

Large Language Models

Course codes: COMP-767 and LING-782

Instructor: [Siva Reddy](#)
Assistant Professor
McGill University and Mila
siva@cs.mcgill.ca

Term: Fall 2024
When: Tuesdays and Thursdays
Time: 11:35 am – 12:55 pm
Where: **WONG 1050**

Office hours: MC 104N Tuesdays 1:20 pm to 2:30 pm (may change)

This page link:

https://docs.google.com/document/d/1r0WpmenrymWFmy_qldkb5vhtl-vgaDBfFnAJMknlj_0/edit

MyCourses link (class announcements, discussions):

<https://mycourses2.mcgill.ca/d2l/home/744362>

Presentation and Panel signup forms:

TODO

Course Description:

This is a seminar-style course, where the class as a whole will work together in running the course. In the first few lectures, I will provide an overview of LLMs and highlight the challenges.

Expected Outcomes:

By the end of the course, you should be able to meaningfully contribute to cutting-edge research in natural language understanding.

Grading (tentative):

This is a demanding course in terms of participation and projects. All deadlines start after the Add/Drop deadline.

- **Review (10%)** You are expected to submit a critical review of the papers that are to be presented in the class, at least an hour before the lecture on MyCourses. You will submit 5 such reviews.
- **Meta-review (5%)** You will critically review at least 2 reviews above by the end of the day. You will do it 5 times.
- **Presentation (20%)** You have to present papers with a presentation partner on the topic of the lecture (at least once). Not the same topic as you wrote a critical summary for.
- **Panelists (10%)**: Sign up as panelists at least twice to critically analyze the presented papers.
- **Project (40%)** You will do a project in groups of two. This involves

- Project Proposal (what is the project about): 4%
- Literature review (8%)
- Baselines/Replication (12%)
- Final paper with new experiments and code submission (16%)
- **Poster (10%)** End of term poster
- **Class participation (5%)**: This is determined at the end of the term depending on how engaged in the lectures you are (asking questions during the lectures, coming to office hours etc).

Topics of interest (subject to change):

1. Ingredients of a Language Model
 - a. Transformer Architecture
 - b. Training recipes
 - c. Evaluation
2. Emergent capabilities
 - a. Scaling laws
 - b. Prompting
 - c. Incontext learning
 - d. Grokking
3. Limitations of Transformers
 - a. Expressivity
 - b. Extrapolation
4. Instruction-following Language Models
 - a. Data
 - b. Evaluation
 - c. Less is more
5. Social considerations
 - a. Hallucinations, Bias, Safety
 - b. LLMEval
6. Alignment and Preference Optimization
 - a. RLHF, PPO
7. RAG
 - a. Retrievers
 - b. Challenges
 - c. Multi-step
8. Adversarial attacks
9. Long context models
10. Parameter-efficient adaptation
 - a. Modularity
 - b. Properties
11. Multi-step reasoning

- a. Chain-of-thought
- b. Self-Instruct
- c. ReAct
- 12. Multimodal
 - a. VLMs
 - b. Diffusion models
- 13. Agents
 - a. Web Agents
 - b. ToolFormer
 - c. LangChain
- 14. Synthetic evaluation and data generation
 - a. LLMEval
 - b. Data generation
- 15. Efficiency considerations
 - a. Quantization
 - b. Flash Attention
 - c. State space models
 - d. Mixture of experts
- 16. Security
 - a. Stealing
 - b. Watermarking

Prerequisites:

You are expected to have done one of the following courses at McGill: natural language processing (COMP/LING 550) or computational linguistics (COMP/LING 445) or applied machine learning (COMP 551). If you have done similar courses at other universities, feel free to take the course. If you are not sure, email me.

Schedule (subject to change):

Lecture	Date	Topic	Presented Papers	Additional Readings
1	Aug 29	Introduction		
2	Sep 3	Ingredients of an LLM: Transformers, Training objectives		
3	Sep 5	Emerging Capabilities		

4	Sep 10 (Add or drop deadline)	Emerging Capabilities		
5	Sep 12	Limitations of Transformers		
6	Sep 17			
7	Sep 19			
8	Sep 24			
9	Sep 26			
10	Oct 1	Workshop		
11	Oct 3	Workshop		
12	Oct 8			
13	Oct 10			
	Oct 15	Reading week		
	Oct 17	Reading week		
14	Oct 22			
15	Oct 24			
16	Oct 29			
17	Oct 31			
18	Nov 5			
	Nov 7			
19	Nov 12			
20	Nov 14			
21	Nov 19			
22	Nov 21			

23	Nov 26			
24	Dec 3			

FAQs:

Other related courses:

McGill Policies:

Assessments in this course are governed by the [Policy on Assessment of Student Learning \(PASL\)](#), which provides a set of common principles to guide the assessment of students' learning. Also see [Faculty of Science-specific rules](#) on the implementation of PASL.

Legally mandated academic accommodations are handled by Student Accessibility and Achievement. For more information see <https://www.mcgill.ca/access-achieve/>

In accord with McGill University's [Charter of Students' Rights](#), students in this course have the right to submit in English or in French written work that is to be graded.

Conformément à la [Charte des droits de l'étudiant](#) de l'Université McGill, chaque étudiant a le droit de soumettre en français ou en anglais tout travail écrit devant être noté, sauf dans le cas des cours dont l'un des objets est la maîtrise d'une langue. (Énoncé approuvé par le Sénat le 21 janvier 2009)

McGill University values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the [Code of Student Conduct and Disciplinary Procedures](#) (See [McGill's guide to academic honesty](#) for more information).

In the event of extraordinary circumstances beyond the University's control, the content and/or assessment tasks in this course are subject to change and students will be advised of the change.