

## Machine Learning Algorithms in Chemistry

Course Workload		Assessment form (examination/ graded test/ ungraded test)
ECTS	Hours	
3	108	Exam

Students will know the principles of development, modification, efficiency evaluation, optimization and interpretation of machine learning algorithms; mathematical description of chemical structures and systems. Students will be able to build interpreted machine learning models of high accuracy; carry out the selection of chemical descriptors; prepare chemical data before submitting to the model; use ready-made technical solutions in the field of data-driven chemistry. Students will have the skills to predict the physico-chemical properties and biological effects of chemical compounds and materials; directional design of molecular structures with predetermined properties.

---

### Course structure:

#### 1. Data analysis

- 1.1. Chemical data visualization and interpretation
- 1.2. Data compression techniques

#### 2. Data preprocessing

- 2.1. Dealing with data sparsity and scarcity
- 2.2. Feature selection and development
- 2.3. Data augmentation
- 2.4. Chemical descriptors

#### 3. Machine learning algorithm choice

- 3.1. Math behind machine learning algorithms
- 3.2. Model precision, under-, and overfitting evaluation
- 3.3. Model interpretability and expandability

#### 4. Machine learning algorithm optimization

- 4.1. Model hyperparameters tuning
  - 4.2. Chemical descriptors correction
-