Homework 1: Deep Learning

Out May 12; Due May 19, beginning of exercise session*

Kristian Kersting, Alejandro Molina, {kersting, molina}@cs.tu-darmstadt.de,

upload link: `https://www.dropbox.com/request/0EyNaA11EGyiexfyadSa`

1. Implement the following network architecture in Python using NumPy.

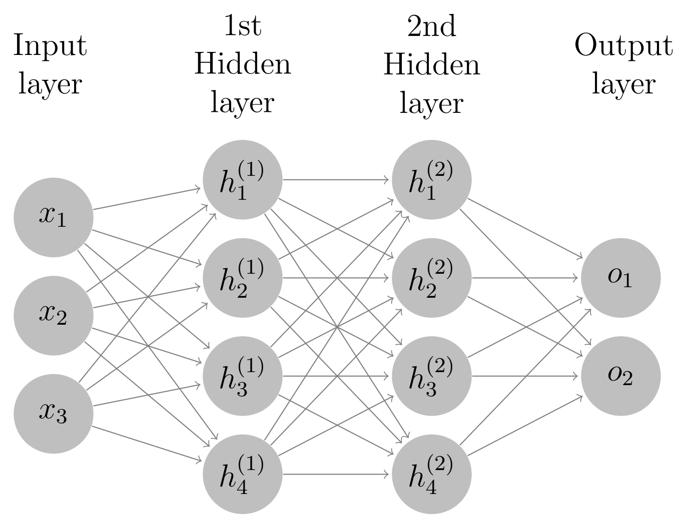

Figure 1: A feed-forward neural network architecture consisting of two hidden layers.

- The sigmoid activation function is used for computing hidden activations.
- No activation function in the output layer.

*We will discuss the solutions in the exercise session. It is my suggestion that you try to address at least 50% of the exercise questions. Simply try hard to solve them. This way, you will get familiar with the technical terms and with the underlying ideas of the lecture.

- Suppose that the network's parameters are initialized as follows:

$$\mathbf{W}_1 = \begin{bmatrix} -0.20 & 0.48 & -0.52 \\ -0.56 & 1.97 & 1.39 \\ 0.10 & 0.28 & 0.77 \\ 1.25 & 1.01 & -1.30 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 0.27 \\ 0.23 \\ 1.35 \\ 0.89 \end{bmatrix} \tag{1}$$

$$\mathbf{W}_2 = \begin{bmatrix} -1.00 & -0.19 & 0.83 & -0.22 \\ -0.27 & 0.24 & 1.62 & -0.51 \\ -0.29 & 0.06 & 0.15 & 0.26 \\ 0.00 & 0.67 & -0.36 & -0.42 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -1.19 \\ -0.93 \\ -0.43 \\ 0.28 \end{bmatrix} \tag{2}$$

$$\mathbf{W}_3 = \begin{bmatrix} -0.13 & 0.01 & -0.10 & 0.03 \\ -0.24 & -0.02 & -0.15 & -0.10 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} -0.13 \\ 0.03 \end{bmatrix} \tag{3}$$

  Note that $\mathbf{W}_1$ is the matrix connecting the input units and the units in the 1st hidden layer, and $\mathbf{b}_1$ denotes a bias term for the 1st hidden layer. The rows are the neurons and the columns are the inputs, e.g. $W_1(1,1)$ is the weight for input X_1 for neuron $h_1^{(1)}$.

- We have 5 input instances, each of which is of a 3-dimensional vector.

$$\mathbf{x} = \begin{bmatrix} 0.13 & 0.68 & 0.80 & 0.57 & 0.97 \\ 0.63 & 0.89 & 0.50 & 0.35 & 0.71 \\ 0.50 & 0.23 & 0.24 & 0.79 & 0.50 \end{bmatrix} \tag{4}$$

2. Given the above setting, compute the output values of every neuron in the hidden and output layers and fill in the following table:

Inputs			Outputs									
x_1	x_2	x_3	$h_1^{(1)}$	$h_2^{(1)}$	$h_3^{(1)}$	$h_4^{(1)}$	$h_1^{(2)}$	$h_2^{(2)}$	$h_3^{(3)}$	$h_4^{(2)}$	o_1	o_2
0.13	0.63	0.50										
0.68	0.89	0.23										
0.80	0.50	0.24										
0.57	0.35	0.79										
0.97	0.71	0.50										