UNIVERSITY OF GOTHENBURG

DEPARTMENT OF PHYSICS

FYM285 Learning from data, 7.5 credits

Bayesiansk dataanalys och maskininlärning, 7,5 högskolepoäng

Second Cycle

Confirmation

This course syllabus was confirmed by Department of Physics on 2019-03-11 and was last revised on 2023-05-08 to be valid from 2024-01-15, spring semester of 2024.

Field of education: Science 100%

Department: Department of Physics

Position in the educational system

The course is part of the master program in physics.

The course can be part of the following programmes: 1) Complex Adaptive Systems, Master's Programme (N2CAS) and 2) Physics, Master's Programme (N2PHY)

Main field of studies	*Specialization*
Physics	A1N, Second cycle, has only first-cycle course/s as entry requirements

Entry requirements

A Bachelor's degree in Physics or similar, including 30 credits mathematics (including linear algebra and analysis), and programming.

Applicants must prove their knowledge of English: English 6/English B from Swedish Upper Secondary School or the equivalent level of an internationally recognized test, for example TOEFL, IELTS.

Learning outcomes

The course introduces a variety of central algorithms and methods essential for performing scientific data analysis using statistical inference and machine learning.

Much emphasis is put on practical applications of Bayesian inference in the natural and engineering sciences, i.e. the ability to quantify the strength of inductive inference from facts (such as experimental data) to propositions such as scientific hypotheses and models.
The course is project-based, and the students will be exposed to fundamental research problems through the various projects, with the aim to reproduce state-of-the-art scientific results. The students will use the Python programming language, with relevant open-source libraries, and will learn to develop and structure computer codes for scientific data analysis projects.

On successful completion of the course the student will be able to:

- integrate knowledge of common statistical distributions and central concepts in Bayesian statistics into the analysis of scientific data;
- explain central aspects of Monte Carlo methods and Markov chains, and numerically apply these methods to sample multivariate probability densities;
- quantify and critically assess uncertainties of model parameters that are statistically inferred from scientific data; perform model comparison using a Bayesian viewpoint;
- understand and numerically implement several basic algorithms used in data analysis and machine learning such as linear methods for regression and classification, simple neural networks and gaussian processes;
- use python to perform scientific data analysis using statistical inference and machine learning and to visualize numerical results.
- write well-structured technical reports where results and conclusions from a scientific data analysis are communicated in a clear way.
- maintain a scientific and ethical conduct in the process of analyzing data and writing computer programs.

Course content

The course has two central parts
1. Bayesian inference and data analysis.
2. Machine learning methods for data analysis.

The following subtopics will be covered
- Basic concepts from statistics, expectation values, variance, covariance, correlation functions and errors; discrete versus continuous probability distributions;
- Review of simple statistics models, binomial distribution, the Poisson distribution, simple and multivariate normal distributions;
- Central elements of Bayesian statistics and modeling;
- Monte Carlo methods, Markov chains, Metropolis-Hastings algorithm;

- Linear methods for regression and classification;
- Gaussian and Dirichlet processes;
- Neural networks;

Form of teaching

Lectures
Supervised computational exercises (group work on numerical projects)
Selected number of small analytical and numerical homework exercises
Two computational projects with written reports.

Language of instruction: English

Assessment

The final grade is based on the performance on homework assignments and the graded numerical projects.

If a student, who has failed the same examined element on two occasions, wishes to change examiner before the next examination session, such a request is to be submitted to the department in writing and granted unless there are special reasons to the contrary (Chapter 6, Section 22 of Higher Education Ordinance).

In the event that a course has ceased or undergone major changes, students are to be guaranteed at least three examination sessions (including the ordinary examination session) over a period of at least one year, though at most two years after the course has ceased/been changed. The same applies to work experience and VFU, although this is restricted to just one additional examination session.

Grades

The grading scale comprises: Pass with Distinction (VG), Pass (G) and Fail (U).
The final grade is based on the performance on homework assignments and the graded numerical projects.

Course evaluation

The results of and possible changes to the course will be shared with students who participated in the evaluation and students who are starting the course.