

19.06.2015

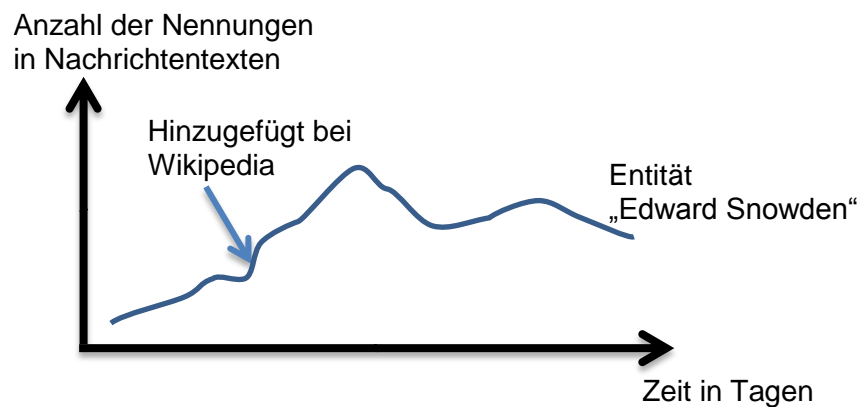
Bachelor-/Masterarbeit

„Trend Detection: Vorhersage der Entstehung von Wikipedia-Artikeln“ zu vergeben!

Um was geht es?

Wann werden Entitäten (Dinge wie „Edward Snowden“) als Seiten in Wikipedia angelegt? Diese Frage soll in dieser Abschlussarbeit geklärt werden. Dazu werden Nachrichtentexte über einen bestimmten Zeitraum (z.B. ein Jahr) analysiert und der Zeitpunkt des Hinzufügens zu Wikipedia (evtl. auch weiterer Edits) betrachtet. Aufbauend auf diesen Daten sollen menschenverständliche Regeln automatisch erlernt werden (etwa: Wenn die Entität in den vergangenen X Tagen Y Mal von der Nachrichtenquelle Z in der Überschrift vorkam, wird sie zu W% Wahrscheinlichkeit in den nächsten 24h zu Wikipedia hinzugefügt).

Zweiter Teil der Abschlussarbeit ist es, die erlernten Regeln auf neue Nachrichtentexte anzuwenden und so die Entstehung von Wikipedia-Seiten live vorherzusagen.



Die Nachrichtentexte und weitere Tipps zum Vorgehen werden im Rahmen der Arbeit zur Verfügung gestellt.

Was sollten Sie mitbringen?

- Interesse an *Data Mining* und *Predictive Analysis*.
- Vorkenntnisse in Data Mining-Tools wie Weka von Vorteil.
- Java-Programmierkenntnisse

Kontaktperson:
Michael Färber
michael.farber@kit.edu
Tel.: 0721/608 479 46