

Bachelor-/Masterarbeit

Prototypische Entwicklung einer offline verwendbaren Konversationellen KI für deutsche Sprache

Aufgabenbeschreibung

Neben dem bekannten ChatGPT [1] gibt es eine Reihe von generativen Sprachmodellen (LLMs), dessen Parameter zum Download verfügbar sind und die damit als Grundlage eigener KI-Systeme eingesetzt werden können [2,4,5]. Wie gut sich die jeweiligen Modelle für spezielle Aufgaben eignen, liegt zu einem Großteil an den Daten, mit denen die Modelle trainiert wurden. Insbesondere bei der Unterstützung von Eingabeaufforderungen in nicht-englischer Sprache gibt es große Unterschiede. Bloom [4] und PolyLM [5] sind Beispiele für LLMs, die unter der Vorgabe von Mehrsprachigkeit trainiert wurden.



In dieser Abschlussarbeit soll ein Konzept erarbeitet werden, mit dem ein vortrainiertes Sprachmodell auf einen Speziellen Anwendungsfall (z.B. Chatten in deutscher Sprache [6]) angepasst werden kann. Die besonderen Herausforderungen bei diesem Projekt sind u.a.:

- Recherche und vergleichende Auswertung aktuell verfügbarer LLMs sowie Datensätze zum Fine-Tuning der Modelle (z.B. [7])
- Aufbau einer Trainingsumgebung auf eine, KI-Cluster
- Trainieren eines Prototypen und Evaluation

Referenzen

- [1] <https://chat.openai.com>
[2] Touvron, Hugo, et al. "Llama 2: Open Foundation and Fine-Tuned Chat Models." arXiv preprint arXiv:2307.09288 (2023).
[3] <https://bigscience.huggingface.co/blog/bloom>
[4] Scao, Teven Le, et al. "Bloom: A 176b-parameter open-access multilingual language model." arXiv preprint arXiv:2211.05100 (2022).
[5] Wei, Xiangpeng, et al. "PolyLM: An Open Source Polyglot Large Language Model." arXiv preprint arXiv:2307.06018 (2023).
[6] <https://opengptx.dfki.de/chat>
[7] <https://open-assistant.io/de>

Iserlohn,
28.07.2023

Fachbereich Informatik und
Naturwissenschaften

Prof. Dr.
Heiner Giefers

Cloud Computing

Telefon
02371 566-5252

E-Mail
giefers.heiner@fh-swf.de

Standort Iserlohn
Frauenstuhlweg 31
58644 Iserlohn

Fachhochschule
Südwestfalen
Sitz: Iserlohn

Hagen
Iserlohn
Lüdenscheid
Meschede
Soest

www.fh-swf.de

Wir geben Impulse

