236756: Introduction to Machine Learning

Technion, Spring 2014

Instructors: Prof. Nathan Srebro-Bartom and Prof. Nir Ailon.

The course will focus on understanding important concepts in machine learning, and will introduce the main paradigms and methods underlying modern machine learning. This will be done by understanding the mathematical formulation of statistical learning and the basics of statistical and computational learning theory, by studying specific learning algorithms, and by empirical experimentation with them.

Course Text: **Understanding Machine Learning: From Theory to Algorithms**, Shalev-Shwartz and Ben-David, Cambridge University Press, *available later in 2014*. The book is not yet publicly available, but a complete electronic version will be made available to students in the course.

All required material will be from the course text, or from supplemental notes posted on the website.

Language of Instruction: All course material, including slides, will be in English. Lectures and recitations will be conducted in Hebrew (subject to change in subsequent years).

Specific topics (not all sections of the indicated chapters will be included):

- What is Machine Learning? Examples of Learning Systems.
- Formal model of Statistical Learning Theory, No-Free-Lunch and the need for inductive bias, PAC Learning, the VC Dimension, and the Fundamental Theorem of PAC Learning (Ch. 2-6)
- Minimum Description Length and using a prior as an inductive bias (Ch 7.3)
- Tractability in Machine Learning; Proper vs Improper Learning (Ch 8)
- Linear predictors, the Perceptron, feature maps, margin and regularization, Kernalization, Support Vector Machines and Boosting as linear prediction (Ch 9, 10, 15, 16)
- The importance of convexity and surrogate loss functions (Ch 12)
- Structured Loss Minimization, Model Selection and Validation (Ch 13.1, 7.1-7.2, 11)
- Stochastic Gradient Descent as an important algorithm for Machine Learning (Ch. 14.1-14.3,15.5)
- Brief introduction of the Online Learning framework (Ch. 21.1,21.4)
- Multi-Layered (Deep) Neural Networks: models, back propagation, feature learning, semi-supervised learning and transfer learning (Ch. 20 and additional material).
- The complexity of decision trees (Ch. 18.1)
- Nearest Neighbour methods (Ch. 19.1)
- Using Generative Models: Naive Bayes, Linear Discriminant Analysis (Ch 24.1-24.3)

Requirements: Homeworks (40%), final exam (60%). Homeworks will include both written questions and 2-3 experimentation exercises, mostly using provided code that students will complete and experiment with on real data.

Prerequisites: The formal prerequisite is “236501 Introduction to Artificial Intelligence”. It will be helpful to take 236501 before this class as it gives an initial introduction to learning, and we will also be referring as examples to some models presented in 236501. However, it is certainly possible to take 236756 without 236501.
Classes in Probability (094412 or alternatives) and Linear Algebra (104167) are very strongly recommended, and classes in Algorithms (234247) and in Numerical Analysis (234107) are also helpful.