CSCI 1952Q: Algorithmic Aspects of Machine Learning (Spring 2024)

Final Project Guidelines

Objectives:

1. **Topics.** Your project must be related to machine learning or theoretical computer science (preferably related to the content of the course but not required).
2. **In-Class Presentation.** Give a 15-minute presentation on your project (12-minute talk and 3-minute Q&A).
3. **Project Report.** Write a research paper/project report of at least 8 pages.

Teams: Each project team should have 1-2 students. All students in the same team will receive the same grade for the final project.

Submission Instructions:

- Submit your project report as a single `.pdf` file via Gradescope. You can include links to other materials (e.g., website, code) in your report.
- Project reports must be written in LaTeX. It is recommended (but not required) that you use style files from conferences such as ICLR, ICML, or NeurIPS.
- The main body of the project report must be **exactly** 8 pages, and there is no limit on the number of pages for references and appendices. Material beyond the first 8 pages will be read at the discretion of the reviewers.
- If any part of your project was done before taking the course or was used (or will be used) as a project for another course, you must clearly state this in your report. All project reports will be shared with the class and possibly future students.

Timeline.

- **Mar 4:** Decide on teams and project titles.
- **Apr 18 – Apr 25:** Present your project in class.
- **May 13:** Submit your project report.
- **May 17:** Review and submit ratings for all projects (optional).

  All deadlines are 11:59pm on the due date.

Grading. Projects will be graded out of 46 points (with up to 9 bonus points):

- *Meeting the deadlines* (6 points):
  - (3 points) Submit teams and project titles by Mar 1.
  - (3 points) All team members must show up on time to give an in-person presentation.
- *Project presentation and report* (40 points):

  We will run a lightweight peer-review process to grade your final projects. We will ask each student to review and rate all projects.

  The presentation and the report each count for 20 points. The following criteria are provided as guidelines. You can discuss/post your criteria on Ed.

  - Is the project relevant to machine learning or theoretical computer science?
  - How original and creative is the idea and the execution of the project?
  - Does the project show a deep understanding of related research areas and topics?
  - How useful is the project to students and researchers in the ML/TCS community?
  - Does the report clearly state the authors' main contributions?
  - Does the report discuss the challenges incurred and highlight the novelty of the results?
  - Is the report well-structured, well-written, and technically sound?
  - Does the report discuss existing literature and provide a comparison with related work?
  - Does the presentation give a good overview of the project's motivation and main results?
  - Does the presentation make the audience want to learn more about the project?
  - Are the speakers well-prepared? Is the presentation delivered at an appropriate pace?
  - Are the slides well-organized and effective?

  In the case of building on existing research projects, the evaluation process should focus on the new components that are developed while taking the course.

- *Peer review* (up to 9 bonus points):
  - Each student can submit a score (22, 28, 34, or 40 points) or declare conflicts of interest for each project. The instructor and all TAs will participate as reviewers (with twice the weight).
  - Scores will be normalized to have the same mean and variance across reviewers. The final score of a project is the average of the normalized scores of all reviewers.
  - (Top projects) The top 20% of projects will receive 3 bonus points. The top 10% of projects will receive 3 additional bonus points (so 6 points in total).
  - (Top reviewers) If your evaluation is among the top 20% most accurate evaluations (measured by ℓ_2-norm of the difference), you will receive 3 bonus points.
  - Names, titles, and reports of the top projects and reviewers will be announced on the public course homepage.

Examples of project ideas. One possible approach is to (1) choose a machine learning problem that excites you, (2) implement state-of-the-art or widely used algorithms for that problem, (3) explore when existing algorithms perform poorly and how they can be improved, and (4) design and analyze new algorithms for this problem.

We provide examples of project topics from previous years to illustrate the variety and possibilities of what you can do. You are highly encouraged to choose a topic that excites you the most, which may not necessarily come from this list.

- A Universal Black-Box Attack on Sentiment Classifiers
- Adversarial Attacks on Image Classifiers
- Column Subset Selection for Entrywise L1-norm Loss
- Comparison of Different Matrix Completion Algorithms
- Comparison of Nonnegative Matrix Factorization Algorithms
- Conditional Generators and Data Hybridization
- Contrastive Learning on Graph Representation Learning
- Convergence of Relativistic GANs
- Different Architectures for Deep Counterfactual Regret Minimization
- Disentangling Casual Mechanisms by Obstructing Classifiers
- Evaluating Random Asymmetric Initialization
- First-Person Video Long-term Action Anticipation
- Gamma-Net for Data-Efficient Image Contour Detection
- Gene Co-expression Estimation for Single-Cell Data
- High Dimensional Density Estimation using Neural Networks
- Identifying Deepfakes
- Low Latency Streaming Speech Selection
- Minimizer Space K-Shingles
- Nonnegative Matrix Factorization for Community Detection
- Nonnegative Matrix Factorization for Drug Discovery
- Predicting the Origin of a DNA Sequence
- Predictive Coding Networks with Free Energy
- Reinforcement Learning Control in Chaotic Environments
- Tensor Decompositions for Dynamic Neural Radiance Fields
- The Strava Art Problem