

Artikel I. Hot topic: KI

Abschnitt 1.01 Künstliche Intelligenz, Algorithmen, Big Data – Aus Daten entstehen individuelle Therapiekonzepte

SEBASTIAN FRITSCH, MORRIS RIEDEL

Künstliche Intelligenz (KI) hat inzwischen praktisch alle Aspekte unseres täglichen Lebens verändert. Durch die Fähigkeit, riesige Datenmengen zu analysieren, darin Muster zu erkennen und so Entscheidungen zu treffen und Aufgaben zu erfüllen, die normalerweise den Einsatz menschlicher Intelligenz erfordern, erreicht KI heute ein Leistungsniveau, das noch vor wenigen Jahren völlig undenkbar schien. Dieses Kapitel soll einige Grundlagen aus dem Gebiet der KI erläutern.

(a) Allgegenwärtig, doch oft unpräzise: Definitionsversuche für typische Begriffe

(i) Künstliche Intelligenz

Der Begriff "*Künstliche Intelligenz*" wird zwar inflationär verwendet, eine allgemein anerkannte Definition gibt es für ihn jedoch bisher noch nicht. Verschiedene Definitionsversuche finden sich in der Box.

Verschiedene Definitionen von „Künstlicher Intelligenz“ (1)

- „Making a machine behave in ways that would be called intelligent if a human were so behaving“ (Dartmouth Research Project, 1955)
- „The science of making machines do things that would require intelligence if done by men“ (M. Minsky, 1968)
- „A system’s ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation“ (Kaplan und Haenlein, 2020)

(ii) Maschinelles Lernen

Die KI-Definition von Kaplan und Haenlein orientiert sich erkennbar am wohl wichtigsten Teilbereich der KI, nämlich dem "*Maschinellen Lernen*" (engl. Machine Learning, ML). ML „erlernt“ aus einem großen Datensatz Muster, um damit auch über neue, bisher unbekannte Daten Aussagen machen zu können. Das Erlernen dieser Muster wird als "Training" eines ML-Modells bezeichnet. Abhängig von Datensatz und Zweck können ML-Verfahren noch weiter unterteilt werden. Besitzen die Datenpunkte eines Datensatzes ein bestimmtes Label, auf dessen Vorhersage das Modell trainiert wird, spricht man von "*Supervised learning*". Fehlt ein solches Label hingegen, wird die Technik als "*Unsupervised Learning*" bezeichnet. In diesem Fall wird z. B. nach Clustern gesucht, deren Verteilung sich von einer Zufallsverteilung unterscheidet. Als drittes ist noch das "*Reinforcement learning*" zu nennen, das stark an das menschliche Lernen angelehnt ist. Bei diesem Verfahren interagiert ein Agent mit seiner Umwelt und erhält für diese Interaktion abhängig von dem erwünschten Soll- und dem erreichten Ist-Zustand eine Belohnung. Der Agent optimiert selbständig seine Strategie, damit seine kumulative Belohnung möglichst groß wird. Neben diesen drei Grundprinzipien der ML gibt es aber inzwischen auch Weiterentwicklungen und Modifikationen, die sich nicht in eine dieser drei Kategorien einordnen lassen, sondern teilweise zwischen ihnen liegen (2).

(iii) Algorithmen und Modelle

Die Nutzung der Begriffe „*Algorithmus*“ und „*Modell*“ erfolgt häufig unscharf. Vereinfacht ausgedrückt, stellt

ein Algorithmus das Verfahren dar, mit dem eine bestimmte Datenverteilung erlernt und ein Ergebnis erzeugt wird. Die Fragestellung, ob ein Patient überlebt oder nicht, kann z. B. gleichwertig durch eine multiple logistische Regression, mithilfe von Entscheidungsbäumen (Random Forest) oder durch die Berechnung einer mehrdimensionalen Ebene, die Überlebende von Nicht-Überlebenden trennt (Support Vector Machine), gelöst werden. Um aber überhaupt konkret in einem Fall anwendbar zu sein, muss ein Algorithmus auf einem spezifischen Datensatz trainiert werden. Der Algorithmus wird dann in Verbindung mit den genutzten Daten zu einem Modell.

(iv) Künstliche Neuronale Netze und Deep learning

Zu den Techniken, die wesentlich zu den Fortschritten der KI in den letzten Jahren beigetragen haben, gehören die „Künstlichen Neuronalen Netze (KNN)“ und das darauf aufbauende „Deep Learning“ (DL). Dieses von der Biologie inspirierte Konzept arbeitet mit „Neuronen“, die über „Dendriten und Axonen“ miteinander verschaltet sind. Ein Neuron empfängt unterschiedlich gewichtete Inputs von mehreren anderen Neuronen und berechnet aus diesen Inputs, ob ein Output (ein „Feuern“ des Neurons) ausgelöst wird. Das Verhalten des KNN wird dabei durch seine Struktur maßgeblich beeinflusst. Diese legt fest, wie viele Neuronen in wie vielen hintereinander liegenden Schichten (layers) wie miteinander verbunden sind. Von außen sichtbar ist in der Regel nur die Ausgabe der letzten Schicht eines KNN, die daher auch als output layer bezeichnet wird. Die Ausgaben vorgelagerter Schichten sind unbekannt und werden daher als hidden layers bezeichnet. Pauschal lässt sich festhalten, dass die Abstraktionsfähigkeit eines Modells umso größer wird, je mehr verdeckte Schichten es enthält.

Modelle mit einer großen Anzahl an hidden layers und zahlreichen Neuronen in den einzelnen layers werden dem Bereich des Deep learning (DL) zugeordnet. Während klassische ML-Techniken strukturierte Daten, also numerische oder diskrete Daten (z. B. Herzfrequenz etc.) benötigen, zeichnen sich DL-Algorithmen vor allem durch die Fähigkeit aus, unstrukturierte Daten, wie z. B. Bilddaten, in numerische Werte umzuwandeln und weiterzuverarbeiten (3). Grundsätzlich können zwar auch unstrukturierte Daten, wie Bild- oder Videodaten durch das sogenannte „Feature engineering“ in ein Format gebracht werden, dass sie von ML-Techniken genutzt werden könnten. Dies erfordert jedoch umfangreichen menschlichen Input (4). DL-Modelle hingegen sind in der Lage, diese Features selbst zu finden. Neben Bilddaten erscheint derzeit z. B. auch die Nutzung von Waveform-Daten, die den zeitlichen Verlauf einer „Kurve“, wie z. B. eines EKG-Signals, darstellen, vielversprechend (5). Auch unstrukturierte Texte, wie Verlaufsdokumentation werden im Rahmen des Natural language processing zunehmend verwendet (6).

(v) Big data

Für das Training von ML-Modellen, vor allem bei DL-Algorithmen, sind extrem große Datenmengen erforderlich. Ist der Datensatz hingegen zu klein, besteht das Risiko, dass der Algorithmus den Datensatz praktisch „auswendig lernt“, ein Phänomen, das auch als „Overfitting“ bezeichnet wird. Die Folge ist, dass das Modell hervorragende Ergebnisse auf den bekannten Daten erzielt – die es ja auswendig kennt – aber eine wesentlich schlechtere Leistung bei unbekanntem Daten erzielt. Sogenanntes „Big data“ wird daher große Bedeutung zugeschrieben. Diese Bezeichnung nutzt man, wenn Daten in großem Umfang (*volume*) und in großer Bandbreite und Vielfalt (*variety*) erzeugt werden und in so hoher Geschwindigkeit anfallen (*velocity*), dass sie für eine herkömmliche Datenverarbeitung zu groß sind. Was im medizinischen Kontext konkret unter Big data zu verstehen ist, hängt auch von Breite und Tiefe der erfassten Daten ab. Administrative Datensätze mit lediglich wenigen Features pro Patient (geringe Tiefe) aber mehreren Millionen enthaltenen Patienten (große Breite) können ebenso Big data darstellen, wie Datensätze mit Sequenzierungs- oder „Omics“-Daten (hohe Tiefe), die nur eine kleine Anzahl Patienten enthalten (geringe Breite) (7).

(b) Der Weg in die Klinik

Die Intensivmedizin gehört wohl zu den datenintensivsten Bereichen der Medizin überhaupt. Die immer breitere Verfügbarkeit von Systemen, die diese Daten elektronisch erheben und speichern, und die weltweit wachsenden Rechenkapazitäten, mit denen diese Daten verarbeitet und analysiert werden können, bilden ideale Voraussetzungen für die Entwicklung und Implementierung KI-basierter Modelle. Zwei typische Bereiche dieser Modelle sind klinische Entscheidungsunterstützung und prädiktives Modelling für Diagnostik und Therapie, sowie Phänotypisierung und Clustering.

(i) Prädiktive Modelle und klinische Entscheidungsunterstützungssysteme

Ein prädiktives Modell wurde entwickelt, um einen bestimmten Endpunkt, also ein Ereignis, einen Zustand oder ein Outcome, vorherzusagen. Während auch klassische Scores (SOFA, APACHE etc.) ein Outcome voraussagen, tun ML-Modelle dies jedoch auf der Basis wesentlich größerer und höher aufgelöster Datensätze und mit deutlich komplexeren Algorithmen als die klassischen Scores.

Werden prädiktive Modelle in eine Form gebracht, dass sie am Patientenbett genutzt werden können, spricht man von klinischen Entscheidungsunterstützungssystemen (engl. Clinical Decision Support Systems, CDSS). Ein CDSS wird definiert als "jedes elektronische System zur direkten Unterstützung der klinischen Entscheidungsfindung, bei dem Merkmale einzelner Patienten verwendet werden, um patientenspezifische Bewertungen oder Empfehlungen zu erstellen, die dann den Klinikern zur Prüfung vorgelegt werden" (8).

Die Bandbreite möglicher Endpunkte und damit die Zweckbestimmung entsprechender KI-basierter CDSS sind dabei äußerst vielschichtig. Der wohl am häufigsten untersuchte Endpunkt ist sicher die Mortalität (9). Auch ein verlängerter Intensivaufenthalt (10), eine verlängerte Beatmung (11) oder eine ungeplante Wiederaufnahme auf die Intensivstation (12), genauso wie die Vorhersage bestimmter komplikativer Krankheitsbilder, wie einer Sepsis (13), eines Delirs (14) oder eines Acute respiratory distress syndrome (ARDS) (15), wurden wiederholt adressiert. Ein weiterer wichtiger Ansatz ist die Vorhersage, ob ein Patient von einer spezifischen Therapie profitiert (16). Es kommen aber auch eher spezielle Fragen vor, wie die Vorhersage eines multiresistenten Erregers bei einem Patienten (17), oder, ob eine Routine-Laboruntersuchung neue Informationen ergibt oder schlicht unterlassen werden kann (18).

(ii) Phänotypisierung und Clustering

Intensivmedizinische Daten zeichnen sich häufig durch eine hohe Komplexität aus, die für Menschen kaum sinnvoll zu erfassen ist. Für Algorithmen des „Unsupervised learning“ stellt dies jedoch keine wirkliche Herausforderung dar. So konnten bereits mehrere Forscher durch entsprechende Algorithmen Phänotypen bestimmter Erkrankungen identifizieren, die sich klinisch oder auf molekularer Ebene deutlich voneinander unterscheiden. So konnten beispielsweise Calfee et al. einen hyper- and hypoinflammatorischen Phänotyp des ARDS identifizieren (19). Auch für ein Sepsis-assoziiertes Multiorganversagen konnte ein Clustering-Algorithmus vier unterschiedlichen Phänotypen ermitteln, die für die Prognose eines Patienten relevant waren (20).

(c) Zusammenfassung

Patienten stehen dem Einsatz von KI in der Medizin grundsätzlich positiv gegenüber (21). Nichtsdestotrotz bringt ihre Anwendung eine Vielzahl technischer, medizinischer, rechtlicher und auch ethischer Fragen mit sich, die bisher nur unzureichend beantwortet wurden. Hierzu zählen unter anderem eine unzureichende oder verzerrte Datenbasis für die Modellentwicklung, eine fehlende Nachvollziehbarkeit der Ergebnisse eines Modells („Black box“-Problem) und soziotechnische Probleme bei der Implementierung (22-24). Die zahlreichen Beispiele, die das große Potenzial KI-basierter Methoden im medizinischen Kontext belegen, sollten Ärzten wie Entwickler motivieren, die aktuell noch bestehenden Probleme der KI-Nutzung anzugehen. Denn sauber entwickelte und klinisch sinnvoll implementierte KI-Tools, die die Ärzte in ihrer täglichen Arbeit unterstützen, haben das Potenzial, die Behandlung von Intensivpatienten deutlich zu verbessern und so dem großen Vertrauensvorschuss der Patienten in diese Technik gerecht zu werden.

(d) Literaturverzeichnis

1. Kaplan A, Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business horizons*. 2019;62(1):15-25.
2. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to machine learning, neural networks, and deep learning. *Translational vision science & technology*. 2020;9(2):14-.
3. Kriegeskorte N, Golan T. Neural network models and deep learning. *Current Biology*. 2019;29(7):R231-R6.

4. Verdonck T, Baesens B, Óskarsdóttir M, vanden Broucke S. Special issue on feature engineering editorial. *Machine Learning*. 2024;113(7):3917-28.
5. Mollura M, Lehman LH, Mark RG, Barbieri R. A novel artificial intelligence based intensive care unit monitoring system: using physiological waveforms to identify sepsis. *Philos Trans A Math Phys Eng Sci*. 2021;379(2212):20200252.
6. Ariño H, Bae SK, Chaturvedi J, Wang T, Roberts A. Identifying encephalopathy in patients admitted to an intensive care unit: Going beyond structured information using natural language processing. *Front Digit Health*. 2023;5:1085602.
7. Sanchez-Pinto LN, Luo Y, Churpek MM. Big Data and Data Science in Critical Care. *Chest*. 2018;154(5):1239-48.
8. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *Bmj*. 2005;330(7494):765.
9. Barboi C, Tzavelis A, Muhammad LN. Comparison of Severity of Illness Scores and Artificial Intelligence Models That Are Predictive of Intensive Care Unit Mortality: Meta-analysis and Review of the Literature. *JMIR Med Inform*. 2022;10(5):e35293.
10. Wu J, Lin Y, Li P, Hu Y, Zhang L, Kong G. Predicting Prolonged Length of ICU Stay through Machine Learning. *Diagnostics (Basel)*. 2021;11(12).
11. Parreco J, Hidalgo A, Parks JJ, Kozol R, Rattan R. Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement. *J Surg Res*. 2018;228:179-87.
12. Desautels T, Das R, Calvert J, Trivedi M, Summers C, Wales DJ, et al. Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: a cross-sectional machine learning approach. *BMJ Open*. 2017;7(9):e017199.
13. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med*. 2020;46(3):383-400.
14. Gong KD, Lu R, Bergamaschi TS, Sanyal A, Guo J, Kim HB, et al. Predicting Intensive Care Delirium with Machine Learning: Model Development and External Validation. *Anesthesiology*. 2023;138(3):299-311.
15. Rashid M, Ramakrishnan M, Chandran VP, Nandish S, Nair S, Shanbhag V, et al. Artificial intelligence in acute respiratory distress syndrome: A systematic review. *Artif Intell Med*. 2022;131:102361.
16. Hellali R, Chelly Dagdia Z, Ktaish A, Zeitouni K, Annane D. Corticosteroid sensitivity detection in sepsis patients using a personalized data mining approach: A clinical investigation. *Comput Methods Programs Biomed*. 2024;245:108017.
17. Liang Q, Zhao Q, Xu X, Zhou Y, Huang M. Early prediction of carbapenem-resistant Gram-negative bacterial carriage in intensive care units using machine learning. *J Glob Antimicrob Resist*. 2022;29:225-31.
18. Li LT, Huang T, Bernstam EV, Jiang X. External Validation of a Laboratory Prediction Algorithm for the Reduction of Unnecessary Labs in the Critical Care Setting. *Am J Med*. 2022;135(6):769-74.
19. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *The Lancet Respiratory Medicine*. 2014;2(8):611-20.
20. Knox DB, Lanspa MJ, Kuttler KG, Brewer SC, Brown SM. Phenotypic clusters within sepsis-associated multiple organ dysfunction syndrome. *Intensive Care Medicine*. 2015;41(5):814-22.
21. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *DIGITAL HEALTH*. 2022;8:20552076221116772.
22. Pinsky MR, Bedoya A, Bihorac A, Celi L, Churpek M, Economou-Zavlanos NJ, et al. Use of artificial intelligence in critical care: opportunities and obstacles. *Critical Care*. 2024;28(1):113.
23. Bleher H, Braun M. Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems. *AI Ethics*. 2022;2(4):747-61.
24. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial intelligence in healthcare: Elsevier; 2020. p. 295-336.*