
Anwendungen der KI / SoSe 2018

Organisatorisches

Prof. Dr. Adrian Ulges

Angewandte Informatik / Medieninformatik / Wirtschaftsinformatik / ITS

Fachbereich DSCM
Hochschule RheinMain

Kurs-Website: www.ulges.de

(Übersicht Kursteilnehmer!)

(brauche Informatik-LDAP-Kennung für Zugriff auf Materialien)



Adrian Ulges

Anwendungen der Künstlichen Intelligenz (SoSe 18)

News

Teilnehmer und Gruppenaufteilung

Die Veranstaltung ist stark nachgefragt, so dass nicht für alle Studierenden Plätze vorhanden sind. Ich habe nun nach folgenden **Regeln** die Plätze vergeben (siehe Tabelle unten):

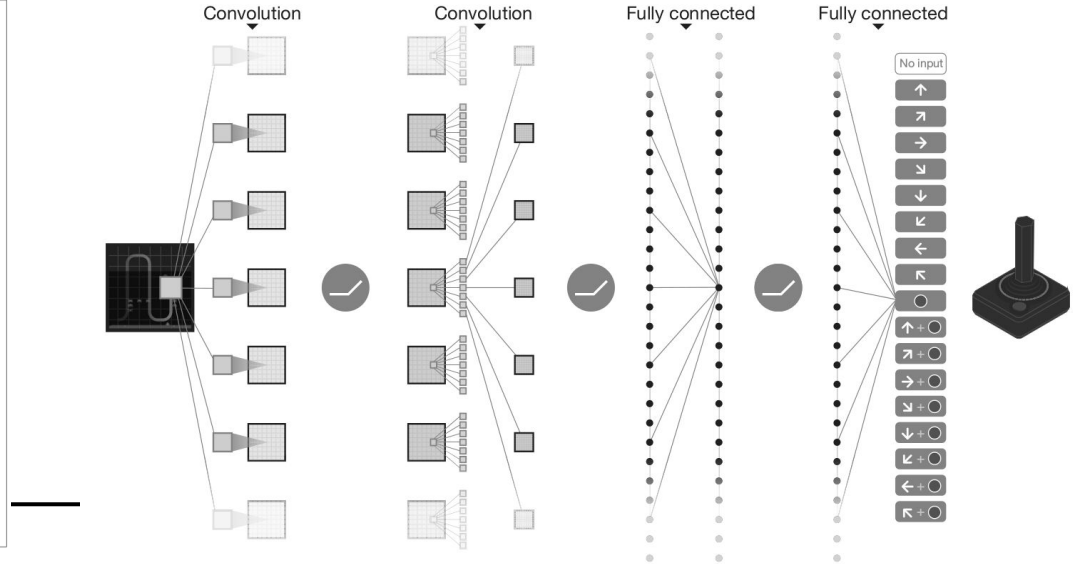
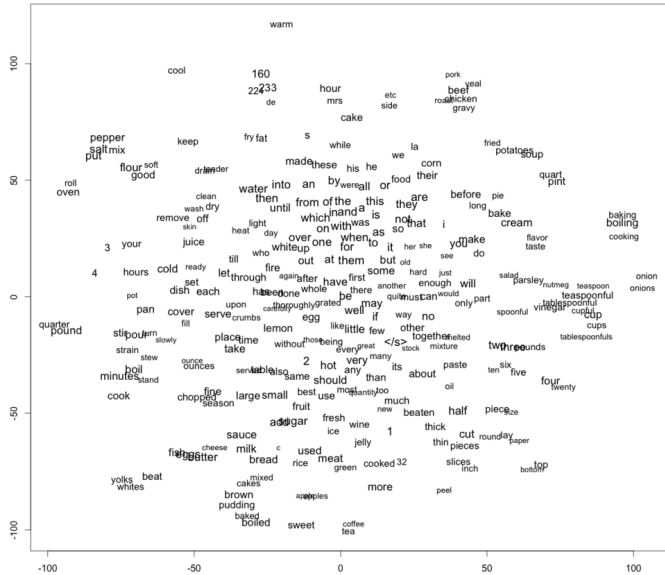
- **Teilnehmer:** Wer das **Praktikum** über IGEL oder AoR belegt hat, darf definitiv teilnehmen.
- **Fairness:** Wenn Sie feststellen, dass das Modul doch nicht das richtige für Sie ist, geben Sie bitte **zeitnah bescheid** - es gibt weitere Interessenten für Ihre Plätze!
- **Gruppenverteilung:** Durch Unregelmäßigkeiten bei den Belegungen wurden leider zu viele Studierende in Gruppe A verteilt. Ich habe per Zufallsauswahl Teilnehmer in Gruppe B verlegt (und hierbei sichergestellt, dass kein Konflikt mit anderen Lehrveranstaltungen existiert). Ein Tausch von Gruppen ist nur möglich, falls ein **einvernehmlicher Tauschpartner** gefunden wird. Zweierteams für das Praktikum können nur **innerhalb derselben Praktikumsgruppe** gebildet werden.
- Die verbliebenen freien Plätze in Gruppe B habe ich per **Zufallsauswahl** verteilt.
- **Warteliste:** Die übrigen sind auf einer **Warteliste**. Ich weise aber darauf hin, dass (nach Auskunft Ihrer Studiengangsleitungen) **in anderen 10CP-Veranstaltungen definitiv noch Plätze frei** sind.
- Sollten **vor Ende der QIS-Deadline Plätze frei** werden, werde ich Nachrücker nach Reihenfolge der Liste kontaktieren und erwarte jeweils **Rückmeldung binnen eines Tages**, sonst ist der nächste dran.
- Für die, die nicht reingekommen sind: Es tut mir leid, ich habe versucht **unter den gegebenen Rahmenbedingungen fairstmöglich** zu verfahren. Ich hoffe Sie verstehen dass ich keine Nachbesserungen am obigen Verfahren vornehmen kann.

Gruppe A	Gruppe B	Warteliste
Heeß, Kevin	Petrak, Dominic	1. Lamkadmi, Mostafa
Lamott, Marcel	Kuhlbrodt, Timo	2. Hemp, Niklas
Diez, Savina	Gehrke, Konstantin	3. Elmas, Onur
Kissel, Jens	Wagner, Julian	4. Shamun, Tiglat
Schönberger, Patrick	Carl, Jan-Luca	5. Sancar, Semih
Krafft, Mathieu	Ramsahye, Yannick	6. Siegert, Jan-Phillip
Koch, Nils	Klonowski, Fabio	
Riegler, Patrick	Wunn, Robin	
Lies, Juliane	Bahrololoomi, Farnod	
Mathis, Vincent	Hanif, Mahmud	
Le, Michelle	Fracella, Adrian	
Leipelt, Dominik	Wehenkel, Lukas	
Weis, Hendrik	Eldags, Dominik	



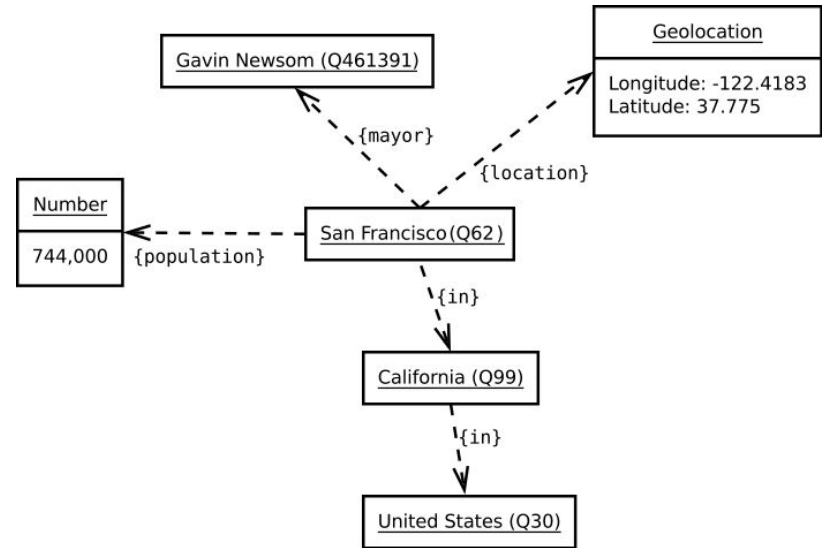
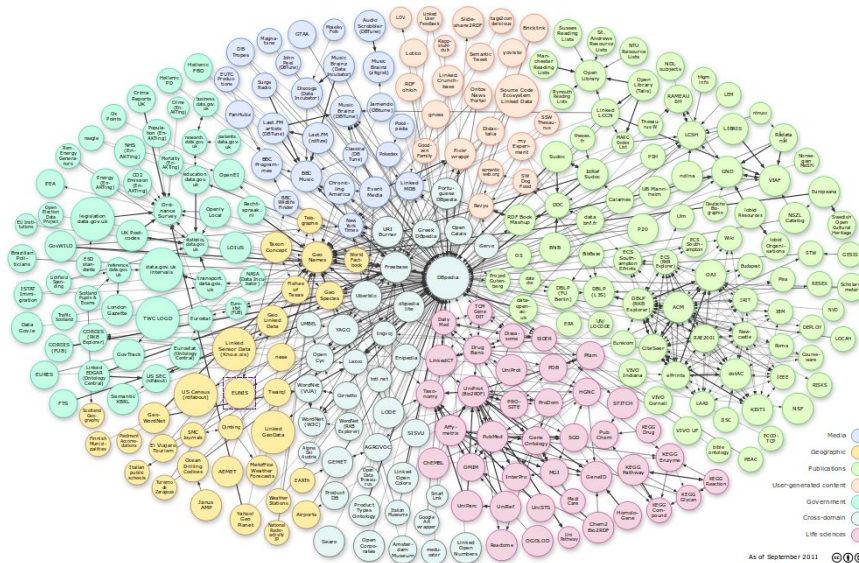
Themen: Machine Learning

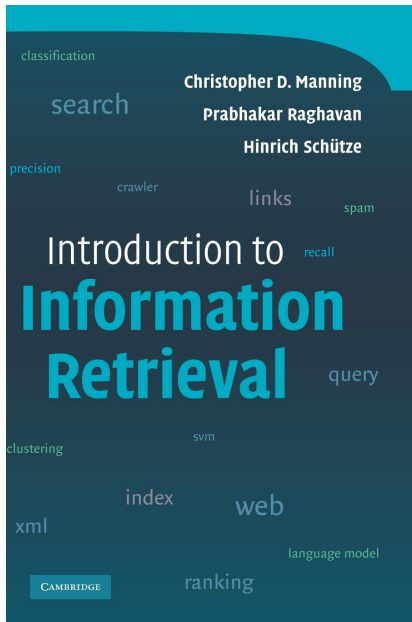
- Grundlagen und Grundbegriffe des Machine Learning
- Neuronale Netze und Deep Learning
- Anwendung: Natural Language Processing (Python)



Themen: Semantic Web

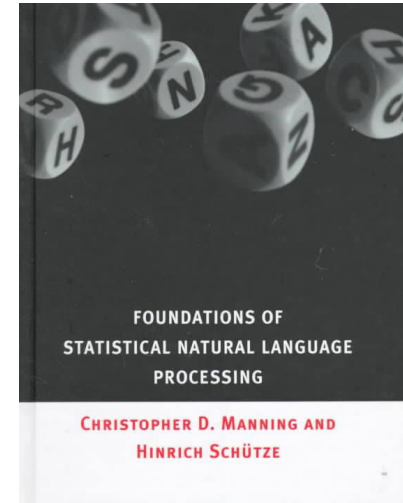
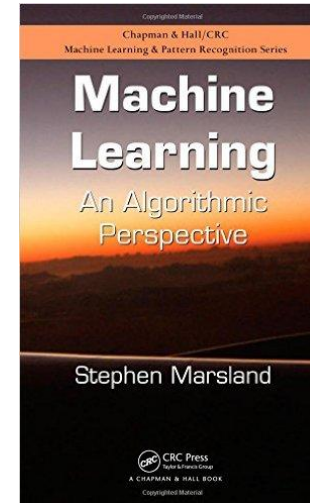
- Wissensgraphen als Sammlungen von Tripeln
- Repräsentation mit RDF, Query-Sprache SPARQL
- **Große Ontologien: Linked Open Data (LOD)**





Literatur

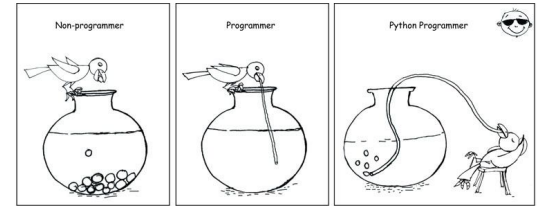
- **Nielsen: Neural Networks and Deep Learning.**
<http://neuralnetworksanddeeplearning.com/>
- **Jurafski, Martin: Speech and Language Processing.**
<https://web.stanford.edu/~jurafsky/slp3/>
- Manning et al.: Introduction to Information Retrieval.
- Marsland: Machine Learning - an Algorithmic Perspective, CRC Press, 2009 (Bibliothek).
- Manning, Schütze: Foundations of Statistical Language Processing (Bibliothek, e-Book).
<https://hds.hebis.de/hdrm/Record/HEB398591768>



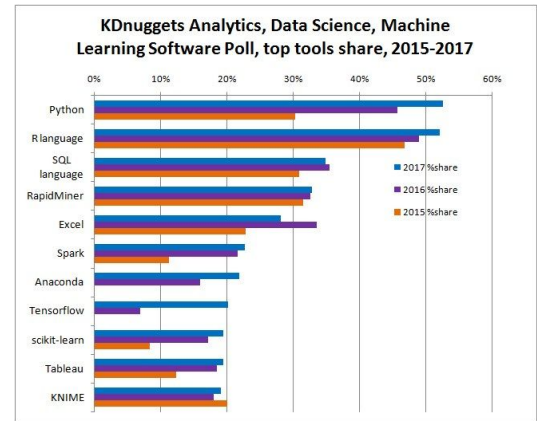
Warum Python?

Weltweit verbreitetste Sprache
im Bereich Data Science

- + Populärste Skriptsprache weltweit
(Rang 4 in TIOBE⁽¹⁾)
- + voll objektorientiert
- + extrem leicht zu lernen
(“Pseudo-Code that runs”)
- + funktionale Elemente
- + Mächtige Bibliotheken (*numpy*,
sklearn, *pandas*, *matplotlib*)
- + Datenintegration (*Web*, *Information
Retrieval*, *NLP*, *Bildverarbeitung*)



Feb 2018	Feb 2017	Change	Programming Language	Ratings	Change
1	1		Java	14.988%	-1.69%
2	2		C	11.857%	+3.41%
3	3		C++	5.726%	+0.30%
4	5	▲	Python	5.168%	+1.12%
5	4	▼	C#	4.453%	-0.45%



(1) <http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html>

Der AnwKI - Technologie-Stack

SPARQL

(Semantische Suche)

Tensorflow

(Neuronale Netze)

ElasticSearch

(Textsuche)

NLTK

(Natural Language
Processing)

Numpy, Sklearn, Matplotlib, Pandas

(Python for Data Science)

Python

Termine	Vorlesung MO, 10:00, D-17	"Praktikum" Block 1 MO, 14:15, D-17	"Praktikum" Block 2 MO, 11:45, D-13 <u>oder</u> MO, 16:00, D-12
16.04.	Einführung (KI & QA)	Information Retrieval	Python-Crashkurs Suchmaschine (ElasticSearch)
23.04.	Information Retrieval	NLP	Suchmaschine (ElasticSearch)
30.04.	Machine Learning	Neuronale Netze	Query Expansion (NLTK), Neuronale Netze (Tensorflow)
07.05.	Neuronale Netze im NLP	Neuronale Netze im NLP	Neuronale Netze (Tensorflow)
14.05.	NLP	Semantic Web	Neuronale Netze (Tensorflow), Wordnet (NLTK)
	Pfingsten	Pfingsten	
28.05.	Semantic Web	Reserve	DBPedia & SPARQL
04.06.	PROJEKT	PROJEKT	
11.06.			
18.06.			
25.06.			
02.07.			

Vorlesung

Praktikum

Projekt

Zeitaufwand



- ❑ Dies ist ein **10-CP-Fach**
→ Rechnen Sie mit hohem Arbeitsaufwand!
 - ❑ 300 h / 15 Wochen = 20 Stunden pro Woche
 - ❑ 6 Stunden Präsenzzeit
 - ❑ 4 Stunden Vor-/Nachbearbeitung
 - ❑ **10 Stunden praktische Tätigkeit**
(berechnet für einen "*Normstudent*")
-

Benotung

- **Prüfungsform:** Praktische Tätigkeit und Fachgespräch
 - 70% Projekt
 - 30% Fachgespräch
- **Frühe Anmeldung** im QIS (keine weitere Warnung)!
- Praktikum: **Anwesenheitspflicht**

QIS-Deadline

23.04.

Benotung (cont'd)

- **Projekt**
 - Aufbau eines Question-Answering-Systems
 - Entwicklung in Python, Teams von 2 Studierenden
 - **Contest**
 - Note auf Grundlage von
 - **Review der Abgabe (Code, Doku, Ergebnisse)**
 - **wöchentlichem Fortschritt**
 - **Fachgespräch**
 - Präsentation + Fragen
 - Fragestellungen ausgehend vom Projekt...
 - decken aber den gesamten Kurs ab
 - vorbereiten wie für eine mündliche Prüfung!
-

Voraussetzung: Englisch

- Vorlesung, Praktikum, Prüfung in **Deutsch**
 - Aber: “Zielsprache” dieser Veranstaltung ist Englisch!
 - Projekt und Übungen zielen darauf ab, **Englische Fragen** mit Hilfe eines **Englischen Korpus** zu beantworten.
 - **Begründung**
 - viele Modelle und Trainingsdaten in Englisch verfügbar
 - besserer Lerneffekt, wenn die Sprache nicht die eigene ist
 - Dementsprechend sind auch viele **Beispiele** in der Vorlesung in Englischer Sprache.
 - **Grundlegende Englisch-Kenntnisse** werden deshalb vorausgesetzt!
-

Who
murdered
Abraham
Lincoln ?

John Wilkes
Booth.

Projekt: Question Answering

- Datengrundlage: "Wikibase" (300,000 Wikipedia-Seiten)
- **Komponenten** werden bereits in den Übungen entwickelt
- Evaluation im Rahmen eines **Contests**

Anwendungen der Künstlichen Intelligenz (SS17)

Scoreboard Question Answering

High Scores

rank		team	score(percent)
1.	🏆	Artificial Stupidity	33.0
2.	🔫	Team Triggered	30.0
3.	🐷	Die eisernen Einhoerner	28.3