Syllabus

Fall 2021

1 Course Goals

The goal of this course is for students to become comfortable tackling mathematically rigorous research papers to learn new material and techniques, within the general area of machine learning. Students in this course will read, present, and discuss both recent and classic papers from machine learning, and related areas in algorithms and computer science theory. This course is aimed at current and potential future graduate students who want to gain technical depth and perspective on the fields of machine learning or CS theory.

The course has no formal prerequisites, though mathematical maturity and previous exposure to computer science proofs at the level of CS48300 or 58000 is strongly recommended.

2 Assignments and Grading

Each class, a pair of students—in consultation with the professor—prepares to present a paper and lead class discussion. Depending on the number of students enrolled in the course, each student will present one, two, or three times over the course of the semester. Presenting a paper requires mastering the line of reasoning leading to the main result(s) of the paper, along with being able to motivate and explain the intuition behind each step. In-class presentations comprise 45% of the grade for this course.

In addition, three times over the course of the semester each student must write a two-page "textbook summary" of the previous week's paper, and this paper must *not* be a paper the student presented to the class. The textbook summary is an opportunity for students to reinterpret complicated ideas and thereby demonstrate a sophisticated understanding of the concepts. This is also an opportunity for students to improve their technical writing skills. The three textbook summaries will comprise 45% of the grade for this course

Finally, this is a discussion-driven class, with discussion both in class and via the online Piazza forum. Each week students must post to Piazza 1) A comment on the previous week's paper (subject to weekly guidelines), and 2) A question about the upcoming paper. These forms of class participation form the remaining 10% of the grade for this course.

3 Preparing and Presenting a Paper

When you sign up (with a partner) to lead a class discussion, reread the following information carefully.

The components of a good presentation (and the basis for your presentation grade) are as follows:

- Introduce and motivate the topic, via a combination of 1) explaining why the results are interesting and/or 2) explaining why the tools used to yield the results are interesting. When appropriate, you should relate the results and tools to other papers in the course.

- Explain, rigorously yet intuitively, one of the main results of the paper. This part should take most of the time, and may require assembling several prerequisite results.

Intuition: Since one of the main purposes of this class is to understand techniques used in research papers, it is crucial that you can intuitively motivate everything you present, including being able to answer questions like "why is this technique the natural technique to try here and how would I recognize situations where this technique is appropriate in my own research?"

Rigor: While you may not have time to give rigorous details of all the steps behind the main result of the paper, your presentation should convey the essential logical structure of the result, where any gaps in the logic are cleanly delineated so that the audience understands why the missing pieces are not important to the overall strategy of the paper. Be prepared to go into details if requested.

- Audience engagement is a crucial component of a good presentation. Make sure your audience is following along; audience inattention may be a sign that you lost the audience by breezing through a crucial step. Keep in mind the crucial takeaway messages you want your audience to remember, and try to convey them in a memorable way, with examples, or via audience interaction.

- Time management is important: keep your eye on the target of your presentation, and on the clock, to make sure you reach the punchline before the end of class.

Preparation and preliminary feedback: To prepare for your presentation, sign up for a 2-hour meeting with the instructor during the week before your presentation. At this meeting, you will go through a practice run of your presentation. Make sure you have a solid grasp of the paper going into this practice run. The meeting will focus on filling in any gaps in your understanding of the intuition of the paper, along with direction about which parts of the paper to emphasize in your final presentation to the class. This is an opportunity to discuss presentation strategies and in general get preliminary feedback about your presentation so you can give a great presentation in class. In rare cases where the presentation must be significantly revised, followup meetings may be necessary.

4 Administrative Information

Instructor: Paul Valiant; `pvaliant@purdue.edu`

Attendance: Because this course is a discussion seminar centered around challenging papers, there is no substitute for attendance, and no feasible way to make up for missing class. However, because this is a graduate seminar, attendance will not be taken, and the course will rely on each student making mature decisions about the use of their time.

Deadlines and Late Policy: The "textbook summaries" will be due Sunday at 6pm, following the Wednesday presentation of each paper. Each student must turn in one summary during each 4-week third of the course, choosing 3 papers that fit their tastes or their schedule (though with 1 fewer summary required for each presentation you do after the 1st). These textbook summaries will be emailed out to the class over the following week to prompt followup discussion in class the

following Wednesday. Because these textbook summaries fuel class discussion, they will not be accepted late; however, students are free to choose which weeks in which to write summaries; so instead of asking for an extension during a busy week, students should instead write a summary another week.

5 Reading List

Papers will be tweaked in response to student interest and class discussions:

Metapapers:

- `http://blizzard.cs.uwaterloo.ca/keshav/home/Papers/data/07/paper-reading.pdf`, "*How to Read a Paper*", S. Keshav
- `https://cs.stanford.edu/~rishig/courses/ref/paper-reading-overview.pdf`, "*Reading in Algorithms: Focus Questions*", T. Roughgarden
- `http://cs.stanford.edu/~rishig/courses/ref/paper-reading-technical.pdf`, "*Paper-Reading Survival Kit*", T. Roughgarden

Class papers:

- `http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf`, "*A Mathematical Theory of Communication*", by Claude Shannon, 1948
- `https://people.mpi-inf.mpg.de/~mehlhorn/SeminarEvolvability/ValiantLearnable.pdf`, "*A Theory of the Learnable*, by Leslie Valiant, in CACM, 1984
- `http://www2.denizyuret.com/ref/blumer/ft_gateway.cfm.pdf`, "*Learnability and the Vapnik-Chervonenkis Dimension*", by Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred Warmuth, STOC1986
- `http://www.face-rec.org/algorithms/Boosting-Ensemble/decision-theoretic_generalization.pdf`, "*A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*", by Yoav Freund and Robert E. Schapire, in JCSS, 1997
- `https://cseweb.ucsd.edu/~yfreund/papers/LargeMarginsUsingPerceptron.pdf`, "*Large Margin Classification Using the Perceptron Algorithm*", by Yoav Freund and Robert E. Schapire, in Machine Learning, 1999
- `https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-pdf`, "*ImageNet Classification with Deep Convolutional Neural Networks*", by Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, NeurIPS 2012; and `https://arxiv.org/pdf/1412.6980.pdf`, "*ADAM: A Method for Stochastic Optimization*", by Diederik P. Kingma, and Jimmy Lei Ba, ICLR2015
- sections 1+2 of `https://arxiv.org/pdf/1204.5721.pdf` "*Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems*", by Sébastien Bubeck and Nicolò Cesa-Bianchi, in Foundations and Trends in Machine Learning, 2012

- `https://www.cs.purdue.edu/homes/pvaliant/RCWC.pdf`, "*Worst-Case Analysis of Randomly Collected Data*", by Justin Y. Chen, Gregory Valiant, and Paul Valiant, NeurIPS 2020.
- `https://arxiv.org/pdf/1712.09203.pdf`, "*Algorithmic Regularization in Over-parameterized Matrix Sensing and Neural Networks with Quadratic Activations*", by Yuanzhi Li, Tengyu Ma, and Hongyang Zhang, published in COLT2018.
- `https://dl.acm.org/doi/10.1145/3186563`, "*Fast Learning Requires Good Memory: A Time-Space Lower Bound for Parity Learning*", by Ran Raz, published in FOCS2016 and then in the JACM in 2018.
- `https://arxiv.org/pdf/1802.05296.pdf`, "*Stronger generalization bounds for deep nets via a compression approach*", by Sanjeev Arora, Rong Ge, Behnam Neyshabur, and Yi Zhang, published in ICML2018.
- `https://arxiv.org/pdf/1806.07572.pdf` "*Neural Tangent Kernel: Convergence and Generalization in Neural Networks*", by Arthur Jacot, Franck Gabriel, and Clement Hongler, published in NeurIPS2018.

6 Policies

Nondiscrimination and diversity: I am committed to making this class a positive experience for all students, and expect all students to help create a communal learning environment that respects and supports each student. Please do not hesitate to contact me if your experience in class does not live up to these goals. In particular, I take demeaning/discouraging comments, and sexual harassment very seriously, whether in or out of the classroom.

Accessibility: Please see the Purdue Disability Resource Center (`https://www.purdue.edu/drc/`) for information on academic accommodations. If course materials are not available in a form that is accessible to you, or you anticipate that this might be the case, please let me know at the start of the semester so that we can arrange a plan to accommodate your needs, in conjunction with the Disability Resource Center.

Academic Integrity: Students in this class will be held to the professional standards of academic computer science. This has two main components: A) Collaboration is expected, and encouraged, and you should feel free to talk to anyone about the concepts in the course; B) However, the actual words you submit in any written assignment **must** be your own (or be a cited quotation) - you should not have any other documents visible to you when you write since this may lead to suspiciously similar text that may appear plagiarized. Further, as part of a healthy academic environment, you should give intellectual credit for ideas that are not your own, and as much as possible, try to give credit to the *originator* of the idea. Finally, there is no tolerance for academic dishonesty or deception.