Diachronic Language Models

Lecture 1 - Kick off

Dr. Wei Zhao 15/10/2024

Lecturers
Who are we

- Dr. Wei Zhao

  - PhD in NLP at TU Darmstadt

  - Postdoc at HITS

  - Lecturer at University of Aberdeen and Heidelberg University

- Yi Fan

  - PhD student at HITS and Heidelberg University

Feedback
It is your course

- Your feedback is important

- Constructive criticisms are welcome

- We will post anonymous feedback form when the time comes.

Organization
About resources and communication

- Github for lecture slides and literature

  - https://github.com/andyweizhao/diaclms

- Discord communication

  - https://discord.gg/FqU8VJ86KS

- Contact

  - diachronic.hd@gmail.com

- Talk to us: in-person, discord, email anytime

Organization
To pass the course, you need

- 50% term paper

- 25% mid-term presentation

- 25% final presentation

- Deadline for term paper:

  - preliminary version: 18, March, 2025

  - final version: 25, March, 2025

Organization
Term paper

- 3-5 students work in group on the relevant topics

  - Joint and independent work

  - Coordination

- 8 pages of content + unlimited pages of references.

- Template: https://www.overleaf.com/1476718452xshrqqpjvkvm#47a63b

- Workload: 2-4 hours per week

- Supervision?

Organization
Presentation

- Each group presents project results for 30 minutes

  - Each project member presents independent work for 8-10 minutes

- 15 minutes for Q&A

Organization
How to structure your presentation

- Motivation

- Research Questions

- Methodology

- Dataset(s)

- Results and data analyses

- Conclusions

Organization
How we evaluate your presentation

- Clarity: clear and easy-to-undertstand

- Entertainment: tables, figures to communicate complex ideas and findings

- Performance in Q&A

  - Answers should be concise and direct.

Organization
How we evaluate your term paper

- Readability: Is the paper easy to follow and well-structured?

- Substance: content is substantial, with your own arguments, ideas and results?

- Soundness: your claims/arguments are factual?

- Novelty: your ideas are simple, clever?

- Structure: intro, research questions, method, results and analyses, conclusions..

- Example 1: https://aclanthology.org/C18-1325.pdf

- Example 2: https://aclanthology.org/W19-4722.pdf

Topics for Term Papers
A list of topics

- Automatic Dictionary Induction

  - Discover novel senses uncovered by OED

  - with Dominik Schlechtweg from Stuttgart University

- Causality for temporal biases in LLMs

  - Relationship between (a) dates formatting, compression rate and (b) model hallucination, answer confidence, next word prediction, etc

  - with Maxime Peyrard from Université Grenoble Alpes

- Temporal code generation

  - Generate changes in codebase for addition, modification and deletion, based on Github comments

Topics for Term Paper
Choose your topic

- Email us with your preferred topic

  - Deadline: 21 Oct, 2024

- We will divide you into groups.

- Let us and your peer students know timely if you want to quit.

Plagiarism
This is Forbidden!

- With any plagiarism, you will fail the course.

- Some examples of plagiarism (https://www.plagiarism.org/article/what-is-plagiarism)

  - Appropriation of words and ideas from other people without giving credit

- Writing supporting tools are permitted, but how and where you use them should be clarified.

- Know why you are here.