

CIS, LMU München
29.05.2017
Michael Strohmayer
11137111
michael.strohmayer@campus.lmu.de

Protokoll zur Sitzung 29.05.2017 – Computerlinguistisches Arbeiten (Repetitorium)

Testen von Signifikanz

In der heutigen Sitzung waren Signifikanztests das Thema. Diese sollen der Evaluierung von verschiedenen Systemen dienen um mit Sicherheit sagen zu können, dass ein System besser performt als ein anderes – hierfür wurde unter anderem der statistische Signifikanztest behandelt.

Außerdem wurden Bewertungseinheiten besprochen. Hier unter anderem Accuracy, Mean Average Precision (z.B. verwendet um Dokumente zu ranken) und Fscore, sowie Möglichkeiten zum Testen der unterschiedlichen Bewertungsschemata.

Einige Tests und Bewertungsmethoden sind bereits in der Bibliothek `scipy.stats` zusammengefasst. Hier ist unter anderem eine Funktion für einen Binomialtest (`scipy.stats.binom_test`), einen paired test (`scipy.stats.ttest_rel`) oder einen Signifikanztest (`scipy.stats.chisquare`) vorhanden.

Man sollte stets alle Testergebnisse mitangeben, auch wenn diese nicht perfekt in die Erwartungen passen. Desweiteren sollte man Tester nicht selbst entscheiden lassen, ob sie ein Ergebnis mitteilen oder nicht und erst ganz am Ende seiner Arbeit mit dem Testen beginnen.