DL4NLP 2022 — Exercise 1

Steffen Eger, University of Bielefeld

April 10, 2022

Task 1 (10min)

Show that the 'AND' problem is linearly separable.

Task 2 (10min)

Softmax is the function $s : \mathbb{R}^n \to \mathbb{R}^n$ that maps a vector $\mathbf{z}$ to $\mathbf{y}$, where

$$y_j = \frac{\exp(z_j)}{\sum_{k=1}^{n} \exp(z_k)}$$

Show that $\mathbf{y}$ is a probability distribution, i.e., $y_j \geq 0$ and $\sum_j y_j = 1$.

Task 3 (20min)

Implement the perceptron learning algorithm from Lecture 1 (slide 89) on the toy problem given in code 'perceptron.py'.

(i) Adjust the code under "YOUR CODE HERE" following the lecture slides.

(ii) What accuracy do you get on the evaluation test set given in the script?

(iii) What do you need to keep in mind when reporting the test set result?

(iv) Change your initial weight parameter guess to '(1,-1,-1)'. What do you observe?

(v) Does the learned weight vector make any sense?

(vi) Error analysis: Where does your model go wrong?

Task 4 (10min)

Play around with the training data size, the learning rate, the number of epochs and activation function. What do you observe?
Also draw the *loss* (how is it defined?) over training epochs for a hyperparameter setting of your choice.

Task 5 (10min)

How would you evaluate a machine translation system and why is evaluation of machine translation difficult?