

ChatGPT bändigen: Sprachmodelle für Innovationen und Technologietransfer nutzen

Yiran Huang und Alexander Studt

Inhalt

- Innovation und Technologietransfer mit LLM
- Warum nicht einfach ChatGPT?
- Exkurs: NLP und LLM
- Überblick. Was bieten wir?
- Hands-On

Innovation und Technologietransfer mit LLM



The image shows a screenshot of the SDSC BW website. The header includes the SDSC BW logo (Smart Data Solution Center Baden-Württemberg) and navigation links: Leistungen, Referenzen, Über uns, Partnernetzwerk, Service, and Aktuelles. There are also buttons for Anbieter, Anwender, and Forscher, along with social media icons for LinkedIn and Twitter. The main content area features an illustration of a man and a woman looking at a large digital display showing data charts and graphs. The display includes a line graph, a bar chart, and a circular gauge showing 60% and 30%. Below the illustration, the text reads: 'Smart Data & KI für den Mittelstand' and 'Finden Sie mit uns den Einstieg'. A paragraph follows: 'Das Thema der Datenanalyse ist so vielfältig, dass die Expertise unserer Forscher und die Vernetzung sowohl in der Anwendercommunity als auch mit den Anbietern von Data Analytics sehr hilfreich sind, wenn es darum geht, möglichst schnell voranzukommen. Das gilt unabhängig davon, ob Sie Anbieter, Anwender oder Forscher sind.'

Idee: Schicken Sie uns Ihre Problemstellung

Wir benutzen LLMs, um Ihnen ihr Problem **zusammenzufassen** und Ihnen **Tools** und **Datensätze** zu empfehlen.

yhuang@teco.edu

Warum nicht einfach ChatGPT?

- Daten/Ideen Souveränität
- Aktuelle Informationen
- Von Experten überprüft

Exkurs: NLP (Natural Language Processing)

- Interaktion zwischen Computern und Menschen durch natürliche Sprachen
- Anwendungen
 - Übersetzungen
 - Spracherkennung
 - Texterstellung
 - Stimmungsanalyse
 - Chatbots
 - Informationsextraktion
 - Text Zusammenfassung
 - ...

Exkurs: NLP (Natural Language Processing)

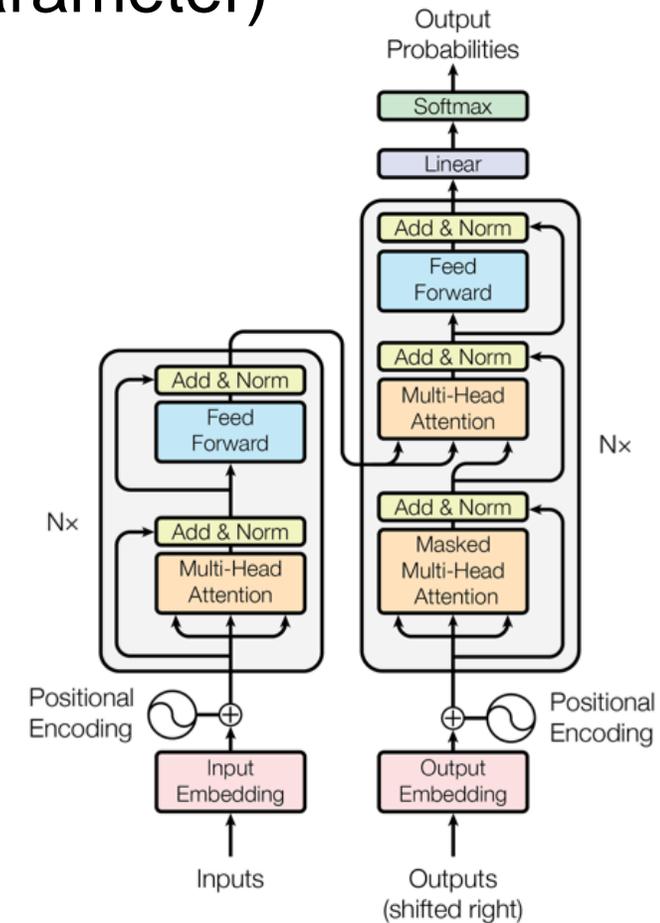
- Interaktion zwischen Computern und Menschen durch natürliche Sprachen
- Schlüsselkomponenten:
 - Syntax: Analyse des Satzbaus
 - Semantik: Bedeutungen von Wörtern/Sätzen
 - Pragmatik: Sprache im Kontext interpretieren
 - Morphologie, Phonetik,...

Exkurs: NLP (Natural Language Processing)

- Interaktion zwischen Computern und Menschen durch natürliche Sprachen
- Herausforderungen
 - Mehrdeutigkeit
 - Kontext und „gesunder Menschenverstand“
 - Vielfalt der Sprachen
 - Intensität der Ressourcen

Exkurs: LLM (Large Language Models)

- Sehr große Modelle (mehr als 100 Millionen Parameter)
- Training mit sehr großen Datensätzen
- Effektive Lösung für NLP Probleme
- Bekanntes Beispiel: ChatGPT
- Es gibt aber auch lokale Modelle



Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

Überblick. Was bieten wir?

- Zusammenfassung
- Github Tools
- OpenML Datensätze

ChatGPT bändigen: Sprachmodelle für Innovationen und Technologietransfer nutzen

Yiran Huang und Alexander Studt