CS1 Repository Entry
Embedded EthiCS @ Harvard Teaching Lab

Overview

Course:	CS1 Great Ideas in Computer Science
Course Level:	Introductory undergraduate
Course Description:	"COMPSCI 1 (or simply CS-1) is a broad introduction to the most important concepts in computer science, and will expose students to the principles and practices of functional and object-oriented programming (OOP), the mathematical, statistical, and computational methods that will enable you to think critically about data as it is employed in fields of inquiry across the Faculty of Arts and Sciences, The landscape of computer science as it exists today, with some reference to its past and future. This will enable us to touch on a variety of really fascinating topics and intellectual paradigms, i.e., some of the "Great Ideas in Computer Science," such as Machine Architecture, Security and Privacy, Computer Communications, Program Execution Time, Noncomputability, Simulation, The Ethics of Algorithms."[1]
Module Topic:	Fairness in algorithmic and human-decision making
Module Author:	Aksel Braanen Sterri
Semesters Taught:	Spring 2023
Tags:	algorithms [CS], machine learning [CS], fairness [both], bias [both], noise [both]
Module Overview:	In this module, we examine the possibilities and drawbacks of using algorithms to guide, supplement or substitute decision-making. Students are introduced to the distinction between two decision-making problems—the *noise problem* and the *bias problem*—and presented with examples of how human experts are vulnerable to both. We explore the relationship between noise and bias on the one hand and fairness on the other. The students get a brief introduction to how machine learning algorithms are trained and operate—and how they can be used to reduce bias and noise. We examine where noise and bias can creep into algorithmic decision-making. Finally, we explore whether a perfectly unbiased and non-noisy algorithm would be fair in an unjust society and ask: Is it a necessary requirement for a fair algorithm that it contributes to a fairer overall distribution of benefits and burdens?

Connection to Course Material:	The course is an introductory computer science course where the students learn basic coding. They are also introduced to machine learning and several ethical aspects pertaining to the development and use of algorithms. This module introduces some of the core issues that are discussed in the course.	The topic of algorithmic fairness is particularly relevant for this class and has been a success in previous years, according to the course instructor. It is a tangible way to examine "fairness," a central ethical notion, and deals with a

[1] https://locator.tlt.harvard.edu/course/colgsas-119953/2022/spring/14566

		topical and important problem that resonates with students from different backgrounds. An alternative topic for this class could be issues pertaining to privacy.

Goals		
Module Goals:	1. Help the students see the benefits and drawbacks of using algorithms. 2. Examine the difference between absolute and comparative ethical assessments. 3. Explore flaws in human decision-making. 4. Distinguish between noise and bias as two different sources of unfairness. 5. Explore whether a decision's contribution to a fair distribution of benefits and burdens in society is a crucial part of fairness.	
Key Philosophical Questions:	1. What is a fair decision? 2. What is the difference between noisy and biased decisions? 3. What are the sources of bias in algorithms? 4. How does background injustice influence the fairness of a particular decision?	Fairness is one of the most widely used normative concepts and has received a substantial interest by computer scientists. It is therefore crucial to examine fairness and how it relates to algorithmic decision-making, where the concern about unfairness comes in. The same is true for bias. How does bias relate to fairness? Is bias sufficient and necessary for fairness? Helping the students tackle these questions will help them deal constructively with these pressing issues in the future.

Materials		
Key Philosophical Concepts:	● Fairness ● Bias ● Noise ● Comparative and absolute fairness	The module instructor presented the distinction between "noise" and "bias" from Kahneman, Sibony, and Sunstein's book *Noise*. Bias is a systematic deviation from the truth, while noise is a non-systematic deviation. If a set of judges disagree on the answer, this is noise, whereas if they tend to agree that White people should get more lenient punishment than Black people for the same crime, they are biased. The module instructor stressed that both noise and bias could make a decision

		unfair. The module instructor also introduced the distinction between comparative and absolute ethical assessments. When we assess algorithmic fairness, should we compare it to an ideal (absolute), or should we compare it to human decision-making (comparative)?
Assigned Readings:	● https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing ● https://hbr.org/2016/10/noise	The two readings were given as part of the assignment. ProPublica's report on COMPAS is a modern classic. It gives the student the impression that algorithms can create unfairness and introduces specific criteria of algorithmic fairness. The article on noise gives a simple introduction to flaws in human decision-making. It is expected to lead students to think about the need for decision tools to improve human decision-making.

	Implementation	
Class Agenda:	1. Students reflect on how algorithms can make fairer and unfairer decisions and make up their minds about whether algorithms will improve or worsen decision-making. 2. Distinguish between two different ways of assessing a practice, either against an ideal or absolute standard or comparatively, against another practice. 3. Examine flaws in human decision-making and distinguish between noise and bias. 4. Introduce a minimal account of fairness as non-biased and non-noisy. 5. Introduce algorithms 6. Explain how algorithms can reduce bias and noise 7. Examine where bias and noise can creep into an algorithm. 8. Discuss the relation between background unfairness and fairness in algorithmic decision-making.	
Sample Class Activity:	The module instructor alternated between getting them to reflect on their own, pen to paper, and discussing in small groups before asking them to share their thoughts with the rest of the class. One sample class activity was to ask the students to reflect on the following questions: Why would	CS1 is both in-person and online, and many students watch only the recorded lecture. The module instructor thus kept these parts relatively short. Interestingly, the students in person were predominantly on the side of

	algorithms help us make fairer decisions? Why would they make unfair decisions? Write 1-2 points on each before coming to an overall assessment: Do you think algorithms more often than not are a source of fairness or unfairness?	human decision-making. They believed algorithmic decision-making was more often than not a source of unfairness.
Module Assignment:	Please read these two pieces: ProPublica's piece on "machine bias" in the COMPAS algorithm: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making, by Daniel Kahneman et al.: https://hbr.org/2016/10/noise Then, respond to the following questions with answers of 1-2 paragraphs each. We don't expect you to do any outside research, though we encourage you to connect to lecture materials, in-class examples we talked about, and the pre-class reading where relevant. Part (a): Explain the concepts of "bias" and "noise" and use them to compare the strengths and weaknesses of human and algorithmic decision-making in the context of risk assessment in the criminal justice system (4 points) Part (b): Suppose COMPAS could be improved to the point where it is no longer "biased." First, what would that entail? Second, in a society where recidivism rates differ between Black people and White people, would this be sufficient to say that COMPAS produces fair results? Or are there other factors that need to be considered? What would they be? (4 points)	The students received this assignment after the module. The idea was for them to read two pieces that would mirror much of what was said in the lecture to reinforce their understanding of the key concepts, bias and noise. They were then asked to show that they understood the key difference between noise and bias, and on noise and bias on the one hand and unfair background circumstances on the other. The responses showed that the students gained a very clear understanding of the key concepts.
Lessons Learned:	The students were overall very positive about the module. 1. It was particularly helpful to distinguish between noise and bias and illustrate with examples. 2. Students want more focus on the “what to do”-part. How can we, in practice, use algorithms to improve fairness? One student wanted simulations.	On point 2, one could draw on “Discrimination in the Age of Algorithms” a paper by Kleinberg et al.