Deep Learning: Exercise Sheet 6 (SoSe2018)

30th of May (due 6th of June at 14:00)

Dr. Josif Grabocka, Rafael Rego Drumond
HiWi: Manish K. Mishra
Information Systems and Machine Learning Lab
University of Hildesheim

Question 10: AdaGrad and Friends! (10 Points)

a) In your own words, explain the fundamental difference between normal gradient descent, AdaGrad and Adam.

b) Assume we have a single layer neural network with a linear activation function with the following configuration and want to perform a regression task, the weights (and therefore the structure) is given through (bias included):

$$X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} Y = \begin{bmatrix} 11 \\ 8 \\ 10 \end{bmatrix}$$

And initial weights

$$w = (1 \quad 1 \quad 1)$$

Do two epochs using stochastic gradient descent with a step size of $\mu = 0.1$ and report the errors and total loss after each epoch. Please go over the instances in order, i.e. first line, second line, third line of X.

c) Repeat the same procedure by using a stochastic gradient descent with Adagrad for an initial step size of $\mu = 0.1$. Does Adagrad help?

Question 11: Solving the XOR Problem with Adagrad (10 Points)

Implement AdaGrad, as well as the normal momentum and the Nesterov's momentum in your implementation of last week's exercise and learn the network again. Which one performs best in your experience?

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.

How to submit?

DO NOT FORGET TO WRITE YOUR NAME ON YOUR SHEET! Nameless files will NOT be graded!

The new submission method is via LearnWeb. https://www.uni-hildesheim.de/learnweb2018/course/search.php?search=3108

Warning!

If we detect **Plagiarism** on your solution, you will receive no points for it. If a second plagiarism attempt is detected, you might fail the class or be expelled from your program. You are allowed to discuss solutions, but if you work on a group, you must indicate on your sheet with whom are you working with. Group submissions earn 0 points, but counts as participation.