

Knowledge Graph Analysis

Exercise Sheet 7

Dr. Hamed Shariat Yazdi, Prof. Jens Lehmann

December 19, 2018

1 IN CLASS

1. Backpropagation

- a) What is the goal of the backpropagation algorithm?
- b) Given is a tiny neural network, with 1 input neuron, 2 hidden neurons and 1 neuron in the output layer. As activation function for all units the sigmoid function is used and the network is trained with the mean squared loss. Calculate the gradient for each of the 4 parameters.

2. MLP-based SRL models

Recall that the E-MLP model is given by

$$\begin{aligned}f_{ijk}^{E-MLP} &= \mathbf{r}_k^T \mathbf{g}(\mathbf{h}_{ijk}) \\ \mathbf{h}_{ijk} &= \mathbf{W}_k^T \mathbf{x}_{ij}^{E-MLP} \\ \mathbf{x}_{ij}^{E-MLP} &= [\mathbf{a}_i, \mathbf{a}_j]\end{aligned}$$

and the ER-MLP model by

$$\begin{aligned}f_{ijk}^{ER-MLP} &= \mathbf{r}^T \mathbf{g}(\mathbf{h}_{ijk}) \\ \mathbf{h}_{ijk} &= \mathbf{W}^T \mathbf{x}_{ij}^{ER-MLP} \\ \mathbf{x}_{ij}^{ER-MLP} &= [\mathbf{a}_i, \mathbf{a}_j, \mathbf{a}_k] .\end{aligned}$$

- a) What are the differences between the two models?
- b) Compare the number of parameters (which depends on the number of entities N_e , the number of relations N_r , the number of latent features for the entities H_e and relations H_r and the dimension H_a of \mathbf{h}_{ijk}) for both models?
- b) How does they compare to the number of parameters used by the NTN? Recall that it is described by

$$\begin{aligned}
 f_{ijk}^{NTN} &= \mathbf{r}^T \mathbf{g}([\mathbf{h}_{ijk}^a, \mathbf{h}_{ijk}^b]) \\
 \mathbf{h}_{ijk}^a &= \mathbf{W}^T [\mathbf{a}_i, \mathbf{a}_j] \\
 \mathbf{h}_{ijk}^b &= [\mathbf{a}_i^T \mathbf{B}_1^k \mathbf{a}_j, \dots, \mathbf{a}_i^T \mathbf{B}_{H_b}^k \mathbf{a}_j]
 \end{aligned}$$

You can use H_a and H_b to denote the dimensions of \mathbf{h}_{ijk}^a and of \mathbf{h}_{ijk}^b , respectively.