CB2310 Applied Machine Learning in Molecular Biotechnology 7.5 credits

Tillämpad maskininlärning inom molekylär bioteknik

This is a translation of the Swedish, legally binding, course syllabus.

Establishment

The official course syllabus is valid from the autumn semester 2026 as decided by the Faculty Board decision PA-2025-0010. Date of decision: 2025-10-01.

Grading scale

A, B, C, D, E, FX, F

Education cycle

Second cycle

Main field of study

Biotechnology

Specific prerequisites

Completed degree project 15 credits, 20 credits in biotechnology, genomics, bio(medical) science, computer science or biostatistics, 6 credits in mathematics and English B/6.

Intended learning outcomes

After completion of the course, the students shall have knowledge to:

- Explain key concepts of machine learning, including supervised and unsupervised learning, neural network architectures, model training and validation, feature selection, and data regularization.
- Apply machine learning to a variety of biological data types - such as genomic sequences, omics datasets, microcopy (or medical) images, and clinical data - to address real-world challenges in biotechnology.
- Evaluate and improve model performance using appropriate metrics (e.g. accuracy, precision, recall), cross-validation, and data handling (e.g. feature selection, normalization).
- Understand concepts and strategies for optimizing model robustness and generalizability.
- Identify and discuss ethical and societal implications of machine learning in life sciences, including bias in data training, model transparency, data privacy, and accountability in automated decision-making.
- Communicate machine learning processes and results effectively to both technical and non-technical audiences, using interpretable visualizations, summary statistics and domain-relevant language.

Course contents

The course provides students with the skills to apply, improve, interpret, and evaluate machine learning methods in biotechnology. The course introduces practical applications of machine learning in genomics, transcriptomics, proteomics, and biomedical research. Students will learn the fundamentals of supervised and unsupervised learning as well as neural network architectures, with an emphasis on real-world applications.

The lectures are combined with hands-on exercises, where theoretical foundations are integrated into biological and biomedical context. Theoretical and practical knowledge is deepened through a project in which students select relevant data and design, optimize, and apply a machine learning model.

- Introduction to machine learning and its applications in biotechnology
- Supervised models in biotechnology I: Classification strategies
- Supervised models in biotechnology II: Regression models
- Model validation and optimization: Key metrics and strategies
- Data normalization and regularization: Limitations, challenges, and the best practices
- Unsupervised models in biotechnology I: Clustering and pattern search
- Unsupervised models in biotechnology II: Dimensionality reduction
- Artificial neural networks in biotechnology: Building networks of algorithms
- Deep learning transforming biotechnology: From structure predictions to functional assays
- Society, ethics, and broader impacts of machine learning.

Examination

- TEN1 - Written exam, 4.0 credits, grading scale: A, B, C, D, E, FX, F
- LAB1 - Computational analyses, 2.0 credits, grading scale: P, F
- PRO1 - Group project, 1.5 credits, grading scale: P, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability. The examiner may apply another examination format when re-examining individual students. If the course is discontinued, students may request to be examined during the following two academic years.

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.