Computational Learning Theory

Staff

Prof. Shafi Goldwasser
`shafi.goldwasser@berkeley.edu`

Jonathan Shafer
`shaferjo@berkeley.edu`

About this Course

This course is an introduction to learning theory. We will present a few definitions of learning, and then proceed to investigate what can (and cannot) be learned according to each definition, and what amount of computational resources (especially data and compute time) are necessary in the cases where learning is possible. Attention will be devoted to understanding definitions and proofs. The course will generally not cover applications and state-of-the-art technologies. See a list of topics on the following page.

Recommended Textbooks

1. Shalev-Shwartz and Ben-David. *Understanding Machine Learning: From Theory to Algorithms.* [SB14]
2. Mohri, Rostamizadeh and Talwalkar. *Foundations of Machine Learning.* [MRT18]
3. Vapnik. *The Nature of Statistical Learning Theory.* [Vap00]

Homework: Six problem sets will be assigned throughout the semester.

Grading

Students may choose between the following options:

1. Six problem sets: 100%
2. Four problem sets: 65%; Final research project: 35%.

Prerequisites: CS 170.

Outline

The following is preliminary and very much subject to change.

Unit	Topic	References
1	Language identification in the limit Learning with queries	[MRT18, Ch. 16] [KV94, Ch. 8] [Gol67, BB75, Ang80, Ang87]
2	PAC learning: – Definitions: PAC, agnostic PAC – Learning via uniform convergence – No free lunch	[SB14, Ch. 3-5] [MRT18, Ch. 2]
3	PAC learning: – VC dimension – Sauer's lemma – Fundamental theorem of PAC learning – Part I	[SB14, Ch. 6] [MRT18, Ch. 3]
4	PAC learning: – Rademacher complexity – Covering numbers – Fundamental theorem of PAC learning – Part II	[SB14, Ch. 26-28] [MRT18, Ch. 3]
5	Boosting	[SB14, Ch. 10] [MRT18, Ch. 7]
6	Runtime complexity of PAC learning Hardness of learning from cryptographic assumptions	[SB14, Ch. 8]
7	Learning parity with noise, and applications to cryptography	[BKW03, KMV08, Reg09]
8	Multiclass learning Natarajan dimension	[SB14, Ch. 29] [MRT18, Ch. 8]
9	Nonuniform learning: – Structural risk minimization – Occam's razor – PAC-Bayes generalization bounds	[SB14, Ch. 7, 31] [MRT18, Ch. 4]
10	Stability and regularization	[SB14, Ch. 13] [MRT18, Ch. 14]
11	Sample compression schemes	[SB14, Ch. 30] [LW86, MY16]
12	Information-theoretic generalization bounds	[XR17, BMN+18, SZ20]
13	Learning algorithms from natural proofs	[CIKK16]
14	Online learning	[SB14, Ch. 21] [MRT18, Ch. 8]
15	Solomonoff induction	[Sol64a, Sol64b, Leg97]

References

[Ang80] Dana Angluin. Inductive inference of formal languages from positive data. *Inf. Control.*, 45(2):117–135, 1980.

[Ang87] Dana Angluin. Queries and concept learning. *Mach. Learn.*, 2(4):319–342, 1987.

[BB75] Lenore Blum and Manuel Blum. Toward a mathematical theory of inductive inference. *Inf. Control.*, 28(2):125–155, 1975.

[BKW03] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, 2003.

[BMN$^+$18] Raef Bassily, Shay Moran, Ido Nachum, Jonathan Shafer, and Amir Yehudayoff. Learners that use little information. In Firdaus Janoos, Mehryar Mohri, and Karthik Sridharan, editors, *Algorithmic Learning Theory, ALT 2018, 7-9 April 2018, Lanzarote, Canary Islands, Spain*, volume 83 of *Proceedings of Machine Learning Research*, pages 25–55. PMLR, 2018.

[CIKK16] Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonina Kolokolova. Learning algorithms from natural proofs. In Ran Raz, editor, *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, volume 50 of *LIPIcs*, pages 10:1–10:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[Gol67] E. Mark Gold. Language identification in the limit. *Inf. Control.*, 10(5):447–474, 1967.

[KMV08] Adam Tauman Kalai, Yishay Mansour, and Elad Verbin. On agnostic boosting and parity learning. In Cynthia Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 629–638. ACM, 2008.

[KV94] Michael J. Kearns and Umesh V. Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, 1994.

[Leg97] Shane Legg. *Solomonoff Induction*. Technical Report 30, Centre for Discrete Mathematics and Theoretical Computer Science, University of Auckland, 1997.

[LW86] Nick Littlestone and Manfred Warmuth. *Relating data compression and learnability*. Unpublished manuscript, 1986.

[MRT18] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. The MIT Press, 2nd edition, 2018.

[MY16] Shay Moran and Amir Yehudayoff. Sample compression schemes for VC classes. *J. ACM*, 63(3):21:1–21:10, 2016.

[Reg09] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34:1–34:40, 2009.

[SB14] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning – From Theory to Algorithms*. Cambridge University Press, 2014.

[Sol64a] Ray J. Solomonoff. A formal theory of inductive inference. Part I. *Inf. Control.*, 7(1):1–22, 1964.

[Sol64b] Ray J. Solomonoff. A formal theory of inductive inference. Part II. *Inf. Control.*, 7(2):224–254, 1964.

[SZ20] Thomas Steinke and Lydia Zakynthinou. Reasoning about generalization via conditional mutual information. In Jacob D. Abernethy and Shivani Agarwal, editors, *Conference on Learning Theory, COLT 2020, 9-12 July 2020, Virtual Event [Graz, Austria]*, volume 125 of *Proceedings of Machine Learning Research*, pages 3437–3452. PMLR, 2020.

[Vap00] Vladimir Naumovich Vapnik. *The Nature of Statistical Learning Theory*. Statistics for Engineering and Information Science. Springer, 2nd edition, 2000.

[XR17] Aolin Xu and Maxim Raginsky. Information-theoretic analysis of generalization capability of learning algorithms. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 2524–2533, 2017.