Neural networks

Final exam – January 29, 2018

1. (6 points) This task considers methods of learning.

1. Explain the idea behind Oja's learning rule, write the expression for it, and describe its properties.
2. Describe the principal component analysis.
3. Describe the AdaBoost algorithm.

2. (6 points) In this task you have to construct a neural network whose inputs are binary (0 or 1) and whose output has the activation function $h(x) = \begin{cases} 1, & x > 0.5 \\ 0, & \text{inače} \end{cases}$. During construction you can use arbitrary number of linear neurons with activation functions $f(x) = a \cdot x + b$ and $g(x) = \max(x, 0)$.

1. Design a network for the AND operator, draw its architecture, note the parameter values, and show that it works.
2. Design a network for the OR operator, draw its architecture, note the parameter values, and show that it works.
3. Design a network for the XOR operator, draw its architecture, note the parameter values, and show that it works.

3. (6 points) Consider a Hopfield network made up of two neurons whose synaptic weight matrix is

$$\mathbf{W} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

The bias applied to each neuron is zero. The four possible states of the network are $s_1 = (1, 1)^T$, $s_2 = (-1, 1)^T$, $s_3 = (-1, -1)^T$ i $s_4 = (1, -1)^T$.

1. Using the stability condition show that states s_2 and s_4 are stable.
2. Using the energy function show that states s_2 and s_4 are stable.
3. Show that states s_1 and s_3 exhibit a limit cycle. What is the length of the limit cycle?

4. (6 points) This task considers convolutional neural networks. The input to the network is a 64×64 image with 3 channels. The first layer consists of 4 filter of size 3×1 with size preservation, the second layer consists of 4 filters of size 1×3 with size preservation, the third layer consists of max pooling of size 2×2 and stride 2 for each dimension, and, finally, there is a fully connected layer with 128 neurons that have an activation function that besides the bias has two additional parameters.

1. Draw a sketch of the network.
2. How many parameters does the network have?
3. Which are some of the main reasons for the success of deep learning in the past 10 years?

5. (6 points) For the network shown in the image the input sample $\mathbf{x} = (1, 2, 3, 4)$ is expected to produce the output value $y = 0.5$. Initial parameter values are $w_1 = 1$, $w_2 = -1$, $w_3 = 1$, $w_4 = -1$, and the loss function is $L = \frac{1}{2}(d - y)^2$. First, the sample $\mathbf{x}$ is fed forward to the network, then the weights are adjusted by applying backpropagation with learning rate $\eta = 0.1$, and, finally, the sample $\mathbf{x}$ is fed forward to the network again. What will then be the output value d?