

Understanding Machine Learning (Shalev-Shwartz and Ben-David)

Course Notes

Alexander Wang (aw576@cornell.edu)

July 8, 2021

Introduction

My goal with this document is to read and learn as many chapters of the textbook, “Understanding Machine Learning” by Shai Shalev-Shwartz and Shai Ben-David as possible. This textbook covers a graduate level introduction to machine learning and is the primary text used by Cornell University’s CS 4780: Introduction to Machine learning, of which I have taken and will be a teaching assistant for in the Fall semester of 2021, and hopefully future semesters. When I took CS 4780, professor Thorsten Joachims provided multiple chapters of this textbook as auxilliary course material to his students, indicating the importance of the concepts in this book to the study of machine learning. In the fall, Professor Kilian Weinberger will be teaching CS 4780, but he will likely still follow closely to what is taught in this book. Seeing that the textbook is meant for a graduate level introduction to the subject, Cornell’s undergraduate class did not cover all aspects of this textbook. However, I will do my best to explore even the more complex topics on my own. My objectives are to be a more knowledgeable person in the field, and to be the teaching assistant I wish I had while taking the course for the first time.

Table of Contents

Part 1: Foundations

- Chapter 1: Introduction [OMITTED]
- Chapter 2: A Gentle Start
- Chapter 3: A Formal Learning Model
- Chapter 4: Learning via Uniform Convergence
- Chapter 5: The Bias-Complexity Trade-off
- Chapter 6: The VC-Dimension
- Chapter 7: Nonuniform learnability
- Chapter 8: The Runtime of Learning

Part 2: From Theory to Algorithms

- Chapter 9: Linear Predictors

- Chapter 10: Boosting
- Chapter 11: Model Selection and Validation
- Chapter 12: Convex Learning Problems
- Chapter 13: Regularization and Stability
- Chapter 14: Stochastic Gradient Descent
- Chapter 15: Support Vector Machines
- Chapter 16: Kernel Methods
- Chapter 17: Multiclass, Ranking, and Complex Prediction Problems
- Chapter 18: Decision Trees
- Chapter 19: Nearest Neighbor
- Chapter 20: Neural Networks

Part 3: Additional Learning Models

- Chapter 21: Online Learning
- Chapter 22: Clustering
- Chapter 23: Dimensionality Reduction
- Chapter 24: Generative Models
- Chapter 25: Feature Selection and Generation

Part 4: Advanced Theory

- Chapter 26: Rademacher Complexities
- Chapter 27: Covering Numbers
- Chapter 28: Proof of the Fundamental Theorem of Learning Theory
- Chapter 29: Multiclass Learnability
- Chapter 30: Compression Bounds
- Chapter 31: PAC-Bayes

Chapter 2: A Gentle Start