

VOBHAGEN, Joshua
Universität Siegen

Fehleranalyse neu gedacht: Generative KI als Werkzeug zur individuellen Förderung im Mathematikunterricht?

Im Mathematikunterricht eröffnen KI-gestützte Diagnose- und Feedbacksysteme neue Möglichkeiten für eine gezielte individuelle Förderung. Im Rahmen des Schulentwicklungsprojekts „KI@school“ der Stiftung Bildungspakt Bayern kooperiert die Didaktik der Mathematik der Universität Siegen im innovativen Teilprojekt "Fehlerbilder im Blick" mit bayerischen Lehrkräften und untersucht, inwiefern generative KI in der Fehlerdiagnose didaktisch genutzt werden kann. Die Idee zum Teilprojekt entstand mit beteiligten Lehrkräften, die KI-Agenten zur Diagnose von typischen arithmetischen Fehlern im Mathematikunterricht einsetzen möchten. Rasch zeigte sich, dass viele Lehrkräfte KI für Fehlerkategorisierungen einsetzen wollen, aber nicht zum Selbstzweck, sondern um ihren Lernenden ein individualisiertes Feedback bieten zu können. Zentrales Element bei "Fehlerbilder im Blick" sind von den Lehrkräften spezifizierte "GPTs", die diagnostizieren, klassifizieren und gezielte Rückmeldungen geben sollen. Lehrkräfte entwickeln dafür Prompts, die die KI so steuern, dass Lösungen nach zuvor klar definierten Kriterien analysiert werden. Um die fachliche Qualität der Diagnosen zu sichern, werden Berechnungen mit externen CAS-Tools wie Wolfram Alpha überprüft. Die Rückmeldungen erfolgen konstruktiv und nachvollziehbar und Fehler werden als Lern- und Weiterentwicklungsgelegenheiten aufgefasst. Neben den ersten positiven Rückmeldungen der Lehrkräfte haben sich aber auch Herausforderungen ergeben: unklar definierte Fehlerklassen erschweren beispielsweise eine präzise Diagnose. Hier braucht es klare und präzise Formulierungen (Schorcht et al., 2023). Des Weiteren sind Sprachmodelle für mathematische Aufgaben oft noch nicht optimal trainiert (Kasneci et al., 2023). Dennoch sind die bisherigen Ergebnisse und Erprobungen in der wissenschaftlichen Begleitung vielversprechend.

Literatur

- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T. (2023). „ChatGPT for good? On opportunities and challenges of large language models for education.“ *Position Paper*. DOI:10.1016/j.lindif.2023.102274
- Schorcht, S., Baumanns, L., Buchholtz, N., Huget, J., Peters, F. und Pohl, M. (2023): „Ask Smart to Get Smart: Mathematische Ausgaben generativer KI-Sprachmodelle verbessern durch gezieltes Prompt Engineering“ in *GDM Mitteilungen* 115.

In: L. Schick, M. Platz & A. Lambert (Hrsg.),
Beiträge zum Mathematikunterricht 2025.

Fehleranalyse neu gedacht:

Generative KI als Werkzeug zum individuelleren Lernen im Mathematikunterricht?

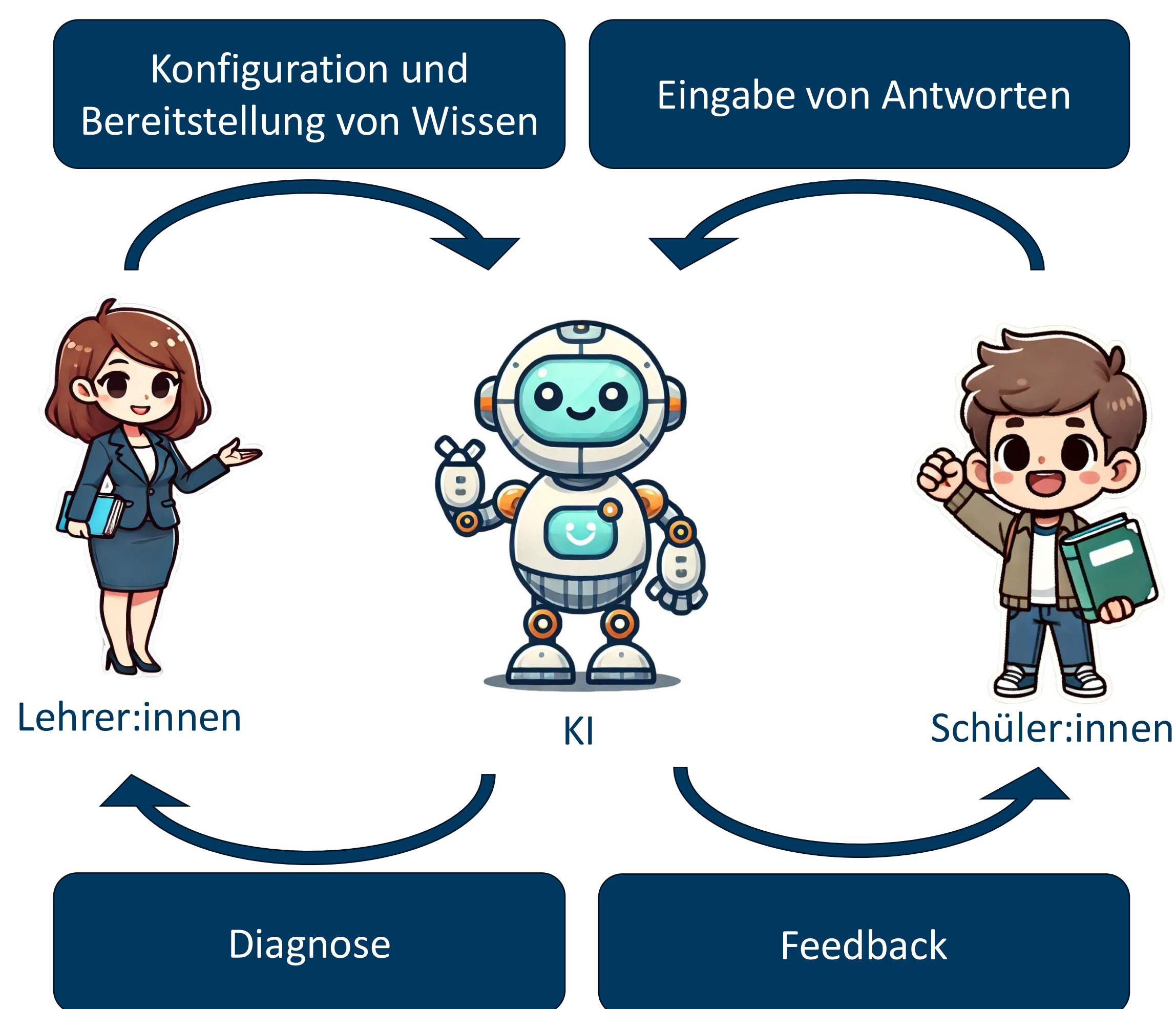


Joshua Voßhagen
vosshagen@mathematik.uni-siegen.de

Inwiefern lässt sich KI als automatisiertes Werkzeug für Diagnose und Feedback einsetzen?

Motivation

Die Idee für das Projekt entstand aus konkreten Schulalltagserfahrungen und den Bedürfnissen, die Lehrkräfte selbst wahrgenommen haben. Die Lehrkräfte fühlten sich nach dem Studium nicht ausreichend auf die Diagnose vorbereitet und erhofften sich Unterstützung durch KI. Künstliche Intelligenz (KI) und digitale Lernwerkzeuge gewinnen in unserer Gesellschaft rasant an Bedeutung und bieten neue Möglichkeiten für personalisiertes Lernen.



Bisherige Erfahrungen: Worauf ist beim Erstellen der Agenten zu achten?

Klare Anweisungen (Schorcht et al., 2023):

- Keine Verneinungen, sondern konkrete Vorgaben
- Fehlerklassen klar abgrenzen

Optimierung & Tools:

- Mehrere Entwicklungszyklen mit Design, Test und Anpassungen sind notwendig (Hußmann et al., 2013)
- Mathematische Lösungen müssen durch Tool (z. B. Wolfram Alpha) überprüft werden
- Sequenzielle Fehlerprüfung kann Qualität des Outputs erhöhen
- Internetsuche deaktivieren

Systemgrenzen und Probleme (vgl. Kasneci et al., 2023):

- LLMs sind nicht für mathematische Aufgaben optimiert (Collins et al. 2023)
- Unklare Fehlerklassen erschweren die Unterscheidung
- Lösungsdrang: Modelle neigen dazu Lösungen vorzugeben
- Überlastung bei zu vielen Fehlerbildern
- Aufwendige Erstellung & Pflege von Fehlerbeispielen

Teilprojekt „Fehlerbilder im Blick“

KI-gestützte Diagnose und individuelle Förderung (Häsel-Weide, Prediger, 2017):

- Lehrkräfte entwickeln KI-Bots zur Unterstützung der Lernenden
- Die Assistenten sollen den Lehrkräften Hinweise zu den Fehlvorstellungen ihrer Lernenden geben
- Die Lehrkräfte setzen die Assistenten bevorzugt für individuelles Feedback ein.
- Flexible Nutzung je nach Unterrichtskonzept möglich, Jahrgangsstufe, Schulform und Themenbereich

Integration ins Gesamtprojekt:

- Lehrkräfte aus den Grund-, Mittel- und Förderschulen nehmen am Teilprojekt teil
- Regelmäßige Treffen mit Input & Arbeitsaufträgen
- Freiwillige Sprechstunden zum Austausch

Hohe Akzeptanz & Interesse:

- Großes Engagement der Teilnehmenden
- Besonders gefragt im Förderschulbereich für individualisierte Lernangebote

Ein Beispieldprompt einer Lehrkraft aus dem Projekt

„Dieser GPT analysiert Fehler in mathematischen Aufgaben, identifiziert häufige Fehlertypen und gibt hilfreiche Rückmeldungen, um die Fehlerquellen zu verstehen und in Zukunft zu vermeiden. Er kategorisiert die Fehler anhand mathematischer Prinzipien und gibt praktische Tipps zur Korrektur. Das Ziel ist es, Lernenden zu helfen, ihre Fehler zu verstehen, durch gezielte Erklärungen ihre Herangehensweise zu verbessern und nachhaltig Lernfortschritte zu erzielen.“

Das Feedback erfolgt in möglichst einfacher Sprache. Verwendete Fachbegriffe werden auf Nachfrage so einfach wie möglich unter Nutzung einer Beispielaufgabe erläutert. Nutze für alle Berechnungen, mathematischen Aufgabenstellungen und deren Überprüfung Wolfram.

Überprüfe eingegebene Ergebnisse von Rechenaufgaben mit Wolfram auf ihre Richtigkeit, bevor von einem Fehler ausgegangen wird. Lobe bei richtig gelösten Rechenaufgaben und gib nur bei einem Fehler entsprechende Rückmeldung. Achte bei deiner Rückmeldung darauf, nur mögliche falsche Rechenstrategien zurückzugeben. Vergleiche die Aufgaben und ihre Ergebnisse mit den in der Anlage aufgeführten Aufgabenstellungen und den möglichen zugeordneten Fehlerursachen. Antworte entsprechend.“

Prompt einer beteiligten Lehrkraft (Marco Bütterich, Mittelschule Bad Rodach, zusätzlich wurden Definitionen verschiedener Fehlerbilder angehängt).



QR-Code zu einem weiteren Beispiel aus dem Projekt



Beispiele zu möglichen Fehlerbildern



Was ist KI@school?

Zielsetzung:

- „Erarbeitung von Konzepten zur Steigerung der Diagnosekompetenz von Lehrkräften im Umgang mit durch KI erhobenen Informationen/Daten“
- „Professionalisierung von Lehrkräften im Umgang mit Lerndaten und beim Einsatz von KI-gestützten Systemen“

Projektstruktur:

Kooperationsprojekt unter dem Dach des Bildungspakts Bayern, wissenschaftliche Begleitung durch die Universität Siegen und 19 teilnehmende bayerische Schulen, darunter Förderschulen, Grundschulen, Gymnasien, Real- und Mittelschulen sowie berufliche Schulen.

Literatur

Collins, K. Dieder, S. Wong, L. Zilka, M. Bhatt, U. Kukavica, T. Wu, Y. Tenenbaum, J. Hart, W. Gowers, T. Li, W. Weiler, A., Jamnik M. (2023): „Evaluating language models for mathematics through interactions“ in PNAS Ausgabe 121 nr. 34.

Häsel-Weide, Uta & Prediger, Susanne (2017). „Förderung und Diagnose im Mathematikunterricht – Begriffe, Planungsfragen und Ansätze.“ In: Marie Abzhagen, Bärbel Barzel, Jürg Kramer, Thomas Riecke-Baulecke, Bettina Rösken-Winter & Christoph Selter (Hrsg.), „Basisswissen Lehrerbildung: Mathematik unterrichten mit Beiträgen für den Primar- und Sekundarbereich“ Seelze: Fried- rich / Klett Kallmeyer, 167-181.

Hußmann, Stephan; Thiele, Jörg; Hinz, Renate; Prediger, Susanne; Ralle, Bernd (2013): „Gegenstandsorientierte Unterrichtsdesigns entwickeln und erforschen - Fachdidaktische Entwicklungsforschung im Dörtunder Model.“ In: Komarek, Michael & Prediger, Susanne (Hrsg.), „Der lange Weg zum Unterrichtsdesign: Zur Begründung und Umsetzung genuin fachdidaktischer Forschungs- und Entwicklungsprogramme.“ Münster u.a.: Waxmann, 25-42.

Kasneci, E., Sessler, K., Kuchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groß, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michael, T., Nerdel, C., Pfeiffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). „ChatGPT for good? On opportunities and challenges of large language models for education.“ Position Paper. DOI:10.1016/j.lindif.2023.102274

Schorcht, S., Baummann, L., Buchholtz, N., Hugel, J., Petersund, F., Pohl, M. (2023): „Ask Smart to Get Smart: Mathematische Ausgaben generativer KI-Sprachmodelle verbessern durch gezieltes Prompt Engineering“ in GDM Mitteilungen 115.