UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR MEDIZINISCHE INFORMATIK
NACHWUCHSGRUPPE MOVEGROUP

Announcement Master, Bachelor and Study Thesis

Approaches for automatic text-summarisation for medical data

Description:

In the field of automatic text summarisation, two main approaches have emerged throughout the years: Extractive and abstractive text summarisation. In extractive summarisation, one tries to extract the most important sentences and paragraphs from the text. There is no paraphrasing, the parts are extracted as is [1]. In contrast, abstractive text summarisation attempts to provide a concise summary by extracting the most important semantic information from the source material and paraphrasing, akin to how humans would summarise texts [2]. In this thesis, we want to examine how we can apply specific approaches of both types to medical data, using an OpenAccess dataset. The student is tasked with:

- Selecting a dataset suitable for the study
- Selecting and comparing abstractive and extractive approaches
- Coming up with a study design for evaluation

In the case of a student research project or similar, a subsequent thesis is possible.

Keywords: AI in medicine, Transformer, Large language models, text summarization, knowledge extraction

Start: Immediately or by arrangement.

[1] A. W. Palliyali, M. A. Al-Khalifa, S. Farooq, J. Abinahed, A. Al-Ansari and A. Jaoua, "Comparative Study of Extractive Text Summarization Techniques," *2021 IEEE/ACS 18th International Conference on Computer Systems and Applications (AICCSA)*, Tangier, Morocco, 2021, pp. 1-6, doi: 10.1109/AICCSA53542.2021.9686867.

[2] Shakil, Hassan and Farooq, Ahmad and Kalita, Jugal, “Abstractive text summarization: State of the art, challenges, and improvements,” 2024 (https://arxiv.org/abs/2409.02413)

UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR MEDIZINISCHE INFORMATIK
NACHWUCHSGRUPPE MOVEGROUP

If you are interested and have any questions about this topic, please feel free to **book an appointment via:** https://calendly.com/fudickar/

Dr. Sebastian Fudickar

Medical Informatics Initiative Junior Research Group

Integration and Analysis of Multimodal Sensor Signals and Clinical Data for Diagnostics and Investigation of Neurological Movement Disorders (MoveGroup)

Further thesis topics at: **move.ulü.de**