

Lehrende/r	Ralf Hartmut Güting Fabio Valdés	Modulbeauftragte/r	Ralf Hartmut Güting
	Dauer des Moduls zwei Semester	ECTS 10	Workload 300 Stunden
			Häufigkeit in jedem Semester
Lehrveranstaltung(en)	01664 Implementierungskonzepte für Datenbanksysteme		SWS 2+1
	01672 Datenbanken II		SS 1.Hälfte SWS 2+1
	01673 Data Mining		WS SWS 2+1
Detaillierter Zeitaufwand	Bearbeiten der Kurseinheiten: 160 Stunden Bearbeitung der Einsendeaufgaben inkl. Verarbeitung des Korrektur-Feedbacks: 80 Stunden Wiederholung und Prüfungsvorbereitung, Prüfung: 60 Stunden		
Qualifikationsziele	<p>01664: Nach erfolgreicher Teilnahme besitzen Studierende einen guten Überblick wie auch Detailkenntnisse der Architektur und Implementierung von Datenbanksystemen. Sie können die Schichtenarchitektur beschreiben und alle wesentlichen Komponenten der Implementierung benennen. Sie können verschiedene Indexstrukturen für Standard- und Geodaten im Detail erklären. Sie können die Schritte der Verarbeitung einer SQL-Anfrage in der Optimierung und Planerzeugung beschreiben.</p> <p>01672: Studierende können erklären, wann verzahnte Abläufe von Transaktionen als korrekt anzusehen sind und wie der Transaktionsmanager des Datenbanksystems sicherstellt, dass nur solche Abläufe zugelassen werden. Sie können erklären, wie nach Systemabstürzen der korrekte Zustand der Datenbank wiederhergestellt wird. Sie sind damit in der Lage, im Beruf etwa als Datenbankadministrator Effizienzproblemen auf den Grund zu gehen.</p> <p>01673: Studierende besitzen einen umfassenden Überblick zum Ablauf von 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 Algorithmen zur Mustersuche vertraut und können entscheiden, ob eine Korrelation zwischen zwei Ereignissen besteht. Sie können erklären, mit welchen Verfahren Daten klassifiziert werden und wie sich die Qualität eines Klassifikationsergebnisses überprüfen lässt. Sie sind mit verschiedenen Methoden der Clusteranalyse sowie deren Evaluation vertraut.</p>		
Inhalte	<p>Kurs 01664 führt zu Beginn eine relativ detaillierte Software-Architektur ein. Die folgenden Kapitel stellen Implementierungsaspekte anhand dieser Schichten dar. Dazu gehören Externspeicher- und Systempufferverwaltung, Indexstrukturen, externes Sortieren und Anfrageverarbeitung. Bei den Indexstrukturen werden neben klassischen B-Baum- und dynamischen Hashverfahren auch Strukturen für die Verwaltung von Geodaten, wie etwa Grid-Files oder R-Bäume behandelt. Zur Anfrageverarbeitung gehört vor allem die Optimierung, d.h. die Ermittlung eines kostengünstigen Plans für die gegebene SQL-Anfrage, auf der Basis von Kostenschätzungen für Planoperationen.</p> <p>Kurs 01672 führt die Studierenden in die wichtigsten Konzepte ein, auf deren Grundlage die Dienste moderner Datenbanksysteme realisiert werden. Ein zentrales Thema ist dabei der Mehrbenutzerbetrieb (Transaktionsmanagement), da in vielen Anwendungen eine große Zahl von Benutzern gleichzeitig auf die Datenbank zugreift. In engem Zusammenhang damit stehen dann Fragen wie Ausfallsicherheit und Wiederherstellung nach Fehlern von Soft- und Hardware (Recovery). Schließlich wird ein Überblick gegeben über neuere Anwendungen und Tendenzen auf dem Gebiet der</p>		

Datenbanksysteme.

Kurs 01673: Das Thema dieses Kurses 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. Der Kurs vermittelt zunächst Kenntnisse zur Vorbereitung von Data-Mining-Methoden hinsichtlich der Charakterisierung (z.B. Berechnung statistischer Maße, Visualisierung) und Vorverarbeitung der Daten (etwa durch Eliminierung von Ausreißern oder Aggregation). Darauf aufbauend, werden typische Techniken wie Mustersuche, Klassifikation oder Clusteranalyse vorgestellt.

Ergänzende Literatur:

- A. Kemper und A. Eickler: Datenbanksysteme: Eine Einführung. 8. Auflage, Oldenbourg-Verlag, 2011.
H. Garcia-Molina, J.D. Ullman, and J. Widom: Database Systems: The Complete Book. 2. Auflage, Prentice-Hall, 2008.
M. Schneider: Implementierungskonzepte für Datenbanksysteme. Springer-Verlag, 2003.
C.J. Date: An Introduction to Database Systems. Pearson Addison Wesley, 2004.
R. Elmasri, S. Navathe: Grundlagen von Datenbanksystemen: Bachelorausgabe, Pearson Verlag, 2009.
Han, J., Kamber, M. und Pei, J. Data Mining: Concepts and Techniques. Morgan Kaufmann, 3. Auflage, 2011.
Hastie, T., Tibshirani, R. und Friedman, J.H. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2. Auflage, 2009.

Inhaltliche Voraussetzung

Kenntnisse der Konzepte von Standard-Datenbanksystemen, z.B. aus Kurs 01671 "Datenbanken I" im Modul "Softwaresysteme" des Bachelor Informatik

Lehr- und Betreuungsformen

Kursmaterial
internetgestütztes Diskussionsforum
Betreuung und Beratung durch Lehrende
Einsendeaufgaben mit Korrektur und/oder Musterlösung

Anmerkung

Von den drei Kursen müssen nur zwei nach Wahl bearbeitet werden.

Die dem Modul Vertiefende Konzepte von Datenbanksystemen zugeordneten Kurse werden letztmalig im SS 2020 angeboten. Eine Prüfungsteilnahme ist nur noch bis einschließlich Wintersemester 2020/21 möglich.

Verwendung des Moduls

B.Sc. Informatik
B.Sc. Wirtschaftsinformatik
M.Sc. Informatik
M.Sc. Praktische Informatik
M.Sc. Wirtschaftsinformatik

Prüfungsformen

Prüfung

Stellenwert
der Note

s. PO

Art der Prüfungsleistung

bestandene benotete mündliche
Modulprüfung

Voraussetzung

keine