

ADVANCED TOPICS IN MULTIMODAL MACHINE LEARNING – SPRING 2024

[Description](#) | [Schedule](#) | [Grades](#)

Instructors: Prof. Daniel Fried, dfried@cs.cmu.edu
Paul Liang, pliang@cs.cmu.edu

Time: Tuesdays and Thursdays, 2:00pm-3:20pm

Classroom: WEH 4709

Recommended preparation: Students are expected to have already taken 11-777 Multimodal Machine Learning course or have equivalent research experience (instructor approval required).

Introduction and Purposes

Multimodal machine learning (MMML) is a vibrant multi-disciplinary research field which addresses some of the original goals of artificial intelligence by integrating and modeling multiple communicative modalities, including language, vision, and acoustic. This research field brings some unique challenges for multimodal researchers given the heterogeneity of the data and the interactions often found between modalities. This course is designed to be a graduate-level course covering recent research papers in multimodal machine learning, including technical challenges with representation, alignment, reasoning, generation, co-learning and quantification. The main goal of the course is to increase critical thinking skills, knowledge of recent technical achievements and understanding of future research directions.

Course format

The course is expected to be in-person given its discussion-based approach. Students are expected to read papers ahead of each course session, and discussions will be moderated to ensure deep understanding of all papers and critical thinking of the paper's strengths and possible avenues for future directions and improvements. The course is planned for 6 credit units. Optionally, students can register for 12 credit units, with the expectation to do a comprehensive research project as part of the semester. These course projects are expected to be done in teams, with the research topic to be in the realm of multimodal machine learning and pre-approved by the course instructors.

Piazza We will be using Piazza for class communication and announcement. The system is highly catered to getting you help fast and efficiently from classmates and the instructors. Rather than emailing questions to the teaching staff, you are encouraged to post your questions on Piazza. You can post privately to the instructor and TAs through Piazza website.

<https://piazza.com/cmu/spring2024/11877/info>

Canvas Students are asked to submit their project assignments through the website Canvas. This platform will be used for grading and to handle any request for re-grading.

<https://canvas.cmu.edu/courses/39063>

Website The main public course website is available on online:

Course Material

Required:

- Reading material will be based on published technical papers available via the ACM/IEEE/Springer digital libraries or freely available online (e.g., arxiv.org). All CMU students have already free access to these digital archives.
- For project assignments, previous experience in Python and deep learning (e.g., Pytorch) programming is expected

Course Topics

*** Exact topics and schedule subject to change, based on student interests and course discussions. ***

- Week 1 (1/16): Introduction
- Week 2 (1/23): Foundations 1: Dimensions of heterogeneity
- Week 3 (1/30): Foundations 2: Multimodal connections
- Week 4 (2/6): Foundations 3: Multimodal interactions
- Week 5 (2/13): Multimodal LLMs 1: Data, pretraining, and scaling laws
- Week 6 (2/20): Multimodal LLMs 2: Fine-tuning, instructing, aligning, model merging
- Week 7 (2/27): Multimodal LLMs 3: Generative models and LLMs
- Week 8 (3/5): No classes – Spring break
- Week 9 (3/12): Interaction 1: Reasoning and large models
- Week 10 (3/19): Interaction 2: Embodiment and planning
- Week 11 (3/26): Interaction 3: Interaction with people
- Week 12 (4/2): Ethics and safety
- Week 13 (4/9): Efficiency
- Week 14 (4/16): Open discussion
- Week 15 (4/23): Project presentations

Course Project Timeline

(This section applies only to 12-unit version of the course. Exact timeline subject to change.)

- **Project preferences** (Due Tuesday 1/23 at 9pm ET) – Online form to share your interests about research projects and help with team matching.
- **Pre-proposal** (Due Tuesday 1/30 at 9pm ET) – You should have selected your teammates, dataset, and task. Submit a 1-page pre-proposal plan.
- **Proposal and Literature Review** (Due Tuesday 2/13 at 9pm ET) - Description of your research ideas, review of relevant papers, and initial results
- **Midterm report** (Due Tuesday 3/19 at 9pm ET) – Intermediate report documenting the updated results in exploring new research ideas.
- **Final report** (Due Tuesday 4/30 at 9pm ET) – Final report describing explored research ideas, with experimental results, analysis, and discussion.

Grades

Grading breakdown for 6-unit version (no course project)

- o Reading assignments 40%
- o Participation and discussions 40%
- o Reading and synopsis leads 20%
- Reading assignments
 - o There are a total of 12 reading assignments planned this semester. Each reading assignment will consist of 2 main parts:
 - Assigned reading paper: Reading the assigned papers and summarizing the main take-away points of each paper.
 - Research question probes: Reflect on the question probes related to the reading papers and prepare discussion points.
 - o The grading for each reading assignment is planned as:
 - **1 point** for scouting relevant papers, blog posts or other resources
 - **2 points** for the take-away points of all assigned reading papers
 - **3 points** for the discussion points related to the question probes
 - o Reading assignments will also contain a section for students to share their clarification questions about the reading papers. This section is optional and is not directly graded. This section will be used to help instructors and discussion leads prepare for the main discussion during lecture time.
 - o The final score will be computed by taking the top 10 scores, out of 12 reading assignments
 - Students are expected to submit all reading assignments. If a reading assignment is not submitted (see late submission section below for details), then 2 points will be removed from the final score for each missing submission.
- Participation and discussions
 - o A core component of this course is centered around live discussions during the course lecture times. Students are expected to be active participants in these discussions. Discussions will usually be performed in smaller groups (8-10 students per group).
 - o Small group discussions will be performed in a round table setting, where all students are given the opportunity to share their observations and discussion points.
 - A first part of the discussion will focus on clarifying any questions or misunderstandings related to the two research papers.
 - The main part of the discussion will focus on the research question probes. Each student is expected to actively participate in this discussion.
 - o The grading for each discussion session is planned as follows:
 - **2 points** for the insight and quality of the shared discussion points
 - **2 points** for interactivity and participation as follow-up to other's questions and suggestions.
 - o The final score will be computed by taking the top 10 scores, out of 12 discussion sessions
 - Given the live nature of the discussions, students are expected to attend all discussion sessions.
 - Although the final grades are computed with the top 10 scores, any absences improperly justified (see Attendance section below for more details about absences) will remove 2 points from the final grade.
- Reading and Synopsis leads

- o 2 or 3 times during the semester (depending on the number of registered students), each student will be scheduled to be leading either the reading or synopsis. Each week, two reading leads and two synopsis leads will be selected.
 - Each discussion group will have one reading lead and one synopsis lead. Both are supposed to work together to create notes from the discussion. The notes are only shared internally with the other students.
- o The main tasks of the reading and synopsis leads are
 - Reading leads are expected to read the assigned papers with extra details, to assist other students with follow-up questions. They are asked to prepare a short presentation at the beginning of the course to address the points that were asked to be clarified by other students. Points that cannot be addressed in this short presentation should be answered on Piazza directly.
 - Synopsis leads are in charge to create a synopsis of both discussion groups. The synopsis should try to summarize the discussions in such a way that an overarching story and structure is present. For example, if the topic was about cross-model interactions, synopsis leads could try to write in a table a short taxonomy of the different cross-modal interactions that were discussed.
- o The grading for reading leads is planned as follows:
 - **4 points** for creating and presenting the short presentation at the beginning of the course addressing the clarification points.
 - **1 point** for helping take notes of observations and points made during small group discussions
- o The grading for synopsis leads is planned as follows:
 - **1 point** for taking notes during small group discussions. These notes should be posted on Piazza for all students.
 - **4 points** for creating the synopsis to summarize the main take-home messages of these discussions
- o The final score will be computed by taking the top 2 scores for both reading and synopsis leads.
 - For example, if the student was synopsis lead twice and reading lead once, then the final score will be the top 2 scores out of these 3 scores.
- o Please refer to the detailed instructions for reading and synopsis leads posted on Piazza.

Grading breakdown for 12-unit version (with course project)

- Grading breakdown of the 6-unit version will be scaled to 50%. The second 50% comes from the course project:
 - o Proposal report 10%
 - o Midterm report 20%
 - o Final report 25%
 - o Final presentation 15%
 - o Bi-weekly written updates 30%
- Project preference form
 - o This form is designed to help students with the team matching process, for the course research project.
- Pre-proposal
 - o This short 1-page pre-proposal is designed to confirm teams and share thoughts about the research project, including key related work, main ideas, and datasets + evaluation.
- Proposal and literature review

- o The proposal report should present the initial research ideas for the course project. Students are expected to explore new research ideas as part of the course research project. The proposal should summarize these research ideas. It should also give an overview of the dataset and main research tasks that will be addressed.
- o An important part of this proposal report will be a detailed literature review, including recent papers and models related to the dataset, research tasks and the new research ideas.
- Bi-weekly updates
 - o Project teams are expected to meet weekly with instructors during Thursday's class.
 - Project meetings will be about 20–30 minutes long.
 - Each meeting will usually be with one instructor.
 - o To help streamline the project meetings, team members need to submit a written update document before each Thursday meeting.
 - Updates are due Tuesdays at 9pm before the meeting, reports (proposal + midterm+final) are also due Tuesdays at 9pm.
 - Each team can decide to use either an online Google Docs or Google Slides for these updates.
 - The goal is to keep these updates informal, with only the main points highlighted in the updates.
 - For example, a bullet list with 3-4 main takeaways, along with a few figures and tables, may be sufficient.
 - The same online document should be used for all weekly meetings, so that instructors can easily review previous updates.
 - o Bi-weekly updates will be graded by instructors, 10 points per update meeting.
 - o The final score will be computed by taking the top 3 scores for the whole semester.
- Schedule for bi-weekly written updates and reports:
 - o Week 3 (1/30): Pre-proposal details with literature review
 - o Week 5 (2/13): **Proposal report:** baseline results and new ideas
 - o Week 7 (2/27): Initial implementation of new ideas
 - o Week 10 (3/19): **Midterm report:** first complete round of results for idea
 - o Week 12 (4/2): Updated results for research idea
 - o Week 14 (4/16): Error analysis, ablations, and visualizations
 - o Week 15 (4/23): **Project presentations**
 - o Week 16 (4/30): **Final report**
- Midterm report
 - o The goal of the midterm report is to summarize the current research progress. Students should have started already exploring new research ideas. The midterm report should summarize these initial results and discuss them. This report should also present the updated list of research ideas that the team plan to explore.
- Final report
 - o The final report should follow a similar structure of a research paper. It should motivate the problem and research ideas. It should present the novel approaches, describe the experiment and discuss the results.
- Final presentation
 - o The final project assignment will also include either an in-person poster or oral presentation, to be confirmed. The poster presentations may be open to the general CMU population.

Notes about absences and late submissions

In general, submitting assignments on time lets the instructional team provide feedback in a more timely and efficient manner. Timely submissions are particularly important for assignments with discussions and peer feedback. Also, it is expected that students will attend the lectures in person (or via Zoom when the course is required to be performed remotely). Live attendance is an essential component of this course, given that it is centered around live discussions. Given the live nature of the discussions, course sessions will not be recorded.

Medical-related absences If for a medical reason you require some extra time for an assignment or may not be able to attend the lecture in person, please contact instructors as soon as possible (the best option is usually via Piazza) and we will help define a new plan that aligns with your constraints.

- Absence requests
 - Students should contact instructors as promptly as possible regarding course absences, with a preference before the course lecture itself.
 - If you plan to be absent for more than one course lecture, it will be important to contact instructors as promptly as possible.

Late submission wildcards We offer students and project teams some late submission wildcards to help deal with potential overlaps with other courses or research deadlines. The details are expressed below:

- Reading assignment wildcards (3 per students)
 - Each wildcard gives the student a 24-hour extension for the reading assignment deadline.
 - Maximum of 1 wildcard per week
 - This constraint is to ensure that sufficient time is made available to prepare for the Friday discussion.
- Project assignment wildcards (2 per team)
 - This gives the project team a 24-hour extension for their project assignment deadline.
 - These can be used for proposal, midterm and/or final deadlines
 - It is possible to use 2 wildcards for the same deadline, giving the team 48 hours extension.
 - Teams are required to message instructors via Piazza before the deadline, to inform that they will be using 1 or 2 wildcards.

Accommodations for Students with Disabilities

If you have a disability and have an accommodation letter from the Disability Resources office, I encourage you to discuss your accommodations and 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.

Statement on Student Wellness

As a student, you may experience a range of challenges that can interfere with learning, such as strained relationships, increased anxiety, substance use, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may diminish your academic performance and/or reduce your ability to participate in daily activities. CMU services are available, and treatment does work. You can learn more about confidential mental health services available on campus at:

<http://www.cmu.edu/counseling/>. Support is always available (24/7) from Counseling and Psychological Services: 412-268-2922.

Diversity statement

Every individual must be treated with respect. The ways we are diverse are many and are fundamental to building and maintaining an equitable and an inclusive campus community. These include but are not limited to: race, color, national origin, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. We at CMU, will work to promote diversity, equity and inclusion not only because it is necessary for excellence and innovation, but because it is just. Therefore, while we are imperfect, we all need to fully commit to work, both inside and outside of our classrooms to increase our commitment to build and sustain a campus community that embraces these core values.

It is the responsibility of each of us to create a safer and more inclusive environment. Incidents of bias or discrimination, whether intentional or unintentional in their occurrence, contribute to creating an unwelcoming environment for individuals and groups at the university. If you experience or observe unfair or hostile treatment on the basis of identity, we encourage you to speak out for justice and support in the moment and/or share your experience using the following resources:

- Center for Student Diversity and Inclusion: csdi@andrew.cmu.edu, (412) 268 2150, www.cmu.edu/student-diversity
- Report-It online anonymous reporting platform: www.reportit.net username: tartans password: plaid

All reports will be acknowledged, documented, and a determination will be made regarding a course of action. All experiences shared will be used to transform the campus climate to be more equitable and just.