Boston University CS 591 S1
Privacy in Machine Learning and Statistical Inference
Fall 2018

How can we learn from a data set of sensitive information while providing meaningful privacy to the individuals whose information it contains? The course explores this question, starting from the problems faced by straightforward solutions and moving on to state-of-the-art rigorous notions, such as differential privacy. Along the way we will see connections to machine learning, game theory, statistics, and optimization.

The class will cover theoretical foundations (involving algorithms, mathematical definitions, and proofs), implementation (involving coding in Python or R), and discussion of the ethical, legal, and social aspects of privacy (involving reading essays and articles from several fields, and writing critical evaluations).

The course is aimed at graduate and advanced undergraduate students in CS, mathematics, statistics, and electrical engineering. Students from other departments are welcome.

Time & Place:
Mondays and Wednesdays, 2:30-3:45 PM
Room MCS B25 (111 Cummington Mall, basement level)

- First lecture: Wednesday, September 5, 2018
- Holidays with no lecture: Mon, October 8; Wed, November 21.
- Unusual lecture: Tuesday, October 9 (BU's logical Monday)

Instructor:
Adam Smith
Office Hours: MCS 135F, M-Tu-W 4-5pm
Email: ads22@bu.edu (but please send course-related questions via Piazza)

Piazza site: http://piazza.com/bu/fall2018/cs591s1 (access code **axani**)

Lecture Schedule with readings:
https://docs.google.com/document/d/1jsZLEd3ZM-ZWdNAjNRI4_bgPysRUsKQDHvy4VKgtzJ8/edit?usp=sharing

Background survey (please fill by September 9)**:** Here

Syllabus: The following gives an idea of the relative balance of topics in the course. The exact schedule will be determined as the course proceeds.

- What does “privacy” mean in learning and statistics? (1 week)
- Attacks on statistical data privacy (2 weeks)
- Defining privacy: differential privacy and its variants (2 weeks)
- Achieving privacy: algorithmic tools for differential privacy (5 weeks)
- Ethical frameworks relating to privacy (1.5 weeks)
- Legal frameworks relating to privacy (1.5 weeks)
- Advanced topics (1 week)

Prerequisites: Students should have a solid grounding in probability and statistics, linear algebra, vector calculus, and algorithms. Students should be comfortable reading and writing mathematical proofs, as well as programming. Specifically, CS591 requires courses equivalent to CAS CS 132 (discrete mathematics), CAS CS 237 (probability and statistics), CAS CS 330 (algorithms), CAS MA 225 (multivariate calculus), or consent of instructor.

Auditors are welcome. In particular, students from other universities are welcome to attend and participate in discussions. I ask that you add yourself to the class Piazza page where course information will be posted. (Non-BU folks have to use the access code “axani”.)
http://piazza.com/bu/fall2018/cs591s1
You’ll get more out of the class if you also do the readings and think about the homework problems.

Coursework and grading: The grade will be based on a course project (50%), written homework (30%), responses to required reading (10%) and class/Piazza participation (10%). There will be no final exam.

Students are expected to attend classes and do all required readings, which will be provided at least 5 days in advance, and fill any reading prompts when they are due (generally the day before lecture). Late homework will not be accepted, except by agreement with the instructor. Make a request on Piazza, at least 48 hours ahead of the due date.

Piazza is a website that allows you to ask questions, either to instructors or course-wide. *We will be using Piazza for almost all course communication outside of the classroom*. Please sign up, and set appropriate email notification options so that you make sure to receive announcements.
http://piazza.com/bu/fall2018/cs591s1
Piazza allows you to ask questions that are visible only to instructors, but it also allows you to ask questions to the entire class, and answer others’ questions. When someone posts a question on Piazza, if you know the answer, please go ahead and post it. However please do not provide answers to homework questions on Piazza. It’s OK to tell people where to look to

get answers, or to correct mistakes; just don't provide actual solutions to homeworks. Also, be polite. See the post "Ethics and Etiquette on Piazza" for detail about our expectations.

Collaboration and academic conduct: You may discuss homework assignments and projects with classmates, but you are solely responsible for what you turn in. Collaboration in the form of discussion is allowed, but all forms of cheating (copying parts of a classmate's assignment, plagiarism from papers or old posted solutions) are NOT allowed. A rough rule of thumb: you should be able to walk away from a discussion of a homework problem with no notes at all and write your solution on your own. Collaboration on any quizzes or exams is forbidden. All of us—staff and students—are expected to adhere to BU's academic code of conduct (and its graduate version). Violations of these codes will be dealt with according to university policy.

Textbook: There is no official textbook for the class. A good reference on (much of) the theoretical material is the free tutorial monograph,

- C. Dwork and A. Roth: The Algorithmic Foundations of Differential Privacy, 2014.

Project: There will be a separate handout with detailed project guidelines.
See this document for details:
https://docs.google.com/document/d/1NCwjiilVQVIyVyRtIOgNqyTxquASh8JNsxfcdn9a078/edit?usp=sharing

~~The key dates are~~

- ~~October 21: Project proposal~~
- ~~November 15: Progress report~~
- ~~December 1: Initial project report submission~~
- ~~December 2: Feedback phase starts~~
- ~~December 8: Feedback to other students due~~
- ~~December 12: Poster session with snacks~~
- ~~December 17: Final report submission~~

Gradescope will be used for submitting homework, project reports, and responses to reading prompts. Instructions on how to sign up will be sent later in the semester.