Deep Learning
Prof. S. Harmeling

Thursday, 14. May 2020
DUE 23:55 Wednesday, 20. May 2020

Exercise set #4

Please upload a scan of your solution or use LaTeX to typeset your solution. Use the same naming scheme as usual and submit a single zip-file to the sciebo dropoff folder for the exercises. The link will be available on the ILIAS page.

1. **Analytic derivatives.**

   (a) Consider the two layer network

   $$f(x, y) = \frac{1}{2}||y - \text{ReLU}(xw)v||_2^2,$$

   where $x, y, w, v \in \mathbb{R}$ and the ReLU activation function defined as $\text{ReLU}(a) = \max(0, a)$ for any $a \in \mathbb{R}$. Do not worry too much about the case $x = 0$ where the ReLU function is not differentiable.

   Compute $\frac{\partial f}{\partial w}$ and $\frac{\partial f}{\partial v}$.

   (b) Consider the following two layer network

   $$g(x, y) = \frac{1}{2}||y - \text{ReLU}(xW)v||_2^2$$

   with weights and inputs

   $$x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \in \mathbb{R}^2, y \in \mathbb{R},$$
   $$W = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2.$$

   Compute $\frac{\partial g}{\partial w_{11}}$ and $\frac{\partial g}{\partial w_{ij}}$.

   30 points

2. **Two-layer neural network.**

   In this exercise you will implement a fully-connected neural network classifier for the CIFAR-10 dataset. For detailed instructions please we refer to the Jupyter notebook *two_layer_net.ipynb*.

   70 points