Advanced Topics in Machine Learning
Assignment # 2
Universität Bern

Meiguang Jin(jin@inf.unibe.ch)

Due date: 28/03/2017

In this assignment you need to upload a zip file to ILIAS which includes: 1) assignment2.lua file and 2) a pdf of your answers to the questions (see next page). The zip file name must be FirstName_LastName.zip. If your implementation requires auxiliary functions, you must implement that function inside the corresponding .lua file. We prefer pdf format (over docx) because it is universally supported on all platforms. Please also indicate your name on top of the first page in your pdf submission. Please do not print or display anything in the code (comment it out before submission).

Maximum grade: 4 Points

1. Answer the following questions.

   - **[0.5 point]** What is RELU?
   - **[0.5 point]** Why do we use that in the network design? Do we have other choices to replace RELU?
   - **[0.5 point]** Why maximum likelihood is almost always the preferred approach to training sigmoid output units?

2. **[2.5 point]** In this exercise, you need to use torch to implement a MLP. Given three inputs $x_1 \in R^{5\times 1}, x_2 \in R^{20\times 1}, x_3 \in R^{15\times 2}$ and $W_{11} \in R^{5\times 10}, W_{12} \in R^{5\times 15}, W_{21} \in R^{25\times 11}, W_{22} \in R^{20\times 12}, W_{31} \in R^{15\times 9}, W_{32} \in R^{15\times 14}$, build a network such that it can have three outputs $h_1, h_2$and h_3. You need to print out the sizes of three outputs $h_1, h_2$and h_3. (Please refer to torch nn.graph module)

$$
\begin{aligned}
h_{11} &= W_{11}^T x_1 + b_{11} \\
h_{12} &= W_{12}^T x_1 + b_{12} \\
h_1 &= [h_{11}; h_{12}] \\
h_{21} &= W_{21}^T h_1 + b_{21} \\
h_{22} &= W_{22}^T x_2 + b_{22} \\
h_2 &= [h_{21}; h_{22}] \\
x_{31}, x_{32} &= split(x_3, 1) \\
h_{31} &= W_{31}^T x_{31} + b_{31} \\
h_{32} &= W_{32}^T x_{32} + b_{32} \\
h_{33} &= [h_{31}; h_{32}] \\
h_3 &= [h_{33}; h_2]
\end{aligned}
$$