

Übungsblatt: 4

Bearbeitung am 14. Juni

Aufgabe 1: Fully Connected

Gegeben ist eine Gewichtsmatrix w , ein Biasvector b und eine Eingabe x einer Fully Connected Layer, wie in der Vorlesung besprochen:

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

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

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

- Berechnen Sie den Forward-Pass
- Berechnen Sie den Vektor δx . Gegeben ist folgender Fehlervektor $\delta y = \begin{bmatrix} 0 & 0.5 & 1 \end{bmatrix}$
- Berechnen Sie die Gewichtsupdates ΔW und Δb für die obigen Angaben.

Aufgabe 2: Convolution-2D

Es sei die Convolution-2D Layer wie in der Vorlesung eingeführt.

- Verändern Sie den Filter f (mit 2 Channels)

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

, so dass er für den Backward-Pass genutzt werden kann.

- Berechnen Sie die Größe der Ausgabe $[o_x, o_y, o_c]$ folgender Convolution:

Shape der Eingabe $[20, 18, 5]$

Shape des Filters $[3, 4, 5, 1337]$

Stride: $s = 2$

Padding: $p = 3$

Aufgabe 3: Rekurrente Netze

Im Folgenden sind die Forward Gleichungen eines sogenannten Peephole-LSTMs gezeigt.

$$f'_t = x_t \cdot W_f + c_{t-1} \cdot W_{Rf} + b_f$$

$$f_t = \sigma(f'_t)$$

$$i'_t = x_t \cdot W_i + c_{t-1} \cdot W_{Ri} + b_i$$

$$i_t = \sigma(i'_t)$$

$$o'_t = x_t \cdot W_o + c_{t-1} \cdot W_{Ro} + b_o$$

$$o_t = \sigma(o'_t)$$

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

$$h_t = \tanh(o_t \cdot c_t)$$

- (a) Berechnen Sie die Gleichung für δc_t

Aufgabe 4: Reverse Mode Automatic Differentiation

Gegeben ist die Funktion $L = \cos(x_1 \cdot x_2) + x_1 \cdot x_2 + x_1$

- (a) Berechnen Sie den Gradienten der Funktion L händisch
- (b) Berechnen Sie den Gradienten von L an der Stelle (1,2)
- (c) Berechnen Sie den Gradienten von L an der Stelle (1,2) mittels Reverse Mode AD

Aufgabe 5: Average Pooling

Average Pooling sei definiert, so dass die Operation des Poolings den Average aller Werte nimmt und in den Output schreibt. Beschreiben Sie, wie Sie die Deltas im Backward-Pass ermitteln können.