

Preliminary Meeting of the NLP Lab Course WS 2026/27

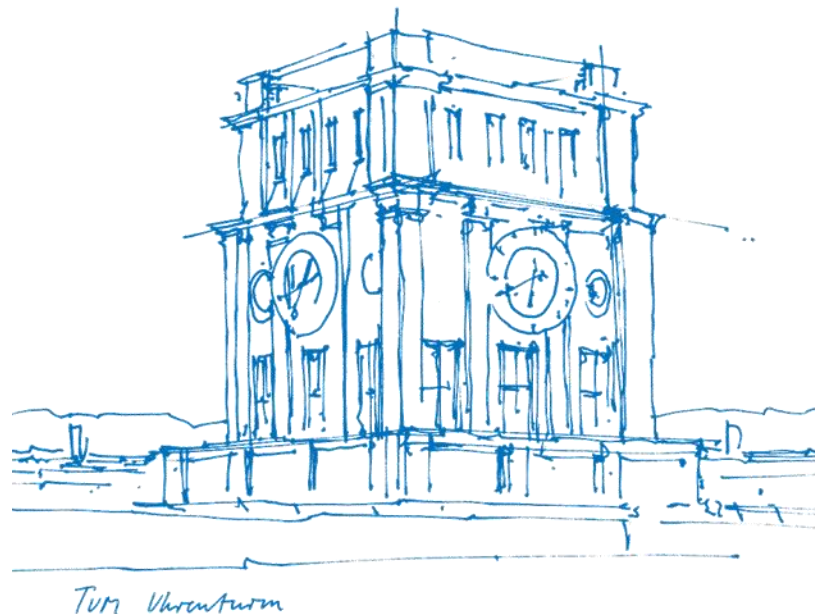
Master Lab Course - Machine Learning for Natural Language Processing Applications (IN2106, IN4249)

Frederic Mrozinski,
and many more

Prof. Dr. Georg Groh

Research Group Social Computing, Department of Informatics,
Technical University of Munich

08.07.2026



Outline

1. Requirements
2. Registration
3. Procedure
4. Project examples
 - Text simplification and summarization
 - Evaluating correctness of generated text
 - Biomedical knowledge enhancement
 - Pattern detection for cancer tissue
 - Green and efficient AI
 - User Knowledge and Need Alignment in LLMs

Requirements

Minimum:

- Master student in computer science, data engineering, or "alike"
- Good enough English skills
- Basic programming and machine learning knowledge

Important:

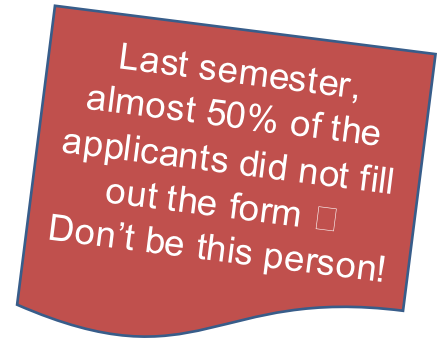
- Hands-on experience in Python, especially Pandas and Numpy
- Basic knowledge about artificial neural networks
- Basic knowledge about natural language processing

Optimal:

- Practical experience with Deep Learning frameworks, such as PyTorch, Tensorflow, Huggingface, etc.

Registration

- Until **16th of July**, fill out the [registration form](#)



- Your entries are considered when ranking the interested students for the course.
- From **10 to 16 July**, you also have to register for the course on the [matching system](#).
- End of **July**, you will (probably) be notified by the matching system about the status of your participation.
- We will get in touch with you in March for the following steps.

Procedure

Project teams:

- You are going to work in teams of 2 or 3 people on one project topic.
- You can choose with whom to work with the project topic.
- Every project member has to report and work equally (no dirty business!).

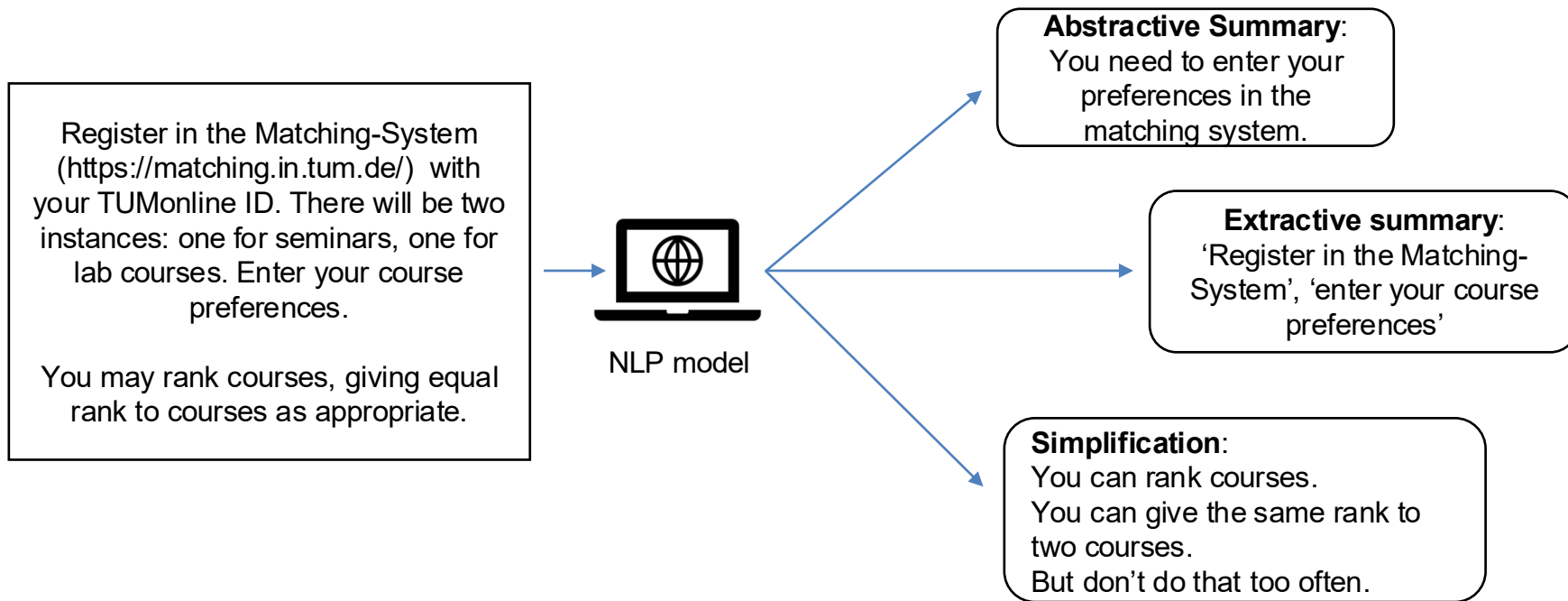
Procedure:

- There will be one kickoff meeting at the beginning of the semester.
- There are going to be bi-weekly consulting and progress report sessions.
- You have to be part of a poster session and hand in a report at the end of the semester.

Everything else will be announced at the beginning of the semester.

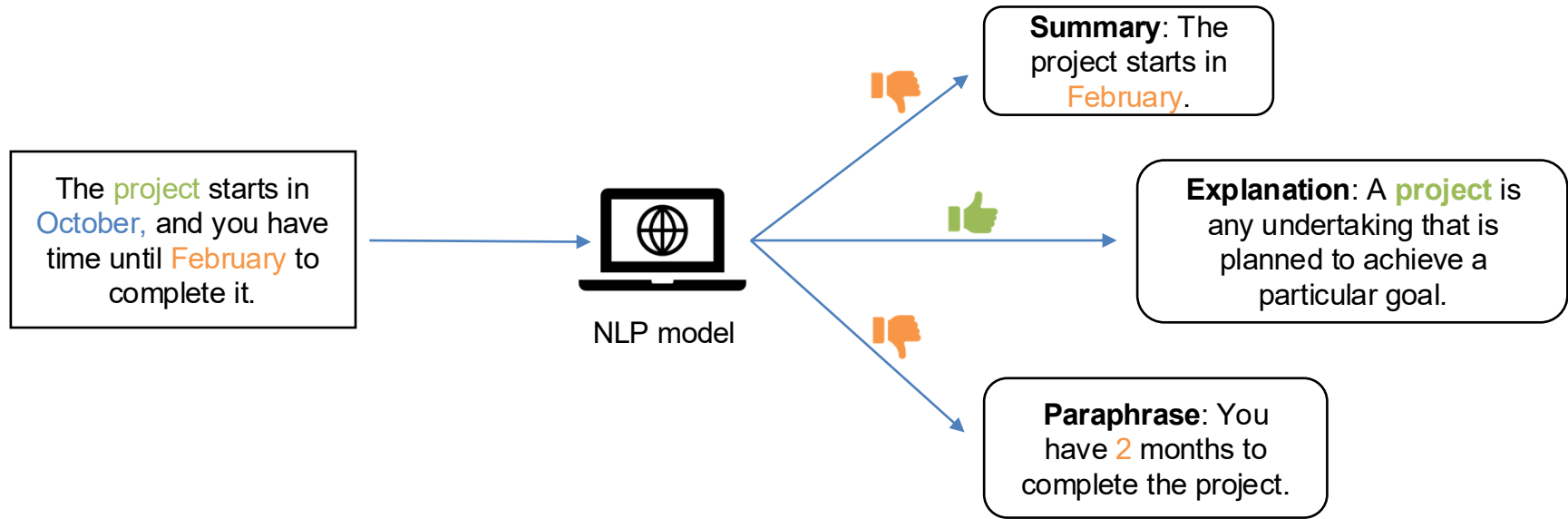
Projects – Text summarization and simplification

Miriam Anschütz, M.Sc.;



Projects – Evaluating correctness of generated text

Miriam Anschütz, M.Sc.



Projects – Biomedical Knowledge Enhancement of LMs

Alexander Fichtl, M.Sc.

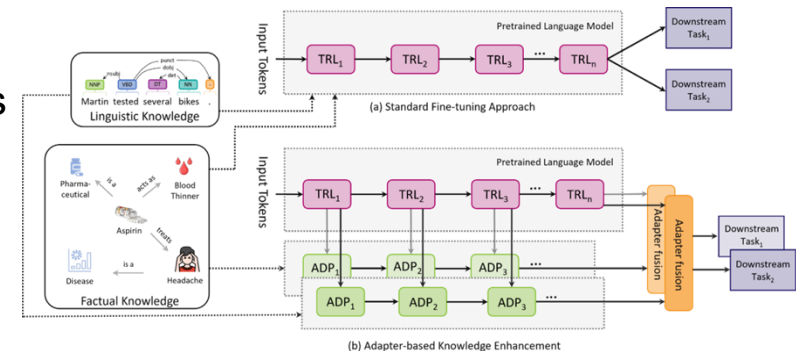
- 60% of medical doctors in Germany spend 3 hours or more per day on documentation and administrative tasks
- Many tasks in bio-medicine can be supported by LMs
- Domain specific LMs profit from explicit knowledge injection

➤ You will be using biomedical knowledge graphs to enhance domain specific LMs that can potentially address these issues



BLURB

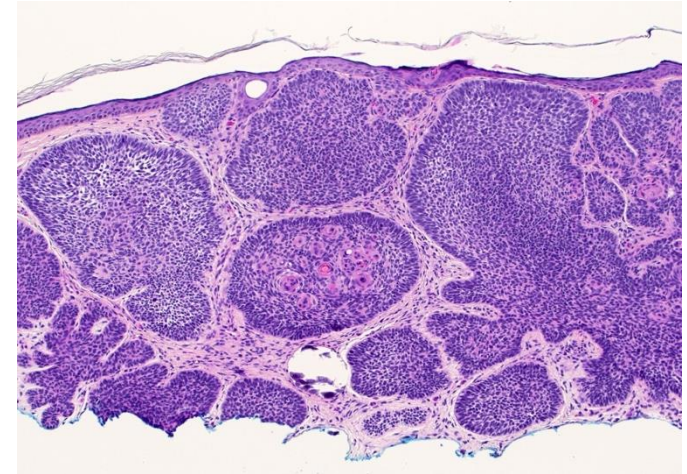
Biomedical Language Understanding
and Reasoning Benchmark



Projects – Few-Shot Pattern Detection for Cancer Tissue

Frederic Mrozinski, M.Sc.

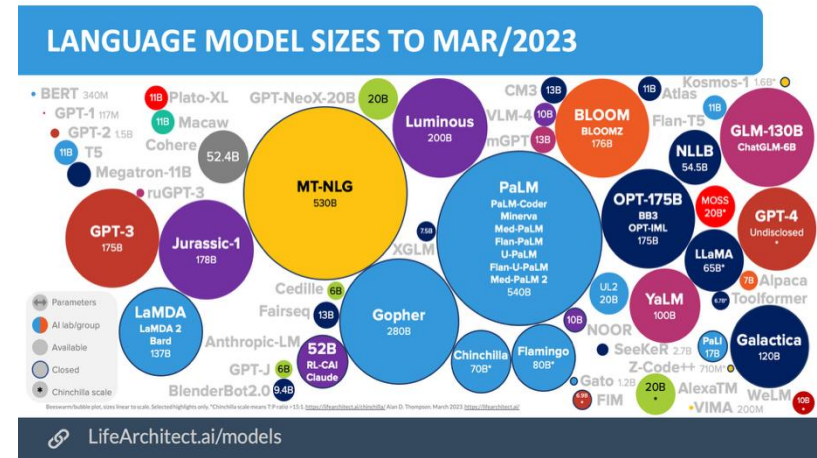
- Medical data is expensive and not publicly available for every disease
- => Can we break down pattern recognition of diseases into more simple subtasks?
- Can we use few-shot capabilities of models like CLIP?



Projects – Green and Efficient AI

Jeremias Bohn, M.Sc.

- Language Model size has exploded over the last years
- Increasing size comes with many disadvantages:
 - Longer training times
 - Increased memory consumption
 - Higher energy consumption (and thus higher CO2 emissions)
 - Independent researchers struggle to contribute, leading roles are taken by big corporations
- We try to focus on reducing model sizes, making training and inference more efficient and less resource-demanding, and reducing power consumption



Projects – User Knowledge and Need Alignment in LLMs

Lukas Ellinger, M.Sc.



Motivation:

- When people talk, they track what is in the others mind (their knowledge, perceptions and thoughts)
- Users ask explicit questions, but often have deeper, unspoken goals

Topics:

Need Inference:

Can the model uncover unspoken needs behind a question?
-> Reveal and support latent goals.

Epistemic Alignment:

How do LLMs model what users know, believe and expect?
-> Maintain shared understanding throughout the conversation.

Answering Strategies:

When should an LLM answer, abstain, or clarify?
-> Match response strategy to ambiguity and user intent.



Questions?

Registration form:

