Syllabus
Learning Theory, 6hp

Issued by the WASP graduate school management group 2021 12 15.

Main field of study

AI/MLX

Course level

PhD student course

Course offered for

PhD Students in the WASP graduate school

Entry requirements

Basic eligibility. Recommended background: Multivariable analysis, Probability theory and statistics, and Numerical methods, basic course, or equivalent knowledge. The participants are assumed to have a background in mathematics corresponding to the contents of the WASP-course “Mathematics for Machine Learning”.

Intended learning outcomes

After passing the course, the student should be able to:

- Derive and apply the basic theoretical tools used in modern machine learning
- Describe known performance guarantees for important machine learning algorithms
- Describe the factors that contribute to the accuracy of learning methods.
- Identify some of the difficulties involved in analyzing current machine learning technology.

Course content

Module 1:

Topic 1. Introduction

Main types of learning: supervised, unsupervised and reinforcement learning, and their mathematical formalization (input and label spaces, hypothesis classes, loss function).

Topic 2. PAC framework and empirical risk minimization

Concept of Probably Approximately Correct (PAC) learnability. Oracle inequalities and bias-variance trade-off. Empirical Risk Minimization Principle. Overfitting and No-Free-Lunch Theorem. Uniform convergence.

Topic 3. Concentration inequalities

Asymptotic versus finite sample probability bounds. Markov, Chebyshev and Chernoff bounds. Sub-Gaussian random variables. Hoeffding's Lemma and Inequality. Bounded difference (McDiarmid) inequality.

Topic 4. Vapnik-Chervonenkis (VC) Theory

PAC learnability of finite hypothesis classes. Shattering and VC dimension. Sauer-Shelah's lemma. Rademacher complexity. Fundamental Theorem of PAC learning.

Module 2:

Topic 5. Linear classification and regression

Linear predictors. Linear classification. Perceptron algorithm. Application of VC theory to multilayer neural networks. Logistic and linear regression.

Topic 6. Regularization, stability and optimization

Regularized risk minimization. Algorithmic stability and its application to generalization bounds for regularized risk minimization. Algorithms for convex learning: gradient descent, sub-gradient descent and stochastic gradient descent.

Topic 7. Support vector machines and kernel methods

Introduction to SVM with hard and soft margins. Performance bounds of hard and soft-margin SVM. Learning algorithms for SVM. Kernel methods; linear separability using embeddings. Kernel trick and the representer theorem; admissible kernels.

Topic 8. Deep neural networks

Neural networks and representation theorems. Training neural nets using back propagation. Dropout as a regularization technique. Recent results about the loss surface and local minima of neural networks. Recent theoretical developments justifying deep learning.

Module 3:

Topic 9. Clustering. Cluster validation and algorithms

Performance metrics for clustering. State-of-the-art clustering algorithms. Cluster evaluation. K-means and its performance guarantees. The EM algorithm and its performance for Gaussian mixtures. Spectral clustering, random matrix theory and concentration.

Topic 10. Active learning, online optimization and sequential decision making

Introduction to bandit problems and reinforcement learning. Exploration-exploitation trade-off. Fundamental limits via the change-of-measure arguments. Examples of algorithms, and their guarantees. Best policy identification vs regret minimization.

Teaching and working methods

Mainly lecture based with some simulation exercises. The course includes three 2-day meetings with intense teaching on-site, typically a mixture of lectures and exercises.

Examination

The examination will consist of three peer-graded individual assignments, one per module.

Grades

Fail or Pass