

Nowadays, machine learning is such an exciting discipline because of its rapid pace of innovation, with new technologies and capabilities emerging seemingly every week, if not every day. However, the speed of progress makes mastery of machine learning incredibly challenging because any frozen skill set risks becoming outdated. In such a dynamic environment, perhaps the most important meta-skill to acquire is to be able to think like a researcher. Thinking like a researcher means (i) accessing new knowledge from a messy scientific literature, pulling out the useful and correct bits from papers while discarding (or correcting) mistakes; (ii) applying new methods and ideas to real problems that you are trying to solve.

The course project is an opportunity for you to build this skill by performing a comprehensive review of a research subfield. Your final output will be a thorough report that synthesizes knowledge from the field, which must include an *implementation component* i.e., simple experiments performed for didactic purposes, to demonstrate understanding and reproducibility. Groups are expected to read deeply into their chosen topic and the final report should paint a comprehensive picture of both the current state of the art, its connections to related areas of research, and its historical context (i.e., connections to ideas in the research literature that preceded it).

The theme of our suggested topics list is generative AI, with a focus on practical implementations and societal impacts. The suggested areas below are all hotbeds of activity and part of the challenge of the project is determining which recent papers are worth focusing on and why. Other topics may be considered provided your group receives explicit instructor approval in advance of the proposal deadline. We've seeded each of the topics below with a couple papers to get you started on your research; to be clear, we expect your final product to go well beyond the short list of references we've provided here. You should also feel free to disregard our suggestions and perform your own research.

- Compression of large language models:
 - Sun et al., (2023) [A Simple and Effective Pruning Approach for Large Language Models](#)
 - Lin et al., (2023) [AWQ: Activation-aware Weight Quantization for LLM Compression and Acceleration](#)
- Efficient fine-tuning of large language models:
 - Hu et al., (2021) [LoRA: Low-Rank Adaptation of Large Language Models](#)
 - Hounsby et al., (2019) [Parameter-Efficient Transfer Learning for NLP](#)
- Ensuring / assessing factuality in machine-generated text:
 - Goodrich et al., (2021) [Assessing The Factual Accuracy of Generated Text](#)
 - Lewis et al., (2021) [Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks](#)
- Robustness of large language models against adversarially perturbed input:
 - Jia et al., (2019) [Certified Robustness to Adversarial Word Substitutions](#)
 - Liu et al., (2020) [Adversarial Training for Large Neural Language Models](#)

- Assessing discrimination/bias in [large language models or image-generation models]:
 - Liang et al., (2021) [Towards Understanding and Mitigating Social Biases in Language Models](#)
 - Luccioni et al., (2023) [Stable Bias: Evaluating Societal Representations in Diffusion Models](#)
- Generative music models:
 - Briot et al., (2019) [Deep Learning Techniques for Music Generation – A Survey](#)
 - Dhariwal et al., (2020) [Jukebox: A Generative Model for Music](#)

Final Project Policies

The following policies will apply to your project:

1. You must work on the course project in groups of 2 or 3; **you may not work on the course project alone.**
2. There are three deliverables associated with the project:
 - a. **A 2-page proposal, due March 22rd at 11:59 PM** - these will largely be graded on completion although you may be asked to resubmit unsatisfactory work; their primary purpose is to catch and correct any early misconceptions about the project task/scope. The proposal is worth 2% of your final grade.
 - b. **A 4-page check-in, due April 8th at 11:59 PM** - these will also largely be graded on completion and serve as a way for us to ensure that you're making sufficient progress on the project. The check-in is worth 3% of your final grade.
 - c. **An 8-page final report, due April 26th at 11:59 PM** - the final report is worth 15% of your final grade. See below for complete details about what is expected in the final report.
3. The page limits for the deliverables defined above are strict, with the exception of pages only containing references/citations, which will not count against this upper limit.
4. Each group will be assigned a course staff member as their project mentor: they will be your primary point of contact for any questions that arise during the course of the project.
 - a. After submitting your proposal, **your group must schedule a meeting with your project mentor where you will collectively review your proposal.**
 - b. Based on this meeting, your group will either receive approval to move forward or be asked to resubmit your proposal, incorporating feedback from the meeting.

Final Report Specifications

Your final report will consist of three deliverables. There will be a separate Gradescope submission for each of the following items:

1. A writeup: this must be a document typeset in LaTeX that uses [this LaTeX template designed for preprint](#). The writeup must have a descriptive title as well as a list of the names and AndrewIDs of all group members who at some point contributed to the project, regardless of whether or not they are still enrolled in the course. Submit the writeup as a group in Gradescope; you should have one submission per group. Your writeup must contain the following elements:
 - a. Title and Author List
 - b. Description of Topic: Briefly describe the focus of your literature review and motivate its importance. For the purposes of completeness, you should also describe any related terms/fields that your group deemed out of scope for your review.
 - c. Literature Review: The bulk of your writeup will be a thorough review of the academic landscape surrounding your chosen topic. It should begin with the historical context: what were some of the pioneering works in the field and how did they influence more recent research? When discussing the current state of research in your chosen topic, **it is crucial that you do more than just list papers and methods**: you should analyze the content of the works that you've read by e.g., drawing connections between different lines of inquiry, comparing and contrasting approaches, finding limitations or weaknesses in one paper that are addressed by another, etc...
 - d. Experimental Results: Briefly describe what method(s) your group implemented and the empirical settings in which you evaluated your implementation. Show plots and/or tables of the performance of your method(s) and interpret what they mean; be sure to label all of your figures and explain the findings. You must also define any performance metrics you used for evaluation.
 - e. Discussion and Analysis: Finally, you should reflect on the overall state of your chosen topic based off of the review your group performed. This could include (but is not limited to) ideas about where the field as a whole is moving towards, what the promising new avenues of research are and conversely which methods do not show a lot of promise or are likely to be subsumed by alternatives, etc... These should not be entirely speculative but should be grounded in your understanding of the state of research in your chosen topic; of course, some (informed) imagination is encouraged here!
 - f. References
2. A statement of individual work: each group member must independently write a short paragraph describing their contributions to the project and submit them individually to

Gradescope. These will not be graded and will only be referenced in the (unlikely) event that we need to assign different grades to separate group members.

3. All code written for the project: every group must submit all the code they wrote to Gradescope. You may submit as many files as you need. Each file must have a meaningful name so that your project mentor can easily identify its purpose. If you wrote code in the form of Python notebooks, please convert those to .py files before submitting them to Gradescope. Submit your code as a group; you should have one code submission per group.

Final Report Rubric

The following rubric will be used to assess the writeups:

- Thoroughness (30 pts) - your review should cover all the seminal works and major sub-areas within your chosen topic, as deemed appropriate by your project mentor. That being said, you should not attempt to include all papers that are even remotely related to your chosen topic; rather, an important skill that will be assessed here is how well your group can identify the influential works in a field. Your literature review should go as far back as possible, trace the development of your chosen topic from the earliest related references you can find and draw connections to contemporary work.
- Analytical Depth (30 pts) - the most important part of any literature review is synthesis i.e., drawing connections and identifying trends in the area. The bulk of your writeup grade will be determined by the level of analysis your group performs when reviewing the research. A simple list of papers is not sufficient: your writeup must demonstrate a deeper understanding of the research area.
- Empirical Quality (20 pts) - the method(s) and experiment(s) you implement must be described in sufficient technical detail such that your project mentor can properly assess your work. Your method(s) should also follow rational machine learning principles/best practices as covered in the course e.g., hyperparameter tuning should be done with a held-out validation dataset and not on the test dataset. A portion of this rubric item will also be how well your submitted code matches your description in the writeup. You should also motivate your experiment(s): what is the connection to the literature review and what insights are you trying to highlight?
- Clarity (10 pts) - this portion of your grade will assess the quality and organization of your writeup; it is crucial that you present your work in a clear and understandable way.
- Formatting (10 pts) - your writeup must adhere to the guidelines we have established above e.g., it respects the 8 page limit and uses [the correct LaTeX template](#).