

**Modulbezeichnung:** Data Warehousing und Knowledge Discovery in Databases (DWKDD) 5 ECTS

(Data Warehousing and Knowledge Discovery in Databases)

Modulverantwortliche/r: Klaus Meyer-Wegener

Lehrende: Thomas Ruf, Klaus Meyer-Wegener

Startsemester: SS 2019

Dauer: 1 Semester

Turnus: jährlich (SS)

Präsenzzeit: 60 Std.

Eigenstudium: 90 Std.

Sprache: Deutsch und Englisch

### Lehrveranstaltungen:

Ausschlussbedingung: Dieses Modul darf nur abgelegt werden, wenn keine der im Modul enthaltenen Lehrveranstaltungen auch noch in einem anderen Modul enthalten ist, das bereits abgelegt wurde.

Knowledge Discovery in Databases (SS 2019, Vorlesung, 2 SWS, Klaus Meyer-Wegener)

Data Warehousing (SS 2019, Vorlesung, 2 SWS, Thomas Ruf)

### Empfohlene Voraussetzungen:

siehe Lehrveranstaltungen

### Inhalt:

siehe Lehrveranstaltungen

### Lernziele und Kompetenzen:

siehe Lehrveranstaltungen

#### Fachkompetenz

##### Wissen

KDD: Die Studierenden

- kennen den typischen KDD-Prozess;
- kennen Verfahren zur Vorbereitung von Daten für das Data Mining;
- wissen, wie ein typisches Data Warehouse aufgebaut ist;
- kennen die Definition von Distanz- bzw. Ähnlichkeitsfunktionen für die verschiedenen Typen von Attributen;
- sind vertraut mit dem Prinzip des Apriori-Algorithmus zur Bestimmung von Mengen häufiger Elemente (frequent itemsets);
- kennen den FP-Growth-Algorithmus zum schnellen Auffinden von Mengen häufiger Elemente;
- können die Definitionen von Support und Confidence für Assoziationsregeln wiedergeben;
- können die Ermittlung von Assoziationsregeln auf der Basis von Mengen häufiger Elemente beschreiben;
- sind in der Lage, die Vorgehensweise bei Klassifikationsaufgaben wiederzugeben;
- können darlegen, wie ein Entscheidungsbaum auf einem Trainingsdatensatz erzeugt wird;
- können das Prinzip der Bayes'schen Klassifikation darstellen;
- können verschiedene Clustering-Verfahren aufzählen;
- können den Ablauf von k-Means-Clustering beschreiben;
- kennen die verschiedenen Arten von Ausreißern.

##### Verstehen

KDD: Die Teilnehmer können

- Distanz- oder Ähnlichkeits-Funktionen auf einem speziellen Datenbestand definieren;
- Attribute eines Datensatzes auf ihre Bedeutung für die Analyse hin überprüfen und ggf. Attributwerte geeignet transformieren.

#### Lern- bzw. Methodenkompetenz

KDD:

Die Studierenden

- müssen ein hohes Maß an Eigeninitiative zur Aneignung des Stoffes mitbringen und werden darin bestärkt;
- müssen die englische Sprache so weit beherrschen, dass sie der Vorlesung folgen können;
- verwenden gezielt ein Lehrbuch zur Vorlesung;

- bekommen zahlreiche Hinweise auf zugrundeliegende Literatur, die sie bei Bedarf heraussuchen und durcharbeiten müssen.

#### *Selbstkompetenz*

KDD:

Die Teilnehmer

- müssen ggf. fehlendes Vorwissen selbständig nacharbeiten (falls sie diese Lehrveranstaltung trotzdem wählen);
- müssen sich selbst die Zeit für die Vor- und Nachbereitung des Stoffs einteilen.

#### *Sozialkompetenz*

KDD:

Die Teilnehmer

- eignen sich den Stoff in einer Gruppe mit extrem verschiedenen Hintergründen an;
- brauchen das Gespräch mit anderen Teilnehmern zur Aneignung des Stoffs;
- müssen sich in der Diskussion auch selbst der englischen Sprache bedienen.

#### **Literatur:**

siehe Lehrveranstaltungen

#### **Verwendbarkeit des Moduls / Einpassung in den Musterstudienplan:**

Das Modul ist im Kontext der folgenden Studienfächer/Vertiefungsrichtungen verwendbar:

##### [1] **Informatik (1. Staatsprüfung für das Lehramt an Gymnasien)**

(Po-Vers. 2007 | TechFak | Informatik (1. Staatsprüfung für das Lehramt an Gymnasien) | Module Fachwissenschaft Informatik | Wahlpflichtbereich Informatik Gymnasium | Wahlpflichtmodul Informatik Gymnasium)

##### [2] **Informatik (Master of Science)**

(Po-Vers. 2010 | TechFak | Informatik (Master of Science) | Wahlpflichtbereich | Säule der softwareorientierten Vertiefungsrichtungen | Vertiefungsrichtung Datenbanksysteme)

##### [3] **International Information Systems (IIS) (Master of Science)**

(Po-Vers. | ReWiFak | International Information Systems (IIS) (Master of Science) | Masterprüfung | Informatics | Informatics Electives | Data Management II)

##### [4] **International Information Systems (IIS) (Master of Science)**

(Po-Vers. 2014w | ReWiFak | International Information Systems (IIS) (Master of Science) | Informatics | Informatics Electives | Data Management II)

##### [5] **Mathematik (Bachelor of Science): ab 5. Semester**

(Po-Vers. 2007 | NatFak | Mathematik (Bachelor of Science) | alte Prüfungsordnungen | Gesamtkonto | Vertiefungsmodule Mathematik (Nebenfach Informatik))

##### [6] **Mathematik (Bachelor of Science): ab 5. Semester**

(Po-Vers. 2009 | NatFak | Mathematik (Bachelor of Science) | Nebenfach Informatik | Module im 2. und 3. Studienjahr | Vertiefungsmodule Mathematik (Nebenfach Informatik))

#### **Studien-/Prüfungsleistungen:**

Data Warehousing und Knowledge Discovery in Databases (Prüfungsnummer: 743223)

(englische Bezeichnung: Data Warehousing and Knowledge Discovery in Databases)

(diese Prüfung gilt nur im Kontext der Studienfächer/Vertiefungsrichtungen [2], [3], [4])

Prüfungsleistung, mehrteilige Prüfung

Anteil an der Berechnung der Modulnote: 100%

weitere Erläuterungen:

Die Bewertung der Prüfungsleistung setzt sich zusammen aus 1/2 Note der Klausur in Data Warehousing (von etwa 60 Minuten Dauer) und 1/2 Note der mündlichen Prüfung in KDD (von etwa 30 Minuten Dauer). Zum Bestehen der Modulprüfungen müssen beide Teilprüfungen bestanden werden.

Erstablingung: SS 2019, 1. Wdh.: WS 2019/2020 (nur für Wiederholer)

1. Prüfer: Meyer-Wegener/Ruf (T10144)

Data Warehousing und Knowledge Discovery in Databases (Prüfungsnummer: 228988)

(englische Bezeichnung: Data Warehousing and Knowledge Discovery in Databases)

(diese Prüfung gilt nur im Kontext der Studienfächer/Vertiefungsrichtungen [5], [6])

Prüfungsleistung, mehrteilige Prüfung

Anteil an der Berechnung der Modulnote: 100%

weitere Erläuterungen:

Die Bewertung der Prüfungsleistung setzt sich zusammen aus 1/2 Note der Klausur in Data Warehousing (von etwa 60 Minuten Dauer) und 1/2 Note der mündlichen Prüfung in KDD (von etwa 30 Minuten Dauer). Zum Bestehen der Modulprüfungen müssen beide Teilprüfungen bestanden werden.

Prüfungssprache: Deutsch oder Englisch

Erstablingung: SS 2019, 1. Wdh.: WS 2019/2020

1. Prüfer: Klaus Meyer-Wegener

Data Warehousing und Knowledge Discovery in Databases (Prüfungsnummer: 824544)

(englische Bezeichnung: Data Warehousing and Knowledge Discovery in Databases)

(diese Prüfung gilt nur im Kontext der Studienfächer/Vertiefungsrichtungen [1])

Prüfungsleistung, mehrteilige Prüfung

Anteil an der Berechnung der Modulnote: 100%

weitere Erläuterungen:

Die Bewertung der Prüfungsleistung setzt sich zusammen aus 1/2 Note der Klausur in Data Warehousing (von etwa 60 Minuten Dauer) und 1/2 Note der mündlichen Prüfung in KDD (von etwa 30 Minuten Dauer). Zum Bestehen der Modulprüfungen müssen beide Teilprüfungen bestanden werden.

Prüfungssprache: Deutsch oder Englisch

Erstablingung: SS 2019, 1. Wdh.: WS 2019/2020

1. Prüfer: Klaus Meyer-Wegener