

## **CHEM 4414/6414 Syllabus**

Intro Data Sci & ML Sci – 3 credits]

Fall 2026

### **Instructor Information**

---

**Instructor: Anh Thu Le**

**Email: [anh.le@chemistry.gatech.edu](mailto:anh.le@chemistry.gatech.edu)**

### **General Course Information**

---

#### **Description**

Data Science has become one of the most rapidly growing fields in all of computing. From advertising targeting individual users on Facebook to predicting future stock market behavior, the principles of Data Science can be applied to solve problems in an incredible diversity of contexts. In this course, students will learn the basic principles of Data Science and develop skills working with the most common tools in the world of Data Science, building from foundational experience with computer programming in the highly versatile Python language. The knowledge and skills developed in this course will therefore be directly transferable to students' future careers in the science, technology, or business sectors.

#### **Course Learning Outcomes**

Upon successful completion of this course, students will be able to...

1. Implement efficient, readable Python programs,
2. Acquire and process datasets for model building and refinement,
3. Apply machine learning/data science principles to solve real-world problems using relevant tools,
4. Utilize various Python libraries common in data science applications, and software projects.

#### **Required Course Materials**

##### **Course Text**

1. J. VanderPlas, Python Data Science Handbook: Essential Tools for Working with Data (O'Reilly, Boston, MA, 2017).

Note: This book should be accessible via the web at

<https://learning.oreilly.com/library/view/python-data-science/9781491912126/> for free with your GT student credentials. Click "Sign In", then enter your GT email address. This

will take you to the GT SSO authorization page and you should have access. There are other python/data science books freely available through O'Reilly books, if more advanced texts are needed for individual students. Additionally, C.M. Bishop, Pattern Recognition and Machine Learning, is an advanced text worth looking into for very advanced students.

### **Course Website and Other Classroom Management Tools**

We will use Canvas as our primary course website. This can be accessed by signing into your account at [www.canvas.gatech.edu](http://www.canvas.gatech.edu). Occasionally throughout the semester, additional notes, videos, solutions, or other relevant resources will be posted to the Canvas website. You will be responsible for their content, unless otherwise indicated. Materials to be utilized in class will be pushed to Canvas in advance of the class meeting.

### **Grading Policy:**

Your final grade will be assigned a letter according to a weighted average of the quality of work submitted and participation throughout the course.

### *Assignments*

There are total of 8 HWs for a total of 80% and a final project that worth 20%.

### **Description of Graded Components**

#### **Grading Scale**

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

*(the percentage listed will guarantee you the letter grade shown)*

A	90-100%
B	80-89%
C	70-79%
D	60-69%
F	0-59%

### **Course Policies**

---

#### **Attendance and/or Participation**

We will meet twice/week in person. It is important that everyone attends class as collaborative coding assignments are performed. This class is oftentimes more of a workshop than a lecture (although ~50% of classes are likely to be lectures). Please let me know if you cannot be in class. Classroom participation will be determined through informal “formative assessments” (FAs) during class. While these FAs (answering questions, participating in coding, helping others, surveys, etc.) are not graded for correctness, you should use them as a tool to assess your own learning by (i) identifying areas or concepts with which you are struggling and (ii) to gauge your progress and improvement as you study those concepts in particular. Just as you will be expected to use these assessments to guide

your learning, the instructor will be using the collective results from these assessments to guide lesson planning and classroom approach.

I do understand, however, that life happens—you may get sick, or some other event may occur which prevents you from attending class. Despite the fact that attendance is not taken, I strongly suggest only to miss class for excused absences, defined by the Institute as “documented reasons of illness, family emergency, or participation in approved Institute activities” (please see <http://catalog.gatech.edu/rules/12/> for more information). I make this recommendation because I design each classroom activity, discussion, and experiment to address a specific goal: to help you make measurable progress towards the course learning objectives. Therefore, your attendance at all class meetings (including lecture and any recitation or review session(s)) and active participation in these activities throughout the semester will be critical for your success in this course

### **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. Review [Georgia Tech’s Honor Code](#) and the student [Code of Conduct](#).

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

### **Core IMPACTS**

[Core IMPACTS](#) is the University System of Georgia’s General Education curriculum. If you are teaching a course that counts towards Core IMPACTS, you should include a syllabus statement about the Core area and associated [career competencies](#). [This resource](#) developed by the Center for Excellence in Teaching and Learning and Online Education at Georgia State University includes template syllabus statements for each of the Core IMPACTS areas that you may adapt for your course.

### **Accommodations for Students with Disabilities**

If you are a student with learning needs that require special accommodation, [contact the Office of Disability Services](#) (404-894-2563) 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. [The Student-Faculty Expectations](#) articulate 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.

### **Pre- &/or Co-Requisites**

*This course is intended for upper-level undergraduate and graduate students in Chemistry & Biochemistry, who want to develop their knowledge and skills in Data Science and Python programming. While it will be helpful to have at least a basic knowledge of computer programming in Python (or any other language), it is not required. Python will be the language of this course, however, so it is expected that students acquire this basic proficiency quickly in the semester as we speed through python basics. Other useful background knowledge: Calculus of one or more variables; Introductory Linear Algebra*

### **Collaboration, Group Work, and Use of Generative AI**

*Effective education is collaborative by nature. I encourage you to consult your fellow students, TA, instructors, other faculty, and/or staff throughout this course. However, all assignment submissions must be your own. This means that you must create your own derivations, solutions, analysis, thoughts, justifications, data tables, results, plots, calculations, etc. on your own. Copying and pasting figures, tables, or text, even from group partners, constitutes plagiarism and is a violation of the GT Honor Code. Please see the Georgia Tech Academic Honor Code (<http://osi.gatech.edu/plugins/content/index.php?id=46>).*

### **Extensions, Late Assignments, & Re-Scheduled/Missed Exams**

*Unless expressly indicated otherwise, all assignments may be submitted on Canvas up to the start of class on the due date. Assignments submitted more than three calendar days late will not be accepted. Extensions on assignments will be given only for documented reasons of illness, family emergency, or participation in approved Institute activities (such as field trips and athletic events, see <http://catalog.gatech.edu/rules/12/> for more information). Presentations and/or demonstrations of capstone projects, scheduled to be given during the final exam period for the course, will be re-scheduled in accordance with Georgia Tech's policy, also found at <http://www.catalog.gatech.edu/rules/12/>.*

### **Student Use of Mobile Devices in the Classroom**

*Electronic devices (e.g., tablets/laptops) are required for each class meeting, as they are necessary for classroom participation. It will often be necessary for students to share their screens during the coding classes. Please contribute to class and try to refrain from use of these electronic devices for other purposes during class (e.g., Facebook, texting, phone calls, etc.). These may disrupt the classroom learning environment for all other students. This class really is a workshop that relies on everyone participating and contributing.*

## Campus Resources for Students

---

### Undergraduate Student Academic Success Resources:

- Academic Support: Academic Success and Advising (a unit in the Office of Undergraduate Education & Student Success) provides free support for your courses. Students can attend scheduled supplemental review (PLUS) sessions, stop by Drop-In Tutoring, or schedule a one-on-one appointment through Knack. To explore what options work best for you, please visit us online at [success.gatech.edu/tutoring](https://success.gatech.edu/tutoring), email us at [tutoring@gatech.edu](mailto:tutoring@gatech.edu), or come see us at Clough Undergraduate Learning Commons, Suite 283.

### Graduate Student Academic and Professional Success Resources:

A list of resources for graduate students is given on the [Office of Graduate and Postdoctoral Education](#) website. Specific information for [current graduate students](#) includes

- [Academic Resources](#) such as the Communications Center, Language Institute, Library, Catalog, Registrar, resources for conducting research, Advocacy and Conflict Resolution resources, and how to manage unexpected situations that may impact your academic performance;
- [Student Resources](#) such as Campus Services, Child Care/Family programs, Health & Wellness, Career Services, and the Student Resource Guide; and
- [Professional Development](#) such as the programming from the Career Center and other professional development resources and events”

### Student Well-Being:

At Georgia Tech, we are concerned about your overall physical, social, and mental well-being. A [comprehensive list](#) of wellness related resources has been compiled and maintained by the Office of the Vice President for Student Engagement and Well-being ([student-resource-guide \(gatech.edu\)](#))

### Practical Advice for Success

Programming and Data Science are challenging subjects, but they are extremely rewarding to study thanks to their growing application and influence. Because I am committed to helping you to succeed in this course, we have determined the following general advice (in no particular order) to help you to do so.