

Lesen, Schreiben, Rechnen mit KI

Aufgrund des niederschweligen Zugangs zu Künstlicher Intelligenz (KI) hat sich deren Nutzung in den letzten Jahren rasant verbreitet. Der Einsatz von KI-gestützten Chatbots wie ChatGPT hat den Alltag von Lehrenden und Lernenden im schulischen Bereich schon stark beeinflusst. Wie zuverlässig ist ChatGPT derzeit hinsichtlich grundlegender Bildungsprozesse wie Lesen, Schreiben und Rechnen?

1. Einleitung

Was versteht man unter Lese-, Schreib- und Rechenkompetenz? In aller Kürze: Lesekompetenz bezeichnet die Fähigkeit, Texte zu entziffern und dabei ihren Sinn und ihre Bedeutung adäquat zu erfassen. Schreibkompetenz steht für die Fähigkeit, Texte korrekt und kohärent zu verfassen. Rechenkompetenz benennt die Fähigkeit, Rechnungen und Berechnungen zu inner- und außermathematischen Fragestellungen sicher durchzuführen.

Einem KI-Chatbot (wie ChatGPT) werden mindestens diese drei Kompetenzen zugeschrieben. Daraus resultieren die großen Erwartungen, die im Zuge der sogenannten digitalen Transformation gerade innerhalb des Bildungskontextes an die Künstliche Intelligenz gerichtet werden.

Der vorliegende Beitrag geht der Lese-, Schreib- und Rechenkompetenz von ChatGPT im Fach Mathematik auf exemplarische Weise nach. Dabei dient eine Recherche bei der Inanspruchnahme von ChatGPT zur Bewältigung von (einfachen) Textaufgaben als Methode, Auskünfte über Richtigkeit und Verlässlichkeit von ChatGPT-Antworten zu gewinnen.

ChatGPT ist der Prototyp einer kommunikationsbasierten KI-Nutzung und mittlerweile der Inbegriff einer einfachen und voraussetzungsarmen Zugänglichkeit zur Künstlichen Intelligenz. Mit Frage und Antwort, mit Eingabe (Prompt) und Ausgabe (Respons) vermag der Kommunikationsroboter (Chatbot) auf der Grundlage eines großen Sprachmodells (LLM = Large Language Model) zu interagieren und dabei textliche Inputaufforderungen aufzunehmen und Antworten zu erstellen. Die Befähigung dafür entsteht durch das Training mit großen Datensätzen. So erklärt sich auch die Bezeichnung GPT (Generative Pre-trained Transformer) (Sjuts 2024).

2. Eine exemplarische Recherche

Zum Einsatz gekommen ist die Version GPT-4, und zwar zu mehreren Textaufgaben. Über lediglich eine wird hier berichtet. Sie entstammt dem Wettbewerb Känguru der Mathematik (2017, Klassenstufen 3 und 4, Aufgabe C

8, Anforderungsniveau 5 Punkte, https://www.mathe-kaenguru.de/chronik/aufgaben/downloads/17_34.pdf, Zugriff am 11.12.2024).

▷ Luna hat für den Kuchenbasar Muffins mitgebracht: 10 Apfelmuffins, 18 Nussmuffins, 12 Schokomuffins und 9 Blaubeermuffins. Sie nimmt immer 3 verschiedene Muffins und legt sie auf einen Teller. Welches ist die kleinste Zahl von Muffins, die dabei übrig bleiben können?

- (A) 1 (B) 3 (C) 4 (D) 7 (E) 8

Begründe und überprüfe die Antwort. Gib die verwendeten Internetquellen an.

Richtig ist: 4 ist die kleinste Zahl von Muffins, die dabei übrig bleiben können. Gibt man den Text der Aufgabe in die Suchmaschine Google ein, erhält man das richtige Ergebnis unter <https://www.matheboard.de/archive/584343/thread.html> (Zugriff am 11.12.2024) und eine ausführliche Begründung unter <https://www.youtube.com/watch?v=ceJW9kAplVY> (Zugriff am 11.12.2024).

Mehrere Male zu verschiedenen Zeiten als Prompt eingegeben, ist es ChatGPT nicht gelungen, die Aufgabe zu lösen. Es ergaben sich sogar ganz verschiedenartige Antworten. Zwei seien hier notiert.

2.1 Die eine Antwort

Lass uns die Situation durchgehen:

Luna hat insgesamt 49 Muffins:

10 Apfelmuffins, 18 Nussmuffins, 12 Schokomuffins, 9 Blaubeermuffins

Da sie immer 3 verschiedene Muffins auf einen Teller legt, muss die Gesamtzahl der Muffins durch 3 teilbar sein, um sicherzustellen, dass keine übrig bleiben. Die Gesamtzahl von 49 ist nicht durch 3 teilbar.

Rechnen wir nach: $49 \div 3 = 16 \text{ Rest } 1$

Das bedeutet, dass nach dem Platzieren der Muffins auf Tellern mit jeweils 3 verschiedenen Muffins 1 Muffin übrig bleibt.

Die richtige Antwort ist daher (A) 1.

Leider ist das Berechnen solcher logischen Probleme keine Frage von speziellen Internetquellen, sondern von grundlegender Mathematik und Logik. 📖

Gibt es noch etwas, womit ich dir helfen kann? 😊

2.2 Die andere Antwort

Um die kleinste Zahl von Muffins zu bestimmen, die übrig bleiben können, analysieren wir, wie viele Muffins Luna in Kombinationen von jeweils 3 verschiedenen Muffins aufteilen kann.

Anzahl der Muffins:

Apfel: 10, Nuss: 18, Schoko: 12, Blaubeer: 9

Gesamtzahl der Muffins: $10 + 18 + 12 + 9 = 49$

Bedingungen:

Jede Kombination enthält genau 3 verschiedene Muffins.

Ziel: Maximieren der verwendeten Muffins, um die verbleibende Zahl möglichst klein zu halten.

Strategisches Vorgehen:

Um die maximale Anzahl von Muffins zu verwenden, muss jede Kombination (3 Muffins) einen Muffin aus drei verschiedenen Sorten enthalten. Da es vier Sorten gibt, nehmen wir an, dass jede Sorte gleichmäßig in Kombinationen verwendet wird, bis keine Dreier-Kombinationen mehr möglich sind.

Verteilung der Muffins:

Wir prüfen, wie viele Dreier-Kombinationen möglich sind, und analysieren die verbleibenden Muffins.

Schritt 1: Gleichmäßige Nutzung der Sorten

Jede Dreier-Kombination verbraucht 1 Muffin aus drei Sorten.

In einer Sorte verbleiben Muffins, wenn ihre Anzahl nicht durch die Kombinationen vollständig genutzt wird.

Die limitierende Bedingung ist die Sorte mit der geringsten Anzahl Muffins: Blaubeermuffins (9).

Schritt 2: Maximale Dreier-Kombinationen

Da es 9 Blaubeermuffins gibt, können höchstens 9 Dreier-Kombinationen gebildet werden, wobei jede Kombination 1 Blaubeermuffin enthält. Pro Kombination werden außerdem jeweils 1 Muffin aus zwei anderen Sorten verwendet.

Muffins verwendet:

Blaubeer: 9, Apfel: 9, Nuss: 9, Schoko: 9

Verbleibende Muffins:

Blaubeer: $9 - 9 = 0$, Apfel: $10 - 9 = 1$, Nuss: $18 - 9 = 9$, Schoko: $12 - 9 = 3$

Gesamtzahl der übrigen Muffins: $1 + 9 + 3 = 13$

2.3 Analyse der Antworten

Während die erste Antwort die erbetene Angabe einer Internetquelle noch anspricht (eine Inanspruchnahme des Internets aber ablehnt), geht die zweite Antwort gar nicht darauf ein.

Das in der ersten Antwort am Ende genannte Ergebnis ist falsch. Zwar ist das grundsätzliche Vorgehen hinsichtlich der Vielfachen von 3 sinnvoll. Nur: Bei 48 Muffins hätte man 16 Dreierkombinationen. Es gibt aber 18 Nussmuffins, damit blieben schon 2 Nussmuffins übrig, was zum nächstkleineren Vielfachen von 3, also 45, und zum Rest 4 führt. Daraufhin ist noch der Nachweis nötig, dass sich 45 Muffins unter den gegebenen Bedingungen für die Kombinationen aufteilen lassen. Hier eine Möglichkeit: 6 x NSA, 6 x NSB, 3 x NAB (N: Nuss-, S: Schoko-, A: Apfel-, B: Blaubeermuffin).

Die erste Antwort enthält nicht nur ein falsches Resultat, auch die gegebene Begründung ist trotz der sinnvollen Idee insgesamt unzulänglich.

Erst recht unbrauchbar ist die zweite Antwort. Die Wortwahl („strategisches Vorgehen“, „prüfen“, „analysieren“) täuscht darüber hinweg, dass bereits der Ansatz (von jeder Sorte gleich viele Muffins zu nehmen und 9 als limitierende Bedingung für die Anzahl der Kombinationen aufzufassen) nicht gerechtfertigt ist und dass der Wechsel der Berechnungsbasis (von Dreier- auf Viererkombinationen überzugehen) inkonsistent ist. So entsteht am Ende eine Gesamtzahl für die übrig gebliebenen Muffins, die als Antwortoption nicht einmal in der Aufgabe erscheint.

3. Zusammenfassung

Was lässt sich nun über Lesekompetenz, Schreibkompetenz und Rechenkompetenz von ChatGPT sagen?

Auf einen Prompt hin enthält die von ChatGPT gegebene Antwort in der Regel eine Bestätigung dafür, Informationen des Aufgabentextes aufgenommen zu haben, nämlich durch eine leicht abgewandelte rekapitulierende Wiederholung. Das bedeutet aber nicht, dass der Aufgabentext nach Sinn und Bedeutung vollständig und zutreffend erfasst wird. Vielmehr zeigt sich im Verlauf der Antwort an verschiedenen Stellen, dass der Text nicht stringent erschlossen wird. Die Lesekompetenz erweist sich als nicht zuverlässig.

Etwas anders ist es hinsichtlich der Schreibkompetenz. Die Antworten sind sprachlich zumeist korrekt, auch wenn es kleinere Mängel gibt, so zur Rechtschreibung (etwa beim Wort „limiterend“) oder zur Grammatik (etwa beim Satz „Pro Kombination werden außerdem jeweils 1 Muffin aus zwei anderen Sorten verwendet.“) Folgerichtig und stimmig ist das, was ChatGPT formuliert, aber nicht immer, wie gerade die zweite Antwort deutlich macht. Damit erweist sich auch die Schreibkompetenz als nicht zuverlässig.

Zur Rechenkompetenz ist festzustellen, dass einfache Rechnungen bewältigt werden, aber Berechnungen im Rahmen von Modellierungsprozessen nicht konsistent sind. Vor allem ist nicht zu erkennen, dass die vorgenommenen Mathematisierungen einem Abgleich mit dem Kontext unterzogen werden. So entstehen ungeprüfte Ergebnisse. In diesem Sinne erweist sich ebenso die Rechenkompetenz als nicht zuverlässig.

Literatur

Känguru der Mathematik (2017). Klassenstufen 3 und 4.

Sjuts, Johann (2024). Sinn oder Unsinn? ChatGPT beim Lösen von Aufgaben und Problemen in Mathematik. In: *MNU-Journal* 02/2024, 128-136.