

## Information retrieval for digital libraries 2

### Informationsåtervinning för digitala bibliotek 2

7.5 credits

7.5 högskolepoäng

---

**Ladok Code:** NLID23

**Version:** 6.0

**Established by:** Education Committee 2015-06-10

**Valid from:** Autumn 2015

**Education Cycle:** Second cycle

**Main Field of Study (Progressive Specialisation):** Informatics (A1N)

**Disciplinary Domain:** Natural sciences

**Prerequisites:** Degree of Bachelor.

**Subject Area:** Informatics/Computer and Systems Sciences

**Grading Scale:** ECTS-credits

---

### Content

The course aims to illustrate development directions in IR research important for digital libraries, namely:

- General theory of automatic classification.
- Supervised machine learning.
- Unsupervised machine learning.
- Feature selection.
- Evaluation of data mining.
- Information visualization for IR.

### Learning Outcomes

After completion of the course the student should:

#### **concerning *knowledge and understanding***

- Be able to analyze the connection between data mining and IR in terms of content representation and content categorization.
- Be able to reason about feature selection for data mining.

#### **concerning *skills and abilities***

- Be able to show a deeper understanding and hands-on competence in data mining and information visualization.
- Be able to explain the basic principles behind the use of IR in a digital library setting.

#### **concerning *professional judgement***

- Be able to evaluate data mining results using standard evaluation measures.

### Forms of Teaching

Tuition is conducted through lectures, demonstrations, practices, project work, independent studies and group projects.

The language of instruction is English.

### Forms of Examination

The course is examined through written examinations and reports.

In the event of changes in course plans students who wish to complete courses can be examined on the basis of the most recent

version of the course plan. For courses that are no longer running, students who wish to complete such courses can read all or part of an equivalent course. (This is in accordance with the Director's decision, Dnr 516-13, 11th June, 2013)

Student rights and obligations at examination are in accordance with guidelines and rules for the University of Borås.

### **Literature and Other Teaching Materials**

Baeza-Yates, R. & Ribeiro-Neto, B. (2011). Modern Information Retrieval: The Concepts and Technology Behind Search. Addison-Wesley: Harlow. (Chapter 2, pp 21-56; Chapter 8, pp 281-336.)

Sebastiani, S. (2005). Text categorization. In Alessandro Zanzi (ed.), Text Mining and its Applications, WIT Press: Southampton, pp. 109-129.

Stavrianou, A., Andritsos, P. & Nicoloyannis, N. (2007). Overview and Semantic Issues of Text Mining. SIGMOD Record, 36(3), pp. 23-33.

Witten, I. & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques 2nd. Ed. Morgan Kaufman Publishers: Amsterdam. pp 3-17.

### **Student Influence and Evaluation**

Students shall be involved in further development of the course and are therefore given the possibility to assess the course in a systematic way by written or oral means. How this assessment falls out and how it affects further development will be reported back to the students. For course evaluation, the standard University College of Borås rules as of 7 June 2005 apply, dnr 56-02- 10.

### **Miscellaneous**

The course is part of Master's Programme Library and Information Science: Digital Library and Information Services