POZNAN UNIVERSITY OF TECHNOLOGY

EUROPEAN CREDIT TRANSFER AND ACCUMULATION SYSTEM (ECTS)

pl. M. Skłodowskiej-Curie 5, 60-965 Poznań

COURSE DESCRIPTION CARD - SYLLABUS

Course name
Data mining project

Course

Field of study	Year/semester
Computing	2/3
Area of study (specialization)	Profile of study
Inteligentne technologie informacyjne	general academic
Level of study	Course offered in
Second-cycle studies	Polish
Form of study	Requirements
full-time	elective

Number of hours

Lecture	Laboratory classes	Other (e.g. online)
30	30	-
Tutorials	Projects/seminars	
-	-	

Number of credit points

4

Lecturers

Responsible for the course/lecturer:
Jerzy Stefanowski, Inst. Informatyki
Wydział Informatyki i Telekomunikacji
ul. Piotrowo 2, 60-965 Poznań
tel: 61 665-2933

Responsible for the course/lecturer:
Krzysztof Dembczyński, Inst. Informatyki
Wydział Informatyki i Telekomunikacji
ul. Piotrowo 2, 60-965 Poznań

Prerequisites

Students should have knowledge concerning machine learning, data mining, statistics and mathematical algebraic calculus. With respect to other competence, they should be sufficiently good at programming (Python and Java), some basic of Big Data processing. Moreover, they should be able to apply analytical and experimental methods, carry out an experimental analysis of algorithms, analyse their results and use statistical tests. Finally students should understand the needs to extend their knowledge and competences.

Course objective

Develop students' skills in solving real life problems related to applying machine learning methods and data mining to engineering and natural life sciences.
In particular, the team working on solving knowledge discovery competition tasks.
Develop students' skills to work with various machine learning and using software libraries.

Furthermore provide more knowledge on advanced data mining approaches and standards of KDD process.

POZNAN UNIVERSITY OF TECHNOLOGY

EUROPEAN CREDIT TRANSFER AND ACCUMULATION SYSTEM (ECTS)

pl. M. Skłodowskiej-Curie 5, 60-965 Poznań

Course-related learning outcomes

Knowledge

Students should have:

1. well organized and theoretically founded knowledge related to the field of computer science related to machine learning and data mining

2. knowledge about supervised, unsupervised and other learning methods + experimental evaluation of classifiers or regression algorithms

3. knows advanced methods, techniques and tools used to solve complex engineering tasks and conduct research in data mining

4. knows advanced methods, techniques and tools used to solve complex tasks and conduct research in a selected area of computer science

Skills

1. can use techniques appropriate for solving complex data mining tasks

2. can integrate knowledge from domains of computer sciences and related ones

3. is able to assess the suitability and the possibility of using new achievements (methods and tools) and new software related to machine learning and data mining

4. is able to assess the usefulness of methods and tools for solving an engineering task, consisting in the construction or evaluation of an IT system or its components, including the limitations of these methods and tools

Social competencies

1. understands that in the field of machine learning the knowledge and skills quickly become obsolete

2. understands the importance of using the latest knowledge in the field of computer science in solving research and practical problems

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Lecture: based on a written or online test including several questions on the scope of course

Laboratory: evaluation of student's knowledge necessary to carry out steps of solving a competition task; monitoring students' activities during classes; evaluation of lab reports (partly started during classes, finished after them).

Programme content

Lecture covers the following topics:

Presenting competitions related to mining non-trivial data. References to KDD cups, kaggle and similar platforms to carrying out such competitions.

Some lectures extend topics in advanced machine learning, data mining and Big Data such as: data pre-processing, dealing with complex data, learning from partially labeled data, advanced time series and data streams. Could be extended by extra lectures provided by representatives of companies related to an industrial practice of implementing knowledge discovery projects or real life complex case studies.

POZNAN UNIVERSITY OF TECHNOLOGY

EUROPEAN CREDIT TRANSFER AND ACCUMULATION SYSTEM (ECTS)

pl. M. Skłodowskiej-Curie 5, 60-965 Poznań

Laboratories – team working on assigned competition tasks. The tasks are divided into steps and milestones – student should regularly report their activity. Additional students presentations on software tools and their ways of solving steps of the task.

Teaching methods

Lecture - learning methods based on multimedia presentations, illustrated with examples or case studies, solving together tasks

Laboratories – tasks, practical exercises also with programming, discussion, teamwork, solving tasks-projects.

Bibliography

Basic

1. Discovering Knowledge in Data. An Introduction to Data Mining. (tłumaczenie polskie Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych, PWN, 2006)., D.Larose, Wiley, 2005.

Additional

1. Mining of Massive Datasets, A. Rajaraman, J. D. Ullman, Cambridge University Press, 2012 (http://www.mmds.org/).

Breakdown of average student's workload

	Hours	ECTS
Total workload	110	4
Classes requiring direct contact with the teacher	60	2.5
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests/exams, project preparation)[1]	50	1.5

---

[1]niepotrzebne skreślić lub dopisać inne czynności