COSC 410B: Applied Machine Learning

Spring 2024, 1 Credit

Time	TR 1:20-2:35 pm
Location	315 McGregory Hall
Instructor	Dr. Grusha Prasad gprasad@colgate.edu 304 McGregory Hall Pronouns: She/Her
Office Hours	T 4-5:30 pm; F 1-2:30 pm Or schedule a meeting via email
Important links	Course website Schedule with links to slides

Course Overview

Catalog description

This course provides a practical introduction to applied machine learning. Students engage in supervised and unsupervised machine learning algorithms, including regression, support vector machines, decision trees, nearest neighbors, clustering, and ensemble methods. Students also learn deep learning techniques, including feedforward, convolutional, and recurrent neural networks. Emphasis is placed on understanding and gaining hands-on experience with machine learning for practical use.
Prerequisite: COSC 202, COSC 290 Co-requisite: COSC 410L

Learning objectives

- Describe different types of Machine Learning algorithms, and explain why a specific algorithm is or is not appropriate for a given problem or task.
- Describe different methods of evaluating Machine Learning algorithms, and identify which evaluation methods are appropriate for different tasks.
- Implement the core building blocks of different Machine Learning algorithms from scratch in Python.
- Describe the standard Machine Learning pipeline and implement this pipeline for different types of models using off-the-shelf Python libraries.
- Design, conduct, and present a Machine Learning research project using self-collected or publicly available datasets.

Textbook

The required text for the course is Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron. The 3rd edition of this book can be ordered through the Colgate Bookstore. If you already happen to have access to the 2nd edition, it is okay for you to use that.

Course organization

Overview of course work and grading scheme

The course has seven programming assignments, two midterm exams and a final project. You will be working on the final project in several stages. The dates for all of these assignments and exams are listed on the course schedule. You will also be expected to do all of the required reading and actively participate in class.

Assignment	**% of Grade**
Programming assignments (6; equally weighted)	30%
Midterm exams (2; equally weighted)	~~25%~~ 23%
Final Project - Project idea 5% - Project proposal (group) 10% - Poster presentation (group) 10% - Final paper 15%	40%
Required reading, Participation and Reflection	~~5%~~ 7%

Letter Grade Assignments

F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A	A+
< 60	60-62	63-66	67-69	70-72	73-76	77-79	80-82	83-86	87-89	90-92	93-99	*

Programming assignments and labs

Each programming assignment is based on a problem or a task that is introduced during a lab session. The problem will be based on the course material and lecture sessions leading up to the lab. The structure of each lab session is as follows:

- You will either work individually or with a partner to understand the problem, and come up with an approach or pipeline to solve the problem. Some labs may involve executing existing code and interpreting the result, or writing new code. If group work is permitted for those labs, you are allowed to execute and/or write the code together.

- As you work, your solution will be checked by the TA and instructor. You'll be given feedback and guidance on how to improve your write-up.
- Once the write-up is complete, you will spend some time individually or with a partner to understand the expectations for the homework assignment associated with the lab, and come up with a detailed plan or checklist for the homework. You can have this checked by the instructor before you leave.

After the lab session, you will complete the implementation of the homework, **on your own** (see Academic Honesty and Collaboration Policy for more details).

Exams

The exams are designed to test your understanding of the core concepts covered in class, and your ability to apply them in different contexts. They will not involve any programming component. The scope of each exam is indicated on the course schedule. The exams will happen during the lab section. **No discussion of exam questions or possible solutions is permitted from the time the exam is first administered until exams have been completed by all students and returned.**

Participation

Active participation involves attending class, doing all of the required readings, participating in class discussions/ problem solving sessions and writing two self-evaluations. These self-evaluations are intended to help you synthesize and contextualize key skills and knowledge acquired through the course of this class, and help me improve as an instructor.

Course Policies

Academic honesty and collaboration

Beyond the discussions that happen in the lab section, you are not allowed to discuss the specifics of programming assignments or share your code with other students in any of the sections of this class. If you get help from other sources (e.g., instructors, TAs, internet, generative AI etc), you should cite the sources and the nature of help that you got at the end of your code in comments. Here are some examples of citations:

- "I ran into a bug with function X in my code. I discussed this with person Y and discovered that the source of the bug was Z."
- "I did not understand concept X. So I conversed with chatbot A to clarify some misconceptions. I verified the insights from this conversation using websites Y (link) and Z (link)."
- "I forgot how the syntax for X worked in Python. I found the solution on website Y (link)"

If you are struggling to complete an activity—due to a lack of understanding, time pressure, personal matters, etc – please contact me. I want you to succeed, and I will do everything I can to help you. Violating the academic honesty expectations for the course (almost) always leads to a bad experience for you and me. If you are unsure what constitutes academic misconduct, please contact me as soon as possible.

Collaborating with peers in the class

You may discuss course concepts, generic aspects of python, and work through the logic of something you don't understand with your peers. However, you should not share code (including pseudo code). Your submitted work should be your own. Here are some concrete rules that exemplify this (but are not intended to be comprehensive):
Do NOT:

- Ask a peer in either section of the course to debug your code.
- Ask a peer for pseudo code for an algorithm needed for a homework assignment.

You CAN:

- Ask clarification questions about the fundamentals of programming (e.g., "How do I create a class in Python?")
- Ask for conceptual clarifications (e.g., "What is the difference between regression and classification?")
- Try to work through the logic of something you don't understand (e.g., "How is data formatted for a neural network?")

Using generative AI tools

Generative AI systems (like ChatGPT), if used correctly, can serve as powerful tools for learning and idea refinement. In this course, you can use generative AI systems to learn about concepts iteratively through a conversation (much like you would have a conversation with a peer, TA or an instructor). However, you cannot ask these systems to directly give you answers or write code for you. One reason for this is because the answers that the system generates can be inaccurate (no matter how confident the system might sound). But more importantly, I believe the intellectual growth you can get from working through a difficult problem and discovering the answer for yourself cannot be replicated by just reading a pre-generated answer. Here are some concrete rules that exemplify this (but are not intended to be comprehensive):

Do NOT:

- Give the model a problem description and ask it to sketch an algorithm for you or write you pseudo code.

- Give the model the homework description and ask it to organize the code for you (e.g., generate the necessary function headers, write the main functions etc).
- Give the model a function description and ask it to generate code for you.
- Have your conversation with the model and your assignment open at the same time. Use your conversation with the AI as a learning experience, then close the interaction down, open your assignment, and let your assignment reflect your revised knowledge.

Using the AI system in ways as described above will count as **cheating** even if you cite the AI system as a source.

You CAN:

- Ask clarification questions about the fundamentals of programming (e.g., "What happens when I try to modify a list within a function in Python?")
- Ask for conceptual clarifications (e.g., "What is the difference between recall, precision and AUC?")
- Try to work through the logic of something you don't understand (e.g., "How does the gradient tell the model which direction to move in during optimization?")
- Given a problem description and your proposed algorithm and "talk" through the potential fallacies.

Note, for any of these models having the correct "prompt" is necessary. So you may have varying levels of success using these models to gain conceptual understanding, and in many cases just talking to your instructors/ TA/ peers or even doing straight up googling is likely to yield better results. If you do decide to use these models, it is your responsibility to also fact check the insights that you gain.

Remember: Policies around the use of Generative AI tools, like any other course policies, vary across different courses both within and outside the department.

Late homework

There are a lot of homework assignments in this course, and therefore once a student gets behind, it becomes increasingly difficult to keep up. Good planning is essential and starting homework early is good practice. Plan to give yourself time to step away from your work and return to it later.

That being said, I understand that life happens and plans do not always pan out. Therefore, if you anticipate requiring additional time on a homework assignment, please reach out to me in advance so we can discuss an alternative deadline. As long as you have made a good faith effort to complete the assignment by the original deadline, I am willing to offer a reasonable

extension. I will be less willing to grant an extension if you wait to start an assignment until two days before it is due, repeatedly ask for extensions, etc.
Any homework that is turned in late without prior approval will receive only a maximum of 50%.

Reflect and retry

The midterm exams are not intended to be points of stress, but rather as learning opportunities where you can discover what you do not yet understand. Your final grade should reflect not just what you know when you take an exam, but also your ability to learn and improve from feedback. Therefore for each midterm you have the option to "reflect and retry". For any questions that you answered incorrectly on the midterm, you can get back up to 50% of the points you lost on that question if you reflect on why your answer was incorrect and explain what the correct answer is. Any re-submission must be submitted **within a week** of getting your exam back.

Attendance

While missing a couple of classes is ok (cf. "life happens" in the Late Homework section), missing classes more frequently without prior approval will negatively impact the participation component of your grade. Please reach out to me in advance if you anticipate missing more than two classes. You are responsible for making up for any of the missed content on the days that you are absent either through office hours or by talking to your peers.

Anonymous feedback

If you have feedback or suggestions for how the course can be improved (or how it is going well), please share them with me at any point in the semester using this form.

Unexpected Circumstances

If unexpected circumstances arise that could impact your involvement in the course (inability to attend class, complete the homework on time, etc.), please let me know as soon as possible so that we may design appropriate accommodations. Usually these accommodations will be made in consultation with your Administrative Dean.

Additionally, any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact their Administrative Dean for support. Furthermore, please notify your instructor if you are comfortable in doing so, so we may provide any resources available.

Non-Discrimination

Colgate and its faculty are committed to supporting our students and seeking an

environment that is free of bias, discrimination, and harassment. Colgate does not discriminate on the basis of race, color, sex, pregnancy, religion, creed, national origin, citizenship status, physical or mental disability, age, marital status, sexual orientation, gender identity and expression, veteran or military status, predisposing genetic characteristics, domestic violence victim status, or any other protected category under applicable local, state, or federal law. If you have encountered any form of discrimination or harassment, including sexual violence (e.g. sexual assault, stalking, dating violence, or stalking), we encourage you to notify the Title IX Coordinator, Marilyn "Lyn" Rugg (mrugg@colgate.edu, 102 Lathrop Hall, 315-228-7288), the Counseling Center, the Chaplain's Office, or Student Health Services.

Privacy Protection

Lecture materials and live and pre-recorded class-related materials that depict the images or voices of students are protected under the Family Educational Rights and Privacy Act (FERPA). It is a violation of federal law to share these materials with anyone outside of the classroom. In addition, all lecture and classroom materials are the sole intellectual property of the instructor. Any use or distribution of these materials requires the explicit written permission of the instructor.

Getting Help

A key to your success at Colgate, and in life in general, is figuring out what resources are available and using them to help you achieve your goals. For any homework problems or other class-related questions that you have, there are several options for getting help. Please take advantage of these opportunities!

1. Come see me during office hours (or send me an email)!
2. Go to the TA office hours.
3. Ask questions in whatever format you feel most comfortable in. If you are usually shy and prefer to ask questions outside of class, let me know if there is anything I can do to make you feel comfortable asking the question in class.
4. Form a study group with other students in the class and work together on a regular basis (note the collaboration policy above).

In addition, please be aware of the great resources that Colgate provides:

Academic Support and Disabilities Services.

If you feel you may need adjustment based on the impact of a disability, you should contact your instructor privately to discuss your specific needs. If you have not already done so, please contact the Director of Academic Support and Disability Services at 315-228-7375 in the Center for Learning, Teaching, and Research. Reasonable and appropriate adjustments for students

with disabilities are determined on a case-by-case basis to ensure that members of the community with disabilities have access to Colgate's programs and services. The director also assists students in identifying and managing the factors that may interfere with learning and in developing strategies to enhance learning.

NASC Liaison Group

NASC liaisons are a group of natural science and mathematics faculty members dedicated to providing science-interested students from underrepresented groups with mentorship, motivation, and individualized support as they navigate their paths in the sciences at Colgate. NASC liaisons do not replace the role of an academic advisor or offer formal academic advising. Rather a NASC liaison may meet one-on-one with a student to give another perspective on their academic plan; give tips on effective studying; or introduce a student to upper-class peers, alumni, or other faculty members that might be able to help them. The roles of NASC liaisons will depend on students' needs, and we encourage students to reach out for mentorship and moral support. To find out more about the group or to contact a member, visit the NASC division webpage.

Information Technology

The Information Technology Service Desk is located on the third floor of Case-Geyer Library and provides services to all students across campus. The help desk consultants assist with problems concerning email, Portal, Moodle, and problems with your personal laptops. Talk to your instructor if problems with your personal computer are affecting your ability to get your work done.

Counseling Center

College is hard and life can be hard. If you are experiencing emotional or personal difficulties, the Counseling Center offers completely confidential and highly professional services.

Administrative Deans

Each student is assigned an Administrative Dean who can advise you regarding personal and/or academic matters. Administrative deans often assist students to understand policies and procedures, navigate personal challenges, work with faculty, and engage with parents.