

Universität Paderborn • Warburger Straße 100 • 33098 Paderborn

An den Ausschuss für Schule und Bildung
Landtag Nordrhein-Westfalen
Platz des Landtags 1
40221 Düsseldorf
Stichwort "A15 — KI-Algorithmen — 12.09.2023"

Institut für Informatik
Didaktik der Informatik

ggf. Akademischer Titel
Carsten Schulte

T +49 5251 60-6343
carsten.schulte@uni-paderborn.de

Sekretariat
Cornelia Wiederhold
T +49 5251 60-6523
F +49 5251 60-6623

Fürstenallee
33102 Paderborn

7. September 2023

Stellungnahme

Welche Rolle können und sollen „KI-Algorithmen im Informatikunterricht“ spielen?
Ausgehend von der Fragestellung werde ich zunächst zu Algorithmen im Informatikunterricht Stellung nehmen und diese Überlegungen dann auf das Thema KI, bzw. genauer, auf das Gebiet des Maschinellen Lernens (ML) übertragen und am Ende drei Schlussfolgerungen ziehen.

Um die Stellung des Themengebiets Algorithmik zu erläutern, zunächst ein kurzer Blick in die ländergemeinsamen Anforderungen der KMK an die Informatik-Lehrkräftebildung (KMK, 2019): Die dort gelisteten fachlichen Studieninhalte (S. 36) gehen vor allem auf die klassischen fundamentalen Grundlagen des Faches ein. Algorithmen und Datenstrukturen bilden einen bzw. den wichtigen inhaltlichen Schwerpunkt. KI und ML sind dagegen nicht enthalten. KI stellt eine echte Erweiterung dar.. Es geht dabei aber nicht allein um eine inhaltliche Erweiterung, sondern auch um eine konzeptionelle bzw. bildungsphilosophisch-didaktische Erweiterung.

Die klassischen fachlichen Grundlagen, die vor allem etwa in den genannten Richtlinien der Lehrkräftebildung gelistet sind, harmonieren sehr schön mit einer klassischen fachdidaktischen Sichtweise und Begründung des Schulfaches Informatik, in dem das algorithmische Problemlösen im Mittelpunkt steht (Vgl. dazu (Schulte, 2001)).

Dazu als ein herausgegriffenes Beispiel ein Zitat aus der Präambel des neunten Kernlehrplans für das Wahlpflichtfach Informatik in NRW, in diesem Fall für die Realschule: „Ein wesentliches Ziel hierbei ist das **selbstständige informatische Problemlösen**.“ (MSB NRW, 2023, S. 7, Hervorhebung im Original) und etwas weiter: „Die Schülerinnen und Schüler erwerben Fähigkeiten zur kritischen und verantwortungsvollen Analyse, Modellierung und Implementierung ausgewählter Informatiksysteme. Informatische Bildung zielt auf eine altersgemäße Auseinandersetzung mit der Gestaltung und der Sicherheit von Informatiksystemen sowie den Folgen und Wirkungen ihres Einsatzes für die Gesellschaft ab. Dabei stehen stets fundamentale und zeitbeständige informatische Ideen, Konzepte und Methoden im Mittelpunkt. In einer Kultur der Digitalität gehört hierzu auch die reflektierte Auseinandersetzung mit textgenerierenden Systemen, die auf künstlicher Intelligenz beruhen.“ (MSB NRW, 2023, p. 7)

In der klassischen fachdidaktischen Perspektive fokussiert sich der Informatikunterricht auf das informatische Problemlösen, wie auch in der Hervorhebung im zuvor zitierten Kernlehrplan zu erkennen ist. Dieses wird nun (siehe das längere Zitat) erweitert, etwa um Aspekte des ML (hier: textgenerierende Systeme). (Nebenbei: Die Analyse von Systemen, die ebenfalls genannt wird, stellt dagegen keine Erweiterung der klassischen algorithmenbezogenen Sichtweise auf Informatikunterricht dar, da die Problem-Analyse zum informatischen Problemlösen dazugehört.)

Um die klassische fachdidaktische Perspektive zu verdeutlichen kann die Ankündigung einer entsprechenden Workshopveranstaltung (19. Informatiktag NRW der Fachgruppe Informatische Bildung NRW der GI, März 2023) dienen:

„Im Workshop setzen wir uns mit dem Wirkprinzip neuronaler Netze auseinander. Wir erstellen ein neuronales Netz mit Neuronen in der Eingangs-, Verborgenen- und Ausgangs-Schicht, die jeweils schichtweise miteinander verbunden sind. Auf dieser Datenstruktur implementieren wir die Forwardpropagation für die Anwendung des neuronalen Netzes und die Backwardpropagation als Lernverfahren. Damit wird das prinzipielle Wirkprinzip neuronaler Netze durchleuchtet und ein Beitrag zu einer realistischen Einschätzung des Potentials künstlicher Intelligenz geleistet.“ (Quelle: <https://indico.uni-wuppertal.de/event/215/contributions/1153/>)

Die bereits in den 1970er Jahren beschriebene Idee des algorithmischen Problemlösens jetzt auch auf die digitale Welt anzusetzen, um etwa Neuronale Netze zu verstehen, ist nur folgerichtig und konsequent. Es birgt aber auch einige Fallstricke und potenzielle Missverständnisse, die im Folgenden erläutert werden sollen.

Maschinelles Lernen selbst stellt eine eigene und neuartige Form des informatischen Problemlösens dar, und kann aufgrund der Bedeutung in der Gesellschaft aber auch der Fachwissenschaft selbst (Rahwan et al., 2019) als wichtige eigenes Thema für den Informatikunterricht angesehen werden (Höper & Schulte, 2023; Tedre et al., 2021).

	<i>Entwicklung klassischer Systeme</i>	<i>Entwicklung datengetriebener Systeme</i>
Problemlöseschritt 1	Formalisierung und Analyse des Problems	Sammeln von Daten für den entsprechenden Kontext
Problemlöseschritt 2	Entwurf einer Problemlösung	Aufbereiten der Daten (Filtern, Bereinigen, Labeln, ...)
Problemlöseschritt 3	Implementierung als algorithmisches Programm	Trainieren eines Modells mit den aufbereiteten Daten
Problemlöseschritt 4	Kompilierung und Ausführung des Programms	Evaluieren und Nutzen des Modells
Problemlösung	Programm ist formal hinsichtlich der Korrektheit prüfbar, transparent und liefert nachvollziehbare Ergebnisse	Modell ist für Testdaten (statistisch) mehr oder weniger gut, (oft) nicht transparent und liefert nicht vollständig nachvollziehbar oder erklärbare Ergebnisse

Abbildung 1: nach (Tedre et al., 2021)

Bei der Entwicklung klassischer, algorithmischer Systeme muss für die Problemlösung das Problem selbst durchdrungen werden, um Regeln aufzustellen, nach denen die Problemlösung funktioniert und diese Regeln (im weitesten Sinne: ein Algorithmus) werden dann als Problemlösung implementiert. Beim Maschinellen Lernen (rechte Seite der Abbildung 1) fehlt diese analytische Durchdringung, die Regeln werden nicht vom Menschen in das System eingegeben und dementsprechend nicht vom Menschen

entworfen, sondern vom System anhand von bereitgestellten Daten als eine Art statistisches Muster schrittweise herausgefunden bzw. wie es im KI-Jargon genannt wird: „gelernt“.

In diesem Prozess spielen die Algorithmen bzw. der Code eine wichtige, aber nur kleine Rolle. Oft wird das mit der folgenden Grafik verdeutlicht:

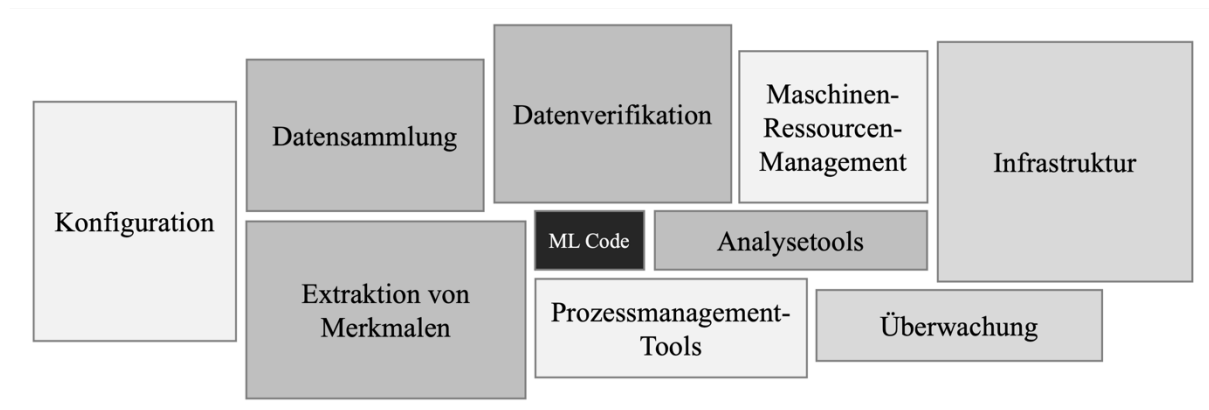


Abbildung 2 aus (Höper & Schulte, 2023) nach: (Sculley et al., 2015). Der schwarze Kasten in der Mitte (ML Code) bezieht sich auf die Algorithmen, die anderen Kästen auf die in ML-Systemen zusätzlichen wichtigen Elemente, die für die Problemlösung ebenso relevant sind.

Hier nun kann ein Missverständnis auftauchen. Ausgehend von der klassischen Idee des informatischen Problemlösens wird man Abbildungen 1 und 2 folgerichtig so interpretieren müssen: Nicht nur ändert sich der Problemlöseprozess, d.h. die Schülerinnen und Schüler müssen zunächst das normale Verfahren kennen und beherrschen, bevor sie dann die neue Variante lernen können (Abbildung 1). Und diese neue Variante hat wiederum ganz viele eigene in sich komplexe Schritte, insbesondere jene die sich auf die Rolle von Daten beziehen, (siehe Abbildung 2) – wie soll das im Informatikunterricht umsetzbar sein?

Erschwerend kommt hinzu, dass das typische klassische informatikdidaktische Erklärmodell nicht mehr funktioniert, welches in der obigen Workshopankündigung ausgedrückt wird: Wir bauen das System nach – d.h. nicht das ganze System, sondern den algorithmischen Kern etwa von Neuronalen Netzen – und so sei dann „das prinzipielle Wirkprinzip neuronaler Netze durchleuchtet und ein Beitrag zu einer realistischen Einschätzung des Potentials künstlicher Intelligenz geleistet“. In der klassischen Problemlösesicht wird damit nicht nur irgendein kleiner Beitrag, sondern der zentrale und entscheidende Beitrag geleistet – aber das stimmt für ML-Systeme so nicht (vgl. dazu Abbildung 2: ML Code (die Algorithmen) ist zwar zentral, die Funktionsweise eines ML-Systems hängt aber noch von diversen weiteren Faktoren ab).

D.h. es braucht eine neue Art von Erklärmodellen und ggf auch von unterrichtlichen Zugängen, die diese klassische Herangehensweise erweitern oder sogar (für das Thema ML) weitgehend ablösen.

Um ein Beispiel für ein solches Erklärmodell anzudeuten:

Abbildung 3 Hannah Bast erklärt das Wirkprinzip eines Neuronales Netzes unter Auslassung des algorithmischen Kerns (Backpropagation). Quelle: <https://www.zdf.de/dokumentation/terra-x/lesch-und-co-live-chatgpt-wie-intelligent-ist-ki-100.html> (ca. min 10ff)

Anhand einer vorbereiteten Lernumgebung wird in dem Beispiel (Abbildung 3) ein Neuronales Netz erkundet mit dem Ziel, die wesentlichen Wirkprinzipien zu erklären – hier gehören die algorithmischen und damit verknüpften mathematischen Grundlagen (also etwa Backpropagation) explizit nicht dazu. Stattdessen wird interaktiv das Wirkprinzip anhand eines sehr einfachen Beispiels schrittweise visualisiert und dabei erläutert und interpretiert. Für Details siehe das Video (Darin – nebenbei – auch interessant, wie Prof. Bast ab und zu korrigiert, wenn Prof. Lesch von Algorithmen spricht und sie darauf verweist, dass ein Neuronales Netz eher als Modell oder als trainiertes Modell bezeichnet wird).

Für den Unterricht bieten sich meines Erachtens solche Erklärmodelle und Beispiele an – sie umgehen einerseits die algorithmische Komplexität der Eigenkonstruktion mit den elementaren Bausteinen einer Programmiersprache und können gleichzeitig auf diese Art und Weise dann auch weitere, nicht-algorithmische Aspekte wie etwa die Rolle der Daten, besser einbeziehen.

Zwei Aspekte gilt es zusätzlich zu beachten:

1. Zwar können so die Wirkprinzipien besser erläutert werden, aber derzeit gibt es tatsächlich nicht wirklich umfassende und theoretisch abgesicherte Erklärmodelle für ML-Systeme – diese Frage ist so komplex, dass sich hier ein eigener, stark interdisziplinär ausgerichteter Teilbereich der Informatik entwickelt hat, die erklärbare KI (XAI). Hier werden vermutlich weitere Erklärmodelle entwickelt werden, die sich ggf. dann auch Ideen für schulische Behandlung liefern können. Zudem könnte die Existenz und Notwendigkeit von XAI das klassische algorithmenzentrierte Weltbild sinnvoll ergänzen.
2. Das Thema KI sollte sich m.E. nicht im Nutzen von Erklärmodellen und/oder Ausprobieren von ML-Systemen erschöpfen, sondern auch die eigene Konstruktion und Entwicklung einschließen (wie in der klassischen Sichtweise auch). Meiner Meinung nach kann hier sinnvoll auf geeignete Teil-Lösungen in Form von Programm-Bibliotheken zurückgegriffen werden, mit denen ein ML-

System in wenigen Code-Zeilen entwickelt werden kann. Programmieren auf einer höheren konzeptuellen Ebene, sozusagen. Die auch didaktische Funktion von diesen Programmier-Bibliotheken (APIs) könnte und sollte stärker genutzt werden, um SuS zu befähigen, eigene KI-Systeme zu entwickeln und auf diese Weise die Wirkungsprinzipien von KI auch auf konstruktivistische Art zu durchdringen.

Zusammenfassung:

Um die ggf. etwas implizit gebliebenen Thesen zu „KI-Algorithmen im Informatikunterricht“ zusammenzufassen:

1. KI-Systeme können nur schwer auf ihren algorithmischen Kern reduziert werden. Für den Informatikunterricht scheinen mir andere Erklärmodelle, die z.B. die Rolle der Daten betonen, geeigneter.
2. ML stellt eine echte Erweiterung des informatischen Problemlöse-Werkzeugkastens dar und ist es auch deshalb wert, im Informatikunterricht als ein eigenständiger Strang thematisiert zu werden (vgl. auch (Höper & Schulte, 2023)). Dabei sollten auch die eigene Entwicklung (z.B. das Trainieren an echten Datensätzen) und das Reflektieren der Rolle von Daten sowie die Entwicklung eines Datenbewusstseins berücksichtigt werden.
3. Hinzu kommt die Bedeutung des Bereichs für die digital vernetzte Welt und den individuellen Alltag, die vermutlich in Zukunft so bleiben oder sogar noch steigen wird. Das Verstehen-Können des Verhaltens dieser „KI-Maschinen“ im Kontext ist daher ein wichtiges Bildungsziel. Das gilt m.E. ebenso für die Fähigkeit mittels entsprechender Bausteine eigene KI-Systeme zusammenstellen (programmieren) zu können. Als drittes Element sollte perspektivisch ein Lernbereich ‚Einarbeiten-Können in neue Technologien‘ entwickelt werden.

Angesichts der Änderungen, sowie der fehlenden Verankerung im Lehramtsstudium (vgl. KMK-Richtlinien zur Lehrkräftebildung) und der Einführung des Pflichtfachs Informatik sind systemetische Lehrkräftefortbildungen – insbesondere auch unter Einbezug der Universitäten – zu verstärken.

Mit freundlichen Grüßen
Carsten Schulte

Zu meiner Person: Prof. für Didaktik der Informatik an der Universität Paderborn. Co-Projektleiter des Prodabi-Projekts (www.prodabi.de), das Fortbildungen und Unterrichtsmaterialien zu den Themen data science und KI entwickelt. Co-Projektleiter von zwei Teilprojekten des ortsverteilten Sonderforschungsbereichs ‚Constructing Explainability‘ (www.trr318.de) zum Thema erklärbare KI und seit August 2023 Co-Projektleiter des CDEC-Projekts, in dem u.a. ML-basierte Auswertungsverfahren für Satellitendaten Schülerinnen und Schülern der Sekundarstufe I zugänglich gemacht werden sollen (<https://cdec.io>).

Quellen

- Höper, L., & Schulte, C. (2023). Paradigmenwechsel vom klassischen zum datengetriebenen Problemlösen im Informatikunterricht. *MNU journal*, 76(4), 314–320.
- KMK. (2019). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung (Beschluss der Kultusministerkonferenz vom 16.10.2008 i. D. F. vom 16.05.2019)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. 'Sandy', ... Wellman, M. (2019). Machine behaviour. *Nature*, 568(7753), 477–486. <https://doi.org/10/gfzvhx>
- Schulte, C. (2001). Vom Modellieren zum Gestalten–Objektorientierung als Impuls für einen neuen Informatikunterricht. *informatica didactica*, 3. <http://ddi.cs.uni-potsdam.de/InformaticaDidactica/Schulte2001.htm>
- Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden Technical Debt in Machine Learning Systems. *Advances in Neural Information Processing Systems*, 28. <https://proceedings.neurips.cc/paper/2015/hash/86df7dcfd896fcdf2674f757a2463eba-Abstract.html>
- Tedre, M., Denning, P., & Toivonen, T. (2021). CT 2.0. In *21st Koli Calling International Conference on Computing Education Research* (S. 1–8). Association for Computing Machinery. <https://doi.org/10.1145/3488042.3488053>