ELEG/CISC 867 - Advanced Machine Learning

Xiugang Wu

Spring 2019

E-mail: xwu@udel.edu
Web: http://www.eecis.udel.edu/~xwu
Office Hours: TR 11:00-12:00 am
Class Hours: TR 9:30-10:45 am
Office: Evans 314
Class Room: Gore 308

Course Description

This advanced course on machine learning features an in-depth treatment of modern learning theory and emphasizes its interplay with real-world learning algorithms. The main goal of this course is to get students started in research, in particular, to help them transition from knowing how to implement towards exploring why to do this and how to do better. Students will carry out research projects, and the hope is that some of these projects will result in research papers that can be published in top machine learning venues.

Prerequisite

Undergraduate-level probability theory and linear algebra; mathematical maturity in general.

Textbook

- Main Textbook: S. S. Shwartz and S. Ben-David, Understanding Machine Learning, Cambridge, 2016. [Free pdf version online]
- Reference: Y. S. Abu-Mostafa, M. Magdon-Ismail and H.-T. Lin, Learning From Data, AMLBook, 2012.

Evaluation

- Attendance: 10%; Project (Presentation and Report): 40%; Final Exam: 50% + 10% bonus

Course Content

Part I: Foundation

- PAC Learning Framework
  - Empirical Risk Minimization (ERM)
  - Uniform Convergence is Sufficient For Learnability
- VC Theory
  - No Free Lunch Theorem
  - Fundamental Theorem of Learning
- Non-Uniform Learnability
  - Structural Risk Minimization (SRM)
  - Minimum Description Length and Occam's Razor

Part II: From Theory to Algorithms

- Linear Predictor and Boosting
  - Linear Regression and Logistic Regression
  - Boosting
- Support Vector Machine
  - SVM: From 0-1 Loss to Hinge Loss
  - Kernel Methods
- Decision Trees
  - Decision Tree Algorithms
  - Random Forests

Part III: Advanced Topics

- Minimax Learning
  - ERM v.s. Minimax Approach
  - Maximum Entropy Machine
- Online Learning
  - Online Classification
  - Online Convex Optimization
- Interpretation of Deep Neural Networks
  - Overview of Feedforward Neural Networks
  - An Information Theoretic Perspective