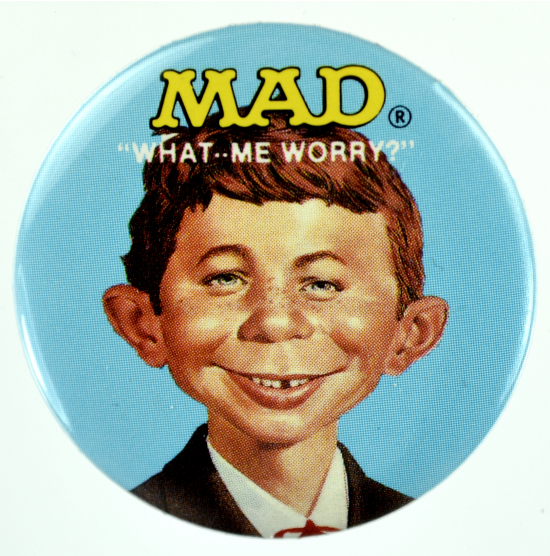
Summary of course

Morten Hjorth-Jensen Email morten.hjorth-jensen@fys.uio.no[1,2]

[1]Department of Physics and Center of Mathematics for Applications, University of Oslo
[2]National Superconducting Cyclotron Laboratory, Michigan State University

Nov 29, 2018

What? Me worry? No final exam in this course!



What did I learn in school this year?

Our ideal about knowledge on computational science

Does that match the experiences you have made this semester?

Topics we have covered this year

The course has two central parts

1. Statistical analysis and optimization of data
2. Machine learning

Statistical analysis and optimization of data

The following topics will be covered

1. Basic concepts, expectation values, variance, covariance, correlation functions and errors;
2. Simpler models, binomial distribution, the Poisson distribution, simple and multivariate normal distributions;
3. Central elements of Bayesian statistics and modeling;
4. Central elements from linear algebra
5. Gradient methods for data optimization
6. Monte Carlo methods, Markov chains, Metropolis-Hastings algorithm;
7. Estimation of errors using cross-validation, blocking, bootstrapping and jackknife methods;
8. Practical optimization using Singular-value decomposition and least squares for parameterizing data.
9. Principal Component Analysis.

Machine learning

The following topics will be covered

1. Linear methods for regression and classification;
2. Boltzmann machines;
3. Neural networks;
4. Decisions trees and nearest neighbor algorithms
5. Support vector machines

Learning outcomes and overarching aims of this course

The course introduces a variety of central algorithms and methods essential for studies of data analysis and machine learning. The course is project based and through the various projects, normally three, you will be exposed to fundamental research problems in these fields, with the aim to reproduce state of the art scientific results. The students will learn to develop and structure large codes for studying these systems, get acquainted with computing facilities and learn to handle large scientific projects. A good scientific and ethical conduct is emphasized throughout the course.

- Understand linear methods for regression and classification;
- Learn about neural network;

- Learn about basic data analysis, Bayesian statistics, Monte Carlo methods, data optimization and machine learning;
- Be capable of extending the acquired knowledge to other systems and cases;
- Have an understanding of central algorithms used in data analysis and machine learning;
- Gain knowledge of central aspects of Monte Carlo methods, Markov chains, Metropolis and Gibbs samplers and their possible applications;
- Work on numerical projects to illustrate the theory. The projects play a central role and students are expected to know modern programming languages like Python or C++.

Other courses on Data science and Machine Learning at UiO

The link here `https://www.mn.uio.no/english/research/about/centre-focus/innovation/data-science/studies/` gives an excellent overview of courses on Machine learning at UiO.

1. STK2100 Machine learning and statistical methods for prediction and classification.
2. IN3050 Introduction to Artificial Intelligence and Machine Learning. Introductory course in machine learning and AI with an algorithmic approach.
3. STK-INF3000/4000 Selected Topics in Data Science. The course provides insight into selected contemporary relevant topics within Data Science.
4. IN4080 Natural Language Processing. Probabilistic and machine learning techniques applied to natural language processing.
5. STK-IN4300 – Statistical learning methods in Data Science. An advanced introduction to statistical and machine learning. For students with a good mathematics and statistics background.
6. INF4490 Biologically Inspired Computing. An introduction to self-adapting methods also called artificial intelligence or machine learning.
7. IN-STK5000 Adaptive Methods for Data-Based Decision Making. Methods for adaptive collection and processing of data based on machine learning techniques.
8. IN5400/INF5860 – Machine Learning for Image Analysis. An introduction to deep learning with particular emphasis on applications within Image analysis, but useful for other application areas too.

9. TEK5040 – Dyp læring for autonome systemer. The course addresses advanced algorithms and architectures for deep learning with neural networks. The course provides an introduction to how deep-learning techniques can be used in the construction of key parts of advanced autonomous systems that exist in physical environments and cyber environments.

Additional courses of interest

1. STK4051 Computational Statistics
2. STK4021 Applied Bayesian Analysis and Numerical Methods

Best wishes to you all and thanks so much for your heroic efforts this semester