

# Vorbesprechung für das Bachelorseminar „Information Theoretic Data Mining“

Prof. Dr. Christian Böhm  
Dr. Bianca Wackersreuther

- Termine:
  - Di, 26.04.2016, 14.00-16.00 Uhr in Raum 065, Oettingenstr. 67  
Vorbesprechung und Themenvergabe
  - Di, 14.06.2016 Abgabe der Vortrags-Slides per E-Mail
  - Fr, 17.06.2016, 12.00-18.00 Uhr in Raum U 127, Oettingenstr. 67  
Block 1: Themen 1-1 bis 1-7 (bzw. s. Themenvergabe)
  - Fr, 24.06.2016, 12.00-18.00 Uhr in Raum U 127, Oettingenstr. 67  
Block 2: Themen 2-1 bis 2-8 (bzw. Ersatztermin)
  - Auf Wunsch bieten wir zusätzlich eine Einführung in LaTeX an
  - Fr, 08.07.2016 Abgabe der Seminararbeit per E-Mail

- Jeder Seminarteilnehmer bekommt heute ein Thema in Form einer wissenschaftlichen Publikation zugeteilt.
- Die Publikation dient als Grundlage für den 25-minütigen Vortrag an einem der beiden Blocktage.
- Es wird erwartet, dass der Bearbeiter eigenständig weitere relevante Arbeiten falls vorhanden mit in die Bearbeitung integriert. Hierzu ist eine eigenständige Recherche im WWW vorgesehen.
- Jeder Teilnehmer präsentiert seinem Betreuer vorab eine Kurzform seines Vortrages um eventuelle Unklarheiten rechtzeitig klären zu können.

1. Teilnahme an der Vorbesprechung
2. Eigener Vortrag, der in das Thema einführt und Methode, Experimente, sowie verwandte Arbeiten vorstellt.
3. Aktive Teilnahme an beiden Seminartagen, d.h. Beteiligung an den anschließenden Diskussionen.
4. Schriftliche Seminararbeit (5-10 Seiten), die das Thema des Vortrags beschreibt. Die Arbeit muss termingerecht als PDF-Datei beim Betreuer eingereicht werden. Für die Erstellung ist das auf der Seminar-Homepage verlinkte LaTeX-Template zu verwenden.

- Gliederung und Foliennummern
- Anschaulich präsentieren - Graphiken verwenden
- Motivation am Anfang ist wichtig
- Vortrag endet mit Zusammenfassung
- Text in Stichpunkten, kurz und bündig
- Formeln und Graphiken wenn möglich selber zeichnen, bei kopierten Graphiken auf die Anzeigequalität achten
- Variablen explizit einführen und dort definieren wo sie gebraucht werden
- Gliederung ruhig mehrmals verwenden
- Animationen mit Vorsicht verwenden

- Zuhörer anschauen
- Hände vor dem Körper
- Sprechtempo beachten, eher langsamer als normal
- Nicht ablesen, sondern versuchen frei zu sprechen
- Kein Daumenkino (ca. 1,5-3 min pro Folie)
- Auf das Wesentliche beschränken und vom Allgemeinen ins Spezielle und nicht anders herum.
- Karteikarten und Notiz-Optionen helfen bei unsicheren Passagen
- Schrift nicht zu klein
- Platz auf der Folie ausnützen, aber Inhalt nicht überladen (Keine großen weißen Ränder)

- Eine Seminararbeit ist eine wissenschaftliche Arbeit.
- Eine Seminararbeit soll das Thema in eigenen Worten wiedergeben und kritisch hinterfragen.
- Sprache ist entweder Deutsch oder Englisch.
- Literatur zu einem Thema muss sich nicht auf die ausgeteilten Materialien beschränken.
- Achten Sie auf korrekte Quellenangaben.

- Anwesenheit und Mitarbeit im Seminar sind Voraussetzung für das Bestehen.
- Bei Krankheit oder anderen entschuldbaren Abwesenheitsgründen vorher per Email Bescheid geben. Wir bemühen uns dann um eine geeignete Lösung.



- 1-1 Robust Information-theoretic Clustering (KDD, 2006)
- 1-2 A Nonparametric Information Theoretic Clustering Algorithm (ICML, 2010)
- 1-3 minCEntropy: a Novel Information Theoretic Approach for the Generation of Alternative Clusterings (ICDM, 2010)
- 1-4 Integrative Parameter-Free Clustering of Data with Mixed Type Attributes (PAKDD, 2010)
- 1-5 INCONCO: Interpretable Clustering of Numerical and Categorical Objects (KDD, 2011)
- 1-6 Dependency Clustering across Measurement Scales (KDD, 2012)
- 1-7 Relevant Overlapping Subspace Clusters on Categorical Data (KDD, 2014)

- 2-1 Fully Automatic Cross-associations (KDD, 2004)
- 2-2 Weighted Graph Compression for Parameter-free Clustering With PaCCo (SDM, 2011)
- 2-3 Summarization-based Mining Bipartite Graphs (KDD, 2012)
- 2-4 Subdue: Compression-based Frequent Pattern Discovery in Graph Data (KDD Workshop, 2005)
- 2-5 Compression-based Graph Mining Exploiting Structure Primitives (ICDM, 2013)
- 2-6 Pics: Parameter-free Identifier of Cohesive Subgroups in Large Attributed Graphs (SDM, 2012)
- 2-7 VoG: Summarizing and Understanding Large Graphs (SDM, 2014)
- 2-8 Mining Connection Pathways for Marked Nodes in Large Graphs (SDM, 2013)