PHYS-231

Data sciences

Zdeborová Lenka

Cursus	Sem.	Type
Physics	BA3	Obl.

Language of teaching	English
Credits	4
Session	Winter
Semester	Fall
Exam	Written
Workload	120h
Weeks	14
Hours	**4 weekly**
Courses	2 weekly
Exercises	1 weekly
Lab	1 weekly
Number of positions	

Summary

This course introduces tools for data analysis in physics: numerical linear algebra, regression, dimensionality reduction, probability, statistics, uncertainty quantification, random walks, Monte Carlo methods and phase transitions, with Python exercises.

Content

Data Analysis for Physics introduces mathematical, statistical and computational tools for extracting information from data in physics and the sciences. The course starts with numerical linear algebra and basic data-analysis methods: eigenvalues and eigenvectors, the power method, singular value decomposition, principal component analysis, linear regression, least squares, regularization, overfitting, validation, gradient descent, logistic regression and simple neural networks.
The second part develops probability and statistics as the language of uncertainty in experiments. Topics include random-number generation, sampling from probability distributions, propagation of errors, correlated and independent uncertainties, the law of large numbers, the central limit theorem, statistical estimators and maximum likelihood estimation. Regression is then revisited from a probabilistic point of view, including error bars, weighted least squares and Bayesian regularization.
The final part turns to stochasticity as a modelling principle in physics. The course covers Brownian motion, diffusion, random walks, maximum entropy, rejection sampling and the curse of dimensionality, Markov chain Monte Carlo, the Metropolis rule, detailed balance and examples from interacting particle systems. It concludes with phase transitions, including percolation, or giant components in random graphs.

Learning Outcomes

By the end of the course, the student must be able to:

- Use basic tools for data analysis
- Compute error bars in physics experiments.

Teaching methods

2 hours of lectures + 2 hours of exercises per week, mostly computer-based.

Assessment methods

Final written exam counting for 70%, and work during the semester counting for the remaining 30%.

Resources

Notes/Handbook

Lecture notes on Moodle.

Moodle Link

- https://go.epfl.ch/PHYS-231