UNIVERSITY OF BORÅS

Information retrieval for digital libraries 2
Informationsåtervinning för digitala bibliotek 2

7.5 credits
7.5 högskolepoäng

Ladok Code: NLID23
Version: 8.0
Established by: Committee for Education in Librarianship, Information, and IT 2017-05-30
Valid from: Autumn 2017

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 present the development in information retrieval research relevant for digital libraries, with a particular focus on the relation between automatic indexing, automatic classification, information retrieval and information visualization in scalable collections. As these four components constitute aspects of advanced access to such collections, while limiting ourselves to text documents only, we cover the prerequisites to and a general theory of automatic classification, supervised and unsupervised machine learning, the role of feature selection, the significance of evaluation in text mining, and visual access to content.

Learning Outcomes

After completion of the course the student should be able to, concerning:

Knowledge and understanding
1.1 Explain the interrelated nature of the components of advanced access to digital collections.
1.2 Explain the basic principles behind of the use of text mining in a digital library setting.
1.3 Analyze the connection between text mining and IR information retrieval in terms of content representation and content categorization.
1.4 Reason about and present arguments for feature selection for text mining.

Competence and skills
2.1 Show hands-on competence in text mining and information visualization.
2.2 Evaluate experimental text mining results using standard evaluation measures.

Forms of Teaching

Tuition is conducted through:

- lectures,
- demonstrations,
- exercises,
- project work,
- independent studies, and
- group projects.

The language of instruction is English.

Forms of Examination

Examination of the course occurs through:

- Assignment: Group assignment
Learning outcome: 2.1 and 2.2
Credits: 3.5
Grading scale: A-F

- Assignment: Take home examination
Learning outcome: 1.1, 1.2, 1.3 and 1.4
Credits: 4.0
Grading scale: A-F

To obtain the grade E for the whole course, the grade E is required for all the assignments on the course. The grade for the whole course is determined by the rounded average of the grades for all the assignments on the course, converted to a numerical scale.

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.

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.)

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems, 26, 3111-3119.

Sebastiani, S. (2005). Text categorization. In Alessandro Zanasi (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.

Student Influence and Evaluation

The course is evaluated in accordance with the current guidelines for course evaluations at the University of Borås, where students' views should be sought. The course evaluation report will be published and disseminated to participating and prospective students in accordance with the current guidelines, and forms the basis for future development of courses and training programs. The course coordinator is responsible for that the evaluation is performed according to current guidelines.

Miscellaneous

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