SYLLABUS

Name: ***Advanced Data Science (InfAAu-IOT>SM2ADS19)***
Name in Polish:
Name in English: ***Advanced Data Science***

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=1041

Short description:

The objective of the course is to make students familiar with chosen issues related to machine learning fields, e.g. data processing and data classification. In particular, there will be presented Bayesian approach to data analysis. It will be also demonstrated that R, RapidMiner, and Weka are effective and efficient tools for data analysis.
Planned invitation of keynote speakers and/or industry experts cause, the examples demonstrated during classes will concern real data processing problems.
Leading foreign professors are invited to cooperate due to their scientific achievements and extensive teaching experience.

Description:

The content of the study programme ensuring learning outcomes (according to the study programme):
• The student knows methods of preparation, cleaning, and improving the quality of data.
• The student knows the rules for creating data model in the given software.
• The student has the ability to assess the quality of the model.
• The student understands the notion of prior and posterior distributions.
• The student is able to apply the appropriate algorithm(s) to solve given data science problems.

ECTS: 3
Total workload: 75 hours (40 contact hours, 35 students' own work hours)
Forms of contact hours:
Lecture 15h
Laboratory 15h
Other (test and reports revision and discussion): 10h
Students' own work: preparation for classes, elaboration of measurement results, writing reports, preparation for tests

Bibliography:

[1] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, D. Rubin, (2014), Bayesian Data Analysis, CRC Press.
[2] Vijay Kotu, Bala Deshpande, Data Science Concepts and Practice, 2nd Edition - November 27, 2018, ISBN: 9780128147610, eBook ISBN: 9780128147627.
[3] Katarzyna Stapor, Paweł Ksieniewicz, Salvador García, Michał Woźniak,
How to design the fair experimental classifier evaluation, Applied Soft Computing, Volume 104, 2021, 107219, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.107219, https://www.sciencedirect.com/science/article/pii/S1568494621001423.
[4] G. Schwarz: Estimating the Dimension of a Model, Ann. Statist., Vol.6, Number 2 (1978), pp.461-464, https://projecteuclid.org/euclid.aos/1176344136.

Learning outcomes:

Knowledge: a student knows and understands ...
- chosen algorithms of machine learning algorithms, a high level programming language, e.g., Python or R, which can be used to complete a calculation or solve a problem (K2A_W03),
- typical engineering technologies in the field of computer science (K2A_W09).

Skills: a student is able to...
- plan and conduct experiments in informatics, sometimes including also the field of computer science, such as clustering or classification (K2A_U07).
- prove the validity of their solution and/or the results achieved in it, by selecting appropriate statistical or machine learning measures/metrics;is also able to illustrate the achieved results, if necessary (K2A_U08).
- solve real-world problems by applying appropriately selected algorithms to real-world data sets, published either in publicly available databases such as UCI or Keel, or provided by the lecturer (K2A_U10).

Assessment methods and assessment criteria:

Depending on the lecturer's choice, the final grade is the assessment of the final test (possible extra points for the students' activity during laboratory and/or remarkable reports) or the grade calculated as an average of partial grades, given for the particular reports.

Method and procedure for making up for student's absence: Absences are possible to make up for the time of homework, or during catching up meetings included in the semester laboratory schedule.

Method and procedure for making up for differences in study programmes for students changing their field of study, changing university, or resuming studies at the Silesian University of Technology: Depending on the arrears, it is determined by the teacher in accordance with the forms of conducting classes and the crediting conditions.

Course consists of two components: lecture and laboratory. According to SUT regulation, lecture attendance is optional (however highly recommended), whereas laboratory exercises are obligatory.

Course credits in various terms:

Informatics, full-time master degree studies 3 sem. (InfAAu-SM3)			
Type of credits	Number	First term	Last term
European Credit Transfer System (ECTS)	3	2020/2021-Z	