EE __: PRIVACY PRESERVING MACHINE LEARNING

Number of credits: 4 credits

Quarter and year: Spring 2023

Instructor's name: Tamara Bonaci

Email Address: tbonaci@uw.edu

EEP 596: Privacy Preserving Machine Learning (4cr)

CATALOG COURSE DESCRIPTION

Focuses on the theoretical and applied aspects of Privacy Preserving Machine Learning. Considers statistical and information-theoretic notion of privacy and privacy attacks against machine learning models. Covers prevention and mitigation of privacy attacks, including multi-party secure computation (MPC), differential privacy (DP), federated learning, robust federated learning, and split learning.

SUMMARY

Driven by recent progress in data science, machine learning and AI, our data is being collected at an ever-increasing pace. Collectors, aggregators, and processors of such data are various private for-profit and nonprofit entities, as well as public organizations. While machine learning models, trained on our data, can be beneficial to us personally, as well as to societies at large, they can also lead to a slew of undesirable, negative, and occasionally catastrophic privacy incidents.

We therefore must balance out two conflicting objectives: the need to maximize accuracy, utility and efficiency of the used machine learning models, while at the same time protecting the privacy of the data used for and by those models, as well as the privacy of the models themselves.

Privacy Preserving Machine Learning (PPML) is an important and very active research area that focuses on this question exactly – how to benefit from machine learning techniques while preserving the privacy of training data and learned models. In this course we will explore the variety of topics related to privacy preserving machine learning, focusing on theoretical and applied aspects of PPML. We will start by considering statistical and information-theoretic notion of privacy. We will then consider privacy attacks against machine learning models. From there, we will examine a variety of topics focused on preventing and mitigating such privacy attacks, including multi-party secure computation (MPC), differential privacy (DP), federated learning, robust federated learning, and split learning.

COURSE OBJECTIVES

By the end of this course, students will demonstrate the ability to:

- Understand statistical and information-theoretic notion of privacy.
- Develop an understanding of privacy attacks against machine learning models.
- Utilize privacy engineering techniques, including secure multiparty computation (MPC), and differential privacy.
- Understand privacy preserving approaches, including federated learning and split learning.
- Gain familiarity with federated learning in adversarial attacks, privacy attacks against machine learning systems, and privacy engineering techniques, including secure multi-party computation (MPC), and differential privacy, and their applications to machine learning systems.
- Utilize state-of-the-art Python libraries and tools used for PPML, such as PySyft.

BACKGROUND KNOWLEDGE

This course assumes a basic familiarity with stochastic mathematics (notion of a random variable and a random process, expectation, variance, independence, Markov process).

TEXTBOOK

There is no required textbook for this course, but some recommended books that you might want to consider include:

• Jin Li, Ping Li, Zheli Liu, Xiaofeng Chen, Tong Li – Privacy Preserving Machine Learning, Springer Briefs on Cyber Security Systems and Networks, Springer 2022

• Muhammad Habib ur Rehman, Mohamed Mdhat Gaber – Federated Learning Systems, Towards Next-Generation AI, Springer, 2021.

COURSE PROGRESSION

Week 1: Lecture 1: Course overview. Why Machine Learning Needs Privacy-Preserving Manner?

Week 2: Lecture 2: Crash Course to Machine Learning

Week 3: Lecture 3: Machine Learning in Adversarial Setting – Privacy Attacks. Statistical and Information-Theoretic Notion of Privacy.

Week 4: Lecture 4: Introduction to Secure Multi-Party Computation (MPC).

Week 5: Lecture 5: MPC and Machine Learning. Introduction to PySyft.

Week 6: Lecture 6: Introduction to Decentralized Privacy-Preserving Machine Learning Algorithms. Federated Learning.

Week 7: Lecture 7: Federated Learning II.

Week 8: Lecture 8: Federated Learning in Adversarial Environment.

Week 9: Lecture 9: Federated Learning in Adversarial Environment II. Differential privacy and federated learning.

Week 10: Lecture 10: Introduction to split learning.

Final exam week: Project presentations

COURSE ELEMENTS AND GRADING CRITERIA

Readings and Discussion:

Readings and discussions of assigned papers are an important part of this course. The goal of this exercise is to work together on understanding broader implications of the foundational privacy-preserving machine learning research. Additionally, we will also try to keep up with the state-of-the-art PPML research. Every week, a single research paper will be assigned, and you will be expected to post to the class discussion board about it. Your post should contain something original beyond what others have already posted. You may consider posting:

• A summary of the paper,

• Evaluation of the paper's strengths and weaknesses,

• Open research question related to the topic of the paper, or

• Question you would like to discuss in class.

All posts are due by 11:59pm PT on a XYZ day, and they will be graded on the scale of 0-2, where:

• 0 means missed or irrelevant post,

• 1 means a relevant post submitted, and

• 2 means a good and interesting post.

Post submitted after the deadline will receive no credit. There will be a total of 11 reading assignments, and we will take 8 best scores when determining your grade.

Labs:

We will have up to four Python-based lab assignments, intended to give you experience with some simple privacy-preserving techniques. The labs will utilize publicly available data sets.

Project:

The final component of this course is a project, and its goal is to give you a deeper understanding of how to think about, and how to solve a real-life problem related to privacy-preserving machine learning. For the project, you will choose a topic related to privacy-preserving machine learning. You can work on the project either individually, or in groups of two. When working in a group, your end-result should reflect the fact that it is a multi-person effort.

Final Grade

Your grade in this course will be based on readings and discussion, homework assignments, and project. The expected grade breakdown is:

• Readings and discussion: 15%

• Labs: 25%

• Project: 60%

COURSE FORMAT

Lectures: lecture slides, relevant reading and additional material will be made available in the following way:

- Lecture notes and relevant reading material will be posted on Canvas, under the appropriate weekly module.
- Pre-recorded video material (if available for the current module) will be posted on Canvas, under the appropriate weekly module.

We will do our best to record lectures, and make it available through Canvas + Panopto.

Important note: if you are unable to attend lectures in person, please reach out to Dr. Bonaci, and I will do my best to accommodate you, and allow you to access lectures remotely.

Classroom recordings: This course, or parts of this course, may be recorded for educational purposes. These recordings will be made available only to students enrolled in the course, instructor of record, and any teaching assistants assigned to the course.

RELIGIOUS ACCOMMODATIONS

Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW's policy, including more information about how to request an accommodation, is available at Religious Accommodations Policy (https://registrar.washington.edu/staffandfaculty/religious-accommodations-policy/). Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request form (https://registrar.washington.edu/students/religious-accommodations-request/).

ACCOMMODATIONS AND ACCESS

If you have already established accommodations with Disability Resources for Students (DRS), please communicate your approved accommodations to me at your earliest convenience so we

can discuss your needs in this course. If you have not yet established services through DRS, but have a temporary health condition or permanent disability that requires accommodations (conditions include but not limited to; mental health, attention-related, learning, vision, hearing, physical or health impacts), you are welcome to contact DRS at 206-543-8924 or uwdrs@uw.edu or disability.uw.edu. DRS offers resources and coordinates reasonable accommodations for students with disabilities and/or temporary health conditions. Reasonable accommodations are established through an interactive process between the student, instructor, and DRS. It is the policy and practice of the University of Washington to create inclusive and accessible learning environments consistent with federal and state law.

ACADEMIC INTEGRITY

Engineering is a profession demanding a high level of personal honesty, integrity and responsibility. Therefore, it is essential that engineering students, in fulfillment of their academic requirements and in preparation to enter the engineering profession, adhere to the College of Engineering Statement of Principles. Any student in this course suspected of academic misconduct (e.g., cheating, plagiarism, or falsification) will be reported to the College of Engineering Dean's Office and the University's Office of Community Standards and Student Conduct to initiate the student conduct process. Any student found to have committed academic misconduct may receive a zero for their grade on the impacted academic work (e.g., assignments, project, or exams), and academic consequences, with the possibility of expulsion

TITLE IX

"UW, through numerous policies, prohibits sex- and gender-based violence and harassment, and we expect students, faculty, and staff to act professionally and respectfully in all work, learning, and research environments. For support, resources, and reporting options related to sex- and gender-based violence or harassment, visit UW Title IX's webpage (https://www.washington.edu/titleix/), specifically the Know Your Rights & Resources guide (https://www.washington.edu/titleix/files/2020/08/KYRR-guide-8-10-2020-LINKED.pdf).

If you choose to disclose information to me about sex- or gender-based violence or harassment, I will connect you (or the person who experienced the conduct) with resources and individuals who can best provide support and options. You can also access those resources directly:

Confidential: Confidential advocates (https://www.washington.edu/sexualassault/support/advocacy/) will not share information with others unless given express permission by the person who has experienced the harm or when required by law.

Private and/or anonymous: SafeCampus (https://www.washington.edu/safecampus/) provides consultation and support and can connect you with additional resources if you want them. You can contact SafeCampus anonymously or share limited information when you call

Please note that some senior leaders and other specified employees have been identified as “Officials Required to Report.” (https://www.washington.edu/titleix/title-ix-officials-required-to-report/) If an Official Required to Report learns of possible sex- or gender-based violence or harassment, they are required to call SafeCampus and report all the details they have in order to ensure that the person who experienced harm is offered support and reporting options (https://www.washington.edu/titleix/resources/).”