

# Classification and Clustering

## Allgemeine Angaben

<b>Kürzel</b>	CLA
<b>Modulverantwortliche</b>	Prof. Dr. Seifert,
<b>Dozenten</b>	Prof. Dr. Seifert,
<b>Lehrsprache</b>	Deutsch
<b>Semester</b>	2
<b>ECTS-Punkte</b>	5
<b>Kontaktstunden</b>	40
<b>Selbststudium</b>	85
<b>Dauer</b>	1 Semester
<b>Art</b>	Pflicht im Rahmen der Spezialisierung
<b>Häufigkeit</b>	jedes Studienjahr
<b>Gewichtung</b>	5/120
<b>Prüfungsleistung</b>	RS

## Stichwörter

- Klassifikation, Clusteranalyse
- Machine Learning

## Zugangsvoraussetzungen

- Forecasting

## Verwendbarkeit

- Knowledge Engineering

## Qualifikations- und Kompetenzziele

Die Teilnehmer besitzen ein gut fundiertes Wissen über Methoden und Verfahren des Data Mining. Dazu gehören Methoden zur Klassifikation und zum Clustering von mehrdimensionalem Datenmaterial. Sie können Techniken im Bereich des Machine Learning einschätzen und erarbeiten, klassische und moderne Methoden wissenschaftlich fundiert darstellen und in neuartigen Praxisbeispielen einsetzen.

## Lehr- und Lernmethoden

Vorlesung, Gruppenarbeiten, Fallstudien, Einsatz einer Statistiksoftware

## Inhalte

- Einordnung
  - Klassifikation vs. Regression
  - Überwachte vs. unüberwachte Verfahren
  - Classification vs. Clustering
- Klassifikations-Methoden, z.B.
  - k-nächste Nachbarn
  - Classification trees
  - Logistische Regression
  - Naive Bayes
- Cluster-Analysis, z.B.
  - k-means
  - Hierarchisches Clustering
- Anwendungen, z.B.
  - Bilderkennung, -analyse

- Assoziationsanalyse
- Bots
- NLP

### Grundlegende Literaturhinweise

GALIT SHMUELI, PETER BRUCE, PETER GEDECK und NITIN PATEL, 2020: *Data Mining for Business Analytics – Concepts, Techniques and Applications in Python*, Hoboken: Wiley.

JÖRG FROCHTE, 2019: *Maschinelles Lernen – Grundlagen und Algorithmen in Python*, München: Hanser.

AURÉLIEN GÉRON, 2020: *Praxiseinstieg Machine-Learning mit Scikit-Learn, Keras und TensorFlow*, Heidelberg: O'Reilly.

### Ergänzende Literaturempfehlungen

SEBASTIAN RASCHKA, VAHID MIRJALILI, 2021: *Machine Learning mit Python und Keras, Tensorflow 2 und Scikit-Learn*, Frechen: mitp.

TREVOR HASTIE, ROBERT TIBSHIRANI, 2008: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, New York: Springer.