



Von Spuren, Treffern und falschen Fährten

Die Digitalisierung des Gesundheitssektors stellt alle bisherigen Prozesse auf den Prüfstand. Kaum ein medizintechnisches Gerät kommt noch ohne digitale Technik daher, am besten noch mit dem Internet der Dinge (IoT) vernetzt oder durch Algorithmen mit Künstlicher Intelligenz (KI) unterstützt. Nach den Putz- und Mährobotern dringen nun auch Pflegeroboter, maschinelle Schmuserobben und KI-Operateure in unser gesundheitliches Lebensumfeld ein. Ungeheure Mengen an Informationen werden digital erhoben, an oft unbekanntenen Orten gespeichert und anschließend durch raffinierte Big-Data-Analysen nach Mustern durchsucht und in wöchentlich 2000 Studien und neuen Geschäftsmodellen ausgewertet. Digitale Zwillinge simulieren in der Medizin Organe und ganze Organsysteme.

Die Künstliche Intelligenz in ihren vielen Ausprägungen erobert also fast explosionsartig das Gesundheitswesen – von Prävention, Früherkennung, Diagnostik, Therapieplanung, Entscheidungsunterstützung bis hin zu Pflege und Rehabilitation. Sie gilt als regelrechter Heilsbringer, hat sie doch während der Corona-Pandemie in Rekordzeit für die Entwicklung von Impfstoffen und Medikamenten gesorgt. Das Machine Learning, insbesondere die neuronalen Netze des Deep Learning sorgen permanent für überraschende Erfolgsmeldungen. So konnte *Prof. Schuller* von der Universität Augsburg bereits im Frühjahr 2020 kurz nach dem Ausbruch in Wuhan eine Corona-Erkrankung durch typische Veränderungen in der normalen Stimme erfolgreich erkennen. Einem dänischen Team gelang es, bei eingelieferten Patienten mit hoher Genauigkeit die Notwendigkeit einer Beatmung oder sogar das Sterberisiko vorherzusagen. Das Bayerische Gesundheitsministerium hat zur Pandemie-Bekämpfung zehn Kliniken mit KI-basierten Computertomographen ausgestattet. Ihre Algorithmen mindern das Bildrauschen, verbessern dadurch die Bildqualität und reduzieren die Röntgendosis um bis zu 40 Prozent. Das KI-System soll anhand bestimmter optischer Details die Corona-Erkrankung von einer Lungenentzündung als Folge eines grippalen Infekts unterscheiden.

So weit, so gut. Wir könnten also den neuen digitalen Assistenten in Diagnose und Therapie dankbar sein, uns beruhigt zurücklehnen und fasziniert den exponentiellen Fortschritt beobachten – wenn nicht ab und zu ein paar verstörende Meldungen von Pannen bei der Digitalisierung des Gesundheitswesens zu uns durchdringen würden. Von den erpresserischen Hacker-Angriffen auf Unternehmen und Kliniken, dem lukrativen Verkauf vertraulicher Patientendaten im dunklen Internet und den alarmierenden Horrormeldungen des Bundesamtes für Sicherheit in der Informationstechnik (BSI) über die Cyberkriminalität ganz zu schweigen ...

Es ist kein Leichtes mit den Daten. Gerade in den letzten zwei Jahren der behördlichen Pandemie-Bekämpfung mussten die verantwortlichen Akteure bei ihren Prognosen aufgrund mangelhafter Datenbasis etliche Blamagen einstecken. Nicht nur die Datenerfasser bekamen das Chaos der Infektions- und Impfmeldungen unzureichend in den Griff, auch die Modellierer von Szenarien und künftigen Verläufen wären besser beraten, mit den ver-

meintlich hochpräzisen Zahlen ebenso die enormen Unsicherheitsfaktoren anzugeben, inklusive der kaum abschätzbaren Dunkelziffer. Diese Kenntnis hätte viele politische Entscheider vor peinlichen Aussagen, Verfügungen und Maßnahmen bewahren können. Denn der Mensch als komplex-dynamisches System reagiert oft unvorhersehbar auf Reize und entwickelt überraschende Verhaltensweisen – siehe die millionenfache Impfverweigerung. Eine uralte Erkenntnis der Statistiker lautet immerhin: „Prognosen sind das Schwierigste, vor allem, wenn sie die Zukunft betreffen.“

Der Biometriker *Prof. Gerd Antes* kritisierte eine „religiös anmutende Verehrung von Daten“, bei der wissenschaftliche Qualitätsstandards über Bord geworfen werden. Er sieht ein Problem in den auftretenden Scheinkorrelationen („spurious correlations“) bei den Big-Data-Variablen [Interview-Gerd-Antes-WirtschaftWeiterbildung-9-19.pdf (mba-journal.de)]. Nur der menschliche Verstand ist in der Lage, tiefergehende Zusammenhänge von scheinbaren Kausalitäten zu unterscheiden. In der Tat: Es braucht die Analyse von zigtausenden von Fotos, um einer KI einigermaßen den Unterschied zwischen Hund und Katze beizubringen. Ein kleines Kind aber lernt das Muster anhand weniger Bilder in kurzer Zeit. Entscheidend ist zudem die Auswahl des Trainingsmaterials, um Trugschlüsse zu vermeiden: Bei einem Referenzdatensatz mit Kindern ohne Covid-Erkrankung lernte die KI, Kinder zu erkennen, aber nicht Corona-Patienten. In einem anderen Fall enthielt der Datensatz einen Mix von Scans mit liegenden und stehenden Patienten. Aus der Lageposition schloss die KI auf ein höheres Corona-Risiko beim Liegen. – Gerade in der Medizin ist es entscheidend, zu verstehen, wie die KI zu ihrem Ergebnis gekommen ist.

So führten Ende 2021 zwei Studien zu einem vernichtenden Ergebnis: Von den untersuchten 232 bzw. 415 KI-Prognosetools war keines für einen klinischen Einsatz geeignet. Vielmehr konstatierte die Bewertung falsche Annahmen und ungeeignete Daten in den Trainings- und Testphasen der Systeme, weil KI-Forschende ohne medizinisches Fachwissen oder Mediziner ohne KI-Erfahrung die Software entwickelt hatten.

Die Akzeptanz der KI-Anwendungen durch Ärzte und Patienten steigt trotzdem langsam, aber stetig. Gemäß einer Umfrage des Bundesverbandes Digitale Wirtschaft (BVDW) e.V. wünschen bereits 57 Prozent der befragten Deutschen vom Arzt die Hinzuziehung einer KI als Zweitmeinung in der Diagnostik. Von den großen Radiologie-Herstellern wird der KI-Einsatz bereits standardmäßig angeboten.

Letztendlich müssen wir uns fragen: Können KI-Programme ohne validierte Datenbasis und Vermittlung des Entscheidungsweges überhaupt vertrauenswürdige Assistenten für Mediziner werden?

Manfred Kindler
Vereidigter Sachverständiger für Medizintechnik
E-Mail: kindler@mt-medizintechnik.de