CSE 6740 A/ISyE 6740, Fall 2023, Georgia Tech

CSE 6740: Project Guidelines

1-page proposal due Nov 2nd, '23 (11:59 pm ET)

- You are free to delve deeper into any subset of topics covered in class, or more broadly, into mathematical machine learning and statistical methods. You are also welcome to find topics that are at the intersection of your research with the course material.
- The project will be evaluated on its computational and theoretical components. You can maximize your score by showing substantial work (2-3 homeworks worth) in the algorithmic/computational aspects or the theoretical aspects or both.
- A self-contained final report of no more than 6 pages is required at the end of term. This carries 65% of the final project score.
- A 5-minute presentation briefly summarizing project components is also required. This carries 25% of the final project score. Sign-up sheets for this have been circulated on Piazza.
- In the 1-page proposal (10% of the final project score), please outline the problem motivation (what problem are you trying to solve or understand), your proposed solution approach, and your inferences (what you hope to learn).
- For maximum points, the proposal should lay out the learning problem you plan to solve in precise mathematical terms. It should briefly explain the proposed solution.
- The final report will be evaluated based on your formulation/modeling of the problem in mathematical terms, your solution strategy and your discussion on your choice of solution strategy. You will not be evaluated on how well your solution works; only how well you have understood your problem and solution method, and how clearly your justify (using statistical methods and theory taught in class) your results.
- Examples of computation-heavy project ideas: statistical learning or sampling problems on datasets from genomics, finance, or climate sciences.
- Examples of theory-heavy project ideas: analyzing the sample complexity and computational complexity of generative modeling algorithms, understanding generalization of supervised and unsupervised learning methods, margin-based generalization or any other topic that falls within classical statistical and data-driven methods.

- You can work individually or in pairs. Only one copy of the deliverables (proposal, report and presentation) is required per team.
- We will read your proposal and provide feedback and suggestions.