Computational Learning Theory - Candidate Papers for Final Project

Fall 2020

Latest version: October 12, 2020 (first version: October 11, 2020)

1 One-Person Projects

1.1 PAC-Learning, General

1. *PAC Learning with Irrelevant Attributes*, Aditi Dhagat and Lisa Hellerstein
2. *Learning boolean functions in an infinite attribute space*, Avrim Blum
   (Happy to discuss this as a two-person project if you intend to provide more details on the proofs.)
3. *Attribute-Efficient Evolvability of Linear Functions*, Elaine Angelino and Varun Kanade
   (I am happy to discuss this as a multi-person project as well.)
4. *Smart PAC-Learners*, Hans U. Simon
   (I am happy to discuss this as a multi-person project as well.)
5. *Lower bounds for PAC learning with queries*, György Turán

1.2 PAC-Learning, Noise

1. *Can PAC Learning Algorithms Tolerate Random Attribute Noise?*, Sally A. Goldman and Robert H. Sloan
2. *On Learning Conjunctions with Malicious Noise*, Yishay Mansour and Michal Parnas
3. *PAC-Learning in the Presence of One-sided Classification Noise*, Hans U. Simon
   (I am happy to discuss this as a multi-person project; an extended version is available here.)
4. *Statistical queries and faulty PAC oracles*, Scott E. Decatur
5. *Characterizing Statistical Query Learning: Simplified Notions and Proofs*, Balázs Szörényi

1.3 PAC-Learning, Distribution-Dependent Learning and Nonuniform Learnability

1. *Learnability with respect to fixed distributions*, Gyora M. Benedek and Alon Itai
2. *Nonuniform Learnability*, Gyora M. Benedek and Alon Itai

1.4 Active Learning

1. *Two faces of active learning*, Sanjoy Dasgupta
2. Coarse sample complexity bounds for active learning, Sanjoy Dasgupta
3. Analysis of a greedy active learning strategy, Sanjoy Dasgupta
4. Diameter-Based Active Learning, Christopher Tosh and Sanjoy Dasgupta
   (I am happy to discuss this as a multi-person project.)

1.5 Teaching

1. *Curriculum Learning*, Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston
2. *Teaching with IMPACT*, Carl Trimbach and Michael L. Littman. There is also a preliminary version of that paper *here*.
3. *Teachability in Computational Learning*, Ayumi Shinohara and Satoru Miyano

1.6 Analogical Reasoning

1. *Analogical classification: A new way to deal with examples*, Myriam Bounhas, and Henri Prade and Gilles Richard
2. *Analogical Classifiers: A Theoretical Perspective*, Nicolas Hug, Henri Prade, Gilles Richard, and Mathieu Serrurier

1.7 Target Changes During Learning

1. *Tracking Drifting Concepts By Minimizing Disagreements*, David P. Helmbold and Philip M. Long
2. *Incrementally Learning Time-Varying Half Planes*, Anthony Kuh, Thomas Petschet, and Ronald L. Rivest

1.8 Local Search Methods

1. *Perhaps Not a Free Lunch But At Least a Free Appetizer*, Stefan Droste, Thomas Jansen and Ingo Wegener
2. *Real royal road functions – where crossover is provably essential*, Thomas Jansen and Ingo Wegener
3. *PAC Learning and Genetic Programming*, Timo Kötzing, Frank Neumann, and Reto Spöhel
4. *The Compact Genetic Algorithm*, Georges R. Harik, Fernando G. Lobo, and David E. Goldberg
5. *On the Choice of the Mutation Probability for the (1+1) EA*, Thomas Jansen and Ingo Wegener
6. *When a genetic algorithm outperforms hill-climbing*, Adam Prügel-Bennett

1.9 Levels-Based Optimization

1. *Methods for the Analysis of Evolutionary Algorithms on Pseudo-Boolean Functions*, Ingo Wegener
2. *General Lower Bounds for the Running Time of Evolutionary Algorithms*, Dirk Sudholt (I am happy to discuss this as a multi-person project as well; an extended version is available on arXiv.)

1.10 Helpful

1. *Sorting and Selection with Imprecise Comparisons*, Miklós Ajtai, Vitaly Feldman, Avinatan Hassidim, and Jelani Nelson
   (I am happy to discuss this as a multi-person project for the extended version that is available on arXiv.)
2. *An Elementary Proof of a Theorem of Johnson and Lindenstrauss*, Sanjoy Dasgupta and Anupam Gupta

1.11 Various

1. *How Boosting the Margin Can Also Boost Classifier Complexity*, Lev Reyzin and Robert E. Schapire
2. *Learning to Classify Incomplete Examples*, Dale Schuurmans and Russell Greiner
3. *A Confidence-Based Approach for Balancing Fairness and Accuracy*, Benjamin Fish, Jeremy Kun and Ádám D. Lelkes
4. *Model Theory and Machine Learning*, Hunter Chase and James Freitag
5. *A Complete and Tight Average-Case Analysis of Learning Monomials*, Rüdiger Reischuk and Thomas Zeugmann
6. *Combining Labeled and Unlabeled Data with Co-Training*, Avrim Blum and Tom Mitchell
   (There is a related video lecture by Tom Mitchell here. Co-training is mentioned in the second part, but the entire lecture is very nice and relevant.)

2 Multi-Person Projects (1-2 People)

2.1 PAC-Learning, General

1. *Crowdsourced PAC Learning under Classification Noise*, Shelby Heinecke and Lev Reyzin
2. *Learning Conjunctive Concepts in Structural Domains*, David Haussler
3. *Efficient PAC Learning from the Crowd*, Pranjal Awasthi, Avrim Blum, Nika Haghtalab and Yishay Mansour
4. *The Optimal Sample Complexity of PAC Learning*, Steve Hanneke
5. *Improved Algorithms for Collaborative PAC Learning*, Huy Lê Nguyên and Lydia Zakynthinou

2.2 PAC-Learning, Noise

1. *PAC learning with nasty noise*, Nader H. Bshouty, Nadav Eiron and Eyal Kushilevitz

2.3 PAC-Learning, Distribution-Dependent Learning and Nonuniform Learnability

1. *The information-theoretic value of unlabeled data in semi-supervised learning*, Alexander Golovnev, Dávid Pál, and Balázs Szörényi
2. *Exact learning of random DNF over the uniform distribution*, Linda Sellie

2.4 Robustness and Security

1. *Robust Decision Trees Against Adversarial Examples*, Hongge Chen, Huan Zhang, Duane Boning, and Cho-Jui Hsieh
2. On the Hardness of Robust Classification, Pascale Gourdeau, Varun Kanade, Marta Kwiatkowska, and James Worrell
   There is also an extended version of the paper on arxiv, which may contain more details on the proofs.
3. *Revisiting Adversarial Risk*, Arun Sai Suggala, Adarsh Prasad, Vaishnavh Nagarajan, and Pradeep Ravikumar

4. *A Simple Explanation for the Existence of Adversarial Examples with Small Hamming Distance*, Adi Shamir, Itay Safran, Eyal Ronen, and Orr Dunkelman
5. *Evading classifiers in discrete domains with provable optimality guarantees*, Bogdan Kulynych, Jamie Hayes, Nikita Samarin, and Carmela Troncoso
6. *Can Adversarially Robust Learning Leverage Computational Hardness?*, Saeed Mahloujifar and Mohammad Mahmoody
7. *Multi-party Poisoning through Generalized p-Tampering*, Saeed Mahloujifar, Mohammad Mahmoody, and Ameer Mohammed
8. *Theoretical evidence for adversarial robustness through randomization: the case of the Exponential family*, Rafael Pinot, Laurent Meunier, Alexandre Araujo, Hisashi Kashima, Florian Yger, Cédric Gouy-Pailler, and Jamal Atif
9. *Learning and inference in the presence of corrupted inputs*, Uriel Feige, Yishay Mansour, and Robert E. Schapire
10. *Adversarial Online Learning with noise*, Alon Resler and Yishay Mansour

2.5 Active Learning

1. Improving Generalization with Active Learning, David Cohn, Les Atlas and Richard Ladner
2. Efficient Semi-supervised and Active Learning of Disjunctions, Maria-Florina Balcan, Christopher Berlind, Steven Ehrlich and Yingyu Liang
3. Teaching Dimension and the Complexity of Active Learning, Steve Hanneke

2.6 Query Learning

1. *Learning with Queries but Incomplete Information*, Robert H. Sloan and György Turán
2. Malicious Omissions and Errors in Answers to Membership Queries, Dana Angluin, Martins Krikis, Robert H. Sloan and György Turán
3. Learning Conjunctions of Horn Clauses, Dana Angluin, Michael Frazier and Leonard Pitt
4. Learning Monotone DNF with an Incomplete Membership Oracle, Dana Angluin and Donna K. Slonim

2.7 Teaching

1. *Learning Using Privileged Information: Similarity Control and Knowledge Transfer*, Vladimir Vapnik and Rauf Izmailov
2. *On the Complexity of Teaching*, Sally A. Goldman and Michael J. Kearns
3. Generalized Teaching Dimensions and the Query Complexity of Learning, Tibor Hegedüs

2.8 Analogical Reasoning

1. Analogical Dissimilarity: Definition, Algorithms and Two Experiments in Machine Learning, Laurent Miclet, Sabri Bayoudh, and Arnaud Delhay

2.9 Target Changes During Learning

1. Learning Switching Concepts, Avrim Blum and Prasad Chalasani
2. *Evolution with Drifting Targets*, Varun Kanade, Leslie G. Valiant and Jennifer Wortman Vaughan

2.10 Local Search Methods

1. Running Time Analysis of the (1+1)-EA for OneMax and LeadingOnes under Bit-wise Noise, Chao Qian, Chao Bian, Wu Jiang, Ke Tang
   (There is an extended version with potentially more information on arXiv.)

2.11 Helpful

1. *The Geometry of Generalized Binary Search*, Robert D. Nowak
2. *Mick Gets Some (the Odds Are on His Side)*, Vasek Chvátal and Bruce A. Reed

2.12 Various

1. *Never-Ending Learning*, Tom M. Mitchell, William W. Cohen, Estevam R. Hruschka Jr., Partha P. Talukdar, Bo Yang, Justin Betteridge, Andrew Carlson, Bhavana Dalvi Mishra, Matt Gardner, Bryan Kisiel, Jayant Krishnamurthy, Ni Lao, Kathryn Mazaitis, Thahir Mohamed, Ndapandula Nakashole, Emmanouil A. Platanios, Alan Ritter, Mehdi Samadi, Burr Settles, Richard C. Wang, Derry Wijaya, Abhinav Gupta, Xinlei Chen, Abulhair Saparov, Malcolm Greaves, and Joel Welling
2. *Solving the multiple instance problem with axis-parallel rectangles*, Thomas G. Dietterich, Richard H. Lathrop, and Tomás Lozano-Pérez
3. *Fuzzy Decision Trees: Issues and Methods*, Cezary Z. Janikow
4. Exploring learnability between exact and PAC, Nader H. Bshouty, Jeffrey C. Jackson and Christino Tamon
5. Learning with Attribute Costs, Haim Kaplan, Eyal Kushilevitz and Yishay Mansour
6. *Learning by Distances*, Shai Ben-David, Alon Itai, and Eyal Kushilevitz
7. *Partial observability and learnability*, Loizos Michael
8. *Distributed Learning, Communication Complexity and Privacy*, Maria-Florina Balcan, Avrim Blum, Shai Fine, and Yishay Mansour
9. *Property Testing and Its Connection to Learning and Approximation*, Oded Goldreich, Shafi Goldwasser, and Dana Ron
   (Only up to and including Section 4; that is, Part I of the paper in the first 23 pages.)