

Protokoll 3 für die Sitzung vom 15.05.2017

Ivana Daskalovska 11139620

Thema: Comparison of transfer methods for low-resource morphology

Student: Kristina Smirnov

Betreuer: Katharina Kann

Am Anfang der Präsentation hat die Frau Smirnov den Titel ihrer Bachelorarbeit und ihre Ideen für die Gliederung ihrer Bachelorarbeit vorgestellt.

Das Ziel Ihrer Bachelorarbeit ist es anhand bestimmter Kriterien flektierte Wortformen im Russischen zu finden. Da die Anzahl der annotierten Daten in diesem Gebiet für das Russische sehr gering ist, versucht die Frau Smirnov das gesamte über das Ukrainische und über nicht annotierten Daten im Russischen zu tun. Das Model mit dem sie arbeitet wurde vom Herrn Hinrich Schütze und Frau Katharina Kann entwickelt. Es handelt sich um einen Morphologischen Encoder-Decoder:

The LMU System for the SIGMORPHON 2016 Shared Task on Morphological Reinflection

Dieses Modell ist erfolgreich bei folgenden Aufgaben:

1. Lemma und Target gegeben → Target finden
2. Sourceform, Sourcetag und Targettag gegeben → Target finden
3. Sourceform und Sourcetag gegeben → Target finden

Man gibt die Sprache und das Lemma ein und bestimmt in welcher grammatikalischen Form das Wort vorkommen soll. Das System soll anhand von diesen Informationen die flektierte Form des Lemma ausgeben.

Frau Smirnov erhöht stufenweise die Anzahl an Trainingsdaten.

Resultate und Evaluation: Das System wurde bereits angewendet aber die Ergebnisse sind noch nicht da, sie folgen in den nächsten Tagen. Frau Smirnov hat vor, zusätzlich zu der statistischen Evaluation, die Ergebnisse manuell zu evaluieren. Ihr Ziel ist es eventuell Regelmäßigkeiten und Ursachen für falsche Zuordnungen zu finden.