Get-Together mit den Erstsemestern des Master Wirtschaftsinformatik

Research Group Data and Web Science
Teaching Overview

The DWS Group offers the following courses for master students

- Semantic Web Technologies
- Web Mining
- Text Analytics
- Data Mining and Matrices
- Web Data Integration
- Data Mining II
- Web Search / IR
- Large-Scale Data Management
- Knowledge Management
- Decision Support
- Databases II
- Data Mining I

Offered this semester.
Data and Web Science Research Group

5 Professors
8 Post-docs
19 PhD students
Overall Research Goal:
Understand heterogeneous data in order to improve applications using knowledge.

- Information Extraction
- Data Integration
- Data Mining & Reasoning
- Integrated Schema.org Data
- Integrated Web Tables and Text Data
- Applications
- Web Search
- Data Analytics
- Recommender Systems
- Process Management
Data and Web Science Research Group – Research Areas

- **Artificial Intelligence** (Prof. Heiner Stuckenschmidt)
  - knowledge representation formalisms and reasoning techniques for information extraction and integration

- **Data Analysis** (Prof. Rainer Gemulla)
  - methods for analyzing and mining large datasets as well as their practical realizations and applications

- **Natural Language Processing** (Prof. Simone Ponzetto)
  - knowledge acquisition from heterogeneous Web sources and its application to text understanding and search
- **Web-based Systems** (Prof. Chris Bizer)
  - technical and empirical questions concerning the evolution of the World Wide Web from a medium for the publication of documents into a global dataspace

- **Web Data Mining** (Prof. Dr. Heiko Paulheim)
  - using web data as background knowledge in data mining, and data mining methods to create and improve large-scale knowledge bases
Lecture contents – the basics of “torturing data”:
- Classification: Will your bank grant you a loan?
- Frequent pattern mining: Which products to place together in a supermarket to maximize customer purchases?
- Text Mining: Do students on Twitter like or dislike this lecture?
- Clustering: How to automatically organize your MP3 collection?

Student project:
- Torture some data of your choice.

Teaching staff:
- Prof. Dr. Christian Bizer (Lectures)
- Oliver Lehmberg, Kiril Gashtoevski, Daniel Ruffinelli (Exercises)
IE 672: Data Mining 2

- Advanced Data Mining methods
  - Regression and Forecasting
  - Dimensionality Reduction
  - Anomaly Detection
  - Time Series Analysis
  - Parameter Tuning
  - Ensemble Learning
  - Online Learning

- Organization:
  - Lectures and Exercises
  - Participation in Data Mining Cup
  - Opportunity to become a certified RapidMiner Data Analyst

- Teaching staff:
  - Prof. Dr. Heiko Paulheim (Lectures), Oliver Lehmberg (Exercises)
Matrices and tensors are powerful representations of data
- Data points, sets, graphs, relational data, knowledge bases, ...

Course goal: Learn how to analyze such data
- Course covers theory and applications of dimensionality reduction, embeddings, denoising, discovery of latent structure, visualization, prediction, clustering, pattern mining, topic modelling, …
- Focus is on unsupervised and semi-supervised learning & matrix decompositions
IE 673: Data Mining and Matrices

- Instructor: Rainer Gemulla
- Tutor: Yanjie Wang
- 2 SWS lecture, 2 SWS tutorium, 6 ECTS
- IE 500 Data Mining I recommended
- Gain hands-on experience
  - Smaller exercises to deepen lecture material
  - Homework assignments to analyze real data
  - Learn R
- Passing requirements
  - Regular assignments
  - Final exam or oral examination
CS 661: Knowledge Management

- Overview of knowledge management strategies and technical foundations of knowledge management systems
  - Information Retrieval and Text Mining
  - Knowledge Representation: Repositories and Query Answering
  - Collective Intelligence: Web 2.0 and Wikis, Social Network

- Teaching staff
  - Prof. Dr. Simone Ponzetto (Lectures)
  - Dr. Goran Glavas, Dr. Stefano Faralli, Petar Ristoski (Exercises)
IE 671: Web Mining

Specific methods for mining knowledge from Web content and Web usage data:
1. Web Usage Mining
2. Recommender Systems
3. Web Structure Mining
4. Social Network Analysis
5. Web Content Mining
6. Information Extraction
7. Sentiment Analysis

Organization:
- Lectures and Exercises
- Student projects (second half of semester)

Tools:
- mahout
- Pajek
- NLP
IE 663: Web Search and Information Retrieval

- Understanding end-to-end search systems
IE 663: Web Search and Information Retrieval

- **Lecture contents**: understanding search systems
  - Boolean and vector space retrieval models
  - Probabilistic information retrieval
  - Evaluation of retrieval systems
  - Web search: Crawling, link-based algorithms

- **Teaching staff**:
  - Dr. Goran Glavas (Lectures)
  - Anne Lauscher (Exercises)
Topic: Reasoning and Learning with Probabilistic Graphical Models

Lecturer:

Prof. Dr. Heiner Stuckenschmidt
Dr. Melisachew Wudage Chekol

Literature:

Daphne Koller & Nir Friedman
Probabilistic Graphical Models
MIT Press, 2009

Registration:

Mail mel@informatik.uni-mannheim.de until March 1st

Preference will be given to successful participants of “Decision Support”
Instructors: Yanjie Wang, Rainer Gemulla

This term: **machine learning libraries and platforms**

In this seminar, you will

- explore and experiment with a popular machine learning platform of your choice,
- solve a small, self-defined machine learning problem with this platform,
- give an overview over the platform, your problem, and your solution

Prerequisites: Data Mining I, suitable programming experience

Check out website and **register today** (Feb 13)
CS 707: Text Analytics Seminar

• In this seminar, you will
  • Learn about recent advancements in text analytics
  • Read, understand, explore, and present scientific literature

• This term: **deep learning for text analytics**
  • Yann LeCun, Facebook AI research lab: "The next big step for Deep Learning is natural language understanding, which aims to give machines the power to understand not just individual words, but entire sentences and paragraphs"
  • Topics: Machine reading, machine translation, question answering, opinion mining, …

• Check out website and **register today** (Feb 13)

• Instructors: Goran Glavas, Simone Ponzetto
The DWS group continuously hires good students.

- To work on:
  - Data and Web Mining projects
  - Information Extraction and Integration projects
  - Knowledge Representation and Reasoning projects
  - Implement open source tools

- 30-60 h/month contracts are possible.

- Contact PostDoc or Professor responsible for the project/area that you are interested in.
  - Include CV and overview about your marks.

- Good start for writing your master thesis within group.