

| | | | |
|----------------------------|---|------------------------|---------------------------------|
| Lehrende/r | Dr. Fabio Valdés | Modulverantwortliche/r | Dr. Fabio Valdés |
| | Dauer des Moduls ein Semester | ECTS 10 | Workload 300 Stunden |
| | | | Häufigkeit in jedem Semester |
| Lehrveranstaltung(en) | Data Mining: Konzepte und Techniken | | |
| Detaillierter Zeitaufwand | Bearbeiten der Lektionen: 160 Stunden Bearbeitung der Einsendeaufgaben inkl. Verarbeitung des Korrektur-Feedbacks: 80 Stunden Wiederholung und Prüfungsvorbereitung, Prüfung: 60 Stunden | | |
| Qualifikationsziele | Nach erfolgreicher Teilnahme besitzen die Studierenden einen umfassenden Überblick zu Wissensentdeckungsprozessen in Datenmengen/-banken. Sie sind in der Lage, verschiedene Attributtypen zu beschreiben und zu visualisieren sowie entsprechende Abstandsmaße zu berechnen. Sie besitzen detaillierte Kenntnisse zur Datenvorverarbeitung. Sie sind mit den Zielen und Methoden der zentralen Data-Mining-Techniken Mustersuche, Klassifikation und Clusteranalyse vertraut. Zudem kennen sie sich mit der Analyse komplexerer Strukturen, etwa Zeitreihen oder Graphen, aus. | | |
| Inhalte | Das Thema dieser Lehrveranstaltung ist Data Mining, grob übersetzbar mit "Wissensentdeckung in Datenmengen/-banken". Die Bedeutung dieses Themengebiets ist in den letzten Jahren rasant gewachsen. Die Zielsetzung besteht darin, Strukturen, Zusammenhänge sowie Gruppen ähnlicher Objekte in sehr großen Datenmengen zu erkennen und zu bewerten. Die Lehrveranstaltung vermittelt zunächst Kenntnisse zur Vorbereitung von Data-Mining-Methoden hinsichtlich der Charakterisierung (z.B. Klassifizierung von Attributtypen, Visualisierung) und Vorverarbeitung der Daten (etwa durch Eliminierung von Ausreißern, Aggregation oder Normalisierung). Darauf aufbauend, werden verschiedene Techniken zur Mustersuche (z.B. Apriori-Algorithmus), Klassifikation (u.a. Entscheidungsbäume, Klassifikation nach Bayes) und Clusteranalyse (beispielsweise k-Means, DBSCAN) sowie passende Evaluationsmethoden vorgestellt. Zudem erläutert die Lehrveranstaltung, wie komplexere Strukturen, d.h. Datenströme, Textdokumente, Zeitreihen, diskrete Folgen, Graphen sowie Webdaten, analysiert werden können. Ein Kapitel mit praktischen Beispielen in Weka bildet den Abschluss der Lehrveranstaltung. | | |
| Inhaltliche Voraussetzung | Keine | | |
| Lehr- und Betreuungsformen | Lehrveranstaltungsmaterial Einsendeaufgaben mit Korrektur und/oder Musterlösung internetgestütztes Diskussionsforum Studientag/e Betreuung und Beratung durch Lehrende | | |
| Anmerkung | | | |
| Formale Voraussetzung | keine | | |
| Verwendung des Moduls | B.Sc. Informatik B.Sc. Mathematisch-technische Softwareentwicklung B.Sc. Wirtschaftsinformatik M.Sc. Data Science M.Sc. Informatik | | |

M.Sc. Praktische Informatik
M.Sc. Wirtschaftsinformatik

Prüfungsformen

Prüfung

Stellenwert 1/8
der Note

Art der Prüfungsleistung

benotete zweistündige
Prüfungsklausur

Voraussetzung

keine