10716: Advanced Machine Learning: Theory and Methods

Instructor:

Pradeep Ravikumar, pradeepr@cs.cmu.edu

Teaching Assistants:

Ananye Agarwal   ananyea@andrew.cmu.edu
Lili Chen   lilic@andrew.cmu.edu

Lectures:

POS 151, Tuesdays and Thursdays, 3:30 – 4:50 PM

Office Hours:

The office hours of the instructor, and TAs, with all the zoom links, will be posted on Piazza.

Key Course Links:

Website: http://www.cs.cmu.edu/~pradeepr/716
Piazza: https://piazza.com/cmu/spring2024/10716
Gradescope: https://www.gradescope.com/courses/691691

Course Description:

Advanced Machine Learning is a graduate level course introducing the theoretical foundations of modern machine learning, as well as advanced methods and frameworks used in modern machine learning. The course assumes that students have taken graduate level introductory courses in machine learning (Introduction to Machine Learning, 10-701 or 10-715), as well as Statistics (Intermediate Statistics, 36-700 or 36-705). The course treats both the art of designing good learning algorithms, as well as the science of analyzing an algorithm's computational and statistical properties and performance guarantees. Theorems are presented together with practical aspects of methodology and intuition to help students develop tools for selecting appropriate methods and approaches to problems in their own research. We will cover advanced machine learning methods such as nonparametric and deep compositional approaches to density estimation and regression; advanced theory such as fundamentals of clustering, classification, boosting; theory and methods at the intersection of statistical and computational efficiency; as well as vignettes of theoretical results on some hot topics such as robustness and explainability.

Pre-requisites:

Basics of probability and statistical inference, and basics of machine learning (such as regression, classification, clustering).

This course is for students who have already taken introductory courses in machine learning and statistics, and who are interested in deeper theoretical foundations of machine learning, as well as advanced methods and frameworks used in modern machine learning.

<u>The course will be very math and theory heavy.</u>

Course Goals:
1. Understand statistical and computational considerations in machine learning methods.
2. Develop the skill of devising computationally efficient and yet statistically rigorous algorithms for solving machine learning problems.
3. Understand the science of modern statistical analysis.
4. Develop the skill of quantifying the statistical performance of any new machine learning method.

Class Website:
http://www.cs.cmu.edu/~pradeepr/716
The class schedule, and lecture materials will be posted there (and on Piazza).
Every other Friday we will have a recitation.

Discussion, Announcements:
We will be using Piazza for class discussion. The system is highly catered to getting you help fast and efficiently from classmates, the TA, and myself. Rather than emailing questions to the teaching staff, I encourage you to post your questions on Piazza. If you have any problems or feedback for the developers, email team@piazza.com.
We will also be using Piazza for announcements, and providing resource materials
Find our class page at: https://piazza.com/cmu/spring2024/10716

Textbooks:
Lecture notes will be posted for each class, which will be largely self-contained.

Course Policies:

Homeworks:
There will be **four** homework assignments, approximately evenly spaced throughout the semester. The assignments will be posted on the course website, and on Piazza. You will get a late day quota of 8 days, which you can distribute among the four homeworks as you wish, subject to a maximum of 3 days per homework. Homeworks submitted after your late day quota will lose all points. The homework schedule is posted right at the beginning of the

semester, so please plan in advance. <u>We expect you to use the (very generous!) late day quota for conference deadlines and events of the like, so we cannot provide an additional extension for such cases.</u> In the case of an emergency (sudden sickness, family problems, etc.), we can give you a reasonable extension. But we emphasize that this is reserved for true emergencies.

Gradescope:

We will use Gradescope for submitting, and grading assignments. If you believe an error was made during manual grading, you'll be able to submit a regrade request on Gradescope. For each homework, regrade requests will be open for only 1 week after the grades have been published. This is to encourage you to check the feedback you've received early!

Exams:

There will one midterm, and one final for the course.

Class project:

There will be a class project. You can form groups of up to 3 students. Further details on the project will be provided on the website (and on Piazza).

Grading:

Homeworks	40%
Project	24%
Midterm	18%
Final Exam	18%

Extensions

In general, we do not grant extensions on assignments. There are several exceptions:
- **Medical Emergencies:** If you are sick and unable to complete an assignment or attend class, please go to University Health Services. For minor illnesses, we expect grace days or our late penalties to provide sufficient accommodation. For medical emergencies (e.g. prolonged hospitalization), students may request an extension afterwards and should include a note from University Health Services.
- **Family/Personal Emergencies:** If you have a family emergency (e.g. death in the family) or a personal emergency (e.g. mental health crisis), please contact your academic adviser or Counseling and Psychological Services (CaPS). In addition to offering support, they will reach out to the instructors for all your courses on your behalf to request an extension.
- **University-Approved Absences:** If you are attending an out-of-town university approved event (e.g. multi-day athletic/academic trip organized by the university), you may request

an extension for the duration of the trip. You must provide confirmation of your attendance, usually from a faculty or staff organizer of the event.

For any of the above situations, you may request an extension **by emailing 10716-instructors@cs.cmu.edu**. The email should be sent as soon as you are aware of the conflict. In the case of an emergency, no notice is needed.

Audit Policy:

Audit Policy Official auditing of the course (i.e. taking the course for an "Audit" grade) is not permitted this semester. Unofficial auditing of the course (i.e. watching the lectures online or attending them in person) is welcome and permitted without prior approval. We give priority to students taking the course for a letter grade, so auditors may only take a seat in the classroom if there is one available 10 minutes after the start of class. Unofficial auditors will not be given access to course materials such as homework assignments and exams.

Pass/Fail Policy:

We allow you to take the course as Pass/Fail. Instructor permission is not required. What grade is the cutoff for Pass will depend on your program. Be sure to check with your program / department as to whether you can count a Pass/Fail course towards your degree requirements.

Accommodations for Students with Disabilities:

If you have a disability and are registered with the Office of Disability Resources, I encourage you to use their online system to notify me of your accommodations and discuss your needs with me as early in the semester as possible. I will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, I encourage you to contact them at access@andrew.cmu.edu.

Important Note!

We want to help you learn the material, and understand unexpected issues arise throughout the semester. If something unexpected comes up, or you are having difficulties that are impacting your learning of the material, come see us! Please, please come. You should also reach out to your academic advisor so they are aware of the situation.

Academic Integrity Policies

Read this carefully!

(Adapted from Roni Rosenfeld's 10-601 Spring 2016 Course Policies.)

Collaboration among Students

- The purpose of student collaboration is to facilitate learning, not to circumvent it. Studying the material in groups is strongly encouraged. It is also allowed to seek help from other students in understanding the material needed to solve a particular homework problem, provided no written notes (including code) are shared, or are taken at that time, and

provided learning is facilitated, not circumvented. The actual solution must be done by each student alone.

- The presence or absence of any form of help or collaboration, whether given or received, must be explicitly stated and disclosed in full by all involved. Specifically, each assignment solution must include answering the following questions:
    1. Did you receive any help whatsoever from anyone in solving this assignment? Yes / No.
        - If you answered 'yes', give full details: _____
        - (e.g. "Jane Doe explained to me what is asked in Question 3.4")
    2. Did you give any help whatsoever to anyone in solving this assignment? Yes / No.
        - If you answered 'yes', give full details: _____
        - (e.g. "I pointed Joe Smith to section 2.3 since he didn't know how to proceed with Question 2")
    3. Did you find or come across code that implements any part of this assignment? Yes / No. (See below policy on "found code")
        - If you answered 'yes', give full details: _____
        - (book & page, URL & location within the page, etc.).
- If you gave help after turning in your own assignment and/or after answering the questions above, you must update your answers before the assignment's deadline, if necessary, by emailing the course staff.
- Collaboration without full disclosure will be handled severely, in compliance with CMU's Policy on Academic Integrity.

Previously Used Assignments

Some of the homework assignments used in this class may have been used in prior versions of this class, or in classes at other institutions, or elsewhere. Solutions to them may be, or may have been, available online, or from other people or sources. It is explicitly forbidden to use any such sources, or to consult people who have solved these problems before. It is explicitly forbidden to search for these problems or their solutions on the internet. You must solve the homework assignments completely on your own. We will be actively monitoring your compliance. Collaboration with other students who are currently taking the class is allowed, but only under the conditions stated above.

Policy Regarding "Found Code":

You are encouraged to read books and other instructional materials, both online and offline, to help you understand the concepts and algorithms taught in class. These materials may contain example code or pseudo code, which may help you better understand an algorithm or an implementation detail. However, when you implement your own solution to an assignment, you must put all materials aside, and write your code completely on your own, starting "from scratch". Specifically, you may not use any code you found or came across. If you find or come across code that implements any part of your assignment, you must disclose this fact in your collaboration statement.

Duty to Protect One's Work

Students are responsible for pro-actively protecting their work from copying and misuse by other students. If a student's work is copied by another student, the original author is also considered to be at fault and in gross violation of the course policies. It does not matter whether the author allowed the work to be copied or was merely negligent in preventing it from being copied. When overlapping work is submitted by different students, both students will be punished. To protect future students, do not post your solutions publicly, neither during the course nor afterwards.

Penalties for Violations of Course Policies
All violations (even first one) of course policies will always be reported to the university authorities (your Department Head, Associate Dean, Dean of Student Affairs, etc.) as an official Academic Integrity Violation and will carry severe penalties.
1. The penalty for the first violation is a one-and-a-half letter grade reduction. For example, if your final letter grade for the course was to be an A-, it would become a C+.
2. The penalty for the second violation is failure in the course, and can even lead to dismissal from the university.

Take care of yourself:
Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle, and this semester is no exception. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at http://www.cmu.edu/counseling/. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:
- CaPS: 412-268-2922
- Re:solve Crisis Network: 888-796-8226

If the situation is life threatening, call the police
- On campus: CMU Police: 412-268-2323
- Off campus: 911