

Wissensgraphen zur Verbesserung von LLMs in der medizinischen Fragenbeantwortung

Forschungsgruppenmodul Datenbanken und Informationssysteme

Adrien Klose; akxeq

April 8, 2024

1 Einleitung

- LLMs advancements and increased usage in population/production; different specialised LLMs
- successful use cases of LLMs
- failed/limited use cases of LLMs
- failure for medical use cases and specific problems for medical question answering
 - give the general workflow of finding entities, resolving ambiguous terms, (creating knowledge graphs), finding relevant information in knowledge graphs, including information in prompts/the prompting process
- discuss possible reasons for limited performance, examples of improvement proposals
- want to examine the impact of triples from knowledge graphs on (medical) question answering

2 Verwandte Literatur

- again short description of the different steps for QA using LLMs and fields of expertise involved
- entity recognition good point (biomedical ER), so that we are not bound to the given triplets in bioasq, but can e
 - extracting questions from natural language questions
 - resolving ambiguous terms
 - ggf. specific features of medical entities
- knowledge graphs
 - creation, maintenance, merging of knowledge graphs and problems involved in those steps
 - extracting/creating knowledge graphs from LLMs
 - knowledge injection into LLMs
 - * direct text injection (problem of creating good representation of KG information), domain specific pre training/problem of closed source for this and other strategies
 - compound relationships/amount and depth of information to include
- LLMs
 - training process of LLM
 - prompting process off LLM, creating good prompts and prompt "frames"

- Basic IO, Chain of Thought, multiple Chain of Thoughts, Tree of Thoughts, Graph of Thoughts
- evaluation metrics for question answering using LLMs
- problem of creating good test data (not to be already included in training process of LLM)
- description of existing knowledge graphs/data (umls, obo foundry, wikidata, bioasq)

3 Experiment

- Bioasq
 - QA Task description
 - problems with existing triples, how we decided which triples to use
 - test different triple configurations/prompting strategies
- reproducing the GoT results with the given examples

4 Evaluation

5 Zusammenfassung

we should additi

one of the planned contributions of FGM is a combined KG tha