SYLLABUS

Name: ***Knowledge Discovery (MakAu-DS>SM2KD19)***
Name in Polish:
Name in English: ***Knowledge Discovery***

Information on course:

Course offered by department: Faculty of Automatic Control, Electronics and Computer Science
Course for department: Silesian University of Technology

Default type of course examination report:

EGZ

Language:

English

Course homepage:

https://platforma2.polsl.pl/rau2/course/view.php?id=911

Short description:

The aim of the course is to introduce students to machine learning methods appropriate for knowledge discovery. The subject of the course focuses on interpretable (also called wihte-boxes) machine learning methods. Interpretable data models can be used both for data description and decision problem solving (prediction). The course also discusses Explainable Artificial Intelligence (XAI) methods for explaining decisions of complex machine learning systems, such as ensembles of classifiers or neural networks (including deep neural networks).

Description:

ECTS: 3
Total workload: 75 hours (40 contact hours / 35 student's own work hours)
Lecture: 15h
Laboratory: 15h
Other (overview of the report): 10h

Student`s own work: preparation for laboratory classes, preparation for the final test

Lecture:
1. Rule learning for knowledge discovery (classification, regression and survival rule discovery; top-down rule induction, sequential covering, search heuristics; itemset and association rule mining)
2. Decision and regression tree induction (methods for tree induction (decision, regression and survival trees); decision tree growing and pruning)
3. Rough Set Theory (RST) in data analysis (RST based knowledge discovery, attribute selection, attribute importance, lower and upper approximations of a concepts).
4. Quality measures in data mining (subjective and objective quality measures; rule based data models simplification with the use of quality measures)
4. Explainable artificial intelligence (instance level explanations, dataset level explanations, explanation of complex machine learning data models)

Laboratory:
1. Rule induction – exemplary datasets analysis (concept learning, regression and survival analysis, rule induction, rule interestingness evaluation, evaluation of the importance of the elementary conditions, ruleset filtering; association rule mining and evaluation)
2. Decision tree induction – exemplary datasets analysis (building tree based interpretable data models)
3. Rough Set theory – exemplary datasets analysis (attribute selection – high dimensional data; concept approximation, decision rule induction, identification of the most important features)
4. Explainable Artificial Intelligence – exemplary datasets analysis (explanation of decisions of global (dataset level) and local (instance level) complex ML models (ensemble models, neural networks))

Bibliography:

1. O. Maimon, L. Rokach: Data Mining and Knowledge Discovery Handbook. Springer 2005.
2. P. Biecek, T. Burzykowski: Explanatory model analysis. CRC Press 2020.
3. J. Fürnkranz, D. Gamberger, N. Lavrac: Foundations of Rule Learning. Cognitive Technologies, Springer 2012.
4. A. Gudyś, M. Sikora, Ł. Wróbel: A comprehensive suite for rule-based learning. Knowledge Based Systems 194 (2020).
5. F.J. Guillet, H.J. Hamilton: Quality Measures in Data Mininig. Springer 2011.
6. M. Sikora, Ł. Wróbel, A. Gudys: GuideR: A guided separate-and-conquer rule learning in classification, regression, and survival settings. Knowl.-Based Syst. 173: 1-14 (2019)
7. R. Agrawal, R. Srikant. Fast algorithms for mining association rules. Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, pages 487-499, Santiago, Chile, September 1994

Learning outcomes:

Course-specific learning outcomes: upon completion of the course, the student:
- knows and understands in-depth issues related to knowledge discovery methods, interpretable machine learning methods, and explainable artificial intelligence (test, final test) K2A_W05
- knows and understands methods for generating models with symbolic knowledge representation and methods for evaluating the quality of predictive models (test, final test) K2A_W08
- is able to use their knowledge to define machine learning tasks for which the interpretability and explainability of the predictive model is important, and is able to select the appropriate tools and libraries to perform the defined task (test) K2A_U01
- is able to implement the process of generating a model of the expected quality by using experimental model evaluation methods and selecting appropriate quality measures (test) K2A_U17

Assessment methods and assessment criteria:

Credit for the course is based on the results of the laboratories and a written final test conducted after all laboratory classes and lectures. The final test takes the form of a multiple choice test but may also include open questions. Marking an incorrect answer to a question may result in a lower score for that question.
In order to pass the course it is necessary to successfully pass each laboratory and to obtain at least 50% of the points from the final test.

Making up missed laboratory classes is possible after contacting the lecturer in advance.

The syllabus is valid from the academic year 2025/26 and its content is not subject to change during the semester.

Course credits in various terms:

<without a specific program>

Type of credits	Number	First term	Last term
European Credit Transfer System (ECTS)	3	2020/2021-Z	