

# Einführung in Data Mining mit R

<b>Termin:</b>	Mittwoch, 18.07.2018, 9:00-17:00 Uhr Donnerstag, 19.07.2018, 9:00-17:00 Uhr
<b>Anmeldefrist:</b>	03.07.2018
<b>Raum:</b>	wird per E-Mail bekannt gegeben
<b>Dozent:</b>	Ferdinand Schirrmeister, eoda, Kassel
<b>Kurs-Nr.:</b>	2018/09_GA
<b>Kurssprache:</b>	Deutsch; verschobener Termin!

Data Mining steht für das hypothesenfreie Extrahieren von Informationen aus Daten. Statistische und mathematische Verfahren werden auf Datenbestände angewendet, um bestehende Muster und Zusammenhänge aufzudecken. Data Mining Verfahren stellen in der Regel geringe Anforderungen an das Messniveau der Daten (kategorial, ordinal, metrisch) und sind in der Lage, komplexe, nicht-lineare Zusammenhänge zu erkennen.

Ziel des Workshops ist es, die Anwendung der Data-Mining-Verfahren anhand von praxisnahen Übungssätzen zu erlernen. Durch die Beispiele werden die zentralen Schritte (wie das vorbereitende Datenmanagement, das Trainieren der Algorithmen sowie die Prognoseerstellung und Validierung) erlernt und direkt in R umgesetzt. Während des Kurses werden R-Skripte erstellt, die den Teilnehmern als Vorlage für eigene Data-Mining-Anwendungen dienen können.

## Inhalte:

- Einführung in die Grundbegriffe des Data Mining
- Einführung in die beliebtesten Datamining Algorithmen, wie z. B. Decision Trees, Random Forest und GBM
- Vorstellung von Validierungstechniken
- Vorstellung des Data-Mining-Frameworks caret

Die Anmeldung erfolgt über das Anmeldeformular auf der Homepage der Graduiertenakademie. Sobald die Graduiertenakademie die Anmeldung bestätigt, gilt sie als verbindlich. Die Rechnung über die Kursgebühr (40,- Euro) geht per Post an die/den Teilnehmer/in und muss bis spätestens zwei Wochen nach dem Kurstermin überwiesen werden. Es gelten die auf dem Anmeldeformular aufgeführten Teilnahmebedingungen.