PHYS-467

Machine learning for physicists

Zdeborová Lenka

Cursus	Sem.	Type
Ing.-phys	MA1, MA3	Opt.
Physicien	MA1, MA3	Opt.

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

Summary

Machine learning and data analysis are becoming increasingly central in sciences including physics. In this course, fundamental principles and methods of machine learning will be introduced and practised.

Content

Reminder of key concept from probability theory. Bayesian inference. Entropy as a measure of information.
K-nearest neighbours and basic concepts of supervised learning. Bias-variance trade-off, concept of overfitting.
Linear regression and least squares. Learning in high dimension and the need and concept of regularization, ridge and Lasso.
Maximum likelihood. Classification problems, logistic regression.
Unsupervised learning, clustering, SVD and PCA, dimensionality reduction. Other spectral methods.
Learning in high dimension aka statistical mechanics, Bayesian inference as sampling from the Boltzmann measure, Maximum likelihood as search for the ground state.
Monte Carlo Markov Chains, Gibbs sampling, simulated annealing in the context of machine learning.
Feature spaces, Kernel methods, support vector machines.
Neural networks as learning features, one hidden layer neural network.
Multilayer (Deep) neural networks, back-propagation aka gradient descent.
Towards modern deep neural networks. Convolution architectures.
Generative models -- Boltzmann machine and maximum entropy principle, restricted Boltzmann machine, auto-encoder, generative adversarial networks (GANs).

Learning Prerequisites

Important concepts to start the course

Basic notions in probability, analysis and basic familiarity with programming. Some notions of statistical physics will be used to support this lecture.

Learning Outcomes

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

- Use basic tools for data analysis and for learning from data
- Explain basic principles of data analysis and learning from data
- List and explain machine learning tools suited for a given problem.

Teaching methods

2h of lecture + 2h of excercise (exercise mostly with a computer)

Assessment methods

Final written exam counting for 50% and several graded homeworks during the semester counting for the other 50%.

Resources

Bibliography

A high-bias, low-variance introduction to Machine Learning for physicists. Pankaj Mehta, Marin Bukov, Ching-Hao Wang, Alexandre G.R. Day, Clint Richardson, Charles K. Fisher, David J. Schwab, https://arxiv.org/abs/1803.08823.

Text book "Information Theory, Inference, and Learning Algorithms" by David MacKay.