Georgia Institute of Technology
Course Syllabus: CS6603 AI, Ethics, and Society

Fall 2021
Delivery: 100% Web-Based on Canvas
Dates course will run: August 23, 2021 – December 13, 2021

Instructor Information
Dr. Mahender Mandala
Email: mandala@gatech.edu

Course Creator
Dr. Ayanna Howard
Email: ah260@gatech.edu

General Course Information

Description
Abuse of big data means your worst fears can come true. Being monitored by your employer? Check. Government intrusions into your daily life? Check. Being turned down by college admissions because you are predicted to not 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, humans that make mistakes, are programming, calibrating, and evaluating their performance. Thus resides the problem. How do we therefore design algorithms that effectively deal with the large amounts of data that are used to train them, while ensuring their outcomes aren't, well, 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
None. Although, it is expected that, throughout the course, you will develop working-knowledge on using Python (which will be helpful for completing the assignments later in the semester). There are a number of optional exercises and additional lecture modules that are provided in order to help you become familiar with Python and using Jupyter notebooks.

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

Course Syllabus: CS6603 AI, Ethics, and Society

Objective: After completing this module, students will be able to understand and apply basic AI/ML techniques to data scenarios, with a focus on identifying fairness and bias issues found in the design of 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 will be able to utilize tools and methods to quantify bias and examine ways to use algorithmic fairness to mitigate this bias, taking into consideration ethical and legal issues associated with it. 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 the potential impacts they may have on society. You will have ample opportunity to critically analyze various situations and viewpoints provided in papers, books, on the web, and from your own observations. You will be able to practice your learned knowledge by writing coherent and well- structured critiques of situations and papers, leading and participating in class discussions, and designing your own algorithmic solutions. The issue of data misuse and abuse is not easily solvable; concrete right or wrong answers are not easily determined until after solutions are typically deployed into society. In view of 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 actually given your approach to a problem some thought, can discuss its various trade-offs and implications, and can be supportive of other viewpoints, even though your personal views may be different.

Course Materials

Course Text
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.
- Piazza Discussion: are located on Canvas.
- Grades: are located on Canvas.
- Exams: are located on Canvas using Honorlock

Course Requirements, Assignments & Grading

Assignment Distribution and Grading Scale

Assignments	Weight
Homework Projects	40%

Georgia Institute of Technology
Course Syllabus: CS6603 AI, Ethics, and Society

Assignments	Weight
Written Critiques	10%
Mid-Term Exam	10%
Final Project	15%
Final Exam	10%
Class Discussion/Exercises *(Cases, lesson exercises, and discussions)*	15%
Total	**100%**

Grading Scale
Your final grade will be assigned as a letter grade according to the following scale:
A   90-100%
B   80-89%
C   70-79%
D   60-69%
F   0-59%

Assignments Due Dates (Time zone)
Note that assignment due dates are all Sundays at 11:59PM Anywhere on Earth time. We will not accept assignments submitted late 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. There are no time extensions provided for the Exams, Final Project, and Class Discussion/Exercises. There will be no make-up work provided for missed assignments. Of course, emergencies (illness, family emergencies) will happen. In those instances, please contact the Dean of Students office. 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 the event of an emergency.

Office Hours
This class uses the chat tool Piazza for its office hours. Piazza office hours are not scheduled at specific times; instead, the teaching team will typically respond quickly on Piazza by end-of-day. You may also feel free to email or post privately on Piazza to set up a chat via an alternate technology.

Feedback
Every semester, we make changes and tweaks to the course formula. As a result, every semester we try some new things, and some of these things may not work. We ask 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 how we can improve and expand the course for future iterations. You can take advantage of the feedback box on Piazza (especially if you want to gather input from others in the class), give us feedback on the surveys, or contact us directly via private Piazza messages.

Technology Requirements and Skills

Georgia Institute of Technology
Course Syllabus: CS6603 AI, Ethics, and Society

Computer Hardware and Software
- High-speed Internet connection
- Laptop or desktop computer with a <u>minimum</u> of a 2 GHz processor and 2 GB of RAM
- Windows for PC computers OR Mac iOS for Apple computers.
- Complete Microsoft Office Suite or comparable and ability to use Adobe PDF software (install, download, open and convert)
- Mozilla Firefox, Chrome and/or 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 with open notes, pen/paper, and calculator allowed. 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 and all that is needed is a computer, a working webcam, and a stable Internet connection. To get started, you will need Google Chrome and to download the Honorlock Chrome Extension. 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, and complete a scan of your room. Honorlock will be recording your exam session by webcam as well as recording 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), and/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 via Piazza), 24 hours after receiving such an email.

Because Piazza announcements are emailed to you as well, you need only to check your Georgia Tech email once every 24 hours to remain up-to-date on new information during the semester. Georgia Tech generally recommends students to check their Georgia Tech email once every 24 hours. So, if an announcement or message is time sensitive, you will not be responsible for the contents of the announcement until 24 hours after it has been sent.

We generally prefer to handle communication via Piazza to help with collaboration among the teaching team, but we understand Piazza is not ideal for having information "pushed" to you. We may contact you via a private Piazza post instead of an email, but if we do so, we will choose to send email notifications immediately, bypassing your individual settings, in order to ensure you're alerted. As such, this type of communication will also spring under #2 above.

Course Syllabus: CS6603 AI, Ethics, and Society

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 Piazza that isn't linked from 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 web site. This also applies in reverse: we do **not** monitor our Canvas message boxes and we may not respond to direct emails. If you need to get in touch with the course staff, please post privately to Piazza (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. In order 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, in order to prevent repeating commentary or asking questions that have already been answered.
- *Avoid language that may come across as strong or offensive*. Language can be easily misinterpreted in written 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 as shouting. Also, the use of emoticons can be helpful when used to convey nonverbal feelings. ☺
- *Consider the privacy of others*. Ask permission prior to 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).
- *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 in nature and/or do not meet the Online Student Conduct and Etiquette guidelines listed above.

University Use of Electronic Email
A university-assigned student e-mail account is the official university means of communication with all students at Georgia Institute of Technology. Students are responsible for all information sent to them via their university-assigned e-mail account. If a student chooses to forward information in their university e-mail account, he or she is responsible for all information, including attachments, sent to any other e-mail account. To stay current with university information, students are expected to check their official university e-mail account and other electronic communications on a frequent and consistent basis. 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; and 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/.

Course Syllabus: CS6603 AI, Ethics, and Society

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

Accommodations for Students with Disabilities
If you are a student with learning needs that require special 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 special needs and to obtain an accommodations letter. Please also e-mail me as soon as possible in order to set up a time to discuss your learning needs.

Student-Faculty Expectations Agreement
At Georgia Tech we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See http://www.catalog.gatech.edu/rules/22/ for an articulation of some basic expectation that you can have of me and that I have of you. In the end, 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 Canvas announcement tool. It is the responsibility of students to check Piazza, email messages, and course announcements to stay current in their online courses.

Course Schedule

Week/Dates	Topics	Deliverables
1 **August 23-30**	**Lesson 1 Data Individuals, and Society Introduction** **Lesson 2 Overview**	**Case Study Loan Denied** **Case Study Facebook Manipulation**
2 **September 6**	**Lesson 3 Ethics vs Law** **Lesson 4 Data Collection**	**FB Assignment Due** **Case Study: Emails Exposed**
3 **September 13**	**Lesson 5 Fairness and Bias**	
4 **September 20**	**Lesson 6 BS of Big Data & Stats 101 Overview** **Lesson 7 Python and Stats 101**	**Ethical Autonomous Vehicles Written Critique Due**
5 **September 27**	**Lesson 8 Descriptive Statistics** **Lesson 9 Inferential Statistics: Sampling Bias**	**Exercise: Anscombe's Quartet** **Exercise: Smoking Sampling Bias Design** **Stats 101 Assignment Due**
6 **October 4**	**Lesson 10 Inferential Statistics: Causation vs Correlation** **Lesson 11 Inferential Statistics: Confidence**	
7 **October 11**	**Lesson 12 AI/ML Techniques: Word Embeddings** **Lesson 13 Bias in Word Embeddings**	**AI/ML Assignment - Part I due** **Case Study Word Embeddings**
8 **October 18**		**Mid-Term Exam (Ethics and Stats)**
9 **October 25**	**Lesson 14 AI/ML Techniques: Facial Recognition** **Lesson 15 Bias in Facial Recognition**	**Exercise: Bias in Word Embeddings**
10 **November 1**	**Lesson 16 AI/ML Techniques: Predictive Algorithms** **Lesson 17 Crime-based Predictive Algorithms**	**Case Study Facial Recognition** **Exercise: Predictive Algorithms**

Georgia Institute of Technology

Course Syllabus: CS6603 AI, Ethics, and Society

Week/Dates	Topics	Deliverables
11 **November 8**	**Lesson 18 Bias in Predictive Algorithms**	**Case Study Predictive Algorithms** **AI/ML Assignment – Part II Due**
12 **November 15**	**Lesson 19 Fairness and Bias**	
13 **November 22**	**Lesson 20 Fairness and Bias Assessment Tools**	**Exercise: AI Fairness 360** **Exercise: What-If Tool**
14 **November 29**	**Lesson 21 AI/ML Techniques for Bias Mitigation**	**Fairness and Bias Assignment Due**
15 **December 6**	**Lesson 22 AI, Society, and Ethics Wrap-up**	**What-If Tool Written Critique Assignment Due**
16 **December 13**		**Final Project Due** **Final Exam**
17 **December 20**		**End-of-Course Survey, CIOS Survey, Peer Feedback**