

SGN-43006 KNOWLEDGE MINING AND BIG DATA, Period I, 2015, 5cr.

Professori Ari Visa, ari.visa@tut.fi
Room TF309
Phone 040 7287969

Lectures 24 h.

- Time and Place:** The Course will be lectured during period I. Schedule is available at the end of this page. The first meeting will take place on the Monday, 24st of August, 12:15 a.m. in the lecture room TB104.
- Topics:** By the increasing popularity of the Internet and large databases the need to knowledge retrieval and management has been increasing. The problem nowadays is volume and variety. On the course the following topics will be treated: data preprocessing, decision trees, rule based reasoning, cluster analysis, incremental learning, A/B testing and so on. The lecture plan is available at the end of this page. The text books are available at the address <http://www.tut.fi/fi/kirjasto/aineistot/kurssikirjat/index.htm>. The aim is to introduce the main approaches in knowledge mining and to create the capability to use and to develop the presented methods.
- Audience:** The course is intended to students who are close to graduation in the fields of signal processing, computer science or telecommunication. The course is also suitable to post-graduate studies.
- Requirements:** The examination is based on the final exam and a wide exercise work. The accepted parts of the exercise work will give points in the final exam. To pass the course 50% of points are required.
Lic. Tech. Riitta Kerminen takes care of exercise work, Email riitta.kerminen@tut.fi. The exercise work is available online from the 31st of August 2015.

Literature:

Data Mining: Concepts and Techniques, Jiawei Han, Micheline Kamber, Morgan Kaufmann Publisher, 2000 (DMCT).

Principles of Data Mining, David, J., Hand, Heikki Mannila, Padhric Smyth, MIT Press 2000 (PDM).

Date	Place	Subject	
24.8.2015	TB104	Introduction to Big Data and Knowledge mining Chapter 1 DMCT	http://www.cs.tut.fi/~avisa/lec1.pdf
28.8.2015	TB104	2. Data Warehousing and OLAP technology for data mining Chapter 2 DMCT	http://www.cs.tut.fi/~avisa/lec2.pdf
31.8.2015	A223	3. Data Preprocessing Chapter 3 DMCT, Models and Patterns Chapter 6 PDM	http://www.cs.tut.fi/~avisa/lec3.pdf
4.9.2015	A223	4. Data mining primitives, languages and system architectures Chapter 4 DMCT	http://www.cs.tut.fi/~avisa/lec3.pdf
7.9.2015	A223	Guest lecturer	
11.9.2015	A223	5. Concept description: Characterization and Comparison Chapter 5 DMCT, Descriptive Modeling Chapter 9 PDM	http://www.cs.tut.fi/~avisa/lec4.pdf
14.9.2015	A223	6. Mining association rules in large databases Chapter 6	http://www.cs.tut.fi/~avisa/lec5.pdf
18.9.2015	A223	7. Classification and prediction Chapter 7 DMCT	http://www.cs.tut.fi/~avisa/lec6.pdf
21.9.2015	A223	Guest lecturer	
25.9.2015	A223	8. Predictive Modeling for Classification Chapter 10 PDM	http://www.cs.tut.fi/~avisa/lec7.pdf
28.9.2015	A223	9. Clustering analysis Chapter 8 DMCT	http://www.cs.tut.fi/~avisa/lec8.pdf
2.10.2015	A223	10. Clustering analysis	http://www.cs.tut.fi/~avisa/lec9.pdf

		+ demands for the exam Chapters 11 PDM	
5.10.2015	A223	11. Mining complex types of data Chapter 9 DMCT	http://www.cs.tut.fi/~avisa/lec10.pdf
9.10.2015	A223	12. Data Mining applications and trends in data mining Chapter 10 DMCT	http://www.cs.tut.fi/~avisa/lec11.pdf