Deep Learning
Exercise Sheet 3

Dr. Josif Grabocka, Nicolas Schilling
Information Systems and Machine Learning Lab
University of Hildesheim

May 17th, 2017

Submission until May 14th, 18.00 via Learnweb (id: 3108)

Exercise 5: Dropout-Regularization in Neural Networks (10 Points)

a) In your own words, briefly explain the dropout regularization scheme for neural networks!

b) Assume we have a neural network with ReLU activation function and want to perform a regression task, the weights (and therefore the structure) is given through:

$$W^{1^T} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -1 \end{pmatrix} \qquad W^{2^T} = \begin{pmatrix} 1 & 2 & -2 \\ 2 & -1 & 1 \end{pmatrix} \qquad W^{3^T} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \qquad v^T = \begin{pmatrix} 1 & -1 \end{pmatrix}$$

and

$$b^1 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \qquad b^2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad b^3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad b^4 = 1$$

Predict for the instance $x^T = \begin{pmatrix} 1 & 1 \end{pmatrix}$ twice, using the following dropout masks:

$$\mu^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mu^1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad \mu^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$\mu^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \mu^1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \qquad \mu^2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

c) Explain, why dropout is not used for bias nodes!

Exercise 6: Solving the XOR Problem (10 Points)

Implement the backpropagation learning algorithm using L2 regularization for a network with one hidden layer and two hidden neurons. The size of the input layer is also set to two. Do not forget biases. Initialize the weights by drawing from a Gaussian distribution centered around zero, and learn the weights for the XOR data:

x_1	x_2	y
1	1	1
-1	1	0
1	-1	0
-1	-1	1

Please print your source code and the convergence of your algorithm and put it to the pdf submission.