Programming with neural networks: Exercise sheet 4 SS 2020

University of Würzburg - Chair for Computer Science VI Prof. Dr. F. Doll, A. Hekalo, M. Krug

Exercise sheet: 4

Edited on June 14th

Task 1: Fully Connected

A weight matrix w , a bias vector b and an input x *of* a fully are given
Connected Layer, as discussed in the lecture:

$$w = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

$$x = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

(a) Calculate the forward pass

(b) Calculate the vector δx . The following error vector δy = is given $\begin{bmatrix} 0 & 0.5 & 1 \end{bmatrix}$

(c) Calculate the weight updates ΔW and Δb for the above information.

Task 2: Convolution-2D

Let the convolution-2D layer be introduced as in the lecture.

(a) Change the filter f (with 2 channels)

$$f_0 = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \end{bmatrix} \quad f_1 = \begin{bmatrix} 13 & 15 & 17 \\ 24 & 16 & 18 \end{bmatrix} \begin{bmatrix} 19 & 21 & 23 \\ 20 & 22 & 24 \end{bmatrix}$$

so that it can be used for the backward pass.

(b) Calculate the size of the output $[o_x , o_y , o_c]$ of the following convolution:

Shape of the input [20 , 18 , 5]
Shape of the filter [3 , 4 , 5 , 1337]
Stride: $s = 2$
Padding: $p = 3$

Task 3: Recurrent networks

The following are the forward equations of a so-called peephole LSTM.
shows.

Page 2

Programming with neural networks: Exercise sheet 4 SS 2020

University of Würzburg - Chair for Computer Science VI Prof. Dr. F. Doll, A. Hekalo, M. Krug

$$f_t = x_t \cdot W_f + c_{t-1} \cdot W_f + b_f$$
$$f_t = \sigma(f_t)$$

$$i_t = x_t \cdot W_i + c_{t-1} \cdot W_{Ri} + b_i$$
$$i_t = \sigma(i_t)$$

$$o_t = x_t \cdot W_o + c_{t-1} \cdot W_{Ro} + b_o$$
$$o_t = \sigma(o_t)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(x_t \cdot W_c + b_c)$$
$$h_t = \tanh(o_t \cdot c_t)$$

(a) Calculate the equation for δc_t

Exercise 4: Reverse Mode Automatic Differentation

Given is the function $L = \cos(x_1 \cdot x_2) + x_1 \cdot x_2 + x_1$

(a) Calculate the gradient of the function L manually

(b) Calculate the gradient of L at (1 , 2)

(c) Calculate the gradient of L at point (1 , 2) using reverse mode AD

Exercise 5: Average Pooling

Let average pooling be defined so that the operation of pooling is the average of all Takes values and writes them to the output. Describe how to create the deltas in the Identify backward pass.