Deep Learning
Exercise Sheet 2

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

May 10th, 2017

Submission until May 17th, 18.00 via Learnweb (id: 3108)

Exercise 3: Multiclass Classification (12 Points)

Given is a neural network with **one** hidden layer. The network takes as input a 3-dimensional vector $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ and outputs a vector $\hat{y} \in \mathcal{Y}$, where $\hat{y}_i = P(y = i \,|\, x)$ using the softmax on the network output. Let $\mathcal{Y} = \{1, 2, 3\}$, meaning that we are facing a classification problem with three classes. The number of neurons in the hidden layer is set to be **two**. The model does **not** use biases, the activation function in the hidden layer is ReLU.

a) Use the parameters

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

and perform a forward pass through the network for the instance $x = (1, 0, 1)$. Which class does it most likely represent?

b) Assume the ground truth label for the instance in a) was $y = 3$. Perform an update of all the model parameters using backpropagation and a step size $\alpha = 0.1$!

Exercise 4: Network Initialization (8 Points)

For the same network structure given above, assume we initialize all weights to zero. What happens if you want to predict for $x = (1, 0, 1)$ now? Which problem occurs if you want to do an update with this instance? How can this problem be circumvened?