine (ELM) and which properties it has. In particular:
   - What kind of structure ELM has?
   - How ELM learns? It suffices to summarize its learning algorithm.
   - Which general properties ELM has?
   - What are the benefits and drawbacks of ELM?

4. Consider the neural network having the structure shown in the figure on the reverse side. Construct a mathematical expression for the output signal y_j of neuron j in the network. You can use a suitable notation for the input signals and weights of the network. For simplicity, we assume that there are no bias terms. The activation function $\varphi(v)$ is a sigmoidal type nonlinear function.



A SCHEMATIC DIAGRAM OF THE NEURAL NETWORK.