KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
Projekt współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego
UNIA EUROPEJSKA
EUROPEJSKI FUNDUSZ SPOŁECZNY

Course title	ECTS code
Machine learning in chemistry	13.3.1293

Name of unit administrating study

null

Studies

faculty	field of study		
Wydział Chemii	Chemia	type	drugiego stopnia
		form	stacjonarne
		specialty	Digital Chemistry
		specialization	wszystkie

Teaching staff

prof. dr hab. Tomasz Puzyn; dr inż. Karolina Jagiełło; dr Agnieszka Gajewicz-Skrętna; dr Alicja Mikołajczyk

Forms of classes, the realization and number of hours

Forms of classes

Laboratory classes, Lecture

The realization of activities

classroom instruction

Number of hours

Lecture: 30 hours, Laboratory classes: 45 hours

ECTS credits

6
Lecture – 30 h
Laboratory classes - 45 h
Student's own work – 30 h
Tutorial classes – 45 h
TOTAL: 150 h – 6 ECTS

The academic cycle

2024/2025 winter semester

Type of course

obligatory

Language of instruction

English

Teaching methods

- computational chemistry experiments and case studies, analysis of obtained results and discussion.
- multimedia-based lecture

Form and method of assessment and basic criteria for eveluation or examination requirements

Final evaluation

- Graded credit
- Examination

Assessment methods

Lecture - written/oral test
Laboratory classes – colloquia and written reports including Python / R codes

The basic criteria for evaluation

exam:
written part (obligatory): single choice test with 15 questions (1 point per question) plus three open questions (5 points per each) – max. 30 points in total. Positive grade if the number of points ≥ 51%. For students having between 41% and 50% from the written part, due to obtain the required number of points (≥ 51%) the oral examination is obligatory. Students with the number of points ≤ 41% do not pass the final test;
oral part (obligatory for students having between 41% and 50% from the written part and facultative for students with ≥ 51%): discussion on three problems related to the topic, selected by the teacher;
students are allowed passing the final test twice (two attempts);
its obligatory to have a positive final grade from the lab exercises before passing the final test.

Method of verifying required learning outcomes

Required courses and introductory requirements

A. Formal requirements
Repetitory in mathematics

- Repetitory in general and inorganic chemistry
- Introduction to Python programming
- Introduction to R programming

Exploratory analysis of multidimensional chemical space

B. Prerequisites

Repetitory in mathematics

- Repetitory in general and inorganic chemistry
- Introduction to Python programming
- Introduction to R programming

Exploratory analysis of multidimensional chemical space

Aims of education

familiarizing the students with machine learning theory and its applications in chemistry
presenting the advantages and disadvantages of various types of machine learning algorithms in chemistry

Course contents

This course is designed to provide students with foundational knowledge of practical aspects of machine learning in chemistry, including:
introduction to machine learning theory and its applications,
overview of various types of machine learning algorithms (supervised and semi-supervised machine learning methods; classification and regression methods; reinforcement learning algorithms; generative versus discriminative models),
challenges in application of machine learning in chemistry (methods for handling uncertain, limited, imbalanced and noisy data; feature selection; model selection and assessment),
open source chemoinformatics software.

Bibliography of literature

Literature required to pass the course
S. D. Brown, R. Tauler, B. Walczak (ed): Comprehensive chemometrics: Chemical and biochemical data analysis. Amsterdam: Elsevier, 2009
R. Kramer: Chemometric techniques for quantitative analysis. New York: Marcel Dekker, Inc, 2005

Extracurricular readings
J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopulos, H. Reis, M. Shukla (ed): Handbook of Computational Chemistry (2nd Edition). Springer 2016. Volume 5: Chemoinformatics, Puzyn T (ed.).

The learning outcomes (for the field of study and specialization)

K_W05:
has extended knowledge in the field of the specialisation studied
K_W06:
applies mathematics to the extent necessary to understand, describe and model chemical processes of extended complexity
K_W08:
demonstrates in-depth knowledge of theoretical computational and IT methods used to solve problems in chemistry
K_W09:
classifies specialist IT tools used in statistical evaluation of experiment results
K_U03
finds necessary information in specialist literature, databases and other sources, lists basic scientific journals in chemistry
K_K02
works in a team taking on various roles in it
K_K03
understands the need for systematic work on various projects of a long-term nature and knows how to set priorities for the implementation of undertaken tasks
K_K06
raises her/his professional and personal competences by

Knowledge

At the completion of this course, the student is expected to be able to:
know and understand the theoretical background of various types of machine learning algorithms, including: multiple linear regression, partial least squares regression, k-nearest neighbors, support vector machines, classification and regression trees and artificial neutral networks,
know the basic division of machine learning methods, and list the application of particular groups of these methods in the analysis of chemical data,
describe the most important challenges for the application of machine learning in chemistry,
describe the benefits and advantages of using machine learning in chemistry,
provide examples of software packages for machine learning in chemistry.

Skills

At the completion of this course, the student is expected to be able to:
choose and apply the appropriate machine learning algorithm to solve a particular problem under consideration in the chemistry science domain,
evaluate the efficacy of the developed model and critically interpret the results obtained with specific machine learning methods,
properly implement strategies for handling limited, imbalanced and noisy data.

Social competence

At the completion of this course, the student is expected to be able to:
describe the benefits of using machine learning methods in her/his daily research practice,
understand the need of deeper learning of the machine learning,
develop interpersonal skills such as communication, cooperation in group (taking different roles), and problem-solving abilities,
understand the social aspects of practical use of knowledge and abilities as well as

using information provided in various sources	connected with them responsibility.

Contact

tomasz.puzyn@ug.edu.pl