

Seminar: 1.07.251 Freiwählbares Modul:
Computational Social Sciences Methods in R
(Lehrsprache englisch/Prüfungsleistung deutsch o.
englisch)

Dr. Taehee Kim

Summer semester 2022

Contact Email: taehee.kim@uol.de
Office: A06 3-317
Office hour: By appointment

Place Prep meeting: Online (BBB)
Block Session: A06 3-313 (OLExS-Labor).

Date and Time This seminar consists of three block appointments.

First Block Friday. 06.05.22 14:15 - 17:45
Saturday. 07.05.22 09:15 - 14:45

Second Block Friday. 10.06.22 14:15 - 17:45
Saturday. 11.06.22 09:15 - 14:45

Third Block Friday. 08.07.22 14:15 - 17:45
Saturday. 09.07.22 09:15 - 14:45

1 Course Description

This course aims to introduce the concepts and techniques of computational social science, especially focusing on machine learning. Nowadays machine learning has become a powerful tool for social science studies. Various libraries and APIs make it easy for researchers to create and utilize machine

learning models. This course introduces supervised machine learning algorithms and its practice especially for the data type heavily used in social science studies, text and images. As for programming languages, we use R for data wrangling and Python for machine learning. The course starts with a basic introduction of R and Python programming. After the introduction of machine learning algorithms including deep learning, students will learn how to implement the algorithms in Python. Along with modeling machine learning by yourself, how to use various machine learning APIs (e.g., Google's Machine Learning APIs) will be introduced as well. The course is held in English. It is designed for students at various levels with a general interest in machine learning focusing on social science context. Previous experience in R/Python is not required.

The course is held in **English**. However, **participants can write a final report either in English or German** (For more detail of a final report, see *Requirements and Grading* section below). The seminar consists of three block appointments, each of which consists of two consecutive days of training and practice (Fri. 14:15 - 15:45 and Sat. 9:15 - 14:45). Previous experience in R/Python is preferable but not required. R refresher sessions will be offered at the first block of the seminar. Basic Python programming required for implementing machine learning algorithms will be provided during the seminar as well.

References

- Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy and Marshall Van Alstyne. (2009) Computational Social Science. *Science* 323(5915):721-723.
- Lubanovic, Bill. (2014) *Introducing Python: Modern Computing in Simple Packages*. O'Reilly Media.
- Raschka, Sebastian and Vahid Mirjalili. (2017) *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow*. 2nd ed. PACKT Publishing.
- Swaroop, C. H.. (2013) A Byte of Python [Webpage](#)
- Kim, A. Y., McConville, K. S. (2019) *Statistical Inference via Data Science: A Modern Dive into R and the Tidyverse*. Chapman & Hall/CRC

The R Series. [Webpage](#)

Wickham, H., Grolemund, G. (2017) R for Data Science. O'Reilly.
[Webpage](#)

2 Requirements and Grading (6 ECTS)

- Solve exercises during the seminar (30%).
- Submit a final report summarizing the result of a machine learning task and corresponding code files (70%). Participants can choose one of the machine learning tasks (text or image classification) given by the lecturer. Instead of a given task, participants can also set up a task for themselves based on their own research interests after a consultation with the lecturer.

The length of the paper should be **6 pages** at maximum. Participants can write it in either English or German. Participants should directly submit it to the lecturer by email.

Deadline: 15th September, 2022

3 Course Outline

First Block: 06.05.22 – 07.05.22

06.05.22 (Friday)

14:15-15:45 Introduction to Computational Social Science

16:15-17:45 Machine Learning in Social Science Studies

07.05.22 (Saturday)

09:15-11:45 R Refresher

12:00-13:00 Lunch

13:15-14:45 Introduction to Python

Second Block: 10.06.22 – 11.06.22

10.06.22 (Friday)

14:15-15:45 Introduction to Machine Learning

16:15-17:45 Machine learning Algorithms for text classification

11.06.22 (Saturday)

09:15-11:45 Machine Learning for Text Classification in Python

12:00-13:00 Lunch

13:15-14:45 Introduction to Google's Machine Learning APIs and Set up

Third Block: 08.07.22 – 09.07.22

08.07.22 (Friday)

14:15-15:45 Deep Learning and Machine Learning Algorithms for image classification 1

16:15-17:45 Deep Learning and Machine Learning Algorithms for image classification 2

09.07.22 (Saturday)

09:15-11:45 Image classification in Python

12:00-13:00 Lunch

13:15-14:45 Q & A
