

Organisatorisches

Katja Markert, Ines Rehbein

Institut für Computerlinguistik
Uni Heidelberg
markert@cl.uni-heidelberg.de

April 23, 2019

Heute

- 1 Zeiten
- 2 Zielgruppe und Voraussetzungen
- 3 Struktur
- 4 Wie besteht man den Kurs?

Termine

Vorlesung/Übung

- Dienstag, 11.15 – 12.45, SR 26, INF 329
- Donnerstag, 11.15 – 12.45, SR26, INF 329

Poolsitzungen sind im CIP-Pool der Computerlinguistik an den folgenden Donnerstagen (anstatt der Vorlesung) mit Unterstützung von Philipp Wiesenbach.

- 1 9 Mai
- 2 23 Mai
- 3 13 Juni
- 4 4 Juli

Voraussetzungen

- Leistungsbewertung: 6 ECTS-Punkte
- Für Bachelorstudierende der Computerlinguistik
- Keine ÜK
- Studierende aus Anwendungsgebieten nach Absprache (mit vollen 6 LP)
- Voraussetzungen:
 - Prog I (nicht verhandelbar)
 - ECL (wenn noch nicht bestanden, aber Klausurzulassung erreicht, dann mit Absprache)
 - Pooltest

Zielgruppe und Niveau

Adressiert an Bachelorstudierende CL 2/3 Semester.

- Lernen “in der Tiefe” über eines der wichtigsten Konstrukte der CL (word embeddings)
- Mit möglichst vollständigen mathematischen Grundlagen zu diesem Thema
- Integration von Grundlagen, die nicht nur für embeddings relevant sind, sondern in der CL insgesamt eine große Rolle spielen (Informationstheorie, Signifikanztests, lineare Algebra . . .).
- Mit praktischen Programmieraufgaben und Poolsitzungen
- Vorträge und Diskussion von Papieren

Struktur

- 1 Erstes Drittel bis ca. 23.5 (Vorlesung/Übung, Markert)
 - 1 Embeddings, die auf Zählverfahren und linearer Algebra basieren
 - 2 Beispiele: sparse embeddings, SVD, PCA
 - 3 Hintergrund und Grundlagen: Assoziationsmaße, Signifikanztests, Entropie, Matrixoperationen, Matrixfaktorisierungen
- 2 Zweites Drittel bis ca 25.6. (Vorlesung/Übung, Rehbein)
 - 1 Embeddings, die auf neuronalen Netzen basieren
 - 2 Beispiele: CBOW, Skip-Gram, FastText
 - 3 Grundlagen: Softmax, (Stochastic) Gradient Descent, ...
- 3 Drittes Drittel (Seminarcharakter, Markert/Rehbein)
 - 1 Anwendungen, Probleme, Erweiterungen: Bias, Multimodale Embedding, Multilinguale Embeddings etc.
 - 2 Studierendenkurzvorträge, Diskussionen etc.

Webseite mit Vorlesungsplan sowie Material: <https://www.cl.uni-heidelberg.de/courses/ss19/emb/material/>

[//www.cl.uni-heidelberg.de/courses/ss19/emb/material/](https://www.cl.uni-heidelberg.de/courses/ss19/emb/material/)

Leistungsnachweise

- Bearbeitung von vier Übungsaufgaben im Paar. Keine benotete Korrektur. Sinnvolle Bearbeitung von allen Übungsaufgaben. Voraussetzung zur Klausurzulassung.
- Kurzreferate im letzten Drittel der Veranstaltung.
- Klausur
- Note wird von Kurzreferat, Diskussionsbeiträgen (Fragenbeantwortung) (30%) sowie Klausur (70%) bestimmt.

Längere Krankheiten etc müssen rechtzeitig mit ärztlichem Attest belegt werden und das Prozedere dann mit uns abgesprochen werden

Anwesendheitspflicht im Referatsteil (bei eigenen und Fremdreferaten re Diskussion).

Übungsaufgaben

- Hauptsächlich praktische Aufgaben.
- Abgabe: Email an `embeddings@cl.uni-heidelberg.de` als pdf oder Code
- Deadlines: Jeweils auf den Übungsblättern angegeben
- Letzte Aufgabe geringer Umfang
- Paarabgaben
- Klausurzulassung
- Hilfe zu den Übungen jeweils **vor** der Abgabe in den Poolsessions!

Kurzvorträge

- Ziel: Einüben von Papiere lesen und präsentieren. Benotet nach Verständnis und Präsentation.
- Nicht so umfangreich wie in sonstigen Seminaren (max. 20 Minuten)
- Wichtige, relativ grundlegende Papiere, angemessen dem 2-4 Semester Bachelorniveau.
- Themen aus angegebener Literaturliste in den *Einführungsfolien* zur Auswahl
- **Abgabe Themenpräferenzen:** Per email an `markert|rehbein@cl...` bis 9.5. mit Inhalt
 - Name, Matrikelnummer, Semester, Genauer Studiengang
 - Mind 2 Themen/Papiere

Klausur

- 50% in der Klausur berechtigen zum Bestehen
- Es gibt keine Probeklausur. Nachklausur nur im Krankheitsfall
- Wahrscheinlich Freitag letzte Vorlesungswoche
- Commitmentfrist (wird später bekanntgegeben): Alle Teilnehmenden müssen sich zur Klausur anmelden
- Klausurthemen beschränken sich (ganz oder fast ganz) auf den Vorlesungs/Übungsteil bis Ende Juni.

Hilfe

- Bitte in der Vorlesung durchaus nachfragen
- Poolsessions
- Sprechstunden Markert/Rehbein: Donnerstag 13.00 – 14.00
- Administrativ: sekretariat@cl.uni-heidelberg.de
- Studienberatung: studienberatung@cl.uni-heidelberg.de