

MACHINE LEARNING FOR CHEMISTRY AND MATERIALS SCIENCE

David Balcells

david.balcells@kjemi.uio.no

April 20, 2021

**Level: BSc & MSc – Points: 10 – Teaching & Exam: Spring
Language: English**

Course Description

This course provides a general overview of machine learning and its application to chemistry and materials science. The students will learn the fundamentals of machine learning and how to implement predictive models, critically assessing their performance in terms of accuracy, transferability, and explainability. The skills acquired in this course are currently in high demand in both industry and academia, including the ability to process, interpret and exploit experimental and computational data.

Learning Goals

- Understanding the fundamental concepts of machine learning
- Collecting, curating, preprocessing, and transforming chemical data
- Understanding molecules and materials by data visualization and analysis
- Choosing the optimal machine learning representation and method for the problem tackled and the goals sought
- Implementing and critically assessing machine learning models for chemistry and materials science
- Python programming for machine learning

Mandatory Courses

You should have passed KJM1101 (General Chemistry), MAT1050 (Applied Mathematics) and IN-KJM1900 (Introduction to Programming for Chemists).

Recommended Courses

The following courses are recommended: KJM1111 (Organic Chemistry I), KJM1121 (Inorganic Chemistry I), KJM1130 (Physical Chemistry I), and KJM2400 (Analytical Chemistry I).

Overlapping Courses

None.

Teaching

The lectures will have two different formats: 1) Theory lectures, for a total of 30 hours, and 2) Computer practicals, for a total of 30 hours. There will be mandatory exercises (5 for the BSc students, 6 for the MSc students). Attendance is mandatory for the first lecture.

Examination

The course will be evaluated in A-F characters (A = highest mark, F = lowest mark). 50% of the final mark will be based on the evaluation of five mandatory exercises. MSc students will have to deliver one extra mandatory exercise. The other 50% of the final mark will be based on an oral exam in which the students will present and defend a project consisting in the implementation and evaluation of a machine learning model. External help resources will not be allowed in the oral exam. The language of the oral exam will be English.