

Homework 2

APPM 4490/5490 Spr 2022 Theoretical ML

Due date: Friday, Jan 28, before midnight, on Canvas
Theme: Bias-variance tradeoff, uniform convergence

Instructor: Prof. Becker

Instructions Collaboration with your fellow students is OK and in fact recommended, although direct copying is not allowed. The internet is allowed for basic tasks (e.g., looking up definitions on wikipedia) but it is not permissible to search for proofs or to *post* requests for help on forums such as <http://math.stackexchange.com/> or to look at solution manuals. Please write down the names of the students that you worked with.

An arbitrary subset of these questions will be graded.

Reading You are responsible for chapters 4 and 5 in “Understanding Machine Learning” by Shai Shalev-Shwartz and Shai Ben-David (2014, Cambridge University Press). Note: you can buy the book for about \$45 on Amazon (or less for an e-book), and the authors host a free PDF copy on their [website](#) (but note that this PDF has different page numbers).

Problem 1: Read the following two papers, and write 1 paragraph *each* about the papers (a response of some sort: your reaction to the paper). Spend no more than 1 hour *total* on the reading. **Your two paragraphs should be written directly in the “HW 2” assignment in Canvas, not printed out!**

- a) *Uniform convergence may be unable to explain generalization in deep learning*, by Nagarajan and Kolter, NeurIPS 2019. [URL](#). This was voted the “Outstanding New Directions Paper Award” at NeurIPS 2019 (NeurIPS is the flagship machine learning conference). Note that they use almost identical notation to our class.
- b) *Reconciling modern machine-learning practice and the classical bias-variance trade-off*, by Belkin, Hsu, Ma, and Mandal; Proceedings of the National Academy of Sciences (PNAS) 2019, vol 116(32). [URL](#). This is short and very readable.

Optional If you want to learn more about some topics:

- The following paper has background on the history of David Wolpert’s No Free Lunch theorems: *No Free Lunch Theorem: A Review* by Adam, Alexandropoulos, Pardalos, Vrahatis. 2019, book chapter. [URL](#)
- [Introduction to Statistical Learning Theory](#) by Olivier Bousquet, Boucheron and Lugosi gives some variants of No Free Lunch theorems in Appendix B. (This is from a 2003 [proceedings of a summer school in machine learning](#)). Also [Learning Theory from First Principles \[DRAFT, 2021\]](#) by Francis Bach gives very similar versions. Both works credit Chapter 7 of **A probabilistic Theory of Pattern Recognition** (Devroye, Györfi, Lugosi; 1996, Springer-Verlag New York) for these results.