

Protokoll zur Sitzung vom 29.05.2017 – Computerlinguistisches Arbeiten

2. Vortrag: Dayyan Smith , “Regularization of Neural Networks for Natural Language Processing”

BA-Betreuer: M. Sc. Katharina Kann

The second presentation of the day was held by Smith whose bachelor's thesis has been supervised by Mrs. Kann. In his bachelor's thesis, Smith is exploring the effect of regularization of a neural network for stance classification in the context of fake news detection. Smith began his presentation with an overview of the topic and presented what fake news is. Fake news is a made-up story with an intention to deceive.

Later, Smith explains how to detect fake news. Assessing the veracity of a news story is a complex and cumbersome task even for trained experts. Automatic fake news can be broken down into multiple stages.

Fake News Challenge is exploring how artificial intelligence technology could be leveraged to combat fake news. It would be helpful to know which news organizations agree with a given claim.

The news organizations which agree with a given claim are the ones with stance detection (find stance of article towards headline). Fake news means a made-up story with an intention to deceive. To detect fake news is a complex and cumbersome task and hence, Automatic Fake News Detection can be broken down into numerous stages. The first stage is stance detection. The possible stances are: agree, disagree, discuss and unrelated. The vectors for each words should be constructed with the help of word2vec. Headlines and bodies separated go through two hidden layers, and are then concatenated to go through classification layer to get the label.

Regularization of a data is imported to get better results. It is an alternative if it is not possible to increase data you have. It is a technique used in an attempt to solve the overfitting problem in statistical models. There are some different ways to regularize data. The first one is L² regularization. Here, the big weights are pushed down more than small weights because the square of weights is penalized. In the other one, L1, both big and small weights are pushed down a little because the absolute value is penalized. This often drives small weights to zero. The third method is dropout regularization: to drop out certain number of neurons (circa 50%), and then restore neurons and drop out other neurons. In this way, neurons learn to detect what matters a lot. This should help cut the noise out. The goal is, with the help of the regularization, get better score than baseline score. Baseline system is SVM system. The results of evaluation is almost the same for all the types of regularization.