Non-Computer Science students welcome!
Prerequisite: CS 4413, or permission of instructor

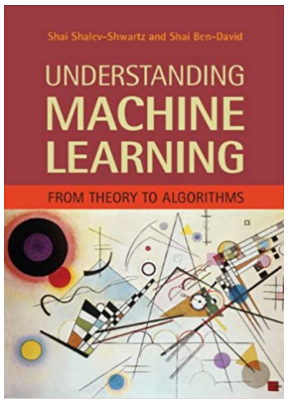
Fall 2023
Tuesdays and Thursdays
10:30am – 11:45am

Kaufman Hall, Rm 140

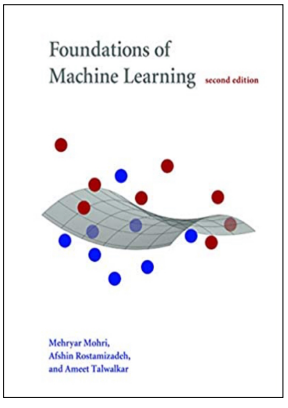
Dimitris Diochnos
School of Computer Science
230 Devon Energy Hall
✉ diochnos@ou.edu
405-325-2978

Computer Science 5713
Computational Learning Theory

Computational learning theory views learning as a computational process and tries to understand the principles that govern learning using tools from computer science and mathematics. The course starts with basic topics such as learning using membership and equivalence queries, version spaces, linear models, and decision trees. Having said that, the focus is on *provable guarantees of machine learning methods*. Along these lines we explore the probably approximately correct (PAC) model of learning, Occam algorithms, the VC-dimension, and the sample complexity of distribution-independent learning. We discuss representation issues, proper learning, reductions, intractability, learning in the realizable case, and agnostic learning. In addition, we explore topics under the broader umbrella of *trustworthy machine learning* such as noise models and adversarial machine learning (poisoning attacks, adversarial examples). This year we plan to expland also upon interpretability and explainability aspects, as well as upon fairness concerns. Other topics include distribution-specific learning, online learning, learning with expert advice, and boosting.



The seminar format will include reading, discussion and rigorous derivations of mathematical properties that characterize various learning problems. We will cover several seminal results and important ideas spanning the last 40 years of research. Several handouts will be distributed and we will maintain our own notes for the course as we make progress. Nevertheless, two books can complement our discussion in class and either one of them can work well as a reference: (A) *Understanding Machine Learning: From Theory to Algorithms* by Shai Shalev-Schwartz and Shai Ben-David, and (B) *Foundations of Machine Learning* by Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. Both books are available for free online. The goals of this course are for students to: (1) develop a comprehensive understanding of the field; (2) understand and be able to apply formal approaches that provide guarantees on the behavior of machine learning algorithms, or argue about the inherent hardness of certain learning problems; (3) understand research papers in this broad field and be able to present the ideas to others.



School of Computer Science
College of Engineering