Repository Entry Template
Embedded EthiCS @ Harvard Teaching Lab

Overview

Course:	CS 249r Tiny Machine Learning
Course Level:	Upper-level undergraduate
Course Description:	"Tiny machine learning (TinyML) is defined as a fast-growing field of machine learning technologies and applications including hardware (dedicated integrated circuits), algorithms and software capable of performing on-device sensor (vision, audio, IMU, biomedical, etc.) data analytics at extremely low power, typically in the mW range and below, and hence enabling a variety of always-on use-cases and targeting battery-operated devices. The pervasiveness of ultra-low-power embedded devices, coupled with the introduction of embedded machine learning frameworks like TensorFlow Lite for Microcontrollers, will enable the mass proliferation of AI-powered IoT devices. The explosive growth in machine learning and the ease of use of platforms like TensorFlow (TF) make it an indispensable topic of study for modern computer science and electrical engineering students."[1]
Module Topic:	Privacy in Context
Module Author:	Susan Kennedy
Semesters Taught:	Fall 2020
Tags:	Embedded ML [CS], IoT [CS], privacy [Phil], autonomy [Phil], fairness [Phil]
Module Overview:	In this module, we first consider how the features of TinyML may pose unique challenges to privacy. More specifically, we discuss: (1) how the small size and discreet nature of TinyML poses problems for informed consent; and (2) how the possibility of deploying this technology in a wide range of contexts requires us to move beyond viewing privacy in terms of public vs. private spheres. As a framework for approaching these challenges, Nissenbaum's argument for privacy as a right to contextual integrity is introduced. According to NIssenbaum, every context is associated with norms which govern the appropriate flow of information and, consequently, our expectations of privacy. Using this lens of analysis, students think through how new devices or practices might disrupt context-relative informational norms and generate privacy violations. Once students are able to identify a violation of privacy, they are then asked to consider what types of ethical considerations might nevertheless justify this violation. To reinforce this lesson, students practice applying the philosophical concepts they have learned to several real-world examples of ML devices. In closing, students reflect on ways they can apply Privacy is often discussed as something that ought to be preserved when working in computer science. In this module, however, students reflect on the kinds of ethical considerations that may nevertheless justify a violation of privacy. This is significant insofar as students become acquainted with useful tools to help them responsibly navigate a path forward in cases where a cutting-edge technology like TinyML might disrupt or redefine informational norms.

[1] https://www.seas.harvard.edu/computer-science/courses

	what they have learned about privacy to make more informed design choices.	
Connection to Course Material:	In this course, students learn to build embedded devices that utilize edge computing. A section of the course focuses on promising applications of TinyML including keyword spotting for personal voice assistants, visual wake words, and arrhythmia detection. For the Embedded EthiCS module, students evaluate relevantly similar applications of embedded devices (including fall detection systems, health wearables, and personal voice assistants) to determine whether they violate privacy while paying special attention to the context in which these devices might be deployed.	This topic was chosen because TinyML can find applications in a wide range of contexts, including the home, hospitals, environment and industry. Thus, discussing a framework for privacy that pays special attention to the context in which devices are deployed is especially fitting. Moreover, given the ways in which TinyML is thought to mitigate security and privacy concerns (insofar as embedded ML reduces transmissions of data up to the cloud and data is primarily stored on the device as opposed to being warehoused in a singular location), this emerging form of ML offers an important opportunity to discuss privacy concerns beyond the threat of malicious actors gaining inappropriate access to data. Other topics that would be useful to cover for this course are the ethical issues surrounding data collection, mitigating bias in datasets, and optimizing a model for fairness.

Goals		
Module Goals:	1. Understand philosophical arguments for privacy, with a special focus on privacy as a right to contextual integrity. 2. Identify violations of privacy. 3. Evaluate ethical considerations that might justify a violation of privacy. 3. Practice applying these concepts to evaluate real-world case studies of ML devices.	Marginal notes
Key Philosophical Questions:	1. What is context? 2. What are informational norms and how do they govern the appropriate flow of information within a given context? 3. In what way might a shift in context-relative informational norms be said to constitute a violation of privacy? 4. What ethical considerations should we take into account when evaluating the desirability of a new device or practice? 5. When might a violation of privacy be justified or outweighed by other ethical considerations?	These questions are listed in ascending order and reflect the order in which they are introduced in class, starting with relatively simple questions and building up to more complex ones. Questions (1) and (2) are essential for answering (3). Question (4) is essential for answering (5), with the latter being the central question students explore in the module and gain practice answering through the class activity.

Materials		
Key Philosophical Concepts:	● Nissenbaum's argument for privacy as a right to contextual integrity ● Context ● Informational norms ● Autonomy ● Fair distributions of costs and benefits ● Power dynamics (Foucault)	Both Nissenbaum's argument for privacy as a right to contextual integrity and context-relative informational norms help students identify when a new device or practice violates privacy. The remaining philosophical concepts are important ethical considerations that can potentially justify a violation of privacy.
Assigned Readings:	● Michael Zimmer, "How Contextual Integrity Can Help Us with Research Ethics in Pervasive Data" (July 2018) *Medium* https://medium.com/pervade-team/how-contextual-integrity-can-help-us-with-research-ethics-in-pervasive-data-ef633c974cc1 ● Helen Nissenbaum, "Contexts, Informational Norms, Actors, Attributes, and Transmission Principles," *Privacy in Context: Technology, Policy, and the Integrity of Social Life*	The *Medium* article offers an accessible overview of Nissenbaum's framework, pulling together aspects of her argument that spans several chapters of her book. Additionally, this article includes a discussion of a published research paper that applies Nissenbaum's framework to evaluate a real-world case, namely Emil Kirkegaard's creation of a public dataset using data from the OkCupid online dating platform. It was useful to reference this discussion in class to explain how viewing privacy in terms of public and private spheres is insufficient for capturing the widespread intuition that the creation of this dataset violated privacy (whereas Nissenbaum's framework can capture this intuition). The chapter from Nissenbaum's book offers an overview of context-relative informational norms and an argument for how disrupting these norms constitutes a violation of privacy. Alternatively, it may be useful to assign sections from this chapter in combination with sections from Chapter 8 "Breaking Rules for Good" which describes how a violation of privacy may nevertheless be justified in light of other ethical considerations.

Implementation

Class Agenda:	1. Explore how the features of TinyML pose unique challenges to privacy 2. Explain Nissenbaum's framework for understanding privacy as a right to contextual integrity 3. Class activity to practice applying this framework to real-world cases 4. Reflect on how contextual integrity can inform decisions about design	Marginal notes
Sample Class Activity:	First, students are presented with background information about a real-world example of an ML device. Then, the Embedded EthiCS TA walks students through defining the context-relative informational norms (the context; the key actors involved including the subject, senders, and recipients of information; the attributes of information shared, and the transmission principle). Next, the Embedded EthiCS TA identifies how the device in question may disrupt the informational norms, thus violating privacy. After the device has been flagged as a violation of privacy, students are asked to perform the second round of evaluation to determine if this violation of privacy is justified in light of other ethical considerations. More specifically, students are given several minutes to consider whether the new device or practice: 1) provides better support for contextual values; 2) promotes autonomy; 3) improves power relations; 4) creates a fair distribution of costs and benefits. Students then complete a Zoom poll that asks them to identify which of the 4 ethical considerations listed above they found most relevant or useful in their analysis. The results of the poll are shared with the whole class and can be used by the module TA to guide the class discussion. This process is repeated for the remaining 2 case studies.	The specific case studies used for this module were: 1) Video surveillance for fall detection of the elderly 2) Health wearable devices used by John Hancock Insurance 3) Amazon employees reviewing users' voice recordings from Alexa (personal voice assistant). In order to generate discussion for a class where students are hesitant to participate, the Zoom polls proved to be extremely effective in engaging students and structuring discussion. For larger classes (>40 students) this activity would work well by having students work in small groups in break out rooms before reconvening in a larger group to debrief.
Module Assignment:	The follow-up assignment asks students to reflect on their final project and write a paragraph explanation for one of the following questions: If you think your application may potentially violate privacy, please write an explanation of why it may still ultimately be desirable with reference to the following questions: Does it provide better support for contextual values? Does it promote autonomy? Does it improve power relations? Does it create a fair distribution of burdens and benefits?	This assignment was designed as an opportunity for students to reflect on their final projects using the philosophical concepts covered in the module. For their final projects, students built their own TinyML devices, covering a wide range of applications from car counting to snoring detection. In one case, a student was working on federated learning (a privacy-

	If you think your application sufficiently protects privacy, explain how it does so by answering the following questions: What is the context? How does this new application maintain or reduce the number of key actors (subjects, senders and recipients of information), how does it maintain or reduce the attributes of the data (the type or nature of information), and how does it maintain the transmission constraints on the flow of information?	preserving technique). In order to capture the different kinds of projects students were working on (and the different implications they might have for privacy), the assignment was split into two options: Students could either work through the four ethical considerations to determine whether their device would be justified even if it violated privacy, or they could explain how their project sufficiently preserves privacy by explaining how it does not result in a shift of context-relative informational norms.
Lessons Learned:	In the follow-up evaluations, quantitative feedback revealed that 100% of students found this class to be both interesting and relevant to their work. In addition, qualitative comments revealed that students found the discussion of power dynamics to be a particularly helpful concept for thinking about the ethical issues we discussed in class. After running this module, there are two lessons that are worth highlighting: 1. Including an explicit discussion during class of how the ethical issues connect to the technical content in the course seemed to improve student engagement with the material. On the evaluation form, one student specifically noted their appreciation for how the module topic was connected to course's focus on embedded ML on devices. 2. Utilizing case studies for the class activity so that students have an opportunity to practice applying what they have learned is important not only for fostering engagement, but for helping students fully understand the philosophical tools and concepts as well. Several students reported feeling more comfortable with the philosophical concepts once we began discussing them in the context of real-world examples.	Marginal notes