Study program: Artificial Intelligence
Name of the subject: Information Retrieval
Teacher(s): Miloš Savić
Status of the subject: elective
Number of ECTS credits: 6
Conditions: none

Subject goal

The main objective of the course is to introduce students to modern information retrieval algorithms, techniques, models, evaluation methodologies, systems and applications.

Outcome of the subject

Minimum: Successful students should be able to use modern information retrieval libraries in the development of applications requiring advanced indexing and searching facilities.

Desirable: At the end of the course it is expected that successful students deeply understands modern information retrieval algorithms, techniques and models, and that they are able to implement and evaluate them in a wide-range of practical applications.

Subject content

Theory

Introduction to modern information retrieval systems and search engines. The Boolean information retrieval model. Basic data structures and algorithms for indexing. Text similarity measures and tolerant information retrieval. Distributed indexing. Index compression techniques. The vector space information retrieval model. Ranking search results. Evaluation of information retrieval systems. Relevance feedback and query expansion. Probabilistic information retrieval models. Language models for information retrieval. Intelligent techniques for ranking, classifying, categorization and filtering of search results. Architecture of web search engines. Web crawling, indexing and ranking of search results.

Practical learning

Practical programming tasks related to indexing and searching data structures and algorithms. Acquaintance with modern Java information retrieval libraries (Lucene and LingPipe).

Literature

Recommended

Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.

Bruce Croft, Donald Metzler and Trevor Strohman. *Search Engines: Information Retrieval in Practice*. Pearson, 2010.

Michael McCandless, Erik Hatcher, Otis Gospodnetic. *Lucene in Action, Second Edition*. Manning Publications. 2010.

Number of active teaching classes	**Theoretical teaching: 2**	**Practical teaching: 3**

Method of carrying out the teaching

Theoretical classes are based on the classical teaching model involving a projector. At theoretical exercises, solutions of practical problems related to information retrieval are presented and discussed with students. Student's knowledge is checked through a written test, practical problems, and preparation of a seminar paper that is defended at the end of the course.

Evaluation of knowledge (maximum number of points 100)

Pre-exam obligations	points	**Final exam**	points
Test	20	Project and seminar paper	50
Practical problems	30		