Aufgabe 1 Mandatory Paper:

Polysemy of Words could lead to wrong mapping, which penalize similarity during capturing a specific sense. Such as bank vs. river and bank vs. money.

Aufgabe 2 Softmax

Aufgabe 2.1:
Softmax function compares each other result with all of them in normalization, which come from all neuron units in previous hidden layer, but conventional activation functions just focus them locally (without comparation with other units).

Aufgabe 2.2:

Aufgabe 2.2:

1) $\forall\, y_i : 0 < y_i < 1$

$\because \exp(z_i/T) > 0 \quad \forall\, z \in \mathbb{R}^+$

$\therefore$ for $n \in \mathbb{N}^+ \wedge n > 1 : \sum_{i=1}^{n} \exp(z_i) > \exp(z_i) > 0$

$\Rightarrow y_i = \frac{\exp(z_i/T)}{\sum_{j=1}^{n} \exp(z_j/T)} \in (0, 1) \Rightarrow$ proved

2) $y_1 + y_2 + \cdots + y_n = 1$

$\Rightarrow y_1 + y_2 + \cdots + y_n = \frac{\exp(z_1/T) + \exp(z_2/T) + \cdots + \exp(z_n/T)}{\sum_{i=1}^{n} \exp(z_i/T)} = \frac{\sum_{i=1}^{n} \exp(z_i/T)}{\sum_{i=1}^{n} \exp(z_i/T)} = 1$

3) $\lim_{T \to \infty} y_i = \frac{\exp(z_i/T)}{\sum_{j=1}^{n} \exp(z_j/T)} = \lim_{k \to 0} \frac{\exp(k)}{\sum_{j=1}^{n} \exp(k)} = \lim_{k \to 0} \frac{1}{n} \cdot \frac{1}{1} = \frac{1}{n}$

$\lim_{T \to 0} y_i = \frac{\exp(z_i/T)}{\sum_{j=1}^{n} \exp(z_j/T)} = \lim_{k \to \infty} \frac{\exp(k)}{\sum_{j=1}^{n} \exp(k)} = \lim_{k \to \infty} \frac{\exp(k)}{n \cdot \exp(k)} = \frac{1}{n}$