

# Präpositionen

SoSe 2020

Gruppe 4

Alexander Mehler

Alexander Henlein

10.07.2020

## 1 Supervised Englisch

Die Aufgabe besteht darin 2-3 verschiedene Tagger auf den verfügbaren Daten zu trainieren. Dabei sollen **systematisch** die wichtigsten Hyperparameter getestet werden, um ein optimales Ergebnis zu erzielen.

## 2 Semi-Supervised

Hier sollte das Projekt ebenfalls Zeitnah abgeschlossen werden. Hier ist besonders wichtig, dass das Projekt entsprechend Dokumentiert wird, unabhängig vom Erfolg. Dazu zählen Projektaufbau, Idee, Ergebnisse und Vorschläge, an welchen Stellen es noch weiter optimiert werden kann.

## 3 Integration TextImager

Zum Abschluss sollen alle aentwickelten Tagger in den TextImager <sup>1</sup> integriert werden. Legen Sie dazu einen entsprechenden Fork an, den Sie später über einen entsprechenden PullRequest mit dem Gesamtprojekt zu mergen können. Achten Sie auch darauf, dass Sie mit der neusten Version arbeiten und als Maven-Projekt importieren. Das Projekt kann unter Umständen in IntelliJ etwas zicken, es wird daher Eclipse empfohlen. Beispiel-Tagger können Sie finden unter z.B.: **textimager-uima-spacy** oder DKPro<sup>2</sup> Test-Beispiele finden Sie unter: **textimager-uima-tool-test**

## Abgabe der Aufgabe

Ihre fertige Projekt schicken Sie dem Dozenten als GitLab oder GitHub Link zu.

Außerdem soll die komplette Abgabe dokumentiert werden. Besonders wichtig sind dabei die Schnittstellen (Intern und Extern) und die verwendeten Tools und Dateiformate.

## Aufgabenzeitraum

Der Aufgabenzeitraum für diese Aufgabe ist **vom 10.07.2020 bis 11.09.2020 (Abgabe)**.

---

<sup>1</sup><https://github.com/texttechnologylab/textimager-uima>

<sup>2</sup><https://github.com/dkpro/dkpro-core/tree/9758c8955ad34afe89522a77a7635c5a23cd4180/dkpro-core-stanfordnlp-gpl/src/main/java/org/dkpro/core/stanfordnlp>