

## Subject card

|                                             |                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                          |                                         |                                     |                                                                          |            |     |
|---------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|-----------------------------------------|-------------------------------------|--------------------------------------------------------------------------|------------|-----|
| Subject name and code                       | Monographic lecture - Machine learning algorithms for small datasets, PG_00051253                                                                                                                                                                                                                                                                                                                                       |                                                          |                                         |                                     |                                                                          |            |     |
| Field of study                              | Chemistry                                                                                                                                                                                                                                                                                                                                                                                                               |                                                          |                                         |                                     |                                                                          |            |     |
| Date of commencement of studies             | October 2025                                                                                                                                                                                                                                                                                                                                                                                                            |                                                          | Academic year of realisation of subject |                                     | 2026/2027                                                                |            |     |
| Education level                             | Master's studies                                                                                                                                                                                                                                                                                                                                                                                                        |                                                          | Subject group                           |                                     | Obligatory subject group in the field of study<br>Optional subject group |            |     |
| Mode of study                               | full-time studies                                                                                                                                                                                                                                                                                                                                                                                                       |                                                          | Mode of delivery                        |                                     | at the university                                                        |            |     |
| Year of study                               | 2                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                          | Language of instruction                 |                                     | English                                                                  |            |     |
| Semester of study                           | 3                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                          | ECTS credits                            |                                     | 3.0                                                                      |            |     |
| Learning profile                            | academic                                                                                                                                                                                                                                                                                                                                                                                                                |                                                          | Assessment form                         |                                     | credit                                                                   |            |     |
| Conducting unit                             | Laboratory of Environmental Chemoinformatics -> Department of Environmental Chemistry and Radiochemistry -> Faculty of Chemistry -> Rector                                                                                                                                                                                                                                                                              |                                                          |                                         |                                     |                                                                          |            |     |
| Name and surname of lecturer (lecturers)    | Subject supervisor                                                                                                                                                                                                                                                                                                                                                                                                      |                                                          | dr hab. Agnieszka Gajewicz-Skrętna      |                                     |                                                                          |            |     |
|                                             | Teachers                                                                                                                                                                                                                                                                                                                                                                                                                |                                                          |                                         |                                     |                                                                          |            |     |
| Lesson types                                | Lesson type                                                                                                                                                                                                                                                                                                                                                                                                             | Lecture                                                  | Tutorial                                | Laboratory                          | Project                                                                  | Seminar    | SUM |
|                                             | Number of study hours                                                                                                                                                                                                                                                                                                                                                                                                   | 30.0                                                     | 0.0                                     | 0.0                                 | 0.0                                                                      | 0.0        | 30  |
|                                             | E-learning hours included: 0.0                                                                                                                                                                                                                                                                                                                                                                                          |                                                          |                                         |                                     |                                                                          |            |     |
| Learning activity and number of study hours | Learning activity                                                                                                                                                                                                                                                                                                                                                                                                       | Participation in didactic classes included in study plan |                                         | Participation in consultation hours |                                                                          | Self-study | SUM |
|                                             | Number of study hours                                                                                                                                                                                                                                                                                                                                                                                                   | 30                                                       |                                         | 5.0                                 |                                                                          | 40.0       | 75  |
| Subject objectives                          | <p>Course objectives are:</p> <ul style="list-style-type: none"><li>Familiarizing the students with real-world machine learning problems where available datasets suffer from small size, noise, missing values, and inconsistency</li><li>Presenting the advantages and disadvantages of various types of machine learning algorithms dealing with limited by the amount (and quality) of the data available</li></ul> |                                                          |                                         |                                     |                                                                          |            |     |

|                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                      |                                   |
|----------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|
| Learning outcomes                                              | Course outcome                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Subject outcome                                                                                                                                                                                                                      | Method of verification            |
|                                                                | [CHEMMU2_K06] Undertakes research tasks consciously and responsibly, understanding the social aspects of the practical application of the acquired knowledge and skills and the responsibility related to it.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | At the end of the course every student understands the social aspects of practical use of knowledge and abilities as well as connected with them responsibility                                                                      | [SK4] test/exam - oral or written |
|                                                                | [CHEMMU2_W09] Classifies specialist IT tools used in statistical evaluation of experiment results.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | At the end of the course every student knows classification of advanced machine learning algorithms to deal with small size of data and provides examples of their applications in the chemistry science domain                      | [SW4] test/exam - oral or written |
|                                                                | [CHEMMU2_W08] Demonstrates knowledge of theoretical computational and IT methods used to solve problems in chemistry.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | At the end of the course every student knows and understands the theoretical background of read-across approach, describes the most important challenges for the application of machine learning for small size and low-quality data | [SW4] test/exam - oral or written |
|                                                                | [CHEMMU2_U03] Finds necessary information in specialist literature, databases and other sources, lists basic scientific journals in chemistry.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | At the end of the course every student understands the need of deeper learning of the risk assessment methods                                                                                                                        | [SU4] test/exam - oral or written |
|                                                                | [CHEMMU2_W05] Has extended knowledge in the field of the specialisation studied.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | At the end of the course every student knows how the small size and low quality of a dataset impact traditional machine learning algorithms and provides few ways to mitigate these issues.                                          | [SW4] test/exam - oral or written |
|                                                                | [CHEMMU2_W06] Applies mathematics to the extent necessary to understand, describe and model chemical processes of medium complexity.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | At the end of the course every student understands and explains theoretical background (algorithm) of the advanced algorithms to deal with small size and low-quality data                                                           | [SW4] test/exam - oral or written |
|                                                                | [CHEMMU2_U02] Critically assesses the results of conducted, performed observations and theoretical calculations and discusses errors.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | At the end of the course every student can critically evaluate the obtained results and understand the necessity of their control                                                                                                    | [SU4] test/exam - oral or written |
| Subject contents                                               | This course is designed to provide students with foundational knowledge on machine learning algorithms dealing with limited by the amount (and quality) of the data available, including: <ul style="list-style-type: none"><li>the effects of data quality on machine learning algorithms (with particular emphasis on small, incomplete, noisy, imbalance or affected by artifacts datasets),</li><li>introduction to the bias-variance trade-off in machine learning,</li><li>overview of various types of machine learning algorithms to deal with small size and low-quality data (including kernel regression methods, such as lasso, elastic net and bridge regression, and their adaptive extensions),</li><li>introduction to the basic concepts and strategies of read-across (including averaging approach, similarity-weighted activity of nearest neighbors, filtering approach, search expansion approach),</li><li>real-world examples that illustrate a successful application of the machine learning algorithms for imperfect datasets in the applied sciences (e.g., chemical and pharmaceutical industries, environmental risk assessment of chemical substances).</li></ul> |                                                                                                                                                                                                                                      |                                   |
| Prerequisites and co-requisites                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                      |                                   |
| Assessment methods and criteria                                | Subject passing criteria                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Passing threshold                                                                                                                                                                                                                    | Percentage of the final grade     |
|                                                                | test                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 51.0%                                                                                                                                                                                                                                | 100.0%                            |
| Recommended reading                                            | Basic literature                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | S. D. Brown, R. Tauler, B. Walczak (ed): Comprehensive chemometrics: Chemical and biochemical data analysis. Amsterdam: Elsevier, 2009                                                                                               |                                   |
|                                                                | Supplementary literature                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopoulos, H. Reis, M. Shukla (ed): Handbook of Computational Chemistry (2nd Edition). Springer 2016. Volume 5: Chemoinformatics, Puzyn T (ed.).                           |                                   |
|                                                                | eResources addresses                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                      |                                   |
| Example issues/<br>example questions/<br>tasks being completed |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                      |                                   |
| Work placement                                                 | Not applicable                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                      |                                   |

Document generated electronically. Does not require a seal or signature.