Repository Entry Template
Embedded EthiCS @ Harvard Teaching Lab

Overview	
Course:	CS 146 Computer Architecture
Course Level:	Undergraduate
Course Description:	Review of the fundamental structures in modern processor design. Topics include computer organization, memory system design, pipelining, and other techniques to exploit parallelism. Discussion of modern topics including GPU architectures, datacenter architecture, mobile/embedded SoC architectures, and machine learning acceleration as time permits. Emphasis on a quantitative evaluation of design alternatives and an understanding of performance and energy consumption issues.[1]
Module Topic:	Sustainable AI
Module Author:	Elís Miller Larsen
Semesters Taught:	Spring 2021
Tags:	natural language processing (NLP) models [CS] processing power [CS] training models [CS] sustainability [both] technological determinism [both] responsibility [phil]
Module Overview:	In this module we discuss sustainable AI in terms of green computing and the future of machine learning. Sustainability generally refers to any form of development that meets the needs of the present without compromising the ability of future generations to meet their own needs. Sustainable AI is a relatively new topic, with most of the discussion focusing on AI for sustainability, i.e., proposals for how AI can help other industries become more sustainable. This module draws an important distinction between AI for sustainability and the sustainability of AI, where the latter is an inward critique towards sustainable processes for AI development and research. One of the major concerns brought out in this module is the level of energy consumption required to train AI, e.g., an NLP model can emit the equivalent of the lifetime of five cars in carbon emissions. Stats like this one raise two important questions: (1) How can AI be sustainable? (2) Who is responsible for sustainable AI? Throughout the module, students are given guidelines and tools to discuss each of these questions.	
Connection to Course Material:	Throughout the course, students are learning about computer architecture and NLPs without specific attention paid to the amount of energy required to train models or maintain large data sets. This module seeks to bring awareness about the energy consumption and carbon footprint emitted by the systems the students are already familiar with.	This course looks primarily at the design, construction and efficiency of computer hardware, including processors, transformers and storage options. As such, energy consumption and economic impact are constant considerations. Considering the environmental impact of different

---

	possible design choices is thus a natural further question to ask.

Goals

Module Goals:	1. Introduce students to the realities of energy consumption of data systems and NLP models. 2. Define sustainability. 3. Draw a distinction between AI for sustainability and the sustainability of AI. 4. Introduce the ethical problems for sustainability of AI: technological determinism and responsibility. 5. Discuss what responsibility might look like for sustainable AI, i.e., *are tech companies and/or researchers responsible for computing sustainably?* 6. Help students to assess three frameworks for ethical responsibility by assisting them to identify the pros and cons of each.	The connection for (4) can be understood as follows: Technological determinism is the idea that technological advancement is determined, and progress becomes part of the fabric of our social and cultural values. This idea seemingly undermines responsibility because it suggests that whatever gains or losses technology poses are inevitable, whereas responsibility suggests that individuals can resist certain advancements, particularly when those advancements might conflict with social values such as privacy or security.
Key Philosophical Questions:	1. What is morally at stake when it comes to sustainable AI? 2. What does sustainability look like for AI? How should we measure it? 3. Who is responsible for sustainability?	The questions cover three main areas of an inquiry into sustainable AI. The first question identifies the problem. The second question aims to define the concept of sustainability. And the third question examines who might be responsible for the problem, or responsible for addressing the problem.

Materials

Key Philosophical Concepts:	● Sustainability ● Responsibility ● Technological Determinism	Discussing criticisms of technological determinism opens up students to the possibility that things could have been otherwise (e.g. designed with ethical considerations in mind).
Assigned Readings:	● Wynsberghe (2021) "Sustainable AI: AI *for* Sustainability and Sustainability *of* AI" ● Strubell et al. (2019) "Energy and Policy Considerations for Deep Learning in NLP"	The Wynsberghe (2021) paper is meant to draw students towards the important distinction between AI for sustainability and the sustainability of AI. It also draws on the Strubell et al. (2019) paper and discusses additional methods for achieving greater sustainability. The Strubell et al. (2019) paper introduces students to the stats of energy consumption by NLP models.

Implementation

Class Agenda:	1. Introduce the problem: carbon emissions and the financial cost of training NLP models 2. Introduce the idea of sustainability as a solution to the problem with a focus on two areas: (a) how we measure sustainability and (b) how we get companies to report accurate measurements 3. Discuss technological determinism and responsibility with respect to sustainable AI solutions 4. Class Activity: develop an [proto] ethical framework for responsibility	
Sample Class Activity:	The students are asked the question: *What responsibility might tech companies and researchers have in terms of improving the sustainability of AI?* They are given three options and asked to (1) provide a reason for and against each option, and (2) weigh their reasons to determine which option is the most ethical. Option #1: Responsibility = cutting carbon emissions by the same % > 0 Option #2: Responsibility = reducing carbon emissions in proportion to the carbon footprint and environmental impact Option #3: Responsibility = reducing carbon emissions in proportion to the social impacts (i.e., risks and benefits)	In this activity, students think about what sustainable AI solutions might look like and how responsibility might be dispersed across tech companies, researchers, and users. With respect to the 3 options listed, environmental impacts of carbon emissions refers to direct pollution, e.g. to air quality and landfills, whereas social impacts refers to the kinds of social goods we might get from a product. For example, whether training an NLP might lead to advances in the diagnosis and treatment of medical conditions such as cancer.
Module Assignment:	Students were asked to come prepared with reading comprehension questions, which were answered throughout the course of the module. At the end of the module, there is a check-in to ensure that all questions have been addressed.	Students didn't have any trouble with reading comprehension, but instead asked about their individual role as researchers in contributing to the negative environmental impacts of AI.
Lessons Learned:	This module worked well as a discussion-based module because of the small class size of five students. Modifications may be needed to accommodate larger class sizes such as having an online discussion forum prior to class for students to brainstorm the main concepts e.g. sustainability, technological determinism, etc. The activity is suitable for large or small classes, as students can be divided into working groups of 3-4.	