Ethics of Artificial Intelligence and Machine Learning

(Image produced using Deep Dream Generator, a computer vision program that utilizes a convolutional neural network to recreate a picture in the style of another)

Course Description

This discussion-based course will provide an introduction to the ethical issues related to artificial intelligence and machine learning. The first half of the semester will focus on concerns existing in the present day, such as bias and transparency. During the semester's second half, we will cover topics that will be increasingly important going forward, from consciousness to the future of labor.

Course Details

- **Course**: CMSC389V
- **Prerequisites**: CMSC216 and CMSC250
- **Credits**: 1
- **Seats**: 30

- **Lecture Time**: Friday, 1:00-1:50 pm
- **Location**: MTH0101
- **Course Materials**: All required course materials will be provided.
- **Course Facilitator**: Anthony Ostuni
- **Faculty Advisor**: Dr. John Dickerson

Class Format

This course will primarily be discussion-based, with the topic of each week's discussion determined by a set of short readings and/or videos. These materials must be completed before class and used to write a response to a short-answer question. Note that the materials do not necessarily reflect the viewpoints of the course facilitator nor faculty advisor; they are simply design to provoke thought on the subject area. Prior to the class discussion, there may be a brief lecture on a specific idea or concept that could be of value to the conversation. The primary assignments for the class will be two papers.

Grading

Grades will be maintained on ELMS.

You are responsible for all material discussed during lecture, as well as readings and other material posted on ELMS or Piazza outside of class.

Percentage	Title
10%	Reading Short-Answer
30%	Participation
25%	First Paper
25%	Second Paper
10%	ML / AI Guidelines

Discussions

The majority of time in the classroom will be spent in group discussion. This will allow for an efficient exchange of diverse ideas and perspectives, as well as forcing you to become more comfortable organizing and articulating technical and philosophical concepts. We will decide discussion rules together as a class during the first week.

Papers

There will be two essays that will compose the majority of your work outside the classroom. The first will be assigned after Week 6, and it will be on one of the topics discussed up to that point in class. The second paper will be assigned after Week 12, and it should represent the culmination of ideas developed throughout the

previous six weeks. The official assignment details for both papers will be released closer to the assignment dates.

All regrade requests must be made within one week of the assignment grade being released.

Final Assignment

In lieu of a final, you will be expected to develop a list of ML / AI Guidelines for a tech company to follow and justify your rules. The official assignment details will be released closer to the assignment date.

Late Policy

All assignments may be turned in up to 24 hours late with a 25% penalty. After 24 hours, no late assignments will be accepted.

Schedule

The schedule is subject to change; students will be notified in such an occurrence.

Week	Topic	Readings and Slides	Assignments
1	Introduction to Ethics	Ground Rules for Discussions, What is Ethics?	
2	Data Privacy	Target Exposed Pregnancy, Netflix "Anonymous" Data	Short Answer Questions
3	Transparency and Interpretability	Importance of Interpretability	Short Answer Questions
4	Machine and System Bias (Part 1)	Recidivism Algorithm	Short Answer Questions
5	Machine and System Bias (Part 2)	China's Citizen Score, Using AI in Hiring	Short Answer Questions
6	Automated System Error	Self-Driving Cars, Autonomous Cars and Surgical Robots	Short Answer Questions
7	War and Military Systems	Why We Must Not Build..., AI is Changing Every Aspect of War	Short Answer Questions
8	Defining Intelligence and the Turing Test	The Turing Test, What is AI?	Short Answer Questions
9	The Future (or End) of Labor	The Rise of the Machines, Universal Basic Income Explained	Short Answer Questions
10	Consciousness and Rights	Do Robots Deserve Rights, The Puzzle of Conscious Experience	Short Answer Questions
	Singularity, Control, and	What Happens When Our Computers Get	Short Answer

11 Week	Unintended Consequences Topic	Smarter Than We Are? Readings and Slides	Questions Assignments
12	Inequality, Unequal Access, and Open-Sourcing	AI and Income Inequality, AI Research Openness	Short Answer Questions
13	Interaction and Affection	AI Companion for the Elderly, Companion Robots are Here	Short Answer Questions
14	Protecting AI	Crippling AI, Protecting AI	Short Answer Questions
15	Artificial Creativity	Can AI Be Creative?, Multi-Agent Hide and Seek	Short Answer Questions

Communication

The primary means of communication outside of class will be Piazza and ELMS. Office hours will be scheduled by appointment. The email addresses below should only be used for important and time-sensitive issues. - **Faculty Adviser:** Dr. John Dickerson: john [at] umd.edu - **Course Facilitator: Anthony Ostuni**: aostuni [at] umd.edu

Excused Absence and Academic Accommodations

See the section titled "Attendance, Absences, or Missed Assignments" available at Course Related Policies.

Disability Support Accommodations

See the section titled "Accessibility" available at Course Related Policies.

Academic Integrity

Note that academic dishonesty includes not only cheating, fabrication, and plagiarism, but also includes helping other students commit acts of academic dishonesty by allowing them to obtain copies of your work. In short, all submitted work must be your own. Cases of academic dishonesty will be pursued to the fullest extent possible as stipulated by the Office of Student Conduct.

It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity or the Student Honor Council, please visit http://www.shc.umd.edu.

Feedback

If you have any suggestions for improving this class, don't hesitate to tell the instructor or facilitators during the semester. At the end of the semester, please don't forget to provide your feedback using the campus-wide CourseEvalUM system. Your comments will help make this class better in future iterations.