

Data Mining und Maschinelles Lernen

Prof. Kristian Kersting
Steven Lang
Felix Friedrich



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Sommersemester 2022
5. April 2022
Übungsblatt 1

Diese Übung wird nicht besprochen oder bewertet.

Benötigte Dateien

Alle benötigten Datensätze und Skriptvorlagen finden Sie in unserem Moodle-Kurs:

<https://moodle.informatik.tu-darmstadt.de/course/view.php?id=1194>

1.1 Einführung in Python

In der Übung zu Data Mining und Maschinelles Lernen werden wir für Programmieraufgaben die Skriptsprache Python (Version ≥ 3.5) verwenden. Ein schneller Weg, Python mitsamt nützlichen Standardbibliotheken zu installieren, ist die Anaconda Plattform (Open-Source Distribution):

- <https://www.anaconda.com/products/individual#Downloads>

Anaconda beinhaltet nach Auswahl ebenfalls die interaktiven Entwicklungsumgebungen *spyder* sowie *jupyter*.

Alternativ können Sie Python auch direkt installieren:

- <https://www.python.org/downloads/>

und ggf. die Entwicklungsumgebung *Pycharm* (Community Edition) kostenfrei herunterladen:

- <https://www.jetbrains.com/pycharm/download/>

Zusätzliche Programmbibliotheken installieren Sie in Python über den Paketinstaller *pip*. Rufen Sie hierfür eine Konsole auf und ersetzen Paket-Name mit dem jeweiligen Paketnamen.

```
$ pip3 install <Paket-Name>
```

a) Installation

Installieren Sie die Programmbibliotheken *numpy*, *matplotlib* und *scikit-learn*.

b) Einstieg in Python

Zu Beginn der ersten Übungsstunde wird eine kurze Einführung in Python und *numpy* stattfinden.

Wenn Sie mit der Skriptsprache Python noch nicht vertraut sind, empfehlen wir Ihnen das Jupyter-Notebook `python_intro.ipynb` zu untersuchen oder der Einführung auf

<https://cs231n.github.io/python-numpy-tutorial/>

zu folgen.