

Allgemeine Information zur Vorlesung
Wissensentdeckung in Datenbanken
Sommersemester 2015

Zur Bereitstellung von Vorlesung und Übungsmaterial werden wir dieses Semester erneut den Electronic Workspace der TU Dortmund (kurz: ews) verwenden. Diesen finden Sie unter der URL <https://ews.tu-dortmund.de/>. Dort können Sie sich mit ihrem Unimail-Account einloggen und müssen anschließend den Arbeitsraum mit dem Titel „*Wissensentdeckung in Datenbanken 2015*“ bzw. dem Ident „*wid2015*“ suchen und sich anmelden. Anschließend werden Sie von uns per Hand freigeschaltet. Nachdem dies geschehen ist, können Sie auf sämtliche Materialien der Vorlesung und der Übung zugreifen. Da wir den durch das ews generierten E-Mail Verteiler benutzen werden, möchten wir Sie bitten, die E-Mails von ihrem Unimail-Konto regelmäßig abzurufen.

Vorlesungszeiten:

Di, 10:00 – 12:00 (HG II/ HS 4, Campus Nord)

Do, 14:00 – 16:00 (HG II/ HS 4, Campus Nord)

Dozenten:

- *Informatik-Teil:* Prof. Dr. K. Morik
Otto-Hahn-Straße 12, R4.007, Campus Nord
Sprechstunde: Mi, 10:00 – 12:00 (nach vorheriger Anmeldung im Sekretariat)
- *Statistik-Teil:* Prof. Dr. C. Weihs
Mathematik, M/719, Campus Nord
Sprechstunde: nach Vereinbarung

Übungen:

Begleitend zur Vorlesung werden Übungen stattfinden. Die Mitarbeit in den Übungen sowie die Bearbeitung der Übungsblätter ist Voraussetzung für die Zulassung zur mündlichen Prüfung (s.u.). Die Übungstermine werden zu den folgenden Zeiten stattfinden:

Mo, 16:00 – 18:00 (OH12 E.003, Campus Nord)

Di, 14:00 – 16:00 (CDI 120, Campus Nord)

Übungsleiter:

- *Informatik-Teil:* Dr. Wouter Duivesteijn, wouter.duivesteijn@tu-dortmund.de,
Otto-Hahn-Straße 12, R4.017, Campus Nord
Sprechstunde: Meist zwischen 9:00 – 12:00 und 13:30 – 17:00 im Büro anzutreffen.
Zur Sicherheit einen Tag vorher per Mail abklären.
- *Statistik-Teil:* M.Sc. Sarah Schnackenberg, schnackenberg@statistik.tu-dortmund.de,
Mathematik, M/E16, Campus Nord
Sprechstunde: Mo, 14:00 – 15:00.
Zu anderen Zeiten am besten nach vorheriger Absprache per Mail.

Anmeldung zu den Übungen:

Die Anmeldung zu den Übungsgruppen erfolgt in der ersten Vorlesung. Dazu können Sie sich in Listen eintragen. Wir werden versuchen die Wunschtermine zu ermöglichen, müssen jedoch darauf achten, dass die Aufteilung zu den zwei Terminen relativ gleichmäßig erfolgt.

Softwareeinführung:

Der erste Übungszettel wird in der zweiten Vorlesungswoche herausgegeben. Die Übungen beginnen ebenfalls in der zweiten Woche. In den ersten beiden Wochen werden dabei die Softwarepakete vorgestellt, die für die Bearbeitung einiger Aufgaben erforderlich sind. Am Montag, den 13.04. (Gruppe 1) und Dienstag, den 14.04 (Gruppe 2) wird die Statistik-Software *R*, am Montag, den 20.04. (Gruppe 1) und Dienstag, den 21.04. (Gruppe 2) das Data-Mining Tool *RapidMiner* für den Informatik-Teil vorgestellt.

Die Programme lassen sich kostenlos runterladen:

- <http://www.r-project.org/> (*R*)
- <http://www.rapid-i.com/> (*RapidMiner*)

Pool-Accounts:

Die Bearbeitung der Statistik-Übungsaufgaben kann in den Rechnerpools der Statistik (M/711 und M/U18) erfolgen. Für Accounts für die Studierenden der Fakultäten für Informatik und Mathematik werden in der ersten Vorlesung Listen herumgegeben und die Accounts gebündelt erstellt. Einzelne Accounts sind danach noch bei der Systemadministration in Raum M/714 erhältlich.

Abgabemodalitäten:

- Die Ausgabe der Übungsblätter erfolgt donnerstags nach der Vorlesung im *ews*-Raum. Sie haben eine Woche Zeit für die Bearbeitung. Die Abgabe muss bis spätestens donnerstags 10:00 Uhr der folgenden Woche an den jeweiligen Übungsleiter erfolgen. Auf den Übungszetteln wird vermerkt, ob es sich um einen Informatik- oder einen Statistik-Zettel handelt.
- Bitte geben Sie nur Dateien der folgenden Formate ab:
 - a) Programmcode: .R- Datei oder als Rapid Miner Process File im .xml-Format,
 - b) Weitere Erläuterungen, plots, Kommentare etc. als .pdf- oder .txt-Datei,
 - c) Handschriftliches und Theorieaufgaben entweder auch als .pdf abschicken, das Original in den Briefkasten (Statistikabgaben: Briefkasten Nummer 126 im Erdgeschoss des Mathematikgebäudes; Informatikabgaben: Briefkasten von Wouter Duivesteijn in Raum 4.005, OH12 (Otto-Hahn-Straße 12, Campus Nord) – der jeweilige Briefkasten ist auf den Übungszetteln vermerkt) einwerfen, oder einscannen. Dabei muss jedoch alles deutlich lesbar und jedes Blatt mit Namen aller Gruppenmitglieder versehen sein.
- Bitte achten Sie darauf, all Ihre Abgaben sinnvoll zu benennen. Vermerken Sie bitte sowohl Ihren Nachnamen als auch die Blattnummer in den Dateinamen. Ein sinnvoller Dateiname ist z.B. *wid2015_schnackenberg_01.R*.
- Die Abgabe der Übungen kann in einmalig zu Beginn festgelegten Gruppen zu maximal 3 Personen erfolgen. Auf der Gruppenabgabe bitte sämtliche Namen vermerken!

Scheinerwerb:

Jedes Übungsblatt wird mit bis zu 10 Punkten bewertet. Für einen Übungsschein müssen Sie 50% der maximal erreichbaren Punkte erhalten. Dieser Schein ist Voraussetzung zur Teilnahme an der mündlichen Prüfung. Im Folgenden sind die einzelnen Module aufgelistet. Für alle Module ist als Prüfungsform eine mündliche Prüfung festgelegt:

- Bachelor Datenanalyse und Datenmanagement (Modul BD XVI – BD 2007/2012)
- Informatikteil im Master Wirtschaftsmathematik
- Master Informatik (Modul INF-MSc-235)
- Master Informatik (Modul INF-MSc-511): „Eine Prüfungsleistung für dieses Modul besteht einschließlich 30.09.2015.“ „Wird das Vertiefungsmodul INF-MSc-511: 'Wissensentdeckung in Datenbanken' gewählt, ist die Wahl des Basismodul INF-MSc-235 'Wissensentdeckung in Datenbanken' nicht möglich.“
- Bachelor Statistik (Modul BS XI „Spezialgebiete“ – BS 2007)