UNIVERSITY OF MANNHEIM

Seminar CS715

Solving Complex Problems with Large Language Models

Presentation Schedule

Presentation Schedule – Morning Session

Slot	Topic	Student	Mentor
10:20	**From Self-consistency to MedPrompt**	Florian Bauer	Alexander Brinkmann
10:37	**Prompt Search / Breeding**	Shivam Suchak	Ralph Peeters
10:54	**Contrastive Prompting**	Ricarda Reiner	Keti Korini
11:11	**LLM-based Agents**	Max Meider	Christian Bizer
11:28	**Agent Cooperation**	Okan Göktepe	Christian Bizer

- Presentation time: 12 minutes. Discussion: 5 minutes
- Please email us your slides after your presentation

Presentation Schedule – Afternoon Session

Slot	Topic	Student	Mentor
13:45	**Limitations of LLMs**	Aaron Koßler	Steffen Eger
14:02	**Task Contamination**	Saman Khursheed	Ralph Peeters
14:19	**LLMs as Evaluation Metrics**	Fabian Rajwa	Jonas Belouadi
14:36	**LLMs with Tools as Evaluation Metrics**	Priscilla Chyrva	Daniil Larionov
14:53	**LLM Self-Evaluation during Fine-tuning**	Sara Koni	Christoph Leiter
Break			
15:30	**Evaluation of Code Writing Ability of LLMs**	Eric Wieland	Ralph Peeters
15:47	**WebAPI Query Planning Using LLM**	Mayte Dächer	Keti Korini
16:04	**Attribute Value Normalization Using LLMs**	Avani Ghandi	Alexander Brinkmann
16:21	**LLM for Literary Translation and Evaluation**	Jiyeon Lee	Ran Zhang
16:38	**Multimodal Reasoning**	Thuy Nghiem	Steffen Eger

Next Steps

Schedule

Date	Session
Thursday, 29.02.2024 (15:30-17:00)	Kick-off meeting and topic/mentor assignment
	Read papers about your topic Search for additional literature Design experimental setup Prepare outline and argumentation for your presentation
Until 20.03.2024	Meet with your mentor to discuss outline and/or experimental setup
	Prepare draft of your presentation
Until 15.4.2024	Send draft presentation to your mentor
	Finalize your presentation
Monday, 29.04.2024 (10:00-12:00) (13:45-17:00)	Presentation and discussion of your topic (30 % of your final grade)
	Write seminar thesis
Friday, 21.06.2024	Submission of your seminar thesis (70 % of your final grade)

Formal Requirements

- Written report (paper)
  - 12-15 pages single column
    - including abstract and appendixes
    - not including bibliography
    - not including the page about generative AI usage
    - every additional page reduces your grade by 0.3
  - written in English
  - use latex template of Springer Computer Science Proceedings
    - http://www.springer.com/de/it-informatik/lncs/conference-proceedings-guidelines
- Final grade
  - 70% written report
  - 30% presentation

Statement About the Tools that You Used

Your report **must include an extra page** about

1. which generative AI tools you used
   - ChatGPT, OpenAI API, Dall3, Perplexity, ….
2. for which purposes
   - structuring your paper
   - summarizing related work
   - writing text for specific chapters
   - improving English grammar and formulations
   - designing experimental setup
   - writing code
   - writing prompts
   - generating training data
   - error analysis
   - ….
3. How useful was each tool for this?

https://www.uni-mannheim.de/infos-fuer/forschende-und-lehrende/lehren/ihre-lehre-im-fokus/

Example of a Generative AI Tools Declaration

Tool	Purpose	Where?	Useful?
ChatGPT	Rephrasing	Throughout	+
ChatGPT	Summarization of related work	Sec 2	-
ChatGPT	Structure of thesis	Sec 1	~
Dall-E	Image generation	Fig 2,3	++
GPT4	Code generation	music.py functions.py	~
GPT4	Training data augmentation	Extended_set.csv	++

Goals of Literature and Experimental Papers

- Goals of Experimental Papers
  1. describe the (prompt engineering) techniques from the selected papers
  2. summarize the evaluation tasks and results from the papers
  3. design experimental setup to evaluate technique on different task
  4. compare your results to the results from the paper

- Goals of Literature Papers
  1. describe the problem / task
  2. describe several existing methods/systems for handling the task,
  3. compare the methods/systems and their evaluation using a systematic set of comparison criteria

How to Structure Your Experimental Paper?

1. Introduction and Problem Statement
   - Which problem is addressed? What is the overall approach for addressing it?
   - Overview of the existing methods/papers and use cases for the evaluation (3 pages+)
   - Structure of your paper
2. Description of Your Experimental Design
   - How to you select examples for which challenges?
   - Which prompt designs and language models do you test?
3. Presentation of Experimental Results
   - Present the results of your experiments (tables containing values and deltas).
   - Present the results of your error analysis (types of errors, frequency of these types)
4. Conclusion
   - What did the experiments and the error analysis show?
   - How to your results compare to the experiments presented in the papers?
5. Bibliography

How to Structure Your Literature Paper?

1. Introduction and Problem Statement
   - Which problem/task is addressed? Why is the problem important?
   - Structure of your paper
2. Description of Existing Approaches
   - Overview of existing methods and features used by the methods
   - Detailed description of selected methods (likely two)
   - Comparison of the selected methods using a set of comparison criteria
3. Evaluation
   - Comparison and discussion of the evaluation tasks, metrics
   - Comparison of the evaluation results using a set of comparison criteria
4. Conclusion
   - What did the comparison of the methods and evaluation results show?
   - Can something be concluded for future work?
5. Bibliography

Learn from Examples

- Read survey articles and previous experimental papers and identify the structure from the previous slides
  - Why can this paragraph be found at that position?
  - What is the purpose of some section / subsection?
- Some relevant surveys
  1. Zhao, et al.: A survey of Large Language Models. arXiv:2303.18223
  2. Mialon, et al.: Augmented Language Models: a Survey. arXiv:2302.0784
  3. Wang, et al: A Survey on Large Language Model based Autonomous Agents. arXiv:2308.11432. 2023.
- Textbook on how to write a thesis
  - Zobel: Writing for Computer Science, 3rd Edition, Springer 2014.

Citing Different Types of Publications

1. Journal article
   - Good to cite, current research results
2. Conference and workshop paper
   - Good to cite, current research results
3. Survey articles
   - Good to cite as overviews for specific topics, but prefer individual papers as reference for specific systems
4. Books (sometimes cited)
   - Textbooks
   - Collections of articles/papers => Cite specific paper in book
5. Websites
   - better not cited, exceptions are, e.g., documents like W3C Specifications
   - Do not cite Wikipedia, ever!
   - Use footnotes to refer to project pages, download pages, or technical documentation
6. Slide sets (especially from our lectures)
   - Never cite!

How to Find Relevant Publications?

1. Start with gathering relevant papers from the **surveys**
   1. Zhao, et al.: A survey of Large Language Models. arXiv:2303.18223
   2. Mialon, et al.: Augmented Language Models: a Survey. arXiv:2302.0784
   3. Wang, et al: A Survey on Large Language Model based Autonomous Agents. arXiv:2308.11432. 2023.
2. **Exploit references:** Given a relevant document x
   - Follow references in the past: papers y that x has cited
   - Follow references in the future: papers y that cited x (**„cited by"** functionality in Google scholar)
3. **Use Google Scholar or Semantic Scholar**
   - we use it a lot ourselves

Goals of Literature and Experimental Papers

- Goals of Literature Papers
  1. describe the problem / task
  2. describe several existing methods/systems for handling the task,
  3. compare the methods/systems and their evaluation using a systematic set of comparison criteria

- Goals of Experimental Papers
  1. describe the prompt engineering technique from the paper
  2. present evaluation task and results from the paper
  3. design experimental setup to evaluate technique on different task
  4. compare your results to the results from the paper

How to Structure Your Literature Paper?

1. Introduction and Problem Statement
   - Which problem/task is addressed? Why is the problem important?
   - Structure of your paper
2. Description of Existing Approaches
   - Overview of existing methods and features used by the methods
   - Detailed description of selected methods (likely two)
   - Comparison of the selected methods using a set of comparison criteria
3. Evaluation
   - Comparison and discussion of the evaluation tasks, metrics
   - Comparison of the evaluation results using a set of comparison criteria
4. Conclusion
   - What did the comparison of the methods and evaluation results show?
   - Can something be concluded for future work?
5. Bibliography

How to Structure Your Experimental Paper?

1. Introduction and Problem Statement
   - Which problem is addressed? What is the overall approach for addressing it?
   - Overview of the existing methods/papers and use cases for the evaluation
   - Structure of your paper
2. Description of Experimental Design
   - What is your How to you select examples for which challenges?
   - Which prompt designs and language models do you test?
3. Presentation of Experimental Results
   - Present the results of your experiments (tables containing values and deltas).
   - Present the results of your error analysis (types of errors, frequency of these types)
4. Conclusion
   - What did the experiments and the error analysis show?
   - How to your results compare to the experiments presented in the papers?
5. Bibliography

Learn from Examples

- Read survey articles and previous experimental papers and identify the structure from the previous slides
  - Why can this paragraph be found at that position?
  - What is the purpose of some section / subsection?
- Important
  - Read survey articles!
  - Read conference or journal papers
- Some relevant surveys
  - Zhao, et al.: A survey of Large Language Models. arXiv:2303.18223
  - Mialon, et al.: Augmented Language Models: a Survey. arXiv:2302.0784
- Textbook on how to write a thesis
  - Zobel: Writing for Computer Science, 3rd Edition, Springer 2014.

Citing Different Types of Publications

- Journal article
  - Good to cite, current research results
  - Survey articles (very good for an overview)
- Conference and workshop paper
  - Good to cite, current research results
- Books (sometimes cited)
  - Textbooks
  - Collections of articles/papers => Cite specific paper in book
- Websites
  - better not cited, exceptions are, e.g., documents like W3C Specifications
  - Do not cite Wikipedia, ever!
  - Use footnotes to refer to project pages, download pages, or technical documentation
- Slide sets (especially from our lectures)
  - Never cite!

How to Find Relevant Publications?

- Use Standard Search Engines
- **Use Google Scholar**
  - we use it a lot ourselves
- Search Engines of the University's library
  - see slides from the library course
- **Exploit references:** Given a relevant document x
  - Follow references in the past: papers y that x has cited
  - Follow references in the future: papers y that cited x (**„cited by"** functionality in Google scholar)