Protokoll zur Sitzung am 29.5.17 (Kolloquium)

Dayyan Smith „Regularisation of neural networks for NLP“ Betreuerin: K. Kann

Mr. Smith presented his bachelor thesis. He explores the effect of regularisation of a neural network for stance classification in the context of fake news detection.

Some people say, fake news is anything they do not agree with, but that is not the definition Mr. Smith used in his work. He uses, rather, the definition of the *New York Time magazine*: „fake news is a made-up story with an intention to decieve“.

But how can one detect fake news? There is now the definition, but it is still very difficult to assess the veracity of a news story. Even human experts find it difficult.

Automatic fake news detection can be diveded into different stages, which were determined in line with the *fake-news challenge program*. The goal of it was to explore artificial intelligence technologies that can be levereged to combat fake-news. One thing that the orginizers of *the fake news challenge* found is that it is important to know, which news organizations agree with a given claim. That was determined through talks with different journalists. To do that, they developed the first task of the challenge, which is *stance detection*: finding the stance of an article towards *any* headline.

Based on some examples, there have been presented 4 stances - *agree, disagree, discuss, unrelated*. Mr. Smith demonstared some more examples and showed the data he was using during his work on the bachelor thesis.

Then Mr. Smith showed, how he encoded these bodies and headlines in order to use them in his neural net. He first used a pretrained *word2vec* to initialize a

vector for each word and then put these vectors into a RNN (Gated Reccurent Unit, which is similar to an LSTM-Network). Tthis gave him a representation of a headline or a body that he could later use for his classification problem.

So this representation used the word vectors and looked at the sequence of them and produced a sentence embedding. These embeddings were then concatenated - the representations for headlines and bodies were put together. He put these concatenated embeddings for his classification problem through two hidden layers and then (an interesting architecture decision) he concatenated the embeddings with the output of the hidden layer *1* and the output of the hidden layer *2* and then used that for representation. The classification decided, if this combination of headline and body had one of those four stances.

A method to prove, wheater a system has a good performance on unseen date is regularization (along with increasing training data, for example). Mr. Smith spoke about *L2 REgularisation*, where big weights are pushed down more than small weights, as well as *L1 Regularisation*, where one looks at the absolute value of the weights, which means means big weights and small weights are both pushed down a little bit and this often results in that, that little weights go down to zero and one has sparse vectors at the end.

Mr. Smith used a dropout regularisation. Dropout regularisation is a different kind of regularisation. Hier you do not change the cost function of your model, but you actually modify a model itself. You *drop out a certain number of neurons*, train them *as is*, *backpropagate the weights* and then you drop out other neurons and train again... In the end the system cuts the noise in your data and your neurons learn to detect, *what actually matters.*

Results: There was a score, which is a weighted accuracy, which shows how good your disambigution of *agree, disagree, discuss and unrelated* is. The result of the *dropout-regularisation* were insignificantly better (There was no significance test yet). There was a schort discussion at the and of the talk.