Georgia Institute of Technology

MGT 6033: Analysis of Unstructured Data Syllabus

Fall 2026

Delivery: 100% Web-Based, Asynchronous

Instructor Information

Robbie Moon
Thomas Williams – Wells Fargo Professor of Management Director,
Business Analytics Certificate robbie.moon@scheller.gatech.edu

Office Hours: Weekly office hours via Zoom (**Day/Time in Canvas**). Attendance is optional but highly encouraged.

General Course Information

Description

The primary goal of this course is to provide students with the tools they need to transform large volumes of unstructured data, such as that appearing on web pages, in social media posts, or in narrative disclosures appearing in firms' financial reports, to an analyzable form. Students will also learn how to analyze this data using techniques from natural language processing and machine learning, including the newest innovations in relying on deep learning models.

Pre- &/or Co-Requisites

There are no official prerequisites for this course. This is a very applied class—we will spend far more time understanding when and how to implement analytic methods than discussing the mechanics of how they work. Nonetheless, some background knowledge is helpful, and I have identified relevant subjects below, including optional pre-class videos/materials.

- Basic understandings of statistics (e.g., statistical distributions, means, medians, etc.) ○ Statistics Foundations 1: The Basics (LinkedIn Learning available through GT) ○ Statistics Foundations 2: Probability (LinkedIn Learning available through GT)

- Core concepts of differential calculus and linear algebra (matrix operations) ○ 3Blue1Brown's Calculus (particularly chapters 1-4) ○ 3Blue1Brown's Linear Algebra (particularly chapters 1-5)

In this class, we will rely on Python for all data analysis. I assume no prior coding experience, so the first module of material is meant to introduce the key aspects of Python, particularly those libraries we will rely on heavily. Below I provide two resources you may wish to review. The first introduces base Python assuming no prior programming experience. The second focuses on using Jupyter Notebooks, which we will use for all class demonstrations.

- Python for Non-Programmers (LinkedIn Learning available through GT)
- Introducing Jupyter (particularly chapters 1-2; LinkedIn Learning available through GT)

Course Goals and Learning Outcomes

Course Learning Outcomes

By the end of the course, students will be able to:

1. Formulate strategies for obtaining and organizing unstructured data and extracting information
2. Analyze an empirical problem that requires unstructured data for analysis and identify an appropriate strategy to conduct such analysis
3. Implement both fundamental and advanced techniques in natural language processing
4. Evaluate the output of advanced analytics techniques related to natural language processing
5. Compare and contrast the appropriateness of various techniques for a given analytical problem

Course Materials

Course Text

None

Required Software and Tools

Python 3 (free, open source) with Jupyter Notebook (available through computing cluster)

Additional Materials/Resources

None

Course Website and Other Classroom Management Tools

We will use **Canvas** as our learning management system. In canvas, you will find all:

- Course videos
- Demonstration materials
- Quizzes
- Experiential Assignments
- Ed Discussion

Course Requirements, Assignments & Grading

Assignment Distribution and Grading Scale

Assignment Weight Distribution and Due Dates

Assignment	Weight
Quiz 1	60% total (10% each, lowest grade will be dropped)
Quiz 2	
Quiz 3	
Quiz 4	
Quiz 5	
Quiz 6	
Quiz 7	
Experiential Task 1	15%
Experiential Task 2	15%

Assignment	Weight
Ed Discussion Engagement	10%

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%

Description of Graded Components

Course assessment will be done across three different types of assignments:

- **Quizzes**: Each module will include a quiz primarily related to the material covered in lecture videos, though completion of demos prior to taking the quiz is encouraged. The quizzes are a mix of True/False, Multiple Choice, and short answer type questions. All quizzes are timed, closed-book, closed-note, and require the use of proctoring software. Once starting the quiz you will have 30 minutes to complete. Your lowest score will be dropped in computing your final grade.
- **Experiential Tasks:** Each experiential task emulates a business setting requiring the use of a combination of the analytic methods we cover in class. You will have approximately two weeks to complete each task, but you should start the task early. You will complete tasks and report results in a Jupyter notebook, which you will submit to via Canvas.
- **Ed Discussion Engagement:** We will use Ed Discussion for asynchronous class discussion. You should view Ed Discussion as a substitute for the discussions we would have in a synchronous lecture format. Each week, I will provide one or more question prompts for you to think about, usually related to material covered in demonstrations. You are also encouraged to use Ed Discussions to ask questions of your instructor and TA related to the material, or start new threads related to your own experiences. The TA will assess discussion periodically and update your grade as the class progresses. Your grade will be based on the number and depth of original contributions as well as responses to classmates' commentary.

Submitting Assignments

All assignments (quizzes, experiential tasks) must be completed and submitted within Canvas. Sending assignments (homework, knowledge checks, exams etc.), whether early, on time, or late to the professors is not permitted and will not be accepted. If there are technical issues, please notify the help desk, as well as the professor immediately. Note that assignments completed on ICE still must be submitted through Canvas.

Assignment Due Dates

All assignments will be due at the times listed above. These times are subject to change so please check back often. Please convert from Eastern Time to your local time zone using a Time Zone Converter.

Late and Make-up Work Policy

Weekly quizzes are due on the dates given on Canvas. The quiz will be left open for 24 hours after the due date, and 10% will be docked if you complete the quiz after the due date. After 24 hours, the quiz will be closed and late submissions will not be possible.

Experiential tasks are similarly due on the dates given on Canvas. Late submissions will be docked 10% per day.

Ed Discussion contributions will be reviewed on an approximately biweekly basis. Given the discussion-based nature of the assignments, late credit is not possible.

If you encounter health or other issues during the semester than make it difficult for you to keep up with the fast pace of the course, please let the instructional team know as soon as possible. Any official medical-related absences are handled by the Dean of Students.

Grading and Feedback

All submissions should be graded within one week of submission, barring unforeseen circumstances. In the event of a grading delay the instructor will notify the class of the issue and estimated time for final grades.

Technology Requirements and Skills

Computer Hardware and Software

- High-speed Internet connection
- Laptop or desktop computer with webcam using either Windows or Mac OS. Suggested minimum requirements:
    - 64-bit multicore processor (quad-core or higher recommended) ○ 16GB RAM
    - 256GB Hard drive (SSD highly recommended) ○ Backup system (GT provides no-cost cloud services)
- Latest versions of Mozilla Firefox, Chrome and/or Safari browsers. Note that Chrome is required for Honorlock.
- Python (I do not recommend 3.14, as some packages are not yet compatible) ○ You are **highly** encouraged to install Python via an Anaconda installation.
    - Jupyter Notebook or JupyterLab (if Anaconda is used, these can be set up with the Anaconda Navigator) ○ Integrated Development Environment (IDE) such as PyCharm or Visual Studio is optional (all demonstrations will use Jupyter)

Note that the demonstrations use large datasets which computers with insufficient RAM may have difficulty handling. You are encouraged to consider using GT's cloud-based computing environment (ICE) in this case. The instructional team will provide instructions on how to access.

In addition, Google Colab environments are available free of charge and should handle any demonstrations. Should you choose to use Google Colab, it is up to you to learn how to set up the environment and how to load data.

Technology Skills

This course is very technology oriented. As noted, you are not expected to have prior programming experience, but general computing proficiency is necessary. For instance, students should be comfortable:

- Navigating a computer operating system & identifying file paths
- Launching and quitting applications (e.g., Jupyter or Jupyter Lab)
- Connecting to the Internet and troubleshooting connection issues
- Downloading, saving, and uploading files
- Accessing remote servers with SSH (instructions will be provided)

Onboarding Quiz and Proctoring Information

All Georgia Tech online degree and certificate students are required to complete the Onboarding Quiz with Honorlock in the first week of the course. Honorlock is utilized for student identity verification and to ensure academic integrity. Honorlock provides student identity verification via facial and ID photos. You may also be asked to scan the room around you. The Onboarding Quiz is needed to help make sure that your identity is verified and that your system is set up to work with Honorlock online proctoring tool. You are required to complete this quiz early in the semester to avoid problems when taking proctored exams.

Technology Help Guidelines

30-Minute Rule: When you encounter struggles with technology, give yourself 30 minutes to 'figure it out.' If you cannot, then post a message to the discussion board; your peers may have suggestions to assist you. You are also directed to contact the Helpdesk 24/7.

When posting or sending email requesting help with technology issues, whether to the Helpdesk, message board, or me use the following guidelines:

- Include a descriptive title for the subject field that includes 1) the name of course 2) the issue. Do NOT just simply type "Help" into the subject field or leave it blank.
- List the steps or describe the circumstance that preceded the technical issue or error. Include the exact wording of the error message.
- When possible, always include a screenshot(s) demonstrating the technical issue or error message.
- Also include what you have already tried to remedy the issue (rebooting, trying a different browser, etc.).

Course Policies, Expectations & Guidelines

Communication Policy

- Email course questions and personal concerns, including grading questions, to me privately using my official GT address (given on first page of syllabus). Do NOT submit posts of a personal nature to the discussion board unless it is a private post.
- Email will be checked at least twice per day Monday through Friday; Saturday and Sunday, email is checked once per day. During the week, I will try to respond to all emails within 24 hours; on weekends and holidays, allow up to 48 hours. If there are special circumstances that will delay my response, I will make an announcement to the class. If you have not heard from me within these windows (24 hours or 48 hours on weekends), feel free to follow up.

- Student Forum/Q&A discussion boards will be checked twice per day Monday through Friday; Saturday and Sunday, these discussion boards will be checked once per day.
- Virtual office hours will be held using Zoom. I will hold **Virtual Office Hours as scheduled in the Canvas course calendar (see there for Zoom link)**, as well as special office hours as needed (such as for special topics, certain assignments, etc.). Special topic hours will be announced in advance. Weekly office hours may be moved due to conflict, which will be announced at least 48 hours in advance. I am also happy to schedule one-on-one office hours via Zoom.
- For questions related to technology, the Digital Learning Support team at https://b.gatech.edu/digitallearningsupport for assistance. You can also reach the Canvas Hotline by phone at 1(877) 259-8498 or by email at support@instructure.com.

Online Student Conduct and (N)etiquette

Although it is not expected to be a problem in a graduate-level class, students are asked to behave in the discussions and other class interactions professionally and civilly. If you are in doubt, do not post it! Instructors reserve the right to remove any postings deemed inappropriate, unprofessional, or otherwise distracting from the course.

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 to 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.

Attendance Policy

This is a 100% asynchronous online course. Regular weekly engagement with course modules is expected; no synchronous attendance is required.

Additional Criteria for Successful Completion

Successful completion of this course is determined solely by the final letter grade as described in the Grading Scale above. There are no additional completion requirements.

Student-Faculty Expectations Agreement

Students and faculty each have a responsibility to maintain an appropriate academic environment as described in the Georgia Tech Student-Faculty Expectations Agreement: https://catalog.gatech.edu/rules/22/

Core IMPACTS Statement

Not applicable.

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/.

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

Generative AI

Generative AI tools (e.g., ChatGPT, Claude, GitHub Copilot) are permitted in this course for the following limited purposes:

- *Self-study and concept review:* You may use AI tools to help clarify course concepts, explore topics discussed in lectures, or generate practice questions for your own learning.
- *Code debugging*: You may use AI tools to help identify errors or understand why existing code is not functioning as expected.

Generative AI tools may **not** be used to develop, write, or generate code submitted as part of any task assignment. All submitted code must reflect your own original work (though you may freely use demonstration code as appropriate). Using AI to produce assignment code, even partially, constitutes a violation of Georgia Tech's Academic Honor Code and will be treated accordingly.

Note that the goal of coding assignments in this course is to build your own skills and understanding. Relying on AI to complete these tasks undermines that purpose and your own professional development. If you are unsure whether a specific use of AI is permitted, ask before submitting.

Collaboration & Group Work

All assignments are individual assignments, so you should not collaborate with others. With that being said, if there is something that is unclear, you are encouraged to pose a question to the class on the discussion board, and the instructor and/or TA may weigh in. You are also encouraged to ask your instructional team (professor, TA(s)) for help with Python.

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 to set up a time to discuss your learning needs.

Copyright

Among the materials that may be protected by copyright law are the lectures, notes, and other material presented in class or as part of the course. Always assume the materials presented by an instructor are protected by copyright unless the instructor has stated otherwise.

Student-Faculty Expectations Agreement

At Georgia Tech we believe that it is important to strive for an atmosphere of mutual respect, acknowledgment, and responsibility between faculty members and the student body. See https://catalog.gatech.edu/rules/21/ for an articulation of some basic expectations 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.

Course Schedule

All assessment due dates are provided in Canvas.