

**Autor:** Meißner, Thomas [ner/809]

**Ressort:** Künstliche Intelligenz

<sup>1</sup> von PMG gewichtet 05-2021

**Mediengattung:** Online News

**Visits (VpD):** 0,145 (in Mio.)<sup>1</sup>

## Skeptische Wissenschaftler

# Der Nutzen der Künstlichen Intelligenz – und ihre Nebenwirkungen

Der Nutzen von Künstlicher Intelligenz sei nicht so groß wie er erscheine, meinen einige Wissenschaftler. Was sie monieren, adressiert die Forschung bereits – wird jedoch noch nicht umgesetzt.

Thomas Meißner

Toronto. Können wir Künstlicher Intelligenz (KI) vertrauen, zum Beispiel, aber nicht nur, in der Medizin? Sind die Software-basierten Algorithmen nicht eine Blackbox, die Ergebnisse ausspuckt, deren Ursprünge für die Nutzer undurchschaubar bleiben, etwa in der bildgebenden Diagnostik oder bei der Analyse multipler medizinischer Befunde? Wer trägt die Verantwortung, wenn am Ende etwas schief geht?

Der gegenwärtige KI-Enthusiasmus sei übertrieben, meinen nordamerikanische Wissenschaftler um Dr. Boris Babic von der University of Toronto in Kanada in der aktuellen Ausgabe des Magazins „Science“. Es sei erforderlich, mögliche Nachteile näher zu beleuchten (Science 2021; 373: 284-286).

Sie fordern die Entwicklung von Methoden, mit denen Blackbox-Entscheidungen nachvollziehbar gemacht werden, bezweifeln jedoch zugleich, dass dies ausreichend möglich sei. Die Folge könnte sein, dass KI in der Hand unzureichend qualifizierter Nutzer ein falsches Sicherheitsgefühl erzeugen und damit letztlich zu Schäden führen würde.

Babic und Koautoren kritisieren vor diesem Hintergrund die aus ihrer Sicht zu laschen Zulassungsprozesse für KI-basierte medizinische Systeme durch die US-Behörde FDA.

Verständnis für Möglichkeiten

„Für eine Patientin kann der Kopf ihres Arztes zunächst genauso eine Blackbox darstellen wie eine Maschine“, hält der Potsdamer Professor für Medizinische Ethik, Dr. Robert Ranisch, dem entgegen. „Wir müssen auch nicht unbedingt

erklären können, warum Flugzeuge fliegen“, heißt es in einer vom Science Media Center Germany (SMC) verbreiteten Stellungnahme. Allerdings hält Ranisch es für nötig, ein breiteres Verständnis für Möglichkeiten und Limitationen von KI in der Medizin zu schaffen.

Nachträglich Erklärungen für Blackbox-Klassifikatoren zu schaffen, sei nicht sinnvoll, meint **Professorin Ute Schmid** vom Fraunhofer-Institut für Integrierte Schaltungen in Erlangen. Mit der **Universität Bamberg** entwickle man bereits seit Jahren Ansätze des maschinellen Lernens für die bildgebende Diagnostik, bei denen „induktives logisches Programmieren (ILP)“, eingesetzt wird. Das gehe über das hinaus, was in „Science“ diskutiert werde: Interpretierbare Befunde werden im Lernprozess mit Erklärungen kombiniert. „Beispielsweise kann visuelles Hervorheben von Bereichen in einem Bild mit sprachlichen Erklärungen kombiniert werden, um zu zeigen, wo sich Tumorgewebe befindet.“

Zusätzlich werde angegeben, aufgrund welcher Parameter wie zum Beispiel Größe, Lage und Form des Gewebearals das KI-System eine bestimmte Tumorklasse identifiziert hat, so Schmid. Gelernte Modelle könnten nie vollständig korrekt sein. „Experten müssen in der Lage sein, Entscheidungen zu korrigieren.“

Angst vor Kontrollverlust

Und der ehemalige Vorsitzende des Deutschen Ethikrats, Professor Peter Dabrock, meint: „Egal wie der Algorithmus entschieden hat, ein Mensch muss die Entscheidung rechtfertigen können.“ Wo dies nicht möglich sei, verdienten KI-Systeme auch kein Vertrauen. Dabrock warnt davor, Algorithmen abstrakt zu betrachten. Sie müssten eingebunden in sozio-technische Systeme bewertet werden.

Hinter der Diskussion um KI in der Medizin wie auch in anderen gesellschaftlichen Bereichen steckt – oft unausgesprochen – die Angst vor Kontrollverlust. Babic und Kollegen sind nicht die ersten, die Transparenz- und Nachvollziehbarkeitsforderungen stellen. Transparenz heißt, die Frage zu beantworten, wie Entscheidungen getroffen werden. Nachvollziehbarkeit bedeutet, dass unabhängige Experten Resultate und das Verhalten von Entscheidungssystemen eigenständig untersuchen können.

Die Informatikerin Professorin Katharina Zweig von der TU Kaiserslautern beschreibt in ihrem populärwissenschaftlichen Buch „Ein Algorithmus hat kein Taktgefühl“ (Heyne 2019) fünf Regulierungsstufen, die anhand des Schadenspotenzials eines algorithmischen Entscheidungssystems eingesetzt werden könnten.

Diese Stufen bieten die Möglichkeit, vom System getroffene Entscheidungen infrage zu stellen und zu verändern. Bewertet werden separat der individuelle Schaden sowie der gesamtgesellschaftliche Schaden. Dabei geht es um allgemeine Forderungen für KI-Systeme, nicht nur solche für medizinische Anwendungen.

Gesellschaftliche Nebenwirkungen

Alle algorithmischen Systeme ohne Schadenspotenzial kommen in Klasse 0. Bereits ab Klasse 1 fordert Zweig Schnittstellen, die Nachvollziehbarkeitsanalysen erlauben. „Zudem sollte die Gesellschaft informiert werden über das Qualitätsmaß, mit dem das System trainiert wurde und welche Methode des maschinellen Lernens verwendet wurde“, heißt es in dem Buch. Ab Klasse 2 seien genaue Kenntnis der Art der Inputdaten zu fordern (Transparenz) sowie eine Qualitätsbewertung vorzunehmen (Nachvollziehbarkeit).

Für Klasse 3 sind die Anforderungen noch einmal schärfer und Klasse-4-Systeme seien rechtlich oder technisch nicht durchsetzbar, sodass „sie nicht existieren sollten.“ Diese differenzierten Forderungen sollen unter anderem auch dazu beitragen, „nicht jede Idee gleich im Keim zu ersticken“.

Ähnlich wie Dabrock betont Zweig die Bedeutung der Interaktion zwischen digitalem System, Individuen, Institutionen und Gesellschaft. Dies ist mit „sozio-technischem System“ gemeint. Für KI-Entwickler ist das tatsächlich

eine neue Perspektive: Ging es bislang bei der Software-Entwicklung hauptsächlich darum, bestimmte Nutzer-Anforderungen zu erfüllen, sollen sich KI-Systeme heute in einen rechtlichen, ethischen und sozialen Kontext einbetten lassen. Dafür gibt es sogar bereits ein neues Berufsbild, nämlich das der Sozioinformatikerin – gefragt immer dann, wenn es um gesellschaftliche Nebenwirkungen von Algorithmen geht. Für eine Patientin kann der Kopf ihres Arztes zunächst genauso eine Blackbox darstellen, wie eine Maschine.

**Professor Robert Ranisch**, Professor für Medizinische Ethik mit Schwerpunkt auf Digitalisierung Universität Potsdam

Egal wie der Algorithmus entschieden hat, ein Mensch muss die Entscheidung rechtfertigen können.

**Professor Peter Dabrock**, Professor für Systematische Theologie mit Schwerpunkt Ethik an der Friedrich-Alexander-Universität Erlangen-Nürnberg.

**Wörter:**

812

**Urheberinformation:**

© 2021 Springer Medizin Verlag GmbH