

# Risiken und Nebenwirkungen von KI in der Bildung - und wie man ihnen begegnet

Prof. Dr. Ute Schmid analysiert in ihrem Gastbeitrag, was es braucht, um künstliche Intelligenz kompetent in den Unterricht zu integrieren.

Prof. Dr. Ute Schmid

12.10.2023 Lesezeit 6 min

**KI muss vorausschauend und didaktisch abgesichert in den Unterricht eingebunden werden, um negativen Konsequenzen vorzubeugen. Wie das gehen kann, erläutert Prof. Ute Schmid, Inhaberin des Lehrstuhls für Kognitive Systeme an der Universität Bamberg, in diesem Gastbeitrag.**

Bereits seit vielen Jahren existieren digitale Werkzeuge, die auch im Bildungsbereich verwendet werden – von der Nutzung von Suchmaschinen, der Online-Enzyklopädie Wikipedia über Plattformen, die Erklärvideos anbieten, bis zur maschinellen Übersetzung. Einige dieser Angebote basieren auf Methoden der künstlichen Intelligenz (KI): Die Suche nach Bildern in Google nutzt neuronale Netze, die es ermöglichen, Objekte in Bildern zu identifizieren. Maschinelle Übersetzung basiert auf großen Sprachmodellen, wie sie auch Textgeneratoren wie ChatGPT zugrunde liegen.

Viele Lehrkräfte, genau wie viele Schülerinnen und Schüler, nutzen diese Werkzeuge ganz selbstverständlich. Aber erst der große Hype um ChatGPT hat dazu geführt, dass das Thema KI in der Bildung sehr breite öffentliche Beachtung findet und dass – erfreulicherweise – dazu eine Debatte zu Chancen und Risiken des Einsatzes von KI-Tools im Bildungskontext angestoßen wurde.

## KI in der institutionalisierten Bildung

Aus meiner Sicht zentrale Leitlinien für institutionalisierte Bildung sind:

- Alle Lernenden sollten bestmöglich unterstützt werden, um in allen Fächern gemäß ihren Voraussetzungen Kompetenzen zu erwerben.
- Individualität und Diversität sollten berücksichtigt und gefördert werden.
- Kompetenzvermittlung sollte über bloßes Reproduzieren hinausgehen. Wissensvermittlung sollte so stattfinden, dass Verständnis, Beurteilungsfähigkeit, Transfer und Problemlösefähigkeiten im fachspezifischen Kontext gefördert werden.

KI-Werkzeuge können, sinnvoll umgesetzt und eingesetzt, diese Ziele unterstützen, bergen aber auch die Gefahr, genau den gegenteiligen Effekt zu erzielen.

## Predictive Analytics – riskante Voraussagen

Ich möchte zunächst den Einsatz von predictive analytics betrachten. Maschinelles Lernen bietet die Möglichkeit, aus Daten Modelle zu erstellen, mit denen für neue Daten Vorhersagen getroffen werden können. So werden im industriellen Kontext Modelle trainiert, die aus Systemdaten die Ausfallwahrscheinlichkeit eines Systems vorhersagen (predictive maintenance). Im medizinischen Kontext können Daten beispielsweise genutzt werden, um bei Nierenpatienten das Risiko für akutes Nierenversagen vorherzusagen.

Auf den schulischen Kontext übertragen, könnten Daten genutzt werden, um etwa die Halbjahresnote eines Schülers im Fach Mathematik vorherzusagen. Inzwischen stehen zahlreiche Tools zur Verfügung, die auch Personen ohne Vorkenntnisse in Statistik und maschinellem Lernen nutzen können. Die naive Anwendung birgt jedoch große Risiken – von schlichten Vorhersagefehlern bis hin zu systematischen Verzerrungen, die bestimmte Personengruppen benachteiligen (unfair biases).

## Was gute Voraussagetools ausmacht

Die Qualität eines gelernten Modells hängt maßgeblich von der Qualität der Daten ab, mit denen es trainiert wurde. Qualität basiert auf

1. einer genügend großen Menge von Daten und deren für die Grundgesamtheit repräsentativen Verteilung,
2. einer kritischen Beachtung von möglichen Verzerrungen in historischen Daten (etwa dass bislang kaum Mädchen in der Oberstufe das Vertiefungsfach Physik gewählt haben),
3. einer sorgfältigen Beurteilung, inwiefern erfasste Daten tatsächlich indikativ für ein bestimmtes, nicht direkt messbares Konstrukt sind (etwa ob die Dauer, in der eine Website geöffnet ist, mit der Aufmerksamkeit, mit der Lernende den Inhalt rezipieren, korreliert).

## Warum es kritische Lehrkräfte braucht

Menschen neigen dazu, den Ausgaben von Maschinen, insbesondere wenn es sich um Zahlenwerte handelt, unkritisch zu vertrauen (automation bias). Dies ist umso bedenklicher, wenn es sich um naiv trainierte Modelle handelt, die nicht kritisch auf systematische Verzerrungen, Robustheit und prädiktive Genauigkeit hin evaluiert wurden. Seit vielen Jahren wird immer wieder moniert, dass viele Ärztinnen und Ärzte ungenügende Statistikkompetenzen haben, um die Ergebnisse von Studien sinnvoll bewerten zu können. Bei Lehrkräften sind die Themen Statistik und empirische Bildungsforschung bislang nur in begrenztem Umfang im Studium integriert. Entsprechend besteht hier die Gefahr, dass predictive analytics tools nicht sinnvoll eingesetzt werden.

## Potenziale und Grenzen von maschinellem Lernen

Der Einsatz von maschinellem Lernen zur Benotung von Schulaufgaben oder Hausarbeiten birgt ähnliche Gefahren. Trainiert man ein Modell mit wenigen und verzerrten Daten – etwa indem man zehn Arbeiten selbst benotet und diese Information dann zum Lernen verwendet, ist es sehr wahrscheinlich, dass das Modell einfache, aber nur teilweise relevante Merkmale wie die Länge einer Antwort extrahiert und zur Vorhersage nutzt. Ein langer Nonsens-Text würde also besser bewertet als eine kurze präzise Antwort.

Maschinelles Lernen kann, sinnvoll eingesetzt, durchaus einen Mehrwert im schulischen Kontext bringen. Beispielsweise könnten Vorhersagemodelle für ganze Klassenverbände Hinweise darauf geben, ob eine Klasse in einem Fach schlechter abschneiden wird als andere. Nutzt man Methoden der erklärbaren künstlichen Intelligenz, also Methoden, die die Vorhersagen von maschinell gelernten Modellen nachvollziehbar machen, könnte man Rückschlüsse auf mögliche Ursachen ziehen, sei es eine andere demografische Zusammensetzung, der man gezielt pädagogisch und didaktisch begegnen müsste, oder seien es höhere Ausfallraten von Unterrichtsstunden.

## Risiken der Predictive Analytics

Damit maschinelles Lernen gut funktioniert, müssen genügend Daten vorhanden sein. Dies könnte dazu führen, dass bestimmte Lerninhalte nicht mehr aus didaktischen Erwägungen digital vermittelt und geprüft werden, sondern nur, um die datenhungrigen Maschinen zu füttern. Zudem besteht die Gefahr, einfache Prüfformate wie Multiple Choice gegenüber freien Formaten zu bevorzugen. Da die Art, wie geprüft wird, einen starken Einfluss auf die Art hat, wie gelernt wird, besteht die Gefahr, dass Schülerinnen und Schüler Inhalte eher so lernen, dass sie einfach wiedererkannt werden können, und weniger mit dem Ziel, eigene Lösungen zu generieren. Predictive Analytics birgt also aus mehreren Gründen die Gefahr, dass Lernprozesse wieder behavioristisch und weniger kognitivistisch und konstruktivistisch gestaltet werden.

Hier können **Intelligente Tutorsysteme (ITS)** eine sinnvolle Alternative bieten. Auch bei solchen KI-Tools kommt oft maschinelles Lernen zum Einsatz. Fokus bei ITS ist jedoch die individuelle Wissensdiagnose im Problemlösekontext; das Ziel dabei ist, Wissenslücken und Fehlkonzeptionen auf individueller Ebene zu diagnostizieren, um gezieltes Feedback zu geben.

## ChatGPT und Co.: Dialogsysteme und das Wissen um ihre Quellen

Als zweites Beispiel für KI-Tools möchte ich **Dialogsysteme wie ChatGPT** betrachten. Lehrkräfte wie Lernende können solche Systeme kreativ nutzen, etwa um sich Ideen generieren zu lassen, Textentwürfe zu bestimmten Themen zu erstellen oder auch Programmcodes zu generieren. Allerdings ist es für einen sinnvollen Einsatz unerlässlich, dass verstanden wird, dass Ausgaben von Sprachgeneratoren nicht über spezielles Wissen oder ein Weltmodell verfügen.

Es ist somit unerlässlich, Ausgaben kritisch zu prüfen und formulierte Sachverhalte anhand anderer, zuverlässiger Quellen zu verifizieren. Entsprechend muss Medienpädagogik möglichst zeitnah um das Thema generative KI erweitert werden. Schülerinnen und Schüler müssen frühzeitig Kompetenzen erwerben, wie Inhalte kritisch hinterfragt und unabhängig geprüft werden können. Die Suche nach geeigneten Informationsquellen sowie die kritische Beurteilung der Vertrauenswürdigkeit von Quellen sollte bereits ab der Sekundarstufe in die Lehrpläne aufgenommen werden.

## Sinnvolle didaktische Begleitung statt Verbot

Die meisten technologischen Neuerungen, seien es der Taschenrechner oder die Rechtschreibprüfung im Textverarbeitungssystem, helfen uns, bestimmte Aufgaben effizienter und oft auch korrekter zu bewältigen. Es wäre unsinnig, solche Werkzeuge aus der Schule zu verbannen – und das gilt auch für ChatGPT und Co. Allerdings sollte die Nutzung sinnvoll didaktisch begleitet werden. Die Nutzung des Taschenrechners darf nicht dazu führen, dass Schülerinnen und Schüler nicht mehr in der Lage sind, einzuschätzen, ob ein Rechenergebnis plausibel ist, oder dass nicht mehr verstanden wird, was es bedeutet, die Wurzel einer Zahl zu ziehen, und in welchem Kontext man diese Operation sinnvoll anwendet.

Entsprechend sollte bei der Nutzung von Sprachgeneratoren und auch bei maschinellen Übersetzern darauf geachtet werden, dass die Fähigkeit zur Strukturierung komplexer Gedankengänge oder das Finden einer sprachlich passenden Übersetzung nicht verloren gehen. Es ist zu beobachten, dass unsere Sprache immer mehr normiert wird und teilweise verflacht. Generierte Sätze folgen einer zwar sehr guten, aber doch auch sehr einheitlichen grammatischen Konvention, seltene Worte werden

nicht genutzt. Diese Merkmale werden auch von KI-Detektoren verwendet, um zu prüfen, ob ein Text maschinell erstellt wurde. Entsprechend rege ich an, Schülerinnen und Schüler gezielt Texte erstellen zu lassen, die von einem Detektor als vom Menschen generiert klassifiziert werden, um sprachliche Vielfalt zu erhalten.

## Für KI in der Schule – gegen die gesellschaftliche Spaltung

Trotz der diskutierten möglichen Gefahren plädiere ich ohne Wenn und Aber für den Einsatz von KI-Werkzeugen im Bildungsbereich. KI-Technologien sind zunehmend aus unserem Alltag nicht mehr wegzudenken und entsprechend ist die Vermittlung eines sinnvollen, sicheren und souveränen Umgangs wichtig – auch, um einem zunehmenden digital divide, also der Kluft in der Gesellschaft zwischen Menschen mit und ohne Möglichkeit, Informations- und Kommunikationstechnologien zu nutzen, entgegenzuwirken.

Damit die Chancen der KI-Tools genutzt und die Risiken ihres Einsatzes minimiert werden, ist es notwendig, dass Lehrende wie Lernende Grundkompetenzen in Informatik und Statistik aufbauen. Die KI-Forschung ist gefragt, die aktuellen, oft rein datengetriebenen Methoden zu erweitern, sodass komplexe Inhalte sinnvoll mit KI-Tools vermittelt und geprüft werden können. Unabdingbar ist es, dass die Diskussion um zukunftsfähige Kompetenzziele in allen Fachdidaktiken geführt werden. KI-Tools, die Lernende gezielt fördern, Individualität erhalten und Diskriminierung vermeiden, können nur im interdisziplinären Dialog zwischen KI-Forschenden und den Fachdidaktiken entwickelt werden.