

SOSE 2017,CIS,LMU,München

Protokoll zur Sitzung am 15.05.2017:

Präsentation von Kristina Smirnov:

Kristina's Thema "Comparison of transfer Methods for low-resource morphology" besteht aus drei Tasks:erster – Kombination die annotierte russische Beispiele mit annotierten ukrainischen Beispielen, zweiter ist die Kombination von annotierten und nicht annotierten russischen Beispiele und letzter ist Kombinierung von aller drei Kategorien. Das Ziel wird durch die Beantwortung durch die wichtigsten Fragen definiert – was kann man machen, wenn es nicht genug Sprachdaten gibt und wie hilft es die ähnlichen Daten hinzufügen.

Die Performance, welche unter diesen Umständen gegeben wird, wird auf der Grundlage der korrekten Erzeugung morphologischer Formen bewertet.

Die nicht annotierte Daten, welche verwenden werden kommen aus Wikipedia und annotierte aus dem 2016 Shared Task on Morphological Reflection.

Das Model, welches verwendet wurde ist MED (das LMU System für SIGMORPHON 2016 Shared Task on Morphological Reflection). Als Daten wurden verschieden Kombinationen von LR und HR benutzt. Für jede Kombination gibt es Training Set für Quelle und Ziel, Development Set, Test Set, Wörterbuch für beide Trainings und die Länge des Wörterbuches. Am Ende wird sowohl als eine statistische als auch linguistische Evaluierung gemacht, um zu finden, wo genau die Fehler entstehen und welche Morphologie nicht erkannt wurde.