UNIWERSYTET GDAŃSKI

Course title
Chemometria / Chemometrics

ECTS code
13.3.0755

Name of unit administrating study
Faculty of Chemistry

Studies

Field of study	Type	Form	
Chemical business	Bachelor / Engineer	Full-time studies	

Teaching staff
Prof. dr hab. Tomasz Puzyn

Forms of classes, the realization and number of hours

A. Forms of classes, in accordance with the UG Rector's regulations
lecture, laboratory classes

B. The realization of activities
in-class learning

C. Number of hours
45 h (15 lecture, 30 laboratory classes)

ECTS credits 3

classes - 45 h
tutorial classes – 5 h
student's own work – 25 h

Total: 75 h - 3 ECTS

The academic cycle
Third year, winter semester

Type of course
obligatory

Language of instruction
Polish

Teaching methods

- Lecture with multimedia presentation
- Students will conduct hands on exercises in the computational laboratory, based on the instructions prepared by the teacher

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

A. Final evaluation, in accordance with the UG study regulations
exam, course completion (with a grade)

B. Assessment methods

- Written/oral final test (as described below)
- Colloquia that must be passed before each laboratory exercise
- Written reports summarizing results achieved during the lab exercises

C. The basic criteria for evaluation or exam requirements

Final test:

- 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 > 50%. For students having between 40% and 50% from the written part, oral part is obligatory. Students with the number of points < 40% do not pass the exam.
- Oral part (obligatory for students having between 40% and 50% from the written part and facultative for students with > 50%): discussion on three problems related to the topic, selected by the teacher
- Students are allowed passing the test twice (two attempts).

Its obligatory to have a positive final grade from the lab exercises before passing the final test

Required courses and introductory requirements

General chemistry

Mathematics (linear algebra)

Statistics

Aims of education

• Familiarizing the students with the possible application of chemometrics algorithms

• Acquiring by the students the skills of collecting, archiving and evaluating of the multivariable data

• Achieving basic skills in chemometric methods by the students (performing basic analyses and interpreting the results)

• Familiarizing the students with the available chemometric software.

Course contents

Lecture:

1. Introduction to chemometrics: specificity of multidimensional data; differences between statistics and chemometrics; area of interest in chemometrics; division of chemometric methods; review of the basic software (including MATLAB, Statistica, Origin, SPSS, QSARINS, KNIME).

2. Methods of analyzing the internal structure of the multidimensional chemical data: similarity of objects in the multivariable feature space; hierarchical cluster analysis (HCA) as an example of a similarity analysis method; principal component analysis (PCA) as an example of the projection search method. Examples of applying these methods in various areas of chemistry.

3. Modeling phenomena and processes with regression and classification methods: linear regression of single and multiple variables (LR and MLR), principal component regression (PCR) and partial least squares regression (PLS); linear discriminant analysis (LDA), a non-linear classifier of k-nearest neighbors (kNN); methods for selecting the optimal set of variables in the model (using the genetic algorithm); validation of regression and classification models. Examples of applying these methods in various areas of chemistry.

B. Computer lab:

1. Introduction to chemometric calculations in the KNIEM software environment. The rules of working with a computer.

2. Hierarchical cluster analysis (HCA).

3. Principal Component Analysis (PCA).

4. Linear regression of single and multiple variables (LR / MLR).

Bibliography of literature

A. 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

B. 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.).
- T. Puzyn, J. Leszczynski, M. T. D. Cronin (ed): Recent Advances in QSAR Studies: Methods and Applications. Springer 2010. ISBN: 978-1-4020-9782-9.
- K. Roy, S. Kar, R. Narayan Das (ed): A Primer on QSAR/QSPR Modeling - Fundamental Concepts. Springer 2015. ISBN: 978-3-319-17281-1.

Knowledge

At the end of the course every student:

understands the need for reliable documentation of the results,

knows the basic division of chemometric methods, lists the use of particular groups of these methods in the analysis of chemical data;

knows basic software packages to be used for chemometric analyses;

knows the theoretical background (algorithm of operation) of the most important chemometric methods, including HCA, PCA, LR, MLR, PCR, PLS.

Skills

At the end of the course every student:

uses the KNIME environment for chemometric analyses;

correctly prepares data for further chemometric analysis;

performs HCA and PCA analyses and correctly interpret the obtained results;

develops regression model (LR / MLR method), validates the models correctly and applies the models for predictions.

Social competence

At the end of the course every student:

1. based on his knowledge and skills, critically evaluate the results of research using statistical methods (research, reports, etc.);

2. is convinced of the benefits of using a computer and introducing chemometric methods to his daily research practice;

3. understands the need of deeper learning of chemometric methods;

4. is aware that every numerical result is burdened with measurement uncertainty.