



ESSENZ AUS DER Reihe KI/ChatGPT des DLH

KI/ChatGPT - wie funktioniert das technisch?

Zum Abschluss der Reihe erfolgte in dieser dritten Veranstaltung am 18.04.2023 der Versuch einer technischen Erklärung: Wie funktioniert ChatGPT eigentlich? Wie wurde das Tool programmiert, wie bearbeitet es Anfragen? Wie lassen sich seltsame Antworten erklären? Welche Rückschlüsse lassen sich aus der technischen Funktionsweise für die Nutzung ziehen?

Wir haben diese Punkte diskutiert.

- Geschichte der KI: Erweiterung des Gehirns ---> Ersetzen des Gehirns?
- Trainingsprozess / Erkenntnis Gewinnung durch „neuronale Netze“, am Beispiel von Bilderkennung (Hund -Katze (etc.) / Wichtigkeit von „Fehlern“ bzw. Ungenauigkeiten zugunsten einer allgemeinen Klassifizierung anstatt nur Einzelfälle zu analysieren.
- Beispiel eines „Nano-GPT“ anhand einer begrenzten Menge an Shakespeare Text (1 Million Zeichen), welche Shakespeare-*ähnliche* Texte produzieren soll, *nicht* Originale.
- Extrapolieren des obigen Modelles auf den „grossen“ ChatGPT: funktioniert gleich, nur alles grösser.

Und wie genau?

- Vorwiegend via Erklärungs-Input von Marcel Waldvogel, ergänzt durch Fragen der Teilnehmenden (übermittelt im Chat und gestellt durch Philipp Wampfler). Zum Schluss gab es noch einen kurzen philosophischen Input von Robin Fürst.

Die wichtigsten Erkenntnisse

- ChatGPT „Wissens-Tool“ oder „Logik-Tool“, es produziert eine *Textvervollständigung* aus geshredderten Texten, die aus dem internet extrahiert wurden. Die Ausgabe wird aufgrund von *Häufigkeiten* erzeugt, mit welcher ein „Token“ (=Sequenz von wenigen Zeichen) auf ein anderes „Token“ folgt, gemäss extrahierten Mustern.
- Der Output wird produziert durch das Zusammenspiel der 3 im Namen enthaltenen Phänomene:
 - G – generativ ---> es wird also Text erzeugt, nicht irgendwoher kopiert.
 - P – pre-trained ---> Häufigkeiten von bestimmten Abfolgen von Zeichen werden gelernt.
 - T – Transformer ---> „Aufmerksamkeit“: der erweiterte Kontext einer Zeichenfolge wird berücksichtigt. Ohne diesen Schritt würden absurde Resultate herauskommen. *Mehrere* Auswertungen von Zusammenhängen und dann eine Abstraktion („Normalisierung“ im mathematischen Sinne) sind notwendig für brauchbare Resultate.
- Jeder Output ist zufällig und anders. ChatGPT hat eine Trainingsdatenmenge von zw. 1.5 – 45 Terrabyte (OpenAI deklariert das nicht). / RLHF (Reinforcement Learning form Human Feedback) hilft Fragen – Antworten besser zu kombinieren.
- ChatGPT kann nicht wirklich rechnen, hat keine „Logik“, löst mathematische Aufgaben auch aufgrund von Text-Mustern.

Für mehr Details: marcel-waldvogel.ch/ki

Rückschlüsse auf die Nutzung im Unterricht aufgrund der technischen Funktionsweise wurden nicht diskutiert. Es gilt: je präziser und detaillierter der Input, desto genauer und brauchbarer wird der Output. Mehr dazu erfahren kann man unter dem Stichwort „Prompt Engineering“.