Course Information

CS 59300ADS – Algorithms for Data Science
Time/Location: TTH 9:00 am at SCHM 116
Instructor: Ruizhe Zhang (`rzzhang@purdue.edu`)
Office hours: By appointment
Prerequisites: Strong mathematical maturity and proficiency with proofs, probability, and linear algebra are required. Familiarity with one of the following at the graduate level will be helpful: CS theory, statistical machine learning, quantum computing, or quantum information.
Webpage: TBD

1 Course Description

This is a graduate-level course that explores the classical and quantum algorithmic foundations of machine learning and data science, with a focus on techniques that provide provable guarantees. The course is divided into two parts. The first half introduces a number of classical techniques for algorithm design, including tensor methods, Sum-of-Squares (SoS), spectral analysis, MCMC, and diffusion. The second half presents a self-contained introduction to quantum algorithms for optimization and machine learning, including quantum linear algebra, quantum sampling, and learning from quantum data. Throughout the course, we will cover applications across a range of domains, including but not limited to robust statistics, generative modeling, statistical and quantum physics. We aim to provide a unified perspective on how advanced algorithmic techniques—both classical and quantum—can be applied to address key challenges in high-dimensional data analysis.

2 Learning Outcomes

By the end of the semester, you should be able to:

1. Understand and analyze several modern classical algorithmic frameworks;

2. Be familiar with current literature on quantum algorithms for optimization and machine learning;

3. Formulate and investigate original research problems in classical or quantum algorithms for machine learning and data science.

3 Course Materials

We will not be following any particular textbook, but the following resources could be helpful for or complementary to various parts of the course:

- Ankur Moitra. `https://people.csail.mit.edu/moitra/cs294.html`. Algorithmic Aspects of Machine Learning.
- Sitan Chen. `https://www.sitanchen.com/cs224/f24/index.html`. Algorithms for Data Science.
- Tselil Schramm. `https://tselilschramm.org/sos-paradigm/sos-paradigm.html`. The Sum-of-Squares Algorithmic Paradigm in Statistics.
- Kevin Tian. `https://kjtian.github.io/cs395t_sp25.html`. Continuous Algorithms.
- Kuikui Liu. `https://kuikuiliu.github.io/teaching/au2023`. Algorithmic Counting and Sampling.
- Lin Lin. `https://math.berkeley.edu/~linlin/math275`. Topics in Applied Mathematics: Quantum algorithms for scientific computation.
- IPAM Tutorial in Fall 2023. Presentations and slides.
- Andrew Childs. `https://www.cs.umd.edu/~amchilds/qa/qa.pdf`. Lecture Notes on Quantum Algorithms.
- Alessandro Luongo. `https://quantumalgorithms.org/`. Quantum algorithms for data analysis.
- Quantum Algorithm Zoo. `https://quantumalgorithmzoo.org/`.

4 Grading

The grading breakdown is as follows:

- Problem sets (15%)
- Participation (10%)
- Scribing (10%)
- Take-home midterm exam (25%)
- Final project (40%)

Problem sets There will be three to four problem sets. All submissions must be in PDF format. You are allowed to discuss problems with other students, but the final submission must be written individually.

Participation In-personal attendance at lectures is **mandatory** to receive full participation credit. If you have any unavoidable conflict (illness, travel, etc.), you should try to let me know in advance whenever possible.

Scribing Each lecture will have one or two scribes who will be responsible for taking notes and writing them up in LaTeX. The deadline for submitting the lecture note is one week after the lecture.

Take-home Midterm Exam The midterm exam will be a take-home exam. You are allowed to reference your lecture notes and problem sets. However, the use of the Internet or phones is prohibited. Discussions with other people are also not allowed.

Final project Students will work individually or in pairs to complete a written final project and an oral presentation of it. The final project should: (1) focus on a research frontier in at least one of the topics discussed in class; (2) include a thorough literature review; (3) have a strong technical component, either in the form of original research or in the form of insightful exposition of existing work.

A list of suggested topics and papers will be provided; however, students are encouraged to propose their own topics, subject to approval by the instructor. We expect each student to come to office hours at least once to discuss their project ideas.

5 Course Schedule

Below is a preliminary schedule. It will be updated as we progress through the course.

- Week 1: *Tensor methods.*
  Introduction to tensors. Tensor decomposition algorithms (Jennrich's algorithm, iterative methods). Applications (Gaussian mixtures, super-resolution).

- Week 2: *Super-resolution.*
  Introduction to spectral estimation. Subspace-based methods (Prony's method, ESPRIT). Sparse Fourier transform.

- Week 3: *Sum-of-Squares.*
  Sum-of-squares proofs, pseudo-distributions. Applications to robust statistics.

- Week 4: *Semi-definite programmings (SDPs).*
  Primal-dual method. Interior-point method. Applications of SDPs.

- Week 5: *Markov chain Monte Carlo*.
  Markov chains fundamentals. Basic methods to prove the mixing time. Advanced techniques (spectral independence). Non-local Markov chains and the Swendsen-Wang algorithm.

- Week 6: *Diffusion models.*
  Langevin diffusion and stochastic differential equations. Girsanov's theorem and discretization analysis.

- Week 7 (October break)

- Week 8: *Welcome to the quantum world.*
  Basic notations for quantum algorithms (qubits, quantum circuits, quantum unitaries, quantum gates, measurements).

- Week 9: *Quantum eigenvalue problem.*
  Early fault-tolerant quantum algorithms. Quantum phase estimation.

- Weeks 10 - 11: *Quantum linear algebra toolkit.*
  Block-encoding. Linear combinations of unitaries (LCU). Quantum singular value transformation (QSVT) and applications (quantum linear system solvers, quantum SDP solvers).

- Week 12: *Quantum sampling.*
  Quantum walk and applications. Quantum Gibbs sampling and open quantum system dynamics.

- Week 13: *Quantum learning theory.*
  Shadow tomography. Hamiltonian learning.

- Week 14 (Thanksgiving)

- Week 15: *Project presentation.*

- Week 16: *Project presentation.*

6 Course Policies[1]

Late Policy Problem sets may be submitted up to 3 days late, with a 10% deduction in credit. Late submissions of the final project report will not be accepted under any circumstances.

[1] Most of the following material is based on the policy followed by Prof. Samuel Wagstaff in his cryptography courses.

Academic Dishonesty Purdue prohibits “dishonesty in connection with any University activity. Cheating, plagiarism, or knowingly furnishing false information to the University are examples of dishonesty.” Furthermore, the University Senate has stipulated that “he commitment of acts of cheating, lying, and deceit in any of their diverse forms (such as the use of substitutes for taking examinations, the use of illegal cribs, plagiarism, and copying during exams) is dishonest and must not be tolerated. Moreover, knowingly to aid and abet, directly or indirectly, other parties in committing dishonest acts is in itself dishonest.” While it is all right to discuss homework in this class with other students in general terms, do not copy another student’s homework or let anyone copy your homework. These discussions should be appropriately acknowledged. When two identical homeworks are found, both are sent to the Dean of Students to determine who copied from whom. The use of generative AI for assignments is strongly discouraged. However, if a student chooses to use it, they must submit a complete transcript of all prompts and ensure that the final solutions are written entirely in their own words. Penalties for academic dishonesty range from a 0 grade on one assignment to a failing grade in the class and even expulsion from the University. A hearing at the Dean of Students’ Office can ruin your whole day.

Violent Behavior Policy Purdue University is committed to providing a safe and secure campus environment for members of the university community. Purdue strives to create an educational environment for students and a work environment for employees that promote educational and career goals. Violent behavior impedes such goals. Therefore, violent behavior is prohibited in or on any University facility or while participating in any university activity.

Students with Disabilities Purdue University is required to respond to the needs of the students with disabilities as outlined in both the Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990 through the provision of auxiliary aids and services that allow a student with a disability to fully access and participate in the programs, services, and activities at Purdue University. The ODOS Testing Center is an excellent place for students with disabilities to take exams for this class. Please tell the instructor at least one week in advance if you wish to take the exams there. If you have a disability that requires other special accommodation, please tell the instructor early in the semester. It is the student’s responsibility to notify the Disability Resource Center of an impairment/condition that may require accommodations and/or classroom modifications.

Emergencies In the event of a major campus emergency (such as a tornado, earthquake, flu epidemic or terrorist attack), course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances beyond the instructor’s control. Relevant changes to this course will be posted onto the course website or can be obtained by contacting the instructor via email or phone. You should read your Purdue email frequently.

Nondiscrimination Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life. Purdue University prohibits discrimination against any member of the University community on the basis of race, religion, color, sex, age, national origin or ancestry, marital status, parental status, sexual orientation, disability or status as a veteran. The University will conduct its programs, services and activities consistent with applicable federal, state and local laws, regulations and orders and in conformance with the procedures and limitations as set forth in Executive Memorandum No. D-1, which provides specific contractual rights and remedies. The instructor agrees completely with all Purdue policies mentioned in this document.

Privacy The Federal Educational Records Privacy Act (FERPA) protects information about students, such as grades. If you apply for a job and wish to use the instructor as a reference, you should tell the instructor beforehand. Otherwise, the instructor cannot say anything about you to a prospective employer who might call. The instructor is happy to provide references and to write letters of recommendation for his students as needed.

Disclaimer.

This syllabus and grading scheme are subject to change.