



Contribution ID: 5

Type: **Poster**

Medical Reasoning in Large Language Models: A Chain of Thought analysis

With the growing popularity of AI chatbots, many people are querying large language models (LLMs) for medical advice. However, the decision-making of these models is hardly understood, and model hallucinations are particularly dangerous in the medical context. This work systematically analyzes medical reasoning in LLMs. We collect decision graphs from the medical literature and construct chain of thought (COT) prompts based on them. Contrary to popular question-answering benchmarks such as MedQA or MMLU, this approach enables a fine-grained analysis of the medical reasoning process.

Primary author: KLEDE, Kai (Friedrich-Alexander Universität Erlangen-Nürnberg)

Co-authors: Prof. ESKOFIER, Björn (Friedrich-Alexander Universität Erlangen-Nürnberg); MAGISTER, Lucie Charlotte (University of Cambridge); Prof. LIÒ, Pietro (University of Cambridge)

Session Classification: Break + Posters session

Track Classification: Poster