MENG 15100 / PHSC 15100
Machine Learning and Artificial Intelligence for Molecular Discovery and Engineering
Autumn 2025

Instructor:
Prof. Andrew Ferguson
ERC 389
andrewferguson@uchicago.edu

Teaching Assistants:
Luna Ramirez Hincapie
Molecular Engineering Modular Offices (MeMo) 125
lunaramirez@uchicago.edu

Hugh Yeh
Molecular Engineering Modular Offices (MeMo) 125
hughy@uchicago.edu

Lecture Time and Location:
Hinds Lab Geo Sci 101 • 3-4:20 pm • Mon, Wed

Lab Time and Location:
Kersten Phys Teach Center (KPTC) 101 • 4:30-5:50 pm • Wed
Kersten Phys Teach Center (KPTC) 101 • 3:30-4:50 pm • Thu

Office Hours Times and Locations:
Eckhardt Research Center (ERC) 381 • 5-5:50 pm • Tue

Dates:
Mon Sep 29 – Sat Dec 13, 2025

Course Overview:
The foundations of artificial intelligence can be dated back to Alan Turing's seminal 1950 paper in which the concept of "thinking machines" was first introduced and the 1956 Dartmouth Summer Conference in which the term artificial intelligence was first coined. The field has since passed through multiple epochs of development, and today artificial intelligence and machine learning are ubiquitous and enabling tools that pervade all corners of engineering discovery and practice. This course will survey the conceptual and historical basis of artificial intelligence and machine learning; the role, integration, and ethics of these approaches in modern engineering practice; the mathematical and algorithmic underpinnings of some popular machine learning techniques; and selected applications of these tools in particular applications in molecular engineering drawn from areas such as energy and sustainability, drug discovery, quantum science and information, materials science, and bioengineering.

Quantitative analysis and numerical calculations will be an important part of the course, but math beyond algebra will not be required. We need some math to understand the theory of machine learning and artificial intelligence, and we need to run code to put it into practice. We will teach you the math and coding that you need for this course (which is not very much). We will only ask you to do simple math and make simple edits to pre-written code.

Learning Objectives:
• A perspective on the role of artificial intelligence and machine learning in molecular engineering
• An understanding of some key machine learning tools and their use cases in engineering practice
• Experience in applying some elementary machine learning tools to solve molecular problems

Topics:

I. History of AI + AI in Molecular Sciences (2 lectures)
II. Google Colab w/ Python (2 lectures)
III. Linear Regression, Regularization, and Cross Validation (4 lectures)
IV. Dimensionality Reduction and Clustering (3 lectures)
V. Artificial Neural Networks (3 lectures)
VI. Ethics of AI (1 lecture)
VII. Adv. Topics: Generative Molecular Design (1 lecture)

Lab Modules:

A series of computational lab modules will give students opportunities to put into practice the machine learning concepts and tools covered in the course lectures.

- **Lab 1 – Google Colab w/ Python**
  Learning how to conduct machine learning using Python within Google Colab environment
- **Lab 2 – Regression**
  Learning how to construct and assess regression models for molecular property prediction
- **Lab 3 – Dimensionality Reduction**
  Learning how to construct reduced-dimensional projections of molecular data
- **Lab 4 – Artificial Neural Networks**
  Learning how to build, train, and run simple artificial neural networks for molecular design

Course Texts:

V. Maini and S. Sabri "Machine Learning for Humans" (2017) – required (free)
https://medium.com/machine-learning-for-humans/why-machine-learning-matters-6164faf1df12
A. Burkov "The Hundred-Page Machine Learning Book" (2019) – required ($20)
https://leanpub.com/theMLbook
E. Alpaydin "Introduction to Machine Learning" (3rd ed.) MIT Press (2014) – optional, advanced
S.J.D. Prince "Understanding Deep Learning" MIT Press (2023) – optional, advanced

Course Components:

60% 4 lab reports
15% 1 midterm exam
25% 1 final exam

Lab Reports:

Lab exercises associated with each lab module will be typically computational assignments designed to give students hands-on experience putting into practice the concepts and tools developed within the course lectures. Lab exercises will typically be conducted in Google Colab (see below) and students will submit a Colab notebook ipynb file containing the analyses and short form answers as their lab report. Provision to work on the lab exercises will be made during the scheduled lab times, but students should also expect to spend time outside of the scheduled class time to complete these assignments

Exams:

Two written exams – a midterm and a final – will be used to assess student understanding of the key concepts and algorithms from the course lectures and lab modules. Exams will be proctored and closed-book, but students will be permitted to bring a **single, handwritten, double-sided, US letter-sized (8.5 x 11 inches) piece of paper** as an examination aid (colloquially known as a "cheat sheet"). Students will also require a pen/pencil, eraser, and straight-edge/ruler. A calculator will not be required – all mathematical operations will be easily handled by mental or simple pencil-and-paper calculation.

Grades:
Letter grades will be based on final aggregate student scores, with numerical cutoffs specified by the instructor. However, students with aggregate scores >95% are guaranteed *at least* an A, >85% *at least* a B, and >75% *at least* a C (i.e. cutoffs for these letter grades will not be higher than these values).

Late Assignments:
Late submissions will not be accepted, but students with legitimate requests for extensions should contact the instructor well in advance of the due date.

Laptop:
A laptop computer with web access is required for the hands-on components of the lectures and labs in this course. If you do not have access to a laptop, the instructor may be able to help make arrangements for a laptop loan.

Canvas:
Course announcements, materials, and grades will be posted via Canvas (https://canvas.uchicago.edu), and course deliverables will be submitted via this portal. It is students' responsibility to check Canvas for announcements and updates. Canvas is best viewed using Chrome or Firefox web browsers; Safari is known to cause problems.

Slack:
We will be using Slack as a virtual and public forum for questions discussion. The system allows for fast responses from the instructor, TA, and classmates, and permits students to see previous questions and answers. Rather than emailing questions to the teaching staff or using Canvas chat, please post your questions publicly on Slack. If you have not already been automatically enrolled, please sign up via https://meng15100.slack.com

Jupyter + GitHub:
Jupyter notebooks are user-friendly interactive programming environments that can be run through web browsers and combine code, text, equations, comments, and multimedia within a single document that is useable, readable, reproducible, and sharable. Hands-on Jupyter notebooks employing Python code will be used to (i) complement lecture notes to promote active learning and immersive experience in the concepts and algorithms presented during class and (ii) form the basis of lab assignments for the course. All notebooks are made available through a dedicated repository housed on the free code sharing platform GitHub at: https://github.com/andrewlferguson/MENG15100

Google Colab:
All Jupyter notebooks are designed to run on Google Colab (short for Google Colaboratory) as a cloud platform that provides free compute resources to host and run Jupyter notebooks: https://colab.research.google.com/ A free Google account will be required to access and use these resources. Students who do not already have a Google account should sign up for one here: https://accounts.google.com/

PollEverywhere:
Poll Everywhere is a web-based application for live interactive polls that we will use for short in-class quizzes. Polls can be completed using smartphones, laptops, and text messages via https://pollev.com/andrewferguson

Collaboration:
Collaboration between students on individual assignments is permissible and encouraged – peer learning is a very effective means to understand and master the course material – but each student must conduct their own calculations and write their own reports. Individual submissions must be the individual effort of each student – work you submit as your own must indeed be your own. Group assignments must be the collective work of the group, and all members of the group must materially contribute to the submitted assignment.

Plagiarism:
Plagiarism will not be tolerated and verified incidents will result in all parties receiving a zero and formal academic sanctions. Students are responsible for awareness of the definition and penalties for plagiarism in the Student Manual (https://studentmanual.uchicago.edu/academic-policies/academic-honesty-plagiarism/). Details on what does and does not constitute plagiarism is available here: https://internationalaffairs.uchicago.edu/page/honest-work-and-academic-integrity-plagiarism

Artificial Intelligence Tools:
AI-based language models (e.g., ChatGPT, Bard, or GitHub Copilot) are now prevalent and useful tools in many facets of modern life. The role and implications of these technologies in academic education and research is still a matter of discussion and debate, but they do present a potential danger to ethical behavior in the classroom. **It is ethical** to use these tools to assist you in understanding materials (e.g., summarizing a technical article, walking you through challenging mathematical developments) and communicating your own work and thoughts (e.g., improving technical writing, informing word choices). If you do choose to use these tools in this manner, you should provide a written acknowledgement of these tools in the same way as you would cite any other external source used to complete an assignment. **It is <u>not</u> ethical** to use these tools to produce responses that you present as your own in the completion of your assignments for this course. This is plagiarism – you are presenting the work of another as your own. This also compromises your learning within the course, misleads the instructor as to your understanding and mastery of the course material, and runs counter to the ethos and goals of education and learning. Furthermore, these tools are far from infallible and may not provide you trustworthy or correct answers. Since you will not have access to these tools during exams, reliance on them in completing the course assignments will likely lead to poor examination performance.

Disability Accommodations:
The University of Chicago is committed to ensuring equitable access to our academic programs and services. Students who have been approved for academic accommodations by Student Disability Services (https://disabilities.uchicago.edu) and need a reasonable accommodation to participate fully in this course should follow the procedures established by SDS. Timely notifications are required to ensure that your accommodations can be implemented. Please set up a time to meet with the instructor to discuss your access needs in this course after you have completed the SDS procedures for requesting accommodations.

Academic Integrity, Civility, and Personal Responsibility:
The University of Chicago is committed to providing a safe and welcoming environment for rigorous learning, inquiry, and research within a climate of respect, civility, and inclusion. The University statements on diversity and non-discrimination are available at https://provost.uchicago.edu/handbook/clause/university-statements-diversity-and-non-discrimination. Every member of the University of Chicago intellectual community is responsible for upholding the highest standards of academic integrity, mutual respect and civility, and personal responsibility. Within the classroom, these values are critical for the success of our educational mission and your growth and development as a scholar. The expectations around integrity, civility, and responsibility are available at https://college.uchicago.edu/student-services/academic-integrity-student-conduct. Resources for reporting instances of bias or discrimination are available here https://cares.uchicago.edu/.

Lecture and Lab Schedule:

Class	Day	Module	Learning Objectives	Readings	Lab	Due
1	M 9/29	I. History of AI & AI in Mol Sci	What is AI, what isn't AI, historical origins of AI, Turing test, AI vs. ML vs. DS, taxonomy of techniques, AI in molecular sciences			
2	W 10/1	I. History of AI & AI in Mol Sci				
					Lab 1: Colab Python	
3	M 10/6	II. Google Colab w/ Python	Scalars, vectors, slopes, familiarity with Google Colab, intro to Python programming, using ML/AI libraries, plotting, data analysis	Maini & Sabri, Ch. 1; Burkov, Ch. 1, 2.1 (1st para only 2.1.9), 2.6, 2.7, 2.8, 2.9		
4	W 10/8	II. Google Colab w/ Python				
					Lab 1: Colab Python	
5	M 10/13	III. Linear Reg., Regularization, and Cross Validation	Linear regression math, introducing nonlinearities, regularization (ridge and lasso), cross-validation	Maini & Sabri, Ch. 2.1; Burkov, Ch. 3.1, 5.3-5		**Lab 1: Colab Python**
6	W 10/15	III. Linear Reg., Regularization, and Cross Validation				
					Lab 2: Regression	
7	M 10/20	III. Linear Reg., Regularization, and Cross Validation				
8	W 10/22	III. Linear Reg., Regularization, and Cross Validation				
					Lab 2: Regression	
9	M 10/27	IV. Dim. Red. and Clustering	Curse of dimensionality, dimensionality reduction, interpretable learning, latent space, compression/decompression, clustering	Maini & Sabri, Ch. 3 (PCA+k-mean only); Burkov, Ch. 9.2.1, 9.3.1		**Lab 2: Regression**
10	W 10/29	IV. Dim. Red. and Clustering				

					Lab 3: Dimension Reduction	
11	M 11/3*	IV. Dim. Red. and Clustering + Midterm Review	Recap of Modules I-IV			
12	W 11/5*	Midterm Exam				**MIDTERM EXAM**
					Lab 3: Dimension Reduction	
13	M 11/10	V. Artificial Neural Networks		Maini & Sabri, Ch. 4; Burkov, Ch. 3.2.1, 6.1, 6.2 (stop after 6.2.1: CNNs)		**Lab 3: Dimension Reduction**
14	W 11/12	V. Artificial Neural Networks				
					Lab 4: Artificial Neural Networks	
15	M 11/17	V. Artificial Neural Networks				
16	W 11/19	VI. Ethics of AI	Fairness, transparency, accessibility, potential for harm, AGI, P(doom), cheating, learning, privacy	https://youtu.be/nB_EjxoP-6w; Marukkula.pdf; NYMag.pdf		
					Lab 4: Artificial Neural Networks	
–	M 11/24		**THANKSGIVING BREAK**			**Lab 4: Artificial Neural Networks**
–	W 11/26		**THANKSGIVING BREAK**			
					NO LABS	
17	M 12/1	VII. Adv. Topics: Generative Molecular Design	Guest Lecture			
18	W 12/3	Course Review	Recap of Modules I-VII			

	T 12/9		Final Exam Hinds Lab Geo Sci 101 3-5 pm Tue Dec 9			**FINAL EXAM**

* Prof. Ferguson may be on travel these dates. Alternative arrangements TBA.