



Beitrag ID: 50 Beitragskennung: 143

Typ: **Poster**

Understanding and Improving Large Language Models

The Machine Learning Group at UHH researches the core mechanisms behind large language models (LLMs). With the rise of ChatGPT, LLMs moved to the focus of public interest. Many research projects analyse the application of LLMs, while comparably little work sheds light on the algorithmic and mathematical foundations behind training, reasoning and abstraction capabilities. We will advance LLM predictability, and provide theoretically sound explanations of their reasoning mechanisms.

The results will make LLMs more transparent, and thus directly help to predict and avoid bias, to make training much more data-efficient, and to compare the reasoning in LLMs with the language understanding in humans.

The project aligns with our commitment to providing more efficient, explainable and transparent algorithms for artificial intelligence.

Find me @ my poster

Keywords

Large Language Models
Transformers
Explainable AI

TentID

Autor: REGNERI, Michaela

Co-Autoren: ABDELHALIM, Alhassan; MOSCH, Merle; LAUE, Sören (Universität Hamburg)