CS 7545 Machine Learning Theory, Fall 2013

Course Information

Lectures: Mon/Wed 3:05-4:25, ES&T L1175.

Instructor: Maria Florina Balcan (KACB 2144 , 404-385-8640).

Course Description: Machine learning studies automatic methods for learning to make accurate predictions or useful decisions based on past observations and experience, and it has become a highly successful discipline with applications in many areas such as natural language processing, computer vision, web mining, or bioinformatics.

This course on the design and theoretical analysis of machine learning methods will cover a broad range of important problems studied in theoretical machine learning. It will provide a basic arsenal of powerful mathematical tools for their analysis, focusing on both statistical and computational aspects. We will examine questions such as: What guarantees can we prove on the performance of learning algorithms? What can we say about the inherent ease or difficulty of learning problems? Can we devise models that are both amenable to mathematical analysis and are successful empirically? In addressing these and related questions we will make connections to statistics, algorithms, complexity theory, information theory, optimization, game theory, and empirical machine learning research.

Prerequisites: Either a good Machine Learning or a good Theory/Algorithms background.

Evaluation and Responsibilities: Grading will be based on 3 or 4 homework assignments, a take-home final, and a class presentation or project.

General structure of the course: We will use roughly 3/4 of the lectures to cover "core" topics in this area, and then will diverge in the remaining 1/4 based on student interest. Here is a short outline of the "core" portion (some bullets correspond to multiple lectures):

- Basic models for passive supervised learning: PAC and Statistical Learning Theory
- Simple algorithms and hardness results for passive supervised learning
- Mistake-bound and Online learning. The Weighted-Majority, Winnow, and Perceptron Algorithms
- VC-dimension, uniform convergence; other modern sample complexity results (e.g., Rademacher complexity, localization)
- Margins and support-vector machines; Kernel methods
- Weak-learning vs. Strong-learning; the AdaBoost algorithm
- Modern learning paradigms: Semi-supervised learning, Interactive learning, Distributed Learning, Transfer Learning.

Textbooks: The recommended (not required) textbooks are *An Introduction to Computational Learning Theory* by M. Kearns and U. Vazirani, and *A Probabilistic Theory of Pattern Recognition* by L. Devroye, L. Gy�rfi, G. Lugosi. Additionally, we will use a number of survery articles and tutorials.