

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

Spring 2024

Delivery: 100% Web-Based on Canvas

Dates course will run: January 08, 2024 – May 02, 2024

Instructor

Dr. Mahender Mandala

Email: mandala@gatech.edu

Course Creator

Dr. Ayanna Howard

Email: ah260@gatech.edu

Head TAs

Vijay Sivanesan, P.E., PMP

Email: vsivanesan3@gatech.edu

&

Jeanette Schofield

Email: jschofield8@gatech.edu

General Course Information

Description

Abuse of big data means your worst fears can come true. Are they being monitored by your employer? Check. Government intrusions into your daily life? Check. Being turned down by college admissions because you are predicted not to donate in 10-20 years? Check. Sounds a bit like the visions in the Minority Report. Alas, machine learning algorithms are already being deployed by industry, government, and, yes, even schools to make decisions that impact us in direct ways. Such programs are typically promoted as fair and free of human biases, but humans who make mistakes are programming, calibrating, and evaluating their performance. Thus resides the problem. How do we design algorithms that effectively deal with large amounts of data to train them while ensuring their outcomes aren't misused? In this course, not only will we examine various AI/ML techniques that can be used to counterbalance the potential abuse and misuse of learning from big data, but we will focus on the effects of these technologies on individuals, organizations, and society, paying close attention to what our responsibilities are as computing professionals.

Pre-Requisites

Although it is expected that, throughout the course, you will develop working knowledge of using Python, basic programming proficiency is required to complete this course. Several optional exercises and additional lecture modules are provided to help you become familiar with Python and using Jupyter notebooks.

A social media account or access to a friend's or relatives' data is also required to complete the first module.

Course Goals and Learning Outcomes

There are several outcomes for the course, based on four primary modules:

Module 1 - Data, Individuals, and Society

Objective: After completing this module, students will be able to understand the power and impact that analytics and AI/ML have on individuals and society, especially concerning issues such as fairness and bias, ethics, legality, data collection, and public use.

Module 2 – The BS of Big Data

Objective: After completing this module, students will be able to understand the underlying components of big data, apply basic statistical techniques to data scenarios, and understand the issues faced when learning from big data, ranging from data biases, overfitting, causation vs. correlation, etc.

Module 3 – Fairness in AI/ML

Objective: After completing this module, students can understand and apply basic AI/ML techniques to data scenarios, focusing on identifying fairness and bias issues in designing decision-making systems. We will work systematically towards understanding technical approaches to current AI/ML applications, such as facial recognition, natural language processing, and predictive algorithms, all while being mindful of its social and legal context.

Module 4 – Bias Mitigation and Future Opportunities

Objective: After completing this module, students can utilize tools and methods to quantify bias and examine ways to use algorithmic fairness to mitigate it, considering ethical and legal issues. Students will apply their knowledge of analytics and AI/ML to transform a current biased dataset into a more objective solution.

In this class, you will be challenged to broaden your understanding of state-of-the-art AI/ML algorithms and solutions, considering their potential impacts on society. You will have ample opportunity to analyze various situations critically and viewpoints provided in papers, books, on the web, and from your observations. You will be able to practice your learned knowledge by writing coherent and healthy-structured critiques of situations and papers, leading and participating in class discussions, and designing your algorithmic solutions. The issue of data misuse and abuse is not easily solvable; concrete right or wrong answers are not easily determined until solutions are typically deployed into society. Given this, you are entitled to your opinions on any topics presented throughout the course, whatever they happen to be. You will not be penalized for your viewpoints; however, you must be able to support your viewpoints and resulting solutions effectively. This means showing that you have given your approach to a problem some thought, can discuss its various trade-offs and implications and can support other viewpoints, even though your personal views may differ.

Course Materials

Course Text

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy by Cathy O'Neil (2016)

Additional Materials/Resources

Additional assigned readings will be included with each assignment.

Classroom Management Tools

- Video Lectures: All video lectures are located on Canvas.
- Projects: are located on Canvas.
- Graded Discussions: are located on Canvas.
- Reading Materials: are located on Canvas.
- **Ed** Discussion: is located on Canvas.
- Grades: is located on Canvas.
- Exams: are located on Canvas using Honorlock

Course Requirements, Assignments & Grading

Assignment Distribution and Grading Scale

| Assignments | Weight |
|---|-------------|
| Homework Projects | 40% |
| Written Critiques | 10% |
| Mid-Term Exam | 10% |
| Final Project | 15% |
| Final Exam | 10% |
| Class Discussion/Exercises (<i>Case Studies and, Exercises</i>) | 15% |
| Total | 100% |

Grading Scale

Your final grade will be assigned as a letter grade according to the following scale:

A 90-100%

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

B 80-89%

C 70-79%

D 60-69%

F 0-59%

Assignments Quality

This is a graduate level course and you are expected to turn in graduate level work. This means properly formatted reports, no one-line student responses on discussion assignments and actual substance to your work. We reserve the right to deduct points for work we do not feel is graduate level.

Assignments Due Dates (Time zone)

Assignment due dates are Sundays at 11:59 PM [Anywhere on Earth time](#). We will only accept assignments submitted on time due to time zone issues. We recommend [changing your time zone in Canvas](#) to show the due date in your local time. There are no exceptions.

Late and Make-up Work Policy

Homework Projects and Written Critique assignments will be accepted with a deduction of 10% per 24-hour period starting after the due date submission time. Assignments over 3 days late (i.e., three 24-hour periods) will not be accepted. No time extensions are provided for the Exams, Final Project, and Class discussions/Exercises. There will be no make-up work provided for missed assignments.

Please verify the information in your assignment before submission. We will NOT make exceptions for submitting the wrong assignment or insufficient or inadequate information.

Of course, emergencies (illness, family emergencies) will happen. In those instances, please get in touch with the Dean. The Dean of Students is equipped to verify emergencies and pass confirmation on to all your classes. For consistency, we ask all students to do this in an emergency.

Regrade Policy

If you consider that your score is incorrect based on the feedback provided, you may request a rescore during the regrade time period, which is 7 full days from the date the grades are released. If the grades are released Monday, you will have till next Monday's End-Of-Day.

This request is only valid if it includes a good explanation of where you think the grader made an error.

The student should **open a private Ed in the appropriate regrade category** post and copy all instructors and provide the following (per question to regrade) in the specified format:

- Question to regrade:
- Why do you consider there is an error in the score?

Questions regarding grades in Canvas will not be considered. Please open a private Ed post in the correct regrade category to avoid disregarding regrade requests.

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

Only one (1) regrade per assignment is allowed. Make sure your initial regrade request provides adequate details and double-check before submitting, as we will consider only the first request and ignore the rest. Once a regrade has been completed by the TA, the grade is final.

Additional General Information on the Re-grading process:

- When submitting a request to have the instructors review your assignment, please provide a detailed list of why you think your work deserves more points. Do not just send us a request to take another look without justification.
- Re-grade requests must be made via private Ed Discussion posts. Please address all of the instructors and the TAs.
- Follow the regrade policy regarding the regrade submission procedure and the timeline for submission of regrades.
- Requesting a re-grade may result in a lower grade than what you were initially given by the TAs.
- It is essential to maintain a professional and formal tone in all communications. Requests that do not adhere to these standards of professionalism may not be considered.
- Once the instructors have reviewed your assignment and issued an updated grade, we will not entertain any further discussion on the grade we have given for that assignment.
- A re-grade is not the same as a request for accommodation due to hardship. If you have a legitimate hardship, please work with the Dean of Students office to have them provide an email to the instructors, and we will happily work with you to allow for additional time, etc., so that you can be successful in this course while working through your life challenges.

AI use Policy

If your work is found to violate this policy, you **will be turned in for Academic Misconduct to the Office of Student Integrity.**

Recent advancements in artificial intelligence—Copilot, ChatGPT, etc.—can be great resources for improving your learning in the course, but it is important to ensure that their benefits are targeted at your learning rather than solely at your deliverables. Toward that end, the same academic integrity policy above applies to AI assistance: you are welcome to consult with AI agents just as you would consult with classmates, discuss ideas with friends, and seek feedback from colleagues. However, just as you would not hand your device to someone else to directly fix or improve your classwork, so also you may not copy anything directly from an AI agent into your document, nor let an AI agent directly generate content for your submission. This rule means you should disable any AI assistance more advanced than a grammar checker inside your word processors and IDEs.

Although you are prohibited from having these tools directly integrated into your workspace or from copying content from these assistants directly into your work, you are nonetheless permitted to use them more generally. The important consideration is to ensure that you are using the AI agent as a learning assistant rather than as a homework assistant, so long as your submission solely reflects your own understanding of the content, you are encouraged to let AI assistants aid in developing your understanding.

Office Hours

This course does not have office hours. Students are encouraged to ask questions via **Ed** discussion forums. The teaching team will respond to the **Ed** posts in an efficient manner.

Feedback

Every semester, we make changes and tweaks to the course formula. As a result, we try new things every semester, and some of these things may need to be revised. We ask for your patience and support as we figure things out, and in return, we promise that we, too, will be fair and understanding, especially with anything that might impact your grade or performance in the class. Second, we want to consistently get feedback on improving and expanding the course for future iterations. You can take advantage of the feedback box on **Ed** especially if you want to gather input from others in the class), give us feedback on the surveys, or contact us directly via private **Ed** messages.

Technology Requirements and Skills

Computer Hardware and Software

- High-speed Internet connection
- Laptop or desktop computer with a minimum of a 2 GHz Processor and 2 GB of RAM
- Windows for PCs OR Mac iOS for Apple computers.
- Complete Microsoft Office Suite or comparable and the ability to use Adobe PDF software (install, download, open, and convert)
- Mozilla Firefox, Chrome, and Safari browsers (*Note: Honorlock requires students to use Chrome*)

Canvas

This class will use Canvas to deliver course materials to online students. All course materials and assessments will take place on this platform.

Proctoring Information

The midterm exam will be proctored as a closed book, closed notes, closed video, and no access to any external media except the exam itself. You are all allowed an 8-½"x11" blank paper and a simple non-scientific calculator as a backup. A proctored exam is similar to the one you would take in the classroom. These exams are delivered via a tool called Honorlock. Honorlock is an online proctoring service that allows you to take your exam from the comfort of your home. You DO NOT need to create an account, download software, or schedule an appointment in advance. Honorlock is available 24/7; a computer, a working webcam, and a stable Internet connection are needed. You will need Google Chrome and download the Honorlock Chrome Extension to start. You can download the extension at www.honorlock.com/extension/install.

When you are ready to take the exam, you would log into CANVAS, go to the course, and click on the exam link. Clicking Launch Proctoring will begin the Honorlock authentication process, where you will take a picture of yourself, show your ID, the plain scratch paper (if used), and complete a room scan. Honorlock will be recording your exam session via the webcam and your screen. Honorlock also has an integrity algorithm that can detect search-engine use, so please do not attempt to search for answers, even if it's on a secondary device. If you encounter any issues, you may contact them by live chat, phone (844-243-2500), or email (support@honorlock.com).

Course Policies, Expectations & Guidelines

Communication Policy

You are responsible for knowing the following information:

1. Anything posted to this syllabus
2. Anything emailed directly to you by the teaching team (including announcements Or Posts via **Ed**) 24 hours after receiving such an email.

Because **Ed** announcements are emailed to you, you should check your Georgia Tech email once every 24 hours to remain updated on new information during the semester. Georgia Tech generally recommends that students check their Georgia Tech email once every 24 hours.

We generally prefer to handle communication via **Ed** to help with collaboration among the teaching team, but we understand **Ed** is not ideal for having information “pushed” to you. We may contact you via a private **Ed** post instead of an email. Still, if we do so, we will send email notifications immediately, bypassing your settings, to ensure you’re alerted. This type of communication will also spring under #2 above.

Note that this means you won’t be responsible for knowing information communicated in several other methods we’ll be using. You aren’t responsible for knowing anything posted to **Ed** that isn’t linked to an official announcement. You don’t need to worry about missing critical information so long as you keep up with your email and understand the documents on this website. This also applies reverse: we do not monitor our Canvas message boxes and may not respond to direct emails. If you need to get in touch with the course staff, please post privately to **Ed** (either to all Instructors or to an instructor individually) or tag the instructor in the relevant post.

Online Student Conduct and (N)etiquette

Communicating appropriately in the online classroom can be challenging. To minimize this challenge, it is important to remember several points of “internet etiquette.”

That will smooth communication for both students and instructors:

- Read first, Write later. Read the ENTIRE set of posts/comments on a discussion board before posting your reply to prevent repeating commentary or asking questions that have already been answered.
- Avoid language that may come across as intense or offensive. Language can be easily misinterpreted in written and electronic communication. Review email and discussion board posts BEFORE submitting. Humor and sarcasm may be easily misinterpreted by your reader(s). Try to be as matter-of-fact and professional as possible.
- Follow the language rules of the Internet. Do not write using all capital letters because it will appear shouting. Also, emoticons can be helpful when used to convey nonverbal feelings. ☺
- Consider the privacy of others. Ask permission before giving out a classmate's email address or other information.
- Keep attachments small. If it is necessary to send pictures, change the size to an acceptable 250kb or less (one free, web-based tool to try is picresize.com).

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

- No inappropriate material. Do not forward virus warnings, chain letters, jokes, etc., to classmates or instructors. The sharing of pornographic material is forbidden.

NOTE: The instructor reserves the right to remove posts that are not collegial and do not meet the Online Student Conduct and Etiquette guidelines.

University Use of Electronic Email

A university-assigned student e-mail account is the official university means of communication with all students at the Georgia Institute of Technology. Students are responsible for all information sent via their university-assigned e-mail account. Students who forward information to their university e-mail account are responsible for all information sent to any other e-mail account, including attachments. To stay current with university information, students must check their official university e-mail accounts and other electronic communications frequently and consistently. Recognizing that some communications may be time-critical, the university recommends that electronic communications be checked minimally twice a week.

Plagiarism & Academic Integrity

Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. All students enrolled at Georgia Tech and all its campuses are to perform their academic work according to standards set by faculty members, departments, schools, and colleges of the university; cheating and plagiarism constitute fraudulent misrepresentation for which no credit can be given and for which appropriate sanctions are warranted and will be applied. For information on Georgia Tech's Academic Honor Code, please visit <http://www.catalog.gatech.edu/policies/honor-code/> or <http://www.catalog.gatech.edu/rules/18/>.

Any student suspected of cheating or plagiarizing on an exam, exercise, or assignment will be reported to the Office of Student Integrity, investigate the incident, and identify the appropriate penalty for violations.

Any attempt to bypass our automated content-checking system will result in a 0 for the assignment.

All graphs shall be generated using Python, Excel, or other software tools, and Hand-drawn graphs, tables, and charts are not allowed. The production of Tables using screenshots is not allowed.

Accommodations for Students with Disabilities

If you are a student with learning needs that require unique accommodation, contact the Office of Disability Services at (404) 894-2563 or <http://disabilityservices.gatech.edu/> as soon as possible to make an appointment to discuss your unique needs and to obtain an accommodations letter. Please also e-mail me as soon as possible to set up a time to discuss your learning needs.

Student-Faculty Expectations Agreement

At Georgia Tech, it is essential to strive for an atmosphere of mutual respect, acknowledgment, and responsibility between faculty members and the student body. See

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

<http://www.catalog.gatech.edu/rules/22/> for an articulation of some basic expectations that you can have of me and that I have of you. Ultimately, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class.

Subject to Change Statement

The syllabus and course schedule may be subject to change. Changes will be communicated via the Ed announcement tool. Students must check **Ed for instructor posts**, email messages, and course announcements to stay current in their online courses.

Course Schedule

| Week/Dates | Topics | Deliverables |
|------------------|---|---|
| 1 January 15 | Lesson 1 Data Individuals, and Society Introduction Lesson 2 Overview | |
| 2 January 22 | Lesson 3 Ethics vs. Law Lesson 4 Data Collection | Case Study: Emails Exposed Case Study: Loan Denied |
| 3 January 29 | Lesson 5 Fairness and Bias | Assignment: FB |
| 4 February 5 | Lesson 6 BS of Big Data & Stats 101 Overview Lesson 7 Python and Stats 101 | Case Study: Facebook Manipulation Assignment: Stats 101 |
| 5 February 12 | Lesson 8 Descriptive Statistics Lesson 9 Inferential Statistics: Sampling Bias | Exercise: Anscombe's Quartet Exercise: Smoking Sampling Bias Design |
| 6 February 19 | Lesson 10 Inferential Statistics: Causation vs. Correlation Lesson 11 Inferential Statistics: Confidence | Written Critique: Ethical Autonomous Vehicles |
| 7 February 26 | Lesson 12 AI/ML Techniques: Word Embeddings Lesson 13 Bias in Word Embeddings | Mid-Term Exam (Ethics and Stats) Case Study: Word Embeddings Exercise: Bias in Word Embeddings |
| 8 March 4 | Lesson 14 AI/ML Techniques: Facial Recognition Lesson 15 Bias in Facial Recognition | Case Study: Facial Recognition Assignment: AI/ML Part I |
| 9 March 11 | Lesson 16 AI/ML Techniques: Predictive Algorithms Lesson 17 Crime-based Predictive Algorithms | Assignment: AI/ML Part II |
| 10 March 18 | Lesson 18 Bias in Predictive Algorithms | Exercise: Predictive Algorithms |

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

| Week/Dates | Topics | Deliverables |
|----------------|---|---|
| | | Case Study: Predictive Algorithms |
| 11 March 25 | Lesson 19 Fairness and Bias | Assignment: Fairness and Bias |
| 12 April 1 | Lesson 20 Fairness and Bias Assessment Tools | Exercise: AI Fairness 360 Exercise: What-If Tool |
| 13 April 8 | Lesson 21 AI/ML Techniques for Bias Mitigation | Written Critique: What-If Tool |
| 14 April 15 | Lesson 22 AI, Society, and Ethics | Final Project |
| 15 April 22 | Wrap-up | Final Exam |
| April 29 | | End-of-Course Survey, CIOS Survey |