Lecture “Advanced Data Analytics”

Problem Set 4

Simon Scheidegger

<u>Exercise 1</u>

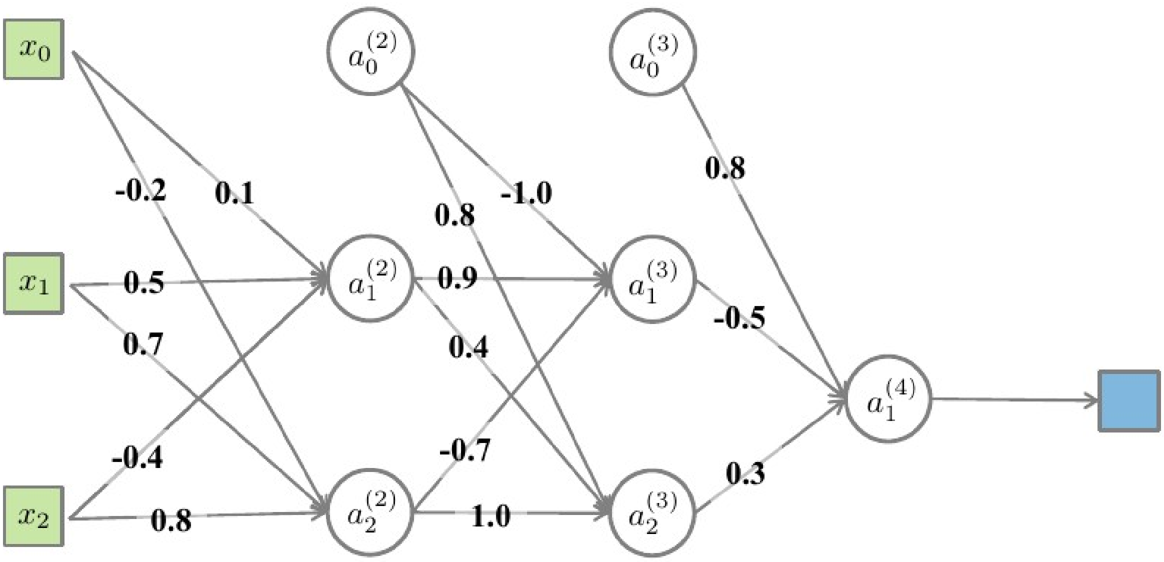
<u>Artificial Neural Nets</u>

You are given the feed-forward network (multi-layer perceptron) below, consisting of sigmoid-perceptrons, and the corresponding weights. Per default, $x_0 = 1$ for the bias.

Write a piece of code that implements the network below with the weights specified.

Determine the output for the following inputs:

a) $x_1 = 0.7$, $x_2 = 1.0$
b) $x_1 = -0.5$, $x_2 = 2.5$
c) $x_1 = -3.0$, $x_2 = 3.0$



Exercise 2

Artificial Neural Nets

Consider the sigmoid activation function for a node in a neural network:

$$g(a) = \frac{1}{1 + exp(-a)}$$

Show that the derivative of this activation function takes the simple form:

$$\frac{d}{da} g(a) = g(a)\,(1 - g(a))$$

Exercise 3

Artificial Neural Nets

Apply the simple multi-layer perceptron class distributed in the class (located **data/mlp.py**) to classify the MNIST data set (located **data/mnist.pkl.gz**)

To do so, proceed as follows.

a) Describe in few words the MNIST data set.
b) Describe every function in mlp.py with few words. What are they doing?
c) Implement the routine below to run a test classification on the data set.
Play with different 3 sizes of test / training sets. What are the respective performances?

Exercise 4

Apply the simple multi-layer perceptron class distributed in the class (located **data/mlp.py**) to the following dataset (taken from Anscombe's quartet):

(x1 , y1) = (10.0, 9.14), (x2 , y2) = (8.0, 8.14), (x3 , y3) = (13.0, 8.74),
(x4 , y4) = (9.0, 8.77), (x5 , y5) = (11.0, 9.26), (x6 , y6) = (14.0, 8.10),
(x7 , y7) = (6.0, 6.13), (x8 , y8) = (4.0, 3.10), (x9 , y9) = (12.0, 9.13),
(x10 , y10) = (7.0, 7.26), (x11 , y11) = (5.0, 4.74).

a) Use the neural network library **mlp.py** to show that a feedforward network with one hidden layer consisting of one unit and a feedforward network with no hidden layers, each using only linear activation functions, do not outperform linear regression based on ordinary least squares (OLS; for OLS, you may use the scikit library).

b) Also demonstrate that a neural network with a hidden layer of three neurons using the tanh activation function and an output layer using the linear activation function captures the non-linearity and outperforms the linear regression.