

# Computerlinguistische Anwendungen

Benjamin Roth

CIS LMU

# Organisatorisches

- 3 Stunden Vorlesung Mi 16:00-17:00 (c.t.) und Do 14:00-16:00, L 155
- 2 Stunden Übung Mi 14:00-16:00 (c.t.), L 155
- Zusatztutorat: Do 16:00-18:00 (c.t.), CIP-Räume Sibirien/Gobi
- ECTS: 9
- Prüfung: Klausur (wird noch festgelegt, voraussichtlich letzte Vorlesungswoche)

# Übungsaufgaben

- Bearbeitung in Teams von 2-3 Personen. (Bitte überlegen Sie sich bis zum Tutoratmorgen die Teams → Webformular)
- Ausgabe der Übungen: Mittwochs nach der Übung
- Abgabe-Deadline jeweils: Mittwochs 14:00 Uhr vor der Übung
- Bewertung der Übungen u.a. automatisiert durch Unit-Tests.
- Klausurbonus: In Abhängigkeit der erreichten Übungspunkte wird ein Klausurbonus von bis zu 10% der maximal erreichbaren Klausurpunkte gewährt, wenn die Klausur auch ohne die Bonuspunkte als bestanden gewertet würde.

# Themen

- Einführung in Machine-Learning und Implementierung einfacher Klassifikatoren
  - ▶ Perceptron
  - ▶ Naive Bayes
- Einführung in die Machine-Learning Bibliotheken Numpy/Scikit-Learn
- Grundlagen von Neuronalen Netzwerken für die Sprachverarbeitung, die Bibliothek Keras
- Implementierung typischer NLP-Anwendungen wie z.B.
  - ▶ Spam-detection
  - ▶ Sentiment-Analyse
  - ▶ Paraphrasen-Erkennung
  - ▶ Wortartenerkennung
  - ▶ Erkennung von Entitäten-Typen
  - ▶ Klassifikation von Relationen
- Evtl. Firmenvorträge gegen Ende des Semesters.

# Plan für heute: Wiederholung

- Repräsentation von Dokumenten/Trainingsinstanzen in Python
- Klassifikation mit Perzeptron
- Unit Tests