ELEG 602 - Advanced Machine Learning

Xiugang Wu

Fall 2021

E-mail: xwu@udel.edu
Web: http://www.eecis.udel.edu/~xwu
Office Hours: TR 10:45-11:45 am
Class Hours: TR 9:30-10:45 am

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

- Exposure to a first course on machine learning (e.g. ELEG/FSAN 815)
- Knowledge on probability theory (e.g. ELEG 310) and linear algebra (MATH 351)
- Mathematical maturity in general

Textbook

- Lectures are based on slides that will be posted on Canvas.
- Main Reference: S. S. Shwartz and S. Ben-David, Understanding Machine Learning, Cambridge, 2016. [Free pdf version online]

Evaluation

- Homework: 50%; Project (Presentation and Report): 50% + 10% bonus

Course Content

Part I: Foundations

- 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
- Research Frontiers
  - PAC-Bayesian Bound
  - Information-Theoretic Analysis of Generalization Capability

Part II: Supervised Learning

- 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
- Research Frontiers
  - Domain Adpation
  - Information Flow in Deep Neural Networks

Part III: Unsupervised Learning

- Dimensionality Reduction
  - Principal Component Analysis
  - Compressed Sensing
- Clustering
  - k-Means
  - Spectral Clustering
- Generative Models
  - MLE, Naive Bayes
  - Latent Variables and the EM algorithm
- Research Frontiers
  - Graph Problems
  - Generative Adversarial Networks

Part IV: Additional Learning Models

- Minimax Learning
  - ERM v.s. Minimax Approach
  - Maximum Entropy Machine
- Online Learning
  - Online Classification
  - Online Convex Optimization
- Reinforcement Learning
  - Markov Decision Processes
  - Multi-Armed Bandits (MAB)
- Research Frontiers
  - Distributed Learning
  - Information-Theoretic Treatment of MAB