Verbesserung automatischer Annotationen durch manueller Korrekturen in Echtzeit

Thema:
Verbesserung automatischer Annotationen durch manueller Korrekturen in Echtzeit
Art:
BA, MA
Betreuer:
Lars Döhling
Status:
Entwurf
Stichworte:
Natural language processing, Evaluation
angelegt:
2017-07-18

Hintergrund

Traditionell werden Texte beim Erstellen annotierter Korpora einmalig automatisch vorannotiert und anschließend manuell korrigiert. Aus den korrigierten Texten wird anschließend ein neues Modell trainiert, welches der Vorannotation neuer Texte dient.

Zielsetzung der Arbeit

Ziel ist, zu prüfen, ob sich erfolgte Korrekturen in Echtzeit gewinnbringend auf die restliche Vorannotation propagieren lassen.

Konkrete Aufgaben

- Anpassen der Viterbi-ähnlichen Inferenz in MarMoT (ein CRF-basierter Taggers) und/oder eines anderen geeigneten (statistischen) Taggers
- Evaluation der Performanz und des erreichbaren Verbesserungspotenzials anhand lateinischer Texte
- Ggf. Einbindung der Echtzeitkorrektur in den Annotationsworkflow (z.Zt. MarMoT mit Excel)

Erwartete Vorkenntnisse

Java

Weiterführende Quellen

Nach Absprache mit dem Betreuer.

From:
https://wiki.mi.ur.de/ - **MI Wiki**

Permanent link:
https://wiki.mi.ur.de/arbeiten/echtzeitannotationsverbesserung?rev=1500382998

Last update: **18.07.2017 13:03**