medizin://dokumentation/informatik/
informationsmanagement/

mdi

ISSN 1438-0900

Heft 3 _ September 2022 _ Jahrgang 24

KÜNSTLICHE INTELLIGENZ IN DER MEDIZIN

BVMI DVMD

Liebe Leserinnen und Leser,

Prof. Dr.-Ing. Oliver J. Bott
Hochschule Hannover
oliver.bott@hs-hannover.de

Prof. Dr. Cord Spreckelsen
Institut für Med Statistik, Informatik und Datenwissenschaften, Universitätsklinikum Jena
Cord.Spreckelsen@med.uni-jena.de

Keine Frage: Künstliche Intelligenz (KI) zieht in den letzten Jahren öffentliche Aufmerksamkeit fast magnetisch an. Sie gilt als strategisch wichtiges Zukunftsfeld und aktiviert damit hohe Fördermittel. An KI in der Medizin knüpfen sich besonders hohe, vielleicht überhöhte Erwartungen. Anders als in früheren Phasen der KI-Euphorie sind aber aktuelle KI-Ansätze im Gesundheitsbereich tatsächlich praxis-, routine- und konkurrenzfähig: Die amerikanische Zulassungsbehörde FDA lässt seit mehreren Jahren im Monatsrhythmus neue KI-gestützte Medizinprodukte zu.

KI in der Medizin steht vor besonderen Herausforderungen: Die Anpassung von KI-Verfahren an die spezifischen Erfordernisse und Anwendungsbedingungen der Medizin ist methodisch herausfordernd. Beispielsweise macht die Forderung nach nachvollziehbarer Datenverarbeitung innovative Ansätze zur erklärbaren KI (Explainable AI - XAI) nötig. Am Beispiel einer KI-Anwendung in der Notfallaufnahme zeigt der Heftbeitrag von Ritter et al. einen solchen Ansatz.

Gerade maschinelle Lernverfahren haben spektakuläre Erfolge der KI erzielt. Strenge regulatorische Vorgaben erschweren ihren Einsatz in der Medizin. Der Beitrag von Johner et al. gibt hier einen präzisen, orientierenden Überblick und macht auf akute Problemfelder einer überbordenden Regulierung insbesondere auf Ebene der EU aufmerksam.

Der Einsatz von KI gerade im klinischen Umfeld führt zu besonderen praktischen Herausforderungen: Hier stehen Fragen der Daten- und Funktionsintegration von KI-Anwendung, ihre Gebrauchstauglichkeit oder ihre Verfügbarkeit und Zuverlässigkeit im Vordergrund. Ein Beispiel dafür ist die oft unzureichende Datenqualität von Versorgungsdaten bei ihrer Nutzung im KI-Kontext. Mit einer Anwendung aus der pädiatrischen Intensivmedizin stellen Wolff et al. ein Entscheidungsunterstützungssystem mit Augenmerk auf dessen Interoperabilität vor.

Die Datenflut aus (mobilen) Sensoren überfordert zunehmend die menschliche Aufmerksamkeit und Aufnahmefähigkeit. KI kann hier bei der Datenzusammenfassung und -bewertung helfen. Besonders chancenreich ist das im Bereich mobiler Gesundheitsanwendungen (mHealth), der Artikel von Kapsecker et al. zeigt ein Beispiel.

Die genannten Herausforderungen zeigen klar: Es ist notwendig, KI-Komponenten systematisch zu entwickeln. Erfolgreiche und sichere KI in der Medizin setzt voraus, dass entwickler-, nutzer- und betreiberseitig KI-spezifische Kenntnisse und Fertigkeiten vorhanden sind. Gerade auch Angehörige der Gesundheitsberufe benötigen Orientierung und gezielte methodisch-technische KI-Einführungen, damit der KI-Einsatz tatsächlich ein Nutzenpotenzial entfaltet und nicht im Gegenteil womöglich schadet. Der Artikel von Mosch & Ritter gibt einen beispielhaften Einblick in Angebote zur gezielten Entwicklung von KI-Kompetenzen.

In der Vergangenheit fehlten weitgehend fachliche, technische und organisatorische Voraussetzungen, um die Herausforderungen des KI-Einsatzes in der Medizin bewältigen zu können. Die Medizininformatik-Initiative (MII) stellt solche Voraussetzungen auf ganz verschiedenen Ebenen her. Der erste Artikel des Themenhefts beschreibt daher den wichtigen Beitrag, den die MII zur Etablierung von KI in der Medizin schon jetzt leistet und noch leisten wird.

Immer klarer zeichnet sich ab: Für den Erfolg von KI in der Medizin ist entscheidend, dass die Zusammenarbeit und Verständigung zwischen sehr unterschiedlichen Fachdisziplinen und Berufsgruppen funktioniert. Der gezielte Aufbau multiprofessioneller Teams ist eine Aufgabe, deren Bedeutung sowohl im Bereich der Entwicklung medizinischer KI-Anwendungen als auch für ihren sicheren und erfolgreichen Einsatz gar nicht zu überschätzen ist.

Medizinische Dokumentation, Medizininformatik und klinisches Informationsmanagement sammeln seit Jahren Erfahrung im Schnittbereich zwischen Gesundheitsberufen und Informationstechnik. Diese Erfahrung kann und sollte gerade für eine erfolgreiche Anwendung von KI im Gesundheitswesen fruchtbar gemacht werden. Wir hoffen, Ihnen, liebe Leserinnen und Lesern, mit diesem Heft Impulse für einen sinnvollen Einsatz von KI in der Medizin mitgeben zu können.

Cord Spreckelsen und Oliver J. Bott

Die nächsten Themenhefte

mdi 4_2022
Tumordokumentation und klinische Register
Hartz, Stein

mdi 1_2023
Forschung und deren Folgenabschätzung
Goldschmidt, Händel

mdi 2_2023
Datenmanagement in Gesundheitsversorgung und medizinischer Forschung
Ose, Katzensteiner, Händel

mdi 3_2023
KHZG, Corona und MII – Impulsgeber für die Digitalisierung im Gesundheitswesen?
Bott, Schmücker

Vorschau

Sie haben zu den genannten Themenheften eine Artikel-Idee? Bitte melden Sie sich bei Markus Stein: mstein@rzv.de

Künstliche Intelligenz in der Medizininformatik-Initiative

Prof. Dr. Cord Spreckelsen[1]
Cord.Spreckelsen@med.uni-jena.de

Prof. Dr. Martin Boeker[2]
martin.boeker@tum.de

Prof. Dr. Andreas Schuppert[3]
aschuppert@ukaachen.de

- Künstliche Intelligenz (KI) in der Medizin stellt hohe Anforderungen an den Umfang, die Qualität sowie die einheitliche Strukturierung, Verständlichkeit und Verfügbarkeit medizinischer Daten.
- Medizinische KI ist vielfach nur unter strengen regulativen Bedingungen zu entwickeln und zu betreiben.
- Die Medizininformatik-Initiative verbessert auf mehreren Ebenen die Bedingungen für erfolgreiche KI-Projekte in der Medizin.
- Besondere Bedeutung haben dabei standortübergreifende Ansätze, die Infrastruktur der Datenintegrationszentren und die Kompetenzentwicklung.

Einführung

Künstliche Intelligenz (KI) benötigt Daten. Das gilt für klassische, logikbasierte KI, und es gilt besonders für KI-Ansätze auf Basis maschineller Lernverfahren, welche Daten benötigen zur Modellierung bzw. Implementierung, bei der Anwendung und zur Evaluation. Gerade KI-Anwendungen in der Medizin setzen viel mehr voraus als nur Daten: Erfolgskritisch sind u.a. die Integration von KI-Anwendungen mittels interoperabler Schnittstellen, Services zur Qualitätssicherung von Daten sowie die Basis einer rechtskonformen Entwicklung und Nutzung. Die Medizininformatik-Initiative liefert hier einzigartige infrastrukturelle und methodische Unterstützung. Sie fördert außerdem den notwendigen Kompetenzaufbau für KI in der Medizin und zwar bei Anwendern und Entwicklern gleichermaßen.

Medizininformatik-Initiative

Förderziele, Förderphasen und Aufbau

Die Medizininformatik-Initiative (MII) ist eine derzeit auf ein Jahrzehnt angelegte Förderinitiative des Bundesforschungsministeriums (BMBF). Das Förderprogramm umfasst drei Phasen: eine Konzeptphase (2016-2017), die Aufbau- und Vernetzungsphase (2018-2021, pandemiebedingt bis Ende 2022 verlängert) sowie die anschließende Ausbau- und Erweiterungsphase (2023-2026).

Die Förderinitiative zielt darauf, Digitalisierungschancen in der Medizin zu ergreifen: IT-Lösungen sollen »den Austausch und die Nutzung von Daten aus Krankenversorgung, klinischer und biomedizinischer Forschung über die Grenzen von Institutionen und Standorten hinweg« ermöglichen und verbessern [1].

Daneben geht es in der MII auch darum, Fachkenntnisse und Fertigkeiten gezielt zu fördern, die zur Erreichung dieser Ziele notwendig sind. Hierzu dient die Stärkung von Aus-, Fort- und Weiterbildungsmöglichkeiten der Medizininformatik und gezielte Nachwuchsförderung.

Die wettbewerbliche Vergabe der Fördermittel führte in der Aufbau- und Vernetzungsphase zur Bildung von vier Projektkonsortien (DIFUTURE, HiGHmed, MIRACUM, SMITH). Inzwischen haben sich alle Universitätsklinika Deutschlands einem dieser Konsortien angeschlossen.

Um neben konkurrierenden Ansätzen der Konsortien gleichzeitig deren Zusammenarbeit und bundesweit einheitliche Lösungen zu erreichen, richtete die MII ein Nationales Steuerungsgremium (NSG) ein, in das jedes Konsortium Vertreter entsendet. Das NSG wird organisiert und moderiert durch die Koordinationsstelle der MII, gebildet von der Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (TMF), dem Medizinischen Fakultätentag (MFT) und dem Verband der Universitätsklinika Deutschlands (VUD).

Funktion der MII-Datenintegrationszentren

Kern der Förderinitiative ist die Einrichtung von Datenintegrationszentren (DIZ) an den Standorten. Die DIZ sorgen dafür, dass Daten aus der Versorgung über Standorte hinweg für Forschungsprojekte verfügbar gemacht werden. Die Daten werden dabei nicht zentral gesammelt. Stattdessen erlaubt es die MII-Infrastruktur, Daten projektbezogen abzurufen und zusammenzuführen. Eine zentrale Rolle spielt dabei, dass die Daten einheitlich strukturiert werden und dass genau dokumentiert ist, was die Daten bedeuten (syntaktische und semantische Interoperabilität).

Eine wichtige Errungenschaft der MII ist daher die Vereinbarung des einheitlich strukturierten Kerndatensatzes der MII. Er nutzt den internationalen HL7-Standard Fast Healthcare Interoperability Resources (FHIR) als Repräsentations- und Datenaustauschformat und verwendet internationale medizinische Terminologiestandards bzw. Klassifikationssysteme (u.a. ICD, LOINC und SNOMED-CT).

Eine weitere Errungenschaft ist der Broad Consent der MII. Dieser formuliert eine einheitliche Einwilligungserklärung zur Datennutzung, welche sich nicht auf ein einzelnes, vorab inhaltlich detailliert festgelegtes Projekt beschränkt, sondern – auf Antrag bei den Use and Access Committees der Standorte – eine Nutzung zu medizinischen Forschungszwecken mit absehbar breitem Nutzen für die Allgemeinheit ermöglicht. Der Broad Consent wurde in einem aufwändigen Ver-

fahren mit allen Datenschutzbeauftragten auf Bundes- und Landesebene abgestimmt.

Die MII erarbeitet und nutzt medizinische Interoperabilitätsstandards, damit Daten über Standorte und Systeme hinweg ausgetauscht und verstanden werden. Sie entwickelt und erprobt außerdem Frameworks und Plattformen für die verteilte Durchführung von Datenanalysen – idealerweise so, dass Daten die einzelnen Standorte gar nicht verlassen müssen, sondern nur Analyseergebnisse mitgeteilt werden.

Unterstützung medizinischer KI durch die MII

Herausforderungen medizinischer KI

Entwicklung und Einsatz von KI in der Medizin gehen mit spezifischen Herausforderungen einher, bei deren Bewältigung die MII hilft. Datenverfügbarkeit ist hier der naheliegendste Aspekt. Dabei ist zu bedenken, dass Datenverfügbarkeit für alle Formen von KI in der Medizin wichtig ist, auch für klassische, logikbasierte KI-Ansätze. Ein kritischer Erfolgsfaktor KI-basierter klinischer Entscheidungsunterstützungssysteme (Clinical Decision Support System – CDSS) ist deren Datenintegration [2]. Patientendaten müssen sich aus Versorgungssystemen direkt übernehmen lassen, damit das CDSS patientenspezifische Unterstützung leisten kann. Hier kann die MII durch verbesserte Schnittstellen auch für Echtzeitdaten eine wesentliche Infrastrukturverbesserung erreichen.

Daten sind auch für die Evaluation von CDSS im klinischen Kontext unverzichtbar. Das gilt besonders dann, wenn der Effekt des CDSS bezogen auf die Ergebnisse klinischer Versorgung (allgemein: klinische Outcomes) gemessen werden soll. In einigen der MII-Anwendungsfälle (MII Use Cases – siehe unten) steht genau dieser Aspekt im Vordergrund. Hier liefert die MII Daten für Evaluationsstudien zum Effekt des CDSS-Einsatzes.

Bei der Bewertung von KI in der Medizin zeigen sich als zusätzliche Herausforderung uneinheitliche Bewertungsergebnisse (Mixed Evidence). Das betrifft die Schwierigkeit, Studienergebnisse zu reproduzieren (z.B. zu positiven Auswirkungen des KI-Einsatzes), und es betrifft gegenläufige Effekte (z.B. die Verbesserung der diagnostischen Entscheidung bei gleichzeitigem Verlust eigenständiger ärztlicher Kompetenz oder bei Störung der Zusammenarbeit im klinischen Team). Es wird deutlich, dass der klinische Mehrwert selbst von Ansätzen, die auf den ersten Blick erfolgreich sind, schwer zu belegen ist, ggf. sogar fehlt. Das macht eine kritische und breit angelegte Erkundung sinnvoller Anwendungsszenarien nötig.

Für KI, die auf maschinellen Lernverfahren basiert, stellt sich das Problem der Datenverfügbarkeit noch drängender: Datengetriebene KI benötigt in aller Regel sehr umfangreiche Datenbestände zum Training und Test ihrer Modelle. Der benötigte Umfang ist meist nur durch Kombination von Datenbeständen mehrerer Standorte überhaupt herstellbar (unmittelbar einsichtig, z.B. im Falle seltener Erkrankungen, aber weit darüber hinaus [3]). Die Zusammenführung von Datenbeständen ist weit mehr als ein Export-/Import-Problem. Die Daten müssen inhaltlich kombinierbar sein. Das fängt bei einheitlichen Maßeinheiten an und hört bei eindeutig identifizierbaren Merkmalen nicht auf. Hier sind Lösungen der MII zur (semantischen) Interoperabilität wichtig [4].

Eng verknüpft mit der Verbesserung der Datenverfügbarkeit ist die Herausforderung einer datenschutzkonformen Nutzung der Daten bei der Entwicklung und der Anwendung von KI in der Medizin [5]. Dabei geht es auch um die Einschätzung und Verminderung des Risikos, dass oberflächlich anonyme Daten durch geeignete Analyseverfahren wieder auf Personen bezogen werden (Re-Identifizierungsrisiken) [6].

Besonders für Versorgungsdaten stellt sich das Problem der Datenqualität. Lücken- oder fehlerhafte Datensätze stören oder verhindern erfolgreiches maschinelles Lernen [7][8]. Datenqualitätsmanagement ist eine wichtige Voraussetzung dafür, dass datengetriebene Verfahren Versorgungsdaten nutzen können. Strukturelle Datenqualität ist nicht der einzige Aspekt. Immer stärker rückt auch die Validität von Daten in den Fokus, spätestens seit einschlägige Beispiele dafür bekannt wurden, dass KI, die z.B. auf Daten der Allgemeinbevölkerung trainiert wurde, Minderheiten systematisch benachteiligte bzw. dort zu ungültigen Ergebnissen kam [9]. In diesen Zusammenhang gehört auch die Frage der standortübergreifenden Übertragbarkeit von KI-Modellen bzw. der Notwendigkeit ihrer standortspezifischen Anpassung [10]. Diese Aspekte lassen sich nur mittels standortübergreifend verfügbaren, qualitätsgesicherten Daten bewerten oder korrigieren.

Eine der größten Herausforderungen sind die regulatorischen Hürden für KI-basierte Entscheidungsunterstützungssysteme. Diese sind sehr hoch (siehe den Beitrag von Johner et al. in diesem mdi-Themenheft) und werden Innovationen im Bereich datengetriebener medizinischer KI und ihren Einsatz als Medizinprodukt in Europa erschweren [11].

Die angesprochenen Herausforderungen machen gezielten Kompetenzaufbau nötig. Angehörige der Gesundheitsberufe benötigen Einblick in die Tragweite, Indikationen und Kontraindikationen der KI-Anwendungen. Hier gilt es, Fehleinschätzungen und Fehlbedienungen zu vermeiden und eine zielführende Arbeitsteilung zu ermöglichen. KI-Entwickler benötigen tiefe Kenntnisse in die spezifischen fachlichen und organisatorischen Besonderheiten des Gesundheitswesens. Hier sind Schulungsangebote und ganze Ausbildungs-/Studiengänge nötig, damit KI in der Medizin nachhaltig erfolgreich sein kann. Jenseits der Fachcommunity wird es zudem wichtig sein, fundierte Beiträge zu einem realistischen gesellschaftlichen Erwartungsmanagement bezüglich künftiger KI-Entwicklungen im Gesundheitswesen einzubringen.

Prof. Dr. Julian Varghese[4]
julian.varghese@uni-muenster.de

Sebastian C. Semler[5]
Sebastian.Semler@tmf-ev.de

[1] Institut für Med. Statistik, Informatik und Datenwissenschaften, Universitätsklinikum Jena
[2] Institut für Künstliche Intelligenz und Informatik in der Medizin (AIIM), Klinikum rechts der Isar, Technische Universität München
[3] Institut für Computational Biomedicine, JRC für Computational Biomedicine, Universitätsklinikum Aachen
[4] Institut für Medizinische Informatik, Universität Münster
[5] TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., Berlin

Strukturelle Unterstützung durch die MII

MII-Datenintegrationszentren

Die MII-Datenintegrationszentren nutzen Interoperabilitätsstandards, mit denen Daten besser zusammengeführt und (auch maschinell) verstanden werden können. Sie erleichtern daher die Integration von KI-Anwendungen. Die Anwendungsfälle der MII (MII-Use Cases, s.u.) zielen darauf, die Nutzbarkeit und Nützlichkeit des Ansatzes zu zeigen. Einige der Use Cases leisten dies für KI-Anwendungen.

Methodenentwicklung

Standortübergreifende Ansätze sind wichtig. Nur so lässt sich in vielen Fällen überhaupt der benötigte Umfang von Trainings-/Testdatensätzen zusammentragen. Und nur durch standortübergreifende Ansätze kann untersucht werden, ob sich KI-Modelle standortunabhängig einsetzen lassen, bzw. wie hoch das Risiko ist, welches durch Übertragung von Modellen auf andere Standorte besteht. Um standortübergreifende/multizentrische Projekte datenschutzgerecht umsetzen zu können hat die MII Ansätze zur föderierten Datennutzung konzipiert (z.B. den Personal Health Train [12]) oder etabliert (z.B. DataSHIELD), mit denen Datenanalysen oder maschinelle Lernprozesse durchgeführt werden können, ohne dass Daten die einzelnen Zentren überhaupt verlassen müssen.

Kompetenzentwicklung

Neue Ausbildungs- und Schulungsangebote: Die MII stärkt die für KI-Entwicklung und -Anwendung nötigen Kompetenzen: So sind in der MII neue Ausbildungsangebote geschaffen worden, u.a. neue Masterstudiengänge für Medizinische Informatik, sowie zwei berufsbegleitende Studiengänge «Medical Data Science", deren Absolventen wichtige Voraussetzungen zur Arbeit an und mit KI-Anwendungen erwerben. Bestehende Studiengänge wurden verbessert und erweitert, Trainingsangebote für neue Mitarbeiterinnen und Mitarbeiter der DIZ geschaffen und spezielle Module für Personen, die Datennutzungsprojekte auf Basis der MII-Infrastruktur durchführen wollen. Die mdi berichtete in einem früheren Beitrag ausführlich über die MII-Ansätze zur Stärkung der Medizininformatik in der Lehre [15].

Tab. 1 MII-Nachwuchsgruppen mit KI-Bezug

Nachwuchsgruppe	Standort	KI-Bezug
Medical Health Data	Berlin	Methoden der Künstlichen Intelligenz
Prospektiv-nutzergerechte Gestaltung klinischer Entscheidungsunterstützungssysteme im Kontext personalisierter Medizin	Dresden	Titel (Entscheidungsunterstützung)
Prädiktive Analyse und datengetriebene künstliche Intelligenz zur logistischen Unterstützung von Versorgungsprozessen (KI-LoV)	Jena	Titel (KI)
Implementierung von Smart-Contract-Technologien zur Analyse-Föderierung in der Intensivmedizin (SAFICU)	München	Algorithmen für eine klinische Entscheidungsunterstützung
Modulare Wissens- und Datengetriebene Molekulare Tumorkonferenz (MoMoTuBo)	Augsburg	Machine-Learning-Verfahren
Entwicklung eines Terminologie- und Ontologiebasierten Phänotypisierungsframeworks (TOP)	Leipzig	Formale Ontologien
KI-gestützte morphomolekulare Präzisions-Medizin in Neuroonkologie (AI-RON)	Gießen	Titel (KI)
Integration & Analyse von multimodalen Sensorsignalen zur Erforschung von neurologischen Bewegungsstörungen (MoveGroup)	Lübeck	Entscheidungsunterstützung und Erkenntnisgewinn mit KI-Methoden
Datenschutz medizinisch relevanter Daten und datenschutzbewusstes Training von Modellen des maschinellen Lernens auf medizinischen Daten (MDPPML)	Tübingen	Titel (Maschinelles Lernen)
Entwicklungen von klinisch-orientierten Entscheidungsunterstützungen für Hochdurchsatzdaten in der personalisierten Medizin (EkoEstMed)	Freiburg	Algorithmen des maschinellen Lernens
Klinische Textanalytik: Methoden für NLP an deutschen Texten (DE.xt)	München	Ermöglichung und Unterstützung von KI

Neue Professuren mit KI-Bezug

Zur Unterstützung der universitären Lehre wurden im Kontext der MII insgesamt 51 Professuren geschaffen. Von diesen 51 verweisen vier explizit im Titel der Professur auf KI oder Entscheidungsunterstützung (Universität Augsburg: »Embedded Intelligence for Healthcare and Wellbeing« sowie »Datenmanagement und Clinical Decision Support«, Universitätsmedizin Göttingen: »Klinische Entscheidungsunterstützung«, Technische Universität München: »Artificial Intelligence in Healthcare and Medicine«). In 18 weiteren Fällen ergibt sich aus Projekten und Publikationen ein ausdrücklicher KI-Bezug (implizite Nähe zu KI-Ansätzen ist hier nicht berücksichtigt).

An einigen Universitäten wurden – wenn auch nicht im direkten Rahmen der Medizininformatik-Initiative – im Zuge einer strategischen Ausrichtung Institute für medizinische KI geschaffen, so 2019 das Institut für KI in der Medizin (IKIM) als Teil der Medizinischen Fakultät der Universität Duisburg-Essen und der Universitätsmedizin Essen sowie 2020 das Institut für Künstliche Intelligenz als Gemeinschaftsinstitution von Universität Marburg und Uniklinikum Gießen und Marburg.

Nachwuchsgruppen mit KI-Bezug

Im Zuge der MII-Förderung konnte die Einrichtung von Nachwuchsgruppen beantragt werden. Sie verfolgen eigene Forschungsvorhaben mit Bezug auf Ziele und Infrastruktur der MII und unterstützen die neu geschaffenen Professuren. Von den insgesamt 21 Nachwuchsgruppen haben 11 einen ausdrücklichen KI-Bezug (explizit im Namen oder in der Beschreibung genannt). Sie entwickeln oder nutzen KI-Ansätze bzw. erarbeiten gezielt Voraussetzungen für KI-Anwendungen. Tabelle 1 gibt einen Überblick.

Nutzung von Versorgungsdaten in KI-Projekten

Die MII sieht vor, dass die Konsortien die Praxistauglichkeit und den Nutzen ihrer Ansätze sowie der DIZ-Infrastruktur anhand von Anwendungsfällen – den MII-Use Cases – zeigen. Dabei wurden auch Anwendungsfälle entwickelt, die KI-Ansätze betreffen. Diese gehen die eingangs beschriebenen Herausforderungen

an. Je nach Projekt geschieht das mit unterschiedlicher Gewichtung. Die Entwicklung der Use Cases trägt auch dazu bei, Anwendungsbereiche mit klinischem Mehrwert zu identifizieren. In der kommenden Ausbau- und Erweiterungsphase der MII werden konsortiumsübergreifende Use Cases eine zentrale Rolle spielen. Sie prüfen die Tragfähigkeit der gemeinsamen, von der MII insgesamt entwickelten Lösungen.

Der Use Case »From Knowledge to Action – Unterstützung für das Molekulare Tumorboard« entstand im MIRACUM-Konsortium. Hier geht es darum Patientinnen und Patienten zu erkennen, bei denen eine seltene Tumorerkrankung vorliegt, oder für die eine herkömmliche Therapie nicht in Frage kommt oder zu geringe Heilungschancen bietet. In diesen Fällen sollen individuelle Therapieansätze der Präzisionsmedizin vorgeschlagen und unterstützt werden. Der Use Case analysiert dazu Hochdurchsatzdaten (insbesondere Sequenzierungsdaten der Tumor-DNA) in einer automatisierten Kette von Analyse-Algorithmen und erzeugt Datenvisualisierungen, welche die Dateninterpretation erleichtern. Auch wenn hier bioinformatische Analyse-Algorithmen im Vordergrund stehen, zeigt der Use Case neue Ansätze zu einer integrierten klinischen Entscheidungsunterstützung durch Datenanalyse und Informationsversorgung, deren regulatorische Bewertung noch vorgenommen wird.

Das HiGHmed-Konsortium entwickelte den Use Case »Infektionskontrolle«. Er dient dazu, frühzeitig herauszufinden, ob sich in einem Krankenhaus multiresistente Erreger in Infektionsketten verbreiten. Das Projekt zielt darauf ab, Häufungen von Infektion automatisch frühzeitig zu erkennen. Das hierzu entwickelte »Smart Infection Control System«, kurz SmICS, implementiert Detektionsalgorithmen für mögliche Erregerhäufungen und nutzt dazu Daten, die in der Routineversorgung entstehen. Es demonstriert so Möglichkeiten zur Datenintegration eines KI-basierten (Infektions-)Monitorings.

Ein weiterer Use Case »Kardiologie« des HiGHmed-Konsortiums zielt auf eine verbesserte Versorgung von Patientinnen und Patienten mit chronischer Herzinsuffizienz. Dabei geht es um die datengestützte Früherkennung – und idealerweise Vermeidung – von Krankheitsschüben. Neben der standortübergreifenden Standardisierung der nötigen Daten spielt dabei auch die Integration von Daten aus Wearables eine Rolle. Ihre algorithmische Interpretation unterstützt die Risikobewertung und Früherkennung und bietet Anknüpfungspunkte für den KI-Einsatz im Bereich der mobil gestützten Gesundheitsversorgung (mHealth).

Das Konsortium DIFUTURE nutzt in zwei Use Cases (»Multiple Sklerose« und »Parkinson«) Ansätze zu verteilten Analysen auf großen multidimensionalen Datenbeständen zur Vorhersage von Krankheitsverläufen sowie zur Vorbereitung von Therapieentscheidungen. Der Ansatz erlaubt es, Erfahrungen im Bereich der föderierten Datennutzung zu sammeln, die sich auf das Training und die Integration von KI-Anwendungen übertragen lassen.

Im MII-Konsortium SMITH wurde im Rahmen des Use Case ASIC der Einsatz von KI-Methoden für die Analyse von Versorgungsdaten aus Intensivstationen intensiv untersucht. Dieser Ansatz war motiviert durch die Beobachtung, dass KI in der medizinischen Diagnostik weitgehend auf den Bereich der Bildanalyse sowie der Analyse von kontinuierlichen Monitoringdaten mit sehr dichten Abtastraten, z.B. EKG oder EEG, fokussiert ist. In diesen Anwendungen wurden in den letzten Jahren mit Hilfe von Deep Learning (DL) Verfahren überzeugende Ergebnisse von KI-Systemen erreicht, die insgesamt eine zu etablierten Protokollen vergleichbare Qualität aufweisen.

Allerdings fällt die Performance von DL-Technologien bei Anwendungen mit Versorgungsdaten oder allgemein mit heterogenen klinischen Datenstrukturen deutlich im Vergleich zu Bildanalysen ab [13][14]. Die Gründe hierfür sind vielfältig, sie reichen von der oft mangelhaften Datenqualität bis hin zur extremen Heterogenität der Datenstrukturen, die konzeptuelle, mathematisch basierte Hürden für den Einsatz von DL-Verfahren darstellen. Die KI-basierte Nutzung von Versorgungsdaten erfordert daher noch erhebliche Anstrengungen sowohl im Bereich der Datenqualität als auch im Aufbau einer ausreichend performanten KI-Maschinerie, die auf diese Datenstrukturen zugeschnitten ist.

Im Use Case ASIC wurde eine Plattform für die KI-basierte Analyse und Prognose des Akuten Respiratorischen Distress Syndromes (ARDS), einer häufig tödlichen Komplikation auf Intensivstationen, entwickelt. Versorgungsdaten aus der Intensivmedizin bieten hierfür erhebliche Vorteile: Standardmäßig werden von den Patienten eine große Zahl von Parametern automatisiert erfasst, so dass die Datenqualität der Versorgungsdaten vergleichsweise hoch ist. Außerdem steht international eine hohe Zahl von intensivmedizinischen Patientendaten für die Forschung zur Verfügung, so dass die MII-Datenstrukturen auch im internationalen Kontext verglichen werden können.

Nach Aufbau einer Softwareplattform, die eine automatisierte Aufbereitung von intensivmedizinischen Daten für das Training von KI-Verfahren unabhängig von der speziellen Spezifikation erlaubt, wurde die Performance einer großen Zahl von KI-Methoden sowohl »klassischen« Typs, wie Random Forests, als auch Deep Learning Verfahren evaluiert. Zielsetzung war dabei die frühe Erkennung von Fällen mit ARDS-Syndrom sowie die Analyse der Übertragbarkeit von KI-basierten Modellen zwischen Krankenhäusern. Letzteres stellt insofern eine erhebliche Herausforderung für die Zukunft der KI in der Versorgung dar, da internationale Anbieter von KI-basierten medizini-

schen Lösungen zum Training der KI typischerweise auf die großen, international verfügbaren Datenbanken zurückgreifen. Ein systematischer Bias sowohl zwischen Daten aus unterschiedlichen Ländern als auch zwischen Krankenhäusern würde daher ein erhebliches Hindernis für die Einführung von KI-basierten Systemen in der Versorgung in Deutschland bedeuten.

In ASIC konnte gezeigt werden, dass für das ARDS eine Übertragbarkeit von KI-basierten Modellen zwischen den Partnerkliniken im Use Case ASIC nur einen leichten Abfall an Performance impliziert. Eine direkte Übertragbarkeit von KI-basierten Modellen zwischen unterschiedlichen Ländern hingegen ist mit zu erwartendem deutlichem Verlust an Performance verbunden.

Als weitere Herausforderung für KI-basierte klinische Entscheidungsunterstützung zeigte sich die hohe Rate an Fehlalarmen bei der Prognose seltener Ereignisse, die bei Einsatz von KI-basierten Modellen mit der heute erreichbaren Performanz zu erwarten sind. In ASIC wurde daher die Verwendung von mehrstufigen KI-basierten Entscheidungsunterstützungssystemen untersucht, die einen ersten Schritt zu einer »personalisierten« KI für die Nutzung von Versorgungsdaten aufzeigt.

In einem weiteren Use Case des SMITH-Konsortiums (HELP) geht es um eine verbesserte Antibiotikatherapie bei Staphylokokken-Blutstrominfektionen. Verbesserungen sollen hierbei durch Entscheidungsunterstützung erzielt werden. In der ersten Aufbaustufe evaluiert das Projekt den Einsatz eines digitalen Handbuches. Dessen Effekt wird standortübergreifend auf Grundlage von Daten aus den DIZ bewertet, ein Beispiel für den Beitrag der DIZ zur Evaluation von KI-Anwendungen. Die zweite Ausbaustufe des Projekts zielt darauf ab, Patientendaten über die DIZ in ein Entscheidungsunterstützungssystem zu integrieren, das als Medizinprodukt eingesetzt werden kann.

Diskussion

Die Eignung von Versorgungsdaten für datengetriebene KI ist problematisch. Notwendig sind ein verbessertes Datenqualitätsmanagement und große Transparenz hinsichtlich der Bedingungen, unter denen die Daten erhoben werden. Die Herausforderung, die Eignung von Versorgungsdaten für KI zu steigern, bleibt bestehen.

MII-weit stehen standortübergreifende Ansätze zur datenschutzgerechten Datennutzung zur Verfügung. Die MII-Datenintegrationszentren stehen bei der Nutzung entsprechender Verfahren oft vor dem Problem, die notwendige Öffnung der IT-Systeme mit den Anforderungen an Datenschutz und Datensicherheit einer kritischen Infrastruktur zu vereinbaren. Nicht überall ist absehbar, ob die notwendigen Anpassungen der IT-Infrastruktur der Krankenhäuser realistische Chancen auf Umsetzung haben.

Der Aufbau eines Kompetenzzentrums für die Bewältigung der regulatorischen Anforderungen steht in der kommenden Förderphase der MII bevor. Es dient dem Kompetenzaufbau und der Etablierung von Austausch- und Beratungsstrukturen in der MII, die – nicht nur für KI-Anwendungen im engeren Sinne – die medizinprodukterechtlichen Herausforderungen der Software-Entwicklungen adressieren. Ein solches Zentrum kann Ressourcen vorhalten, die einzelne Standorte nicht lokal aufbauen können und deren mehrfacher paralleler Aufbau auch ineffizient wäre. Die stärkere Einbindung der klinischen Studienzentren in die Arbeit der MII verspricht weitere Verbesserungen. Einzelne Standorte integrieren in diesem Sinne Datenintegrationszentren und Studienzentren bereits organisatorisch. Gleichzeitig ist insbesondere durch die anstehende Verabschiedung des europäischen Artificial Intelligence Act (AIA) das Risiko gestiegen, dass konkurrierende Normen die Hürden für KI-Anwendungen in der Medizin weiter erhöhen und dass Fehlregulation die Entwicklung auf diesem Gebiet in Europa bremst. Die MII hat die Chance, mit vereinten Kräften auf verbesserte gesetzliche Rahmenbedingungen hinzuwirken.

Fazit

Die MII hat unverzichtbare Voraussetzungen für die Implementierung und Nutzung künstlicher Intelligenz in der Medizin geschaffen und setzt dies fort. Das betrifft die Bereitstellung von Daten, die Vermittlung von Kompetenzen und die Entwicklung von Methoden gleichermaßen. Die MII kann und sollte zur Verbesserung der regulatorischen Rahmenbedingungen und zu einer realistischen KI-Agenda im Gesundheitwesen beitragen. Sie arbeitet auch daran, Unterstützungsstrukturen aufzubauen, die es erlauben, medizinische KI überhaupt rechtskonform zu entwickeln.

Quellen

[1] TMF – Technologie- und Methodenplattform: Ziele – Medizininformatik-Initiative [Internet]. [zitiert 26. August 2022]. Verfügbar unter: https://www.medizininformatik-initiative.de/de/ueber-die-initiative/ziele.
[2] Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. Int J Med Inform. Februar 2017;98:56–64.
[3] Hersh WR, Cohen AM, Nguyen MM, Bensching KL et al. Clinical study applying machine learning to detect a rare disease: results and lessons learned. JAMIA Open. Juli 2022;5(2):ooac053.
[4] Marco-Ruiz L, Bellika JG. Semantic Interoperability in Clinical Decision Support Systems: A Systematic Review. Stud Health Technol Inform. 2015;216:958.
[5] Warnat-Herresthal S, Schultze H, Shastry KL, Manamohan S et al. Swarm Learning for decentralized and confidential clinical machine learning. Nature. Juni 2021;594(7862):265–70.
[6] Rocher L, Hendrickx JM, de Montjoye YA. Estimating the success of re-identifications in incomplete datasets using generative models. Nat Commun. 23. Juli 2019;10(1):3069.
[7] Holmes JH, Bilker WB. The effect of missing data on learning classifier system learning rate and classification performance. In: International Workshop on Learning Classifier Systems. Springer; 2002. S. 46–60.
[8] Nijman S, Leeuwenberg A, Beekers I, Verkouter I et al.Missing data is poorly handled and reported in prediction model studies using machine learning: a literature review. Journal of Clinical Epidemiology. 1. Februar 2022;142:218–29.
[9] Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 25. Oktober 2019;366(6464):447–53.
[10] Kim E, Caraballo PJ, Castro MR, Pieczkiewicz DS et al. Towards more Accessible Precision Medicine: Building a more Transferable Machine Learning Model to Support Prognostic Decisions for Micro- and Macrovascular Complications of Type 2 Diabetes Mellitus. J Med Syst. 17. Mai 2019;43(7):185.
[11] Vokinger KN, Gasser U. Regulating AI in medicine in the United States and Europe. Nat Mach Intell. September 2021;3(9):738–9.
[12] Welten S, Mou Y, Neumann L, Jaberansary M et al. A Privacy-Preserving Distributed Analytics Platform for Health Care Data. Methods Inf Med. Juni 2022;61(S 01):e1–11.
[13] MišiĐ VV, Rajaram K, Gabel E. A simulation-based evaluation of machine learning models for clinical decision support: application and analysis using hospital readmission. NPJ Digit Med. 14. Juni 2021;4(1):98.
[14] Subudhi S, Verma A, Patel AB, Hardin CC et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. NPJ Digit Med. 21. Mai 2021;4(1):87.
[15] Schmücker P, Schemmann U, Winter A, Bott OJ et al. Aus-, Fort- und Weiterbildung in der Medizininformatik-Initiative. mdi – Forum der Medizin, Dokumentation und Medizin-Informatik, 2019;21 (4/2019):102-105.

Ein wissensbasiertes, interoperables Entscheidungsunterstützungssystem für die pädiatrische Intensivmedizin

Prof. Dr. Antje Wulff
Big Data in der Medizin, Department für Versorgungsforschung, Fakultät VI Medizin und Gesundheitswissenschaften, Carl von Ossietzky Universität Oldenburg,
Peter L. Reichertz Institut für Medizinische Informatik der TU Braunschweig und der Medizinischen Hochschule Hannover,
antje.wulff@uni-oldenburg.de

Louisa Bode, B. Sc.
Peter L. Reichertz Institut für Medizinische Informatik der TU Braunschweig und der Medizinischen Hochschule Hannover,
bode.louisa@mh-hannover.de

- Die Standardisierung heterogener Routinedaten ist für klinische Entscheidungsunterstützungssysteme (CDSS) wertvoll.
- Interoperabilitätsstandards in CDSS vereinfachen ihre institutionsübergreifende Nutzung, sind aber selten.
- Das Projekt ELISE erforscht die Potentiale interoperabler CDSS für die pädiatrische Intensivmedizin.
- ELISE nutzt wissensbasierte Modelle sowohl klassisch als auch für die Generierung routinebasierter Trainingsdaten.
- Für eine nachhaltige Ergebnisverwertung kombiniert ELISE den Einbezug eines Systemherstellers und Open Science.

Trotz des steigenden Interesses an digitalen Lösungen für die medizinische Versorgung haben sich klinische Entscheidungsunterstützungssysteme (Clinical Decision Support Systems, *CDSS*) noch nicht in den klinischen IT-Landschaften etabliert. Die Digitalisierung in der Medizin geht mit steigenden Datenbeständen aus der Routine und Forschung einher, die durch aktuelle methodische und infrastrukturelle Entwicklungen in der Medizininformatik immer besser genutzt werden können. Vor diesem Hintergrund eröffnen sich auch für die Erforschung interoperabler CDSS neue Chancen. Auch heute noch können dabei klassische, wissensbasierte Ansätze einen wichtigen Beitrag leisten. In diesem Artikel möchten wir daher einen Einblick in diesen Themenkomplex aus Interoperabilität sowie wissens- und datengetriebenen CDSS-Ansätzen geben und insbesondere ein Forschungsprojekt vorstellen, das sich diesen Aspekten widmet.

Klinische Entscheidungsfindung

Das Lösen komplexer Entscheidungsfragen ist eine Kernaufgabe der medizinischen Versorgung. Fachpersonal ist in der Lage, vielfältige Entscheidungssituationen zu bewältigen, indem es wissenschaftliche Erkenntnisse aus Studien und Leitlinien sowie ihre persönlichen Erfahrungen nutzt. Herausfordernde Arbeitsbedingungen können die Entscheidungsgrundlage destabilisieren und zu unsicheren oder falschen Entscheidungen führen [1]. Insbesondere bei zeitkritischen, komplexen oder seltenen Entscheidungen könnten CDSS eine Unterstützung bieten, indem sie Routinedaten wiederverwenden, sie zielgerichtet zusammenführen und analysieren. Durch die situationsgerechte Bereitstellung entscheidungsrelevanter Informationen und Empfehlungen können CDSS einen wertvollen Beitrag in der Patientenversorgung leisten [2].

Wiederverwendung klinischer Daten und Interoperabilität

Medizinische Datenbestände sind breiter verfügbar als bisher. Damit eröffnen sich Möglichkeiten zur Zusammenführung, Analyse und Interpretation dieser Daten und damit ihrer Nutzung über den ursprünglichen Dokumentationszweck hinaus (»multiple use of data«). Gleichzeitig ist die digitale Versorgungslandschaft von nicht vernetzten Datensilos geprägt, sodass Routinedaten noch nicht effizient genutzt werden. Die Entwicklung von Methoden, die eine effektive Wiederverwendung heterogener, verteilter Datenbestände und eine Umsetzung in mehrwertbringende Anwendungen ermöglichen, ist eine der großen Herausforderung der Medizininformatik [3]. Demnach ist auch eine konsequente Implementierung von auf Routinedaten basierenden CDSS selten. Zudem sind CDSS häufig als herstellerabhängige, institutionsspezifische Standalone-Anwendungen umgesetzt. Eine breite CDSS-Nutzung und ihre Verankerung in der Routine sind ressourcenintensiv.

Der Einsatz von Interoperabilitätsstandards könnte ermöglichen, Anwendungen wiederverwendbar und einfacher institutionsübergreifend einsetzbar zu machen. Die Nutzung offener Interoperabilitätsstandards muss dabei zunächst grundsätzlich verankert werden. Dieses Bestreben wird u. a. in der Medizininformatik-Initiative (MII) verfolgt, in der Universitätskliniken am Aufbau nationaler Infrastrukturen zur Standardisierung und einrichtungsübergreifenden Nutzung medizinischer Routinedaten arbeiten. Das HiGHmed-Konsortium als eines der geförderten Verbünde strebt den Aufbau einer Plattform an, die auf international standardisierten Datenmodellen basiert. Es sollen Daten aus verschiedenen Quellsystemen am Standort kombiniert werden, um auf harmonisierten und institutionsübergreifend standardisierten Datenbeständen u. a. die Entwicklung klinischer, interoperabler Anwendungs- und Forschungssysteme zu ermöglichen [4].

Wissensbasierte Entscheidungsmodelle und standardisierte Wissensrepräsentation

CDSS sind Computersysteme, die das medizinische Fachpersonal bei ihrer Entscheidungsfindung unterstützen, indem Daten und Wissen unter Nutzung von

Marcel Mast, M. Sc.
Peter L. Reichertz Institut für Medizinische Informatik der TU Braunschweig und der Medizinischen Hochschule Hannover, Hannover
mast.marcel@mh-hannover.de

Entscheidungsmodellen verarbeitet und entscheidungsrelevante Informationen dargestellt werden [5]. Eine Kategorie solcher Modelle sind jene, die nicht auf vorhandenes Wissen zurückgreifen, sondern eigenes Wissen generieren. Bei diesen *datengetriebenen Entscheidungsmodellen* erfolgt die Wissensgenerierung durch eine Analyse vorliegender Daten mittels maschineller Lernalgorithmen (induktiver Ansatz, *Lernen durch Daten* [6]). *Wissensbasierte Entscheidungsmodelle* verfolgen einen deduktiven Ansatz (*Lernen durch Wissen* [6]).

Ein *wissensbasiertes System* repräsentiert das zur Problemlösung notwendige Wissen in einer Wissensbasis und trennt es von einer Wissensverarbeitungskomponente. Häufig bilden die früher als *Expertensysteme* bezeichneten Anwendungen das vorher aufwendig akquirierte und formalisierte Fachwissen in Form von *Regeln* ab. Die digitale, regelbasierte Diagnose- und Therapieunterstützung hat ihren Ursprung in den 1970/80er-Jahren und wird noch immer erfolgreich eingesetzt (z. B. 2021 von Sahoo et al. zur Identifikation von COVID-19 Symptomen [7]).

Für eine regelbasierte Wissensrepräsentation existieren verschiedene Formalismen und Technologien. Die *HL7 Arden-Syntax* ist ein medizinisch-fokussierter, standardisierter Formalismus, der die Repräsentation klinischen Wissens und ihrer Entscheidungslogik in Medical Logic Modules ermöglicht. Ein weiterer Ansatz aus der HL7-Familie ist die *Clinical Quality Language (CQL)*, die eine standardisierte Abbildung klinischen Wissens für die Bereiche CDSS und klinische Qualitätsmaßnahmen ermöglicht ([8] liefert einen Vergleich von Arden-Syntax und CQL). Auch die *Guideline Definition Language* (GDL) ist eine formale Sprache zur maschinenlesbaren Beschreibung klinischer Regeln, Leitlinien und Logik. GDL nutzt das Referenzmodell des semantischen Interoperabilitätsstandards openEHR. Nan et al. [9] nutzen GDL z. B. für eine institutionsübergreifend austauschbare Repräsentation medizinischer COVID-19-Leitlinien. Auch domänenübergreifende Formalismen und Tools sind im medizinischen Kontext verbreitet (z. B. Drools (Red Hat)). Ein aktuelles Literaturreview von Papadopoulos et al. zeigt viele der aktuell eingesetzten regelbasierten CDSS-Standards [10]. Dass die Wahl des passenden Formalismus von vielen Faktoren abhängig ist und wie ein Kriterien-Framework aussehen kann, präsentieren Igleasias et al. am Beispiel der Implementierung von Clinical Guidelines für Antibiotika [11].

Das Projekt ELISE

Hintergrund

Die pädiatrische Intensivmedizin muss lebensbedrohliche Erkrankungen für eine Vielzahl krankheits- und altersspezifischer Abhängigkeiten und unter besonderen Arbeitsbedingungen rechtzeitig und verlässlich antizipieren. Das Erkennen des systemischen inflammatorischen Response-Syndrom (SIRS), eine schwere Entzündungsreaktion des Organismus, und der Sepsis sind Beispiele für herausfordernde, klinische Routineaufgaben [12]. Die WHO schätzte für das Jahr 2017 20 Mio. globale pädiatrische Sepsis-Fälle und 2,9 Mio. Todesfälle in der Gruppe unter fünf Jahren [13]. Eine frühe Therapie ist entscheidend, um ein Voranschreiten hin zu Organversagen und Tod zu vermeiden: Han et al. zeigten, dass jede Stunde ohne angemessene Behandlung die Mortalitätswahrscheinlichkeit um 50% steigerte (pädiatrisches Kollektiv mit septischem Schock) [14]. CDSS könnten hier unterstützen, sind derzeit aber nicht verfügbar, nicht realitätsnah evaluiert oder durch fehlende Standards und Schnittstellen nur sehr ressourcenaufwändig in die Routine überführbar [15].

Zielsetzung

Das Forschungsprojekt ELISE (»Ein Lernendes und Interoperables, Smartes Expertensystem für die pädiatrische Intensivmedizin«) widmet sich der Entwicklung interoperabler CDSS-Konzepte zur Detektion und Prädiktion kritischer und die Mortalität und Morbidität bestimmender Zustände (SIRS/Sepsis und assoziierter Organdysfunktionen) in der pädiatrischen Intensivmedizin.

Methoden

Im *ersten Schritt* werden die Quellsysteme und Routinedaten identifiziert und extrahiert. Die Daten werden mittels semantischer Interoperabilitätsstandards des HiGHmed-Plattformansatzes und international standardisierter Datenmodelle in eine openEHR-

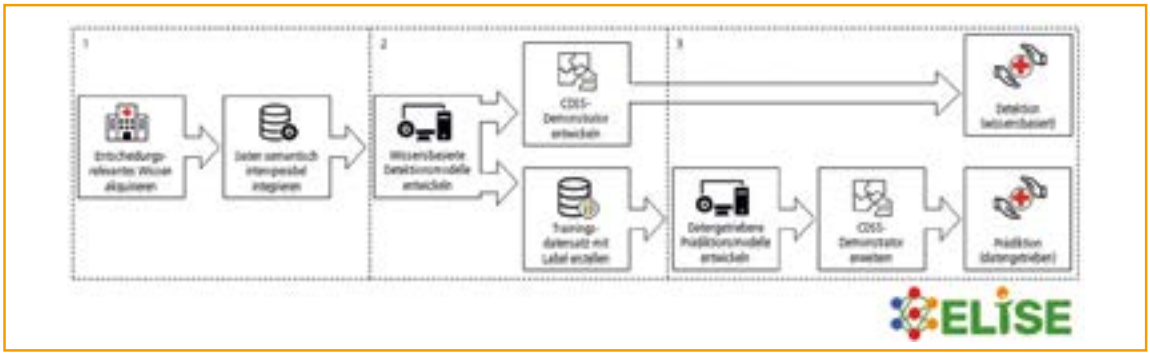


Abb. 1 *Methodisches Vorgehen im Forschungsprojekt ELISE*

Plattform integriert. Der *zweite Schritt* umfasst die Entwicklung wissensbasierter Detektionsmodelle sowie die Implementierung in einen interoperablen CDSS-Demonstrator. In Abwesenheit eines objektiven Goldstandards, werden die Modelle zur Evaluation mit den Einschätzungen des Fachpersonals (imperfekter Goldstandard) verglichen. Bei hoher Güte werden die wissensbasierten Modelle dann genutzt, um dem historischen Datensatz computerbasiert und effizient Outcome-Label hinzuzufügen, da die Routinedaten keine solche Kennzeichnungen liefern. Unter Nutzung dieses Trainingsdatensatzes werden im *dritten Schritt* datengetriebene Prädiktionsalgorithmen entworfen (siehe Abb. 1).

Ergebnisse

Die im *ersten Schritt* durch ausführliche Wissensakquisition identifizierten und mittels openEHR standardisierten Routinedaten umfassen u. a. administrative Daten, Diagnosen, Prozeduren, Medikationen sowie Vital- und Laborparameter und stammen primär aus dem Patientendatenmanagementsystem (PDMS), dem Clinical Data Warehouse und unterschiedlichen Laborinformationssystemen. Im *zweiten Schritt* wurde ein interoperabler, wissensbasierter Ansatz für die SIRS-Detektion entwickelt, der auf einer openEHR-Plattform, standardisierten Datenabfragen und Drools basiert [16]. Zur Modellentwicklung wurden internationale Leitlinien und implizites (Erfahrungs-)Wissen herangezogen. Der Ansatz wurde als CDSS in einer klinischen Studie anhand eines Kollektivs (n=168) der Pädiatrischen Intensivmedizin der Medizinischen Hochschule Hannover evaluiert und zeigte vielversprechende Ergebnisse für die SIRS-Detektion auf Tagesebene (Sensitivität 97,5%, Spezifität 91,5%). Eine vergleichende Auswertung mit den Einschätzungen vom klinischen Personal unter Routinebedingungen (Sensitivität 62,0%, Spezifität 83,3%) zeigte die Potentiale des Konzepts [17]. Der Ansatz wird aktuell auf die Detektion der weiteren Zustände ausgeweitet. Die bereits evaluierten wissensbasierten Ansätze wurden zudem auf den zusammengeführten, historischen Daten (n=4200) angewendet, um automatisiert Outcome-Label hinzuzufügen. Dieser anonymisierte Datensatz steht auch anderen Forschenden zeitnah zur Verfügung [18]. Für die Entwicklung datengetriebener Algorithmen zur Prädiktion (*dritter Schritt*) wurde zunächst die Eignung dieses Datensatzes und seiner präzisen, aber dennoch imperfekten Label als Trainingsgrundlage für die Entwicklung maschineller Lernalgorithmen im Vergleich zur Nutzung des aufwändig erstellten, manuellen Labelings des erfahrenen Fachpersonals (siehe oben, imperfekter Goldstandard) erfolgreich gezeigt [19]. Aktuell wird durch den zertifizierten Systemhersteller zudem an der Einbindung der Ansätze in das genutzte PDMS gearbeitet. Parallel wird die Veröffentlichung eines offenen, standardbasierten Demonstrators vorbereitet.

Fazit

Ein möglicher Grund für die mangelnde Translation von CDSS in die Routine ist ihr Design als Insellösung ohne Berücksichtigung standardisierter Schnittstellen und semantischer Interoperabilität. Insbesondere das Letztgenannte gewinnt bei der Entwicklung von CDSS zwar stetig an Bedeutung, ist aber in aktuellen Lösungen nicht abgebildet, wie auch Papadolous et al. resümieren: *»Interoperability has to be addressed from the start of the development to ensure that the CDSS integrates with the EHR as well as other healthcare information systems […] the findings from this review suggest these recommendations are not being addressed in the development of CDSSs.«* [10]. ELISE verfolgt daher die Umsetzung interoperabler CDSS. Die konsequente Etablierung digitaler Dokumentation und semantischer Interoperabilitätsstandards in den IT-Landschaften ist und bleibt dabei eine notwendige Bedingung. Nicht zuletzt die MII stärkt das Bewusst-

Quellen

[1] Zavala AM, Day GE, Plummer D, Bamford-Wade A. Decision-making under pressure: medical errors in uncertain and dynamic environments. Aust Health Rev. 2018;42(4):395-402.
[2] Castaneda C, Nalley K, Mannion C, Bhattacharyya P, et al. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. J Clin Bioinforma. 2015;5:4.
[3] Prince K, Jones M, Blackwell A, Simpson A, et al. Barriers to the secondary use of data in critical care. J Intensive Care Soc. 2018;19(2):127-31.
[4] Haarbrandt B, Schreiweis B, Rey S, Sax U, et al. HiGHmed – An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. Methods Inf Med. 2018;57(S 01):e66-e81.
[5] Shortliffe EH. Biomedical Informatics: Defining the Science and Its Role in Health Professional Education. In: Hutchison D, Kanade T, Kittler J, Kleinberg JM, et al, editors. Information Quality in e-Health. Lecture notes in computer science. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 711–14.
[6] Beierle C, Kern-Isberner G. Methoden wissensbasierter Systeme: Grundlagen, Algorithmen, Anwendungen. 6th ed. Wiesbaden: Springer Vieweg; 2019.
[7] Sahoo HS, Silverman GM, Ingraham NE, Lupei MI, et al. A fast, resource efficient, and reliable rule-based system for COVID-19 symptom identification. JAMIA Open. 2021 Aug 7;4(3):ooab070.
[8] Soares A, Jenders RA, Harrison R, Schilling LM. A Comparison of Arden Syntax and Clinical Quality Language as Knowledge Representation Formalisms for Clinical Decision Support. Appl Clin Inform. 2021 May;12(3):495-506.
[9] Nan S, Tang T, Feng H, Wang Y, et al. A Computer-Interpretable Guideline for COVID-19: Rapid Development and Dissemination. JMIR Med Inform. 2020 Oct 1;8(10):e21628.
[10] Papadopoulos P., Soflano M., Chaudy Y., Adejo W., et al. A systematic review of technologies and standards used in the development of rule-based clinical decision support systems. Health Technol. 12, 713–727 (2022).
[11] Iglesias N, Juarez JM, Campos M. Comprehensive analysis of rule formalisms to represent clinical guidelines: Selection criteria and case study on antibiotic clinical guidelines. Artif Intell Med. 2020 Mar;103:101741.
[12] Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics*. Pediatr Crit Care Med. 2005;6(1):2-8.
[13] World Health Organization [Internet]. Sepsis. [cited 2022 Aug 02]. Available from: https://www.who.int/news-room/fact-sheets/detail/sepsis.
[14] Han YY, Carcillo JA, Dragotta MA, Bills DM, et al. Early reversal of pediatric-neonatal septic shock by community physicians is associated with improved outcome. Pediatrics. 2003;112(4):793-9.
[15] Wulff A, Montag S, Marschollek M, Jack T. Clinical Decision-Support Systems for Detection of Systemic Inflammatory Response Syndrome, Sepsis, and Septic Shock in Critically Ill Patients: A Systematic Review. Methods Inf Med. 2019;58(S 02):e43-e57.
[16] Wulff A, Haarbrandt B, Tute E, Marschollek M, et al. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. Artif Intell Med. 2018;89:10-23.
[17] Wulff A, Montag S, Rübsamen N, Dziuba F, et al. Clinical evaluation of an interoperable clinical decision-support system for the detection of systemic inflammatory response syndrome in critically ill children. BMC Med Inform Decis Mak. 2021;21(1):62.
[18] Wulff A, Mast M, Bode L, Rathert H, et al. Towards an Evolutionary Open Pediatric Intensive Care Dataset in the ELISE Project. Stud Health Technol Inform. 2022 Jun 29;295:100-103.
[19] Mast M, Marschollek M, Jack T, Wulff A, et al. Developing a Data Driven Approach for Early Detection of SIRS in Pediatric Intensive Care Using Automatically Labeled Training Data. Stud Health Technol Inform. 2022 Jan 14;289:228-231.

sein für diese Perspektive und fördert ihre Umsetzung. Auch die Wahl der CDSS-Entscheidungsmodelle und Formalismen ist herausfordernd und sollte anwendungsfallspezifisch erfolgen. Die Relevanz datengetriebener Methoden nimmt domänenübergreifend, aber auch für den skizzierten Anwendungsfall zu [15]. Wissensbasierte Methoden liefern dennoch hochwertige und – insbesondere im Vergleich zu den noch immer häufig durch Black-Boxes generierten, intransparenten Ergebnissen datengetriebener Ansätze – erklärbare Ergebnisse. ELISE zeigt zudem ein weiteres Potential regelbasierter Methoden: Die Generierung gelabelter Trainingsdatensätze aus Routinedaten. Besondere Vorteile bieten auch standardisierte Regelformalismen, die eine direkte Verknüpfung mit einer standardisierten Datenbasis erlauben. ELISE nutzt Drools in Verbindung mit openEHR, aber insbesondere HL7 Arden-Syntax und GDL könnten vielversprechende Alternativen sein.

Auch die Einbeziehung von Stakeholdern und Systemherstellern sowie die breite Verfügbarmachung von Forschungsergebnissen sind weitere Aspekte für einen nachhaltigen Erfolg entwickelter Ansätze. ELISE verfolgt dies sowohl durch die Einbindung eines zertifizierten PDMS-Herstellers als auch durch die Bereitstellung eines offenen CDSS-Demonstrators. Eine konsequent standardisierte Daten- und Modellrepräsentation sowie -veröffentlichung (Open Science) leistet einen wertvollen Beitrag hin zu einer nachhaltigen Ergebnisverwertung.

Funding und Danksagung

ELISE wird gefördert vom Bundesministerium für Gesundheit, FKZ2520DAT66A. Wir danken unserer Study Group (Pädiatrische Kardiologie und Intensivmedizin und Peter L. Reichertz Institut für Medizinische Informatik der Medizinischen Hochschule Hannover, Institut für Epidemiologie und Sozialmedizin der Westfälische Wilhelms-Universität Münster, Arbeitsgruppe Bioinformatik des Fraunhofer Institut für Toxikologie und Experimentelle Medizin, medisite GmbH). ■

Prof. Dr. Christian Johner
Johner Institut GmbH
christian.johner@johner-institut.de

Regulatorische Anforderungen an Verfahren des maschinellen Lernens bei Medizinprodukten

- Medizinprodukteher steller setzen zunehmend Verfahren der künstlichen Intelligenz (KI), insbesondere des maschinellen Lernens (ML) ein.
- Das Medizinprodukterecht kennt derzeit keine Anforderungen, die spezifisch auf KI-/ML-basierte Medizinprodukte zugeschnitten sind.
- Behörden und Benannte Stellen stehen vor der Herausforderung, die bestehenden regulatorischen Anforderungen auf diese Produktklasse übertragen und dabei den sich rasch wandelnden Stand der Technik berücksichtigen zu müssen.
- Leitfäden sollen Herstellern und Benannten Stellen zu einem gemeinsamen Verständnis bei den Audit- und Zulassungsprozessen verhelfen.

Einführung

Bereits 2018 hat die amerikanische Gesundheitsbehörde FDA das erste Medizinprodukt zugelassen, das Verfahren der künstlichen Intelligenz (KI) verwendet. [1] Seither haben die FDA, andere Behörden sowie die europäischen Benannte Stellen viele weitere KI- bzw. ML-basierte Medizinprodukte zugelassen.

Dazu müssen sie wie bei allen Medizinprodukten beurteilen, ob die regulatorischen Anforderungen erfüllt sind. Dabei stehen sie vor den folgenden Herausforderungen:

- Es existieren keine gesetzlichen Vorgaben, die spezifisch für diese Produktklasse sind.
- Allerdings wurde die Entwicklung einer Vielzahl an Normen und Leitlinien angestoßen, die nur teilweise spezifisch für das Gesundheitswesen sind.
- Auditoren und Prüfer technischer Dokumentationen mit Software-Kenntnissen stehen nur eingeschränkt zur Verfügung. KI-Experten (z.B. Data Scientists) sind besonders rar.
- Der Stand der Technik ändert sich schnell. Das betrifft die ML-Verfahren ebenso wie deren Anwendung in der Medizin. Der entsprechende Publikationsausstoß ist in den Jahren 2014 bis 2019 im Schnitt um über 45% pro Jahr gewachsen. [2]

Um diese Herausforderungen zu bewältigen, verwenden die deutschen Benannten Stellen einen eigenen Leitfaden zur künstlichen Intelligenz, der auf der »AI Guideline«[3] des Johner Instituts basiert. Dadurch entstehen de facto untergesetzliche Regelungen.

Qualifizierung und Klassifizierung von ML-basierter Software

Software, die Verfahren des maschinellen Lernens verwendet, zählt in den meisten Fällen als »Medical

Device Software«, weil sie der Diagnose und Therapie dient. Ein Gegenbeispiel ist eine Software, die den Netzwerkverkehr analysiert und KI-basiert entscheidet, ob ein Cyberangriff auf das Medizinprodukt erfolgt.
Bei der Klassifizierung für Software, die der Diagnose und Therapie dient, die Regel 11 der Medical Device Regulation (MDR) anzuwenden. Damit ist für Software, die Informationen liefert, »die zu Entscheidungen für diagnostische oder therapeutische Zwecke herangezogen werden«, eine Einteilung in die Klasse I ausgeschlossen. [4]
Wendet man zusätzlich die Vorgaben der Medical Device Coordination Group (MDCG) [5] an, die auf der Risikoeinteilung des Internation Medical Device Regulation Forums (IMDRF) [6] basieren, fällt solche KI-/ML-basierte Software typischerweise in die Klassen IIa und IIb.

Forderungen der MDR und IVDR

Die MDR und weitgehend buchstabenidentisch die In Vitro Diagnostic Regulation (IVDR) stellen Anforderungen an die Hersteller, die zwar nicht spezifisch für KI-/ML-basierte Produkte sind, die die Hersteller aber auf diese Produkte übertragen müssen. Diese Anforderungen lassen sich aufteilen in:

- Grundlegende Sicherheits- und Leistungsanforderungen gemäß Anhang I
- Anforderungen an die Technische Dokumentation gemäß Anhang II
- Anforderungen an die Post-Market Surveillance gemäß Anhang III
- Anforderungen an die klinische Bewertung gemäß Artikel 61 und Anhang XIV

Dieser Artikel geht vorwiegend auf die erste Gruppe an Anforderungen ein.

Allgemeine Anforderungen

Die zentrale Forderung der EU-Verordnungen besteht darin, dass die Sicherheit der Produkte gewährleistet sein muss und »etwaige Risiken im Zusammenhang mit ihrer Anwendung gemessen am Nutzen für den Patienten vertretbar und mit einem hohen Maß an Gesundheitsschutz und Sicherheit vereinbar sein müssen; hierbei ist der allgemein anerkannte Stand der Technik zugrunde zu legen.«[7]

Damit besteht eine wesentliche Aufgabe der Hersteller darin, den Stand der Technik zu ermitteln. Dazu zählt es, alternative Verfahren und Technologien und die damit einhergehenden Nutzen und Risiken zu bestimmen.

Somit müssen die Hersteller von Produkten, die Verfahren des maschinellen Lernens verwenden, jeweils den spezifischen Nutzen und die spezifischen Risiken diskutieren für:

- Verfahren ohne Software (z.B. Erkennung von Hautkrebs mit dem Auge)
- Software ohne maschinelles Lernen z.B. mit klassischen Algorithmen
- Alternative KI-/ML-Verfahren, z.B. eine andere Architektur eines Deep Neural Networks (DNN) oder der Einsatz von Ensemble Trees anstatt DNNs

Der Stand der Technik ist von den Herstellern im Rahmen der klinischen Bewertung fortlaufend zu ermitteln.

Leistungsfähigkeit

Leistungsparameter spezifizieren

Die Hersteller müssen die »vorgesehene Leistung« der Produkte gewährleisten. [8] Das bedingt, dass sie die vorgesehene Leistung spezifizieren. Bereits bei dieser Aufgabe scheitern viele Hersteller:

Die »Data Scientists« experimentieren mit verschiedenen Verfahren und Architekturen und bestimmen dabei Werte wie die Genauigkeit [9], Sensitivität [10] und Spezifität [11]. Genau diese Werte definieren die Hersteller dann als die Leistungsparameter des Produkts. Korrekter wäre es hingegen, die Wahl dieser Leistungsparameter und deren Zielwerte aus der Zweckbestimmung und aus dem Risikomanagement herzuleiten.

Beispielsweise wäre die Festlegung, dass ein Produkt 99,9995% der Werte richtig klassifiziert (Genauigkeit) bei einem Ebola-Test wenig hilfreich. Denn wenn dieser Test bei einer Million Patienten alle Patienten als negativ einstuft, auch die fünf infizierten Patienten, wäre zwar der Leistungsparameter erfüllt, der Test wäre aber wertlos. Hier wäre die Sensitivität das geeignetere Maß, wenn die Zweckbestimmung darin besteht, möglichst alle mit Ebola infizierten Personen zu identifizieren.

Leistungsparameter nachweisen: Das Test-Dilemma

Der Nachweis, dass diese Leistungsparameter erfüllt sind, erfolgt meist im Rahmen der Verifizierung durch entsprechende Tests. Das Dilemma dabei besteht darin, dass Tests nur Stichprobenprüfungen sind und daher nur Fehler nachweisen können, aber nicht die Korrektheit des Produkts – in diesem Fall des Algorithmus.
Daher sollten die Hersteller solche Testfälle spezifizieren, die geeignet sind, Fehler (hier Abweichungen von spezifizierten Leistungswerten) mit einer besonders hohen Wahrscheinlichkeit zu entdecken. Dazu zählen Tests mit verschiedenen Input-Daten, die sich beispielsweise unterscheiden bezüglich:

- Patientenpopulation
  - Demographische Daten wie Alter, Geschlecht, Rasse
  - zu diagnostizierende oder therapierende Erkrankung bzw. Verletzung
  - Co-Morbiditäten
  - Behandlungen (z.B. unterscheiden sich Lungengewebe, die bereits einer Bestrahlung ausgesetzt wurden, von nicht bestrahltem Gewebe.)
  - Laborwerte

Prof. Dr. Jens Prütting, LL.M. oec.
Bucerius Law School gGmbH, Hochschule für Rechtswissenschaft
Lehrstuhl für Bürgerliches Recht, Medizin- und Gesundheitsrecht,
Institut für Medizinrecht,
Jens.Pruetting@law-school.de

- Verfahren, um die Daten zu generieren
  - Bei Labordaten Messverfahren
  - Bei Bildern Formate, Auflösung, Aufnahmeverfahren (z.B. MRT, CT), Aufnahmeparameter

Die Hersteller sollten nicht nur die tatsächlichen Leistungswerte dokumentieren, sondern auch die zugehörigen Konfidenzintervalle. Insbesondere bei seltenen Wertekombinationen (z.B. in Deutschland Patienten mit Hautkrebs und dunkler Hautfarbe) offenbaren diese Angaben die Unsicherheiten, mit denen die Hersteller die Leistungsparameter gewährleisten können.

Leistungsparameter nachweisen: Ground Truth

Beim Training und der Überprüfung [12] besteht eine große Herausforderung darin, festzulegen, was die richtigen Werte (Soll-Werte) sind. Beispielsweise ist es unter mehreren Ärzten häufig nicht unstrittig, ob auf einem CT-Bild ein Krebs zu erkennen ist. Hier würde der histologische Befund die sogenannte »Ground Truth« bilden, der bestimmt, was die richtigen Werte sind, mit denen die Algorithmen trainiert und überprüft werden.

Bei einem KI-basierten Blutdruckmessgerät (z.B. in einer SmartWatch) ist der »Ground Truth« die invasive Blutdruckmessung. Als »Goldstandard«, mit dem der Hersteller seine Messungen vergleichen würde, käme hingegen eine nicht-invasive Blutdruckmessung mit einer Oberarm-Manschette zum Einsatz.

Risikomanagement

Die MDR und IVDR verpflichten die Hersteller, Risiken zu identifizieren und zu beherrschen.

Beispiele für Risiken

Die meisten ML-Algorithmen dienen entweder der Klassifizierung oder der Regression. D.h. die Risiken entstehen dadurch, dass die Algorithmen entweder falsch klassifizieren oder falsche Werte berechnen z.B.

- auf CT-Bildern irrtümlich ein Gewebe als Krebs klassifizieren, was zu invasiven Biopsien führt,
- auf dermatologischen Bildern eine Läsion nicht als malige identifizieren, weshalb eine zeitnahe und möglicherweise lebensrettende Behandlung unterbleibt,
- die Wahrscheinlichkeit einer Erkrankung falsch berechnen, weshalb die Ärzte keine engmaschige Überwachung empfehlen,
- die Verträglichkeit und Wirksamkeit eines Zytostatikums [13] falsch einschätzen, weshalb der Krebs nicht bestmöglich bekämpft wird und die Patienten unter unnötig starken Nebenwirkungen leiden.

Beispiele für Ursachen

Die Ursachen für diese Risiken sind vielfältig:

- Die Trainingsdaten sind nicht repräsentativ. Dies führt zu einem sogenannten Bias.
- Das Training erfolgt nicht mit einer ausreichend großen Menge an Daten.
- Beim »Pre-Processing«, d.h. der Vorbereitung der Trainings-, Test- und Validierungsdaten unterlaufen Fehler z.B. bei der Konvertierung von Einheiten, beim Umgang mit fehlenden oder unsinnigen Werten, bei der Umwandlung nummerischer in kategoriale Werte, bei der Skalierung usw.
- Beim Daten-Labeling passieren Fehler. Beispielsweise übersieht eine Ärztin eine Zelle in einem histologischen Schnitt mit Krebszellen.
- Der Hersteller wählt nicht das leistungsfähigste ML-Verfahren. Beispiele für diese Verfahren sind Deep Neural Networks (DNN), die logistische Regression sowie komplexe Ensemble Trees wie XGBoost.
- Der Hersteller findet für das gewählte ML-Verfahren nicht die beste Architektur bzw. für die gewählte Architektur nicht den besten Satz an Hyperparametern. Bei einem DNN zählen zu den Hyperparametern die Anzahl der Layers, die Anzahl der Neuronen pro Layer, die Werte für die Initialisierung der Fit-Parameter, die Wahl der Loss-Funktion sowie die Wahl des »Optimizers«.
- Der Hersteller bemerkt nicht, dass der Algorithmus nur zufällig die richtige Vorhersage trifft. Beispielsweise reagiert der Algorithmus, der Hautkrebs erkennen soll, nicht auf die Hautläsion, sondern auf ein Lineal, das im Bild neben der Läsion zu erkennen ist.
- Dem Software-Entwickler unterläuft ein Programmierfehler.

Dass ein Algorithmus falsche Vorhersagen trifft, ist den Verfahren inhärent: So haben Algorithmen zur Klassifizierung fast immer eine Sensitivität und Spezifität kleiner 100%.

Beispiele für die Risikominimierung

Die Hersteller sind verpflichtet, alle Risiken, die aus den oben genannten Ursachen (und vielen weiteren) erwachsen nach Stand der Technik zu identifizieren, zu bewerten und zu beherrschen.

Zum Stand der Technik zählt beispielsweise, dass auch bei komplexen Algorithmen wie DNNs die Hersteller eine Interpretierbarkeit gewährleisten. Diese Interpretierbarkeit umfasst zwei Elemente [14]:

1. Erklärbarkeit (Explainability): Algorithmen können sich erklären, indem sie beispielsweise für ein Bild eines Augenhintergrunds nicht nur angeben, dass eine Retinopathie vorliegt, sondern die Stellen markieren, die sie zu diesem Schluss geführt haben bzw. die dazu geführt haben, dass der Algorithmus eine andere Klassifizierung als unwahrscheinlich verworfen hat.
2. Transparenz (Transparency): Die Algorithmen arbeiten nicht als Blackbox, sondern erlauben einen Einblick in ihr Innenleben. Einige Algorithmen verfügen über eine intrinsische Transparenz, wie einfache Entscheidungsbäume. Bei anderen Algorithmen wie DNNs helfen Verfahren, wie LIME [15].

Anforderungen (harmonisierter) Normen

Die EU-Medizinprodukteverordnungen kennen ebenso wie die EU-Medizinprodukterichtlinie das Konzept der harmonisierten Normen.[16] Demnach dürfen die Hersteller die Konformität ihrer Produkte mit den Teilen der Verordnungen als nachgewiesen annehmen, für die sie die entsprechenden harmonisierten Normen erfüllen.

ISO 13485:2016

Die DIN EN ISO 13485:2016 trägt den Titel »Medizinprodukte – Qualitätsmanagementsysteme – Anforderungen für regulatorische Zwecke«. Mit Bezug zu ML-basierten Produkte sind unter anderem die Anforderungen relevant, die in den folgenden Teilkapiteln genannt werden.

Kompetenzen des Personals

Die Hersteller müssen spezifisch für jedes (!) Entwicklungsprojekt einen Entwicklungsplan erstellen, der »die erforderlichen Ressourcen, einschließlich der notwendigen Kompetenzen des Personals« dokumentiert. [17] Das bedingt, dass die Hersteller die Rollen und deren Aktivitäten identifizieren müssen, die im jeweiligen Entwicklungsprojekt notwendig sind wie z.B.:

- Ableiten der Leistungsmetriken aus der Zweckbestimmung
- Spezifizieren weiterer Software-Anforderungen z.B. an die Benutzungsschnittstelle
- Festlegen der »Ground Truth« bzw. des »Gold Standards«
- Spezifizieren der Art und Menge der Daten für das Trainieren, Validieren und Testen des ML-Modells. Das beinhaltet die Festlegung der Patientenpopulation ebenso wie der Aufteilung der Daten zum Trainieren, Validieren und Testen des Modells.
- Labeling der Daten
- Trainieren, Testen und Validieren des Modells
- Verifizieren und Valideren der Software.

Die Hersteller müssen nicht nur die Kompetenzen für diese Tätigkeiten festlegen, sondern diese auch nachweisen.

Begründung statistischer Verfahren

Die Norm verlangt neben dem Entwicklungsplan auch einen Validierungsplan. Dieser muss »die Methoden, Annahmekriterien und, soweit angemessen, statistische Methoden mit Begründung für den Stichprobenumfang enthalten«. [18]

Beim Machine Learning bedingt dies eine Begründung für die Anzahl der Test-, Validierungs- und Trainingsdaten und für die Auswahl dieser Daten. Hier müssen die Hersteller darlegen, weshalb diese Daten repräsentativ für die spezifizierte Patientenpopulation sind und weshalb Werte außerhalb dieser erwarteten Wertebereiche bei der Validierung einbezogen werden müssen oder eben nicht.

Validierung des Labeling-Prozesses

Ein Teil der Entwicklung besteht im Trainieren des Modells. Dieses Trainieren setzt zumindest beim »Supervised Learning« »gelabelte« Daten voraus. Das sind Daten, bei denen meist Menschen zu Input-Daten den erwarteten Output bestimmen. Beispielsweise müssen Radiologen bei radiologischen Aufnahmen dokumentieren, ob und falls wo auf dem Bild eine Läsion vorliegt, die auf einen Krebs schließen lässt. Diese Kennzeichnung der Daten nennt man das Labeling. Den Prozess des Labelings müssen die Hersteller validieren, um ein fehlerhaftes Labeling z.B. aufgrund der folgenden Probleme möglichst zu vermeiden:

- Unklarheit der für das Labeling verantwortlichen Personen, nach welchen Kriterien die Daten zu bewerten sind
- Müdigkeit dieser Personen
- Mangelnde Kompetenz dieser Personen
- Mangelnde Zeit für das Labeling (z.B. wegen Bezahlung pro Bild)
- Unzureichende Informationen, die für das Labeling benötigt werden.

Validierung computerisierter Systeme

Ein großer Teil der entwickelten Software wird nicht Teil des Medizinprodukts, sondern wird benötigt für das Pre-Processing und für das Trainieren und Validieren des Modells. Für diese Software ist nicht die IEC 62304 anwendbar, sondern die ISO 13485.
Die Norm fordert die Validierung dieser computerisierten Systeme.[20] Diese Validierung muss auch die ML-Bibliotheken umfassen, die meist in großem Umfang zum Einsatz kommen. Dabei müssen diese Bibliotheken gedanklich in zwei Teile aufgeteilt werden:

1. Teil, der der Datenvorverarbeitung (Pre-Processing), dem Trainieren, Validieren und Testen dient und nicht während der Nutzung des Medizinprodukts benötigt wird.
2. Teil, der als Teil des Medizinprodukts die Vorhersage trifft, z.B. Input-Daten klassifiziert (Bild zeigt Krebs oder nicht) oder Werte vorhersagt (Dosis für Medikament).

Diese konzeptionelle Trennung ist den Herstellern meist nicht bewusst. Diese ist aber relevant, auch weil unterschiedliche Normen anwendbar sind: Für den ersten Teil müssen die Hersteller die Anforderungen an die Computerized System Validation der ISO 13485 erfüllen, für den zweiten Teil die Anforderungen der IEC 62304.

IEC 62304

Die DIN EN IEC 62304 nennt sich »Medizingeräte-Software – Software-Lebenszyklus-Prozesse«. Sie ist anwendbar auf Software, die entweder Teil eines Medizinprodukts oder die selbst ein Medizinprodukt ist. Eine Anforderung dieser Norm betrifft den Einsatz von Bibliotheken Dritter. Diese Bibliotheken zählen zur »Software of Unknown Provenance« (SOUP). Üblicher-

weise verwenden die Hersteller vornehmlich SOUP, um die Verfahren des Machine Learnings zu implementieren. Entsprechend wichtig ist es, diese SOUP zu validieren. Eine Mindestanforderung wäre, dass der Hersteller alle Anforderungen, die er an die SOUP stellt, dokumentiert und im Rahmen der Validierung dieser SOUP überprüft – typischerweise mit Hilfe von Software-Tests. Bei vortrainierten Modellen folgt eine weitere Schwierigkeit: Die SOUP besteht nicht nur aus Code, sondern auch aus Daten, mit denen das Modell bereits vom SOUP-Hersteller »gefittet« wurde. Die IEC 62304 verpflichtet die Hersteller dazu, alle SOUP-Komponenten auch nach der Inverkehrbringung, d.h. im Rahmen der Post-Market Surveillance zu überwachen. [21]

Diskussion und Zusammenfassung

Spezifische regulatorische Anforderungen an ML-basierte Medizinprodukte fehlen, und sowohl die Hersteller als auch die Benannten Stellen tun sich schwer damit, bestehende regulatorische Anforderungen präzise auf die Domäne des maschinellen Lernens zu übertragen. Mit der geplanten KI-Verordnung [22] der EU droht ein Gesetz, das die Hersteller weiter belastet [23], u.a. weil dessen Anwendungsbereich sehr weit gefasst ist, weil es teilweise redundante Anforderungen enthält und weil es den Einsatz der KI in Bereichen zu beschränken droht, in denen er für die Patienten besonders (überlebens)wichtig wäre. Kurzfristig gilt es, den drohenden Wildwuchs an Leitlinien und Best Practices zum Machine Learning einzudämmen und die Inhalte zu konsolidieren. Diesen Ansatz verfolgt auch die Arbeitsgruppe AI4H [24] der WHO. Die Leitfäden der WHO sollen den Herstellern und Benannten Stellen zu einem gemeinsamen Verständnis und zu Sicherheit bei den Audit- und Zulassungsprozessen verhelfen. Völlig unabhängig davon gilt es, die rechtlichen Konsequenzen aus anderen Rechtsbereichen zu untersuchen, wie dem Haftungsrecht. ■

Quellen bzw. Fußnoten

[1] https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye
[2] Guo Y, Hao Z, Zhao S, Gong J, Yang F; Artificial Intelligence in Health Care: Bibliometric Analysis; J Med Internet Res 2020;22(7):e18228
[3] https://github.com/johner-institut/ai-guideline/
[4] Die Risikoklassen der MDR reichen von Klasse I (geringes durch das Produkt induziertes Risiko für Patienten) über IIa und IIb bis Klasse III (hohes Risiko).
[5] MDCG 2019-11: https://ec.europa.eu/docs-room/documents/37581, Seite 26ff.
[6] http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf
[7] MDR, Anhang I, Absatz 1, Satz 2
[8] MDR, Anhang I, Absatz 1, Satz 2
[9] Anteil der richtig vorhergesagten Ergebnisse
[10] Anteil der richtig vorhergesagten positiven Befunde
[11] Anteil der richtig vorhergesagten negativen Befunde
[12] Beim Machine Learning spricht man von Training, Validation & Testing. Diese Validierung entspricht nicht der Validierung im Sinne der MDR.
[13] Krebsmedikament
[14] TOMSETT, Richard, et al. Interpretable to whom. A role-based model for analyzing interpretable machine learning systems. arXiv, 2018.
[15] Local Interpretable Model agnostic Explanations. Siehe auch Ribeiro, Marco & Singh, Sameer & Guestrin, Carlos. (2016). »Why Should I Trust You?«: Explaining the Predictions of Any Classifier. 97-101. 10.18653/v1/N16-3020. (http://dx.doi.org/10.1145/2939672.2939778)
[16] MDR bzw. IVDR, Artikel 8
[17] DIN EN ISO 13485:2016 Kapitel 7.3.2.f)
[18] DIN EN ISO 13485:2016 Kapitel 7.3.7, Absatz 2
[19] Kapitel IV.1.c nannte typische Verarbeitungstätigkeiten beim Pre-Processing
[20] DIN EN ISO 13485:2016 Kapitel 4.1.6
[21] DIN EN IEC 62304:2007 Kapitel 7.1.3
[22] https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52021PC0206&from=EN
[23] Eine kritische Bewertung dieser Verordnung findet sich in 1.b) ii) dieses Artikels https://www.johner-institut.de/blog/regulatory-affairs/regulatorische-anforderungen-an-medizinprodukte-mit-machine-learning/
[24] https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/default.aspx

Dr. Zully Ritter[1]
zully.ritter@med.uni-goettingen.de

Digitale Assistenz zur Entscheidungsunterstützung in der Notaufnahme mit Hilfe von erklärbaren KI-Verfahren (ML)

- In der Notfallversorgung ist die primäre Diagnosestellung von essentieller Bedeutung für das Patienten-Outcome und stellt aufgrund des Zeitdrucks insbesondere bei vital bedrohlichen Notfällen eine besondere Herausforderung für das ärztliche Personal dar.
- Es ist daher sinnvoll, effektive und effiziente klinische Entscheidungsunterstützungssysteme in diesem Setting zu implementieren.
- Für die Erstellung maschineller Lernmodelle als Teil eines klinischen Entscheidungsunterstützungssystems für die diagnostische Vorhersage in Notaufnahmen können die Notfallbehandlungsdaten des AKTIN-Notaufnahmeregisters genutzt werden.
- Der Aspekt der Erklärbarkeit von maschinellen Lernmodelle ist für die Anwendung, die Akzeptanz, Usability und Utility der Modelle im klinischen Umfeld wichtig.

In klinischen Versorgungsprozessen – zum Beispiel in der Notaufnahme – werden Entscheidungen getroffen, die den Genesungsprozess des Patienten maßgeblich beeinflussen. Im Rahmen dieser klinischen Entscheidungsunterstützung spielt die fortschreitende Digitalisierung im Gesundheitswesen eine immer wichtigere Rolle, da zunehmend relevante Informationen in Form von Behandlungsdaten zur Verfügung stehen. Während die Verfügbarkeit von Behandlungsinformationen für Ärzte und ihre Entscheidungsprozesse grundsätzlich positiv ist, stellt der Umgang mit der Datenmenge gerade in komplexen Behandlungsszenarien eine zunehmende Herausforderung dar [1]. Künstliche Intelligenz (KI) verspricht, mit dieser Vielzahl und Vielfalt an Daten und Informationen umgehen zu können. Während die Erwartungen an KI sehr hoch sind, zeigt sich im klinischen Kontext, dass diese Erwartungen aufgrund unzureichender oder unstrukturierter Daten, nicht verfügbarer Informationen oder mangelnder

Erklärbarkeit der vorgeschlagenen KI-Entscheidungen noch nicht ausreichend erfüllt werden. Dadurch fehlt es den Anwendern oft an Vertrauen in die Systeme.

Diese Herausforderungen wurden im Projekt ENSURE adressiert [2,3]. ENSURE: ENtwicklung Smarter Notfall-Algorithmen dURch Erklärbare KI-Verfahren wird vom Bundesministerium für Gesundheit gefördert (FKZ ZMVI1-2520DAT803) und hat zum Ziel, ein klinisches Entscheidungsunterstützungssystem (Clinical Decision Support System, CDSS) für die Notaufnahme zu entwickeln und zu evaluieren. Teil des Projekts ist die Integration erklärbarer KI-Ansätze in das Entscheidungssystem.

Das Szenario der klinischen Notfallversorgung

Das Setting der Notfallversorgung in der Zentralen Notaufnahme der Krankenhäuser ist im Gegensatz zu anderen klinischen Versorgungsbereichen insbesondere dadurch charakterisiert, dass eine rasche, aber vor allem präzise Behandlung des Patienten insbesondere bei vital bedrohlichen Notfällen erfolgen muss. Grundvoraussetzung hierfür ist, dass der Diagnoseprozess möglichst akkurat und zeiteffizient durchlaufen wird. Insgesamt kann das Entscheidungsszenario im Setting der Notaufnahme als herausfordernd beschrieben werden. Hierbei spielen der Zeitdruck und die z.T. hohe Komplexität einzelner Notfälle eine große Rolle. Darüber hinaus werden in den Notaufnahmen häufig Rotationsassistent*innen verschiedenster Fachdisziplinen zur Weiterbildung eingesetzt, die noch nicht über das erforderliche notfallmedizinische Fachwissen verfügen. Die Nutzung von nicht-evidenzbasierten Wissensquellen als Hilfsmittel stellt dabei ein potentielles Risiko dar [4], weshalb die Entwicklung von vertrauenswürdigen und digital verfügbaren Hilfsmitteln für die Notfallversorgung erstrebenswert ist. ENSURE wurde konzipiert, um eine strukturierte und evidenzbasierte Vorgehensweise in der Notfalldiagnostik zu implementieren und auf diese Weise die Qualität der Notfallversorgung zu verbessern.

CDSS und ENSURE

Klinische Entscheidungsunterstützungssysteme (CDSS) sind komplexe IT-Systeme [5], welche Anwender, etwa Ärzte, Pflegekräfte oder Patienten, in ihrem Entscheidungsfindungsprozess im klinischen Kontext assistieren [6]. Die Vielfalt von klinischen Entscheidungen spiegelt sich in der Diversität der CDSS wider, sodass die Anwendungsfelder von einfachen Erinnerungs- und Alarmsystemen bis hin zu komplexen Diagnoseunterstützungssystemen (DDSS) reichen können. Die Komplexität von CDSS verlangt ein interdisziplinäres Team aus klinischen Domänenexpert*innen, Data Scientist*innen, Informatiker*innen und vieler weiterer Professionen. Die Beschaffenheit eines CDSS wird durch unterschiedliche Komponenten charakterisiert. Im Fokus dieses Artikels liegt die Anwendung von künstlicher Intelligenz, welche durch die Wissensbasis (Knowledge Base) und die Inferenz implementiert wird [7]. Im Projekt ENSURE kommt zum einen der KI-Ansatz »Expertensystem« (regelbasiertes System) und zum anderen »Machine Learning« zum Einsatz, um Notfallmediziner*innen bei der Diagnosestellung zu unterstützen. Die dafür notwendigen Falldaten werden im Entscheidungsszenario in eine mobile Applikation per Tablet eingegeben, um daraus eine Empfehlung durch ENSURE berechnen zu lassen. Die Berechnung der Empfehlung erfolgt randomisiert entweder per Expertensystem oder durch Machine Learning. Mit dieser Arbeitsweise ist die Fragestellung verbunden, welcher der Ansätze die beste Performanz in der Versorgung leisten kann.

Entwicklung einer diagnostischen Entscheidungsunterstützung (DDSS) für den Einsatz in der Notfallversorgung

Die Erprobung des DDSS ENSURE soll im Rahmen einer sonstigen klinischen Prüfung erfolgen. Für die prospektive Anwendung von ENSURE im klinischen Umfeld der Notaufnahme sind zwei Aspekte von übergeordneter Bedeutung: 1. Hochqualitative Daten und 2. die Erklärbarkeit der KI-Algorithmen. In den folgenden Abschnitten werden diese beiden Aspekte detailliert beschrieben.

Auswahl der Wissensbasis für Expertensysteme und KI-Modelle

In ENSURE werden zwei unterschiedliche Ansätze für die Wissensbasis verwendet, um zum einen das »Expertensystem« (regelbasiertes System) und zum anderen das »Machine Learning« Modell mit entscheidungsrelevanten Informationen zu versorgen. Expertensysteme verwenden evidenzbasierte Wissensquellen, um eine Entscheidungslogik abzubilden. Bei maschinellen Lernmodellen werden Daten, im besten Fall Patientendaten, direkt aus den Produktivsystemen verwendet, um die Zielpopulation für die Entscheidungsfindung zu repräsentieren.

Die maßgebliche Wissensquelle für das Expertensystem von ENSURE ist das sog. SOP-Handbuch Interdisziplinäre Notaufnahme [8]: In diesem Handbuch sind für alle notfallmedizinischen Leitsymptome und Leitdiagnosen die standardisierten Vorgehensweisen in Notfalldiagnostik und -therapie auf der Basis der aktuell gültigen medizinischen Leitlinien und im Konsens mit den federführenden Fachgesellschaften dargestellt. Die Translation des SOP-Handbuchs in eine Regelrepräsentation für das Expertensystem erfolgt durch die enge Zusammenarbeit zwischen Notfallmedizinern und IT-Experten.

Stefan Rühlicke, M.Sc.[1]
stefan.ruehlicke@med.uni-goettingen.de

Prof. Dr. Sabine Blaschke[2]
sabine.blaschke@med.uni-goettingen.de

Prof. Dr. Tibor Kesztyüs[1]
Kerstin Pischek-Koch[1]
Stefanie Wache[1]
Frank Schultze[2]
Katrin Esslinger, B.A.[2]
Michael Schmucker[3]
Andreas Reiswich[3]
Prof. Dr. Martin Haag[3]
Wiebke Schirrmeister[4]
Prof. Dr. Dagmar Krefting[1]

[1)] Institut für Medizinische Informatik, Universitätsmedizin Göttingen
[2)] Zentrale Notaufnahme, Universitätsmedizin Göttingen
[3)] GECKO-Institut, Hochschule Heilbronn
[4)] Klinik für Unfallchirurgie, AKTIN-Notaufnahmeregister Universitätsklinik Magdeburg

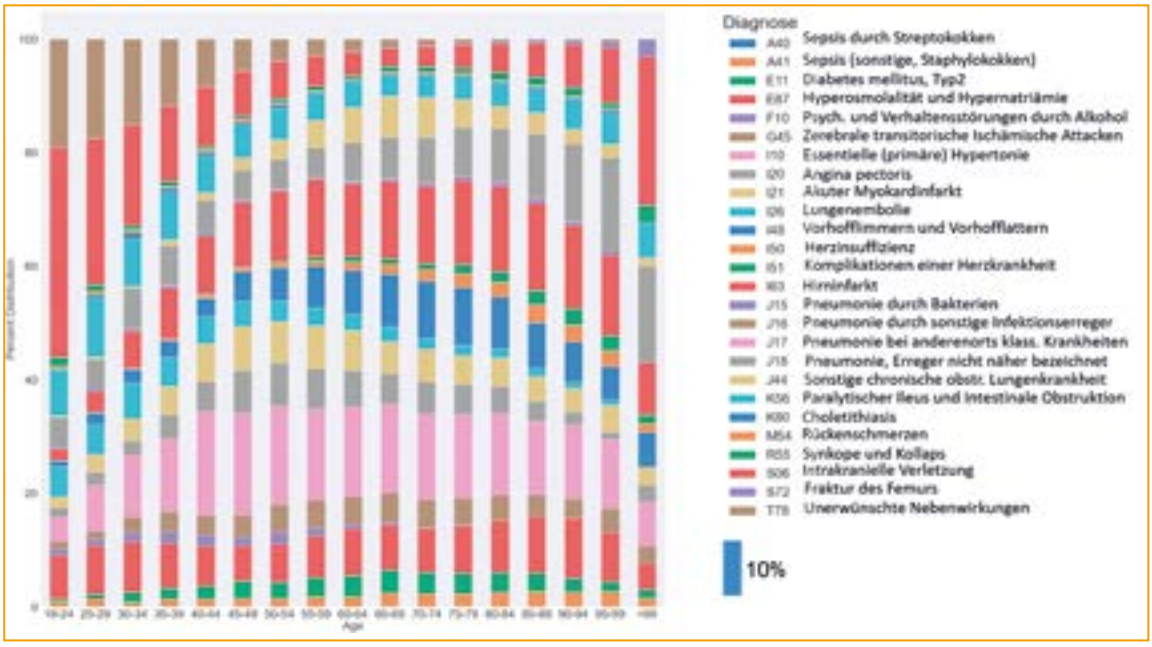


***Abb. 1** Analyse der häufigsten Notfall-Diagnosen in Prozent in Abhängigkeit vom Alter im AKTIN-Notaufnahmeregister-datensatz im Zeitraum 2017-2020 in n=13 teilnehmenden Kliniken*

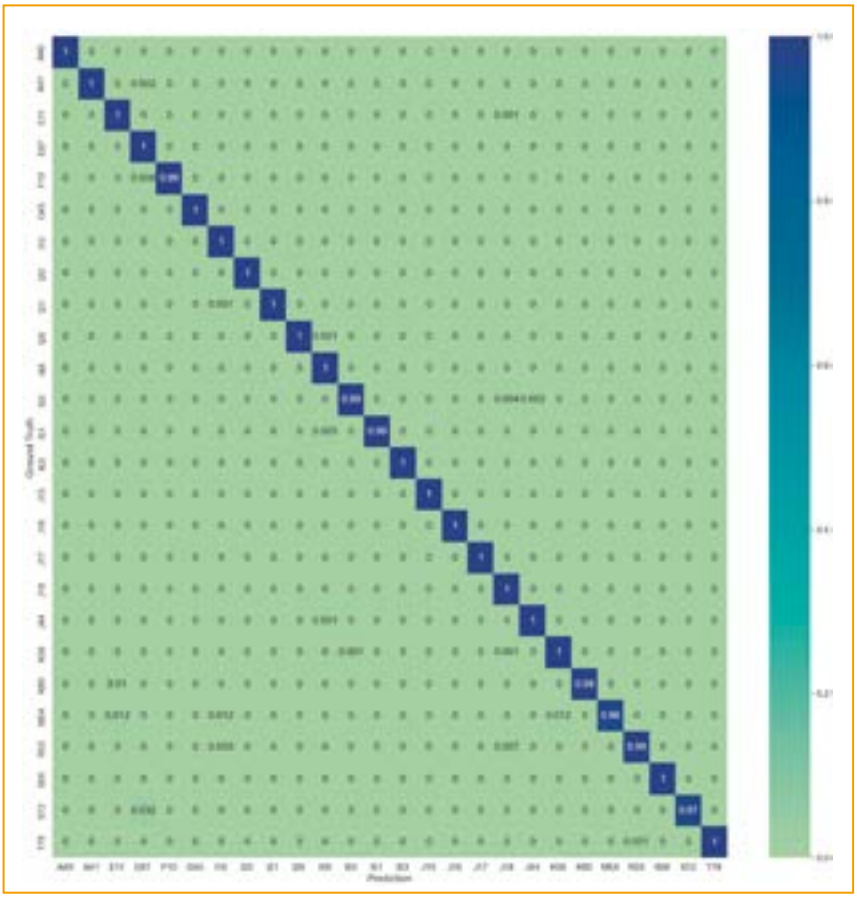


***Abb. 2** Normalisierte Mehrklassen-Konfusionsmatrix eines trainierten Random Forest Models zur Diagnoseklassifikation (ICD Codes in alphabetischer Reihenfolge sortiert).*

Für das maschinelle Lernmodell von ENSURE werden zwei unterschiedliche Datenquellen einbezogen: Zum einen werden Falldaten von 2017 bis 2020 aus dem AKTIN-Notaufnahmeregister verwendet [9,10], welche von 13 angeschlossenen Notaufnahmen in Deutschland bereitgestellt wurden. Insgesamt liegen 51 Merkmale inklusive der abschließenden Fachabteilungsdiagnose von 137.152 Fällen als Zielvariablen in der Registerdatenbank vor. Abb. 1 zeigt beispielhaft die Verteilung der häufigsten Diagnosen (ICD Code) mit Bezug zum Alter der Patient*innen. Aus der Grafik wird deutlich, dass bestimmte Diagnosen in Abhängigkeit vom Alter wahrscheinlicher werden, wohingegen andere Diagnosen relativ konstant bleiben. Beispielhaft ist der ICD-10 E87 (Hyperosmolarität und Hypernatriämie) nahezu konstant in allen Altersgruppen vertreten, wohingegen das Auftreten des ICD-10 I63 (Hirninfarkt) im Alter zunimmt und im Alter von 75-79 das Maximum von 15 % erreicht.

Als zweite Datenquelle werden Notfallbehandlungsdatensätze aus dem ZNA-Informationssystem der Universitätsmedizin Göttingen (UMG) in Verknüpfung mit den sog. §21-Daten verwendet, die im medizinischen Datenintegrationszentrum (MeDIC) der UMG zusammengeführt werden. Während das ZNA-Informationssystem die gesamten Notfallbehandlungsdaten liefert, wird der §21-Datensatz verwendet, um die abschließende Fachabteilungsdiagnose zu den jeweiligen Fällen einzubeziehen.

Mit der Nutzung der beiden beschriebenen Datenquellen sind die Ziele verbunden, einerseits das Risiko von möglichem Bias in den Trainingsdaten zu reduzieren und andererseits den Grad der Modellgeneralisierbarkeit zu erhöhen. Beide Aspekte limitieren häufig die Anwendbarkeit von KI-Modellen in der Praxis.

Auswahl und Entwicklung der Machine Learning Modelle

Maschinelles Lernen (ML) als Verfahren eines ganzen Spektrums von unterschiedlichen Methoden der KI erlernt Muster in Datensätzen, um Klassifikations- und Prädiktionsprobleme zu lösen. Die im Projekt ENSURE

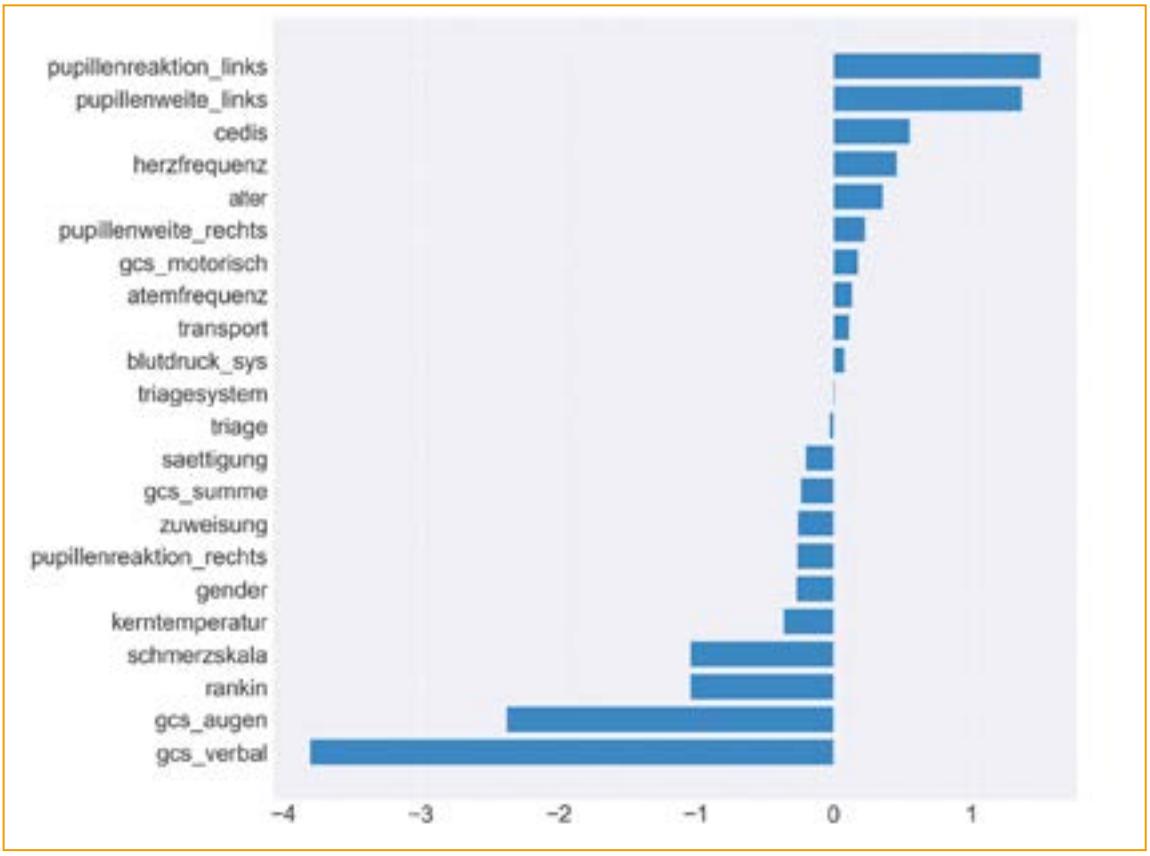


Abb. 3 Diagnose-Klassifikationsergebnis eines Random Forest Models erklärt mit Shap [12] durch Darstellung der Merkmalsgewichtung

verwendeten Datensätze der Notfallbehandlung umfassen dabei unterschiedliche Parameter, so u.a. demographische Daten, Vitalparameter, Symptome sowie Diagnosen. Um diese Daten für das Training von ML-Modellen verwenden zu können, wurden zunächst die üblichen datenwissenschaftlichen Prozessschritte umgesetzt. Hierzu gehört, das Bereinigen der Daten, das Imputieren fehlender Daten und die Prüfung der Plausibilität etwa von physiologischen Werten auf Basis von Literatur und klinischer Erfahrung.

Während sich viele Studien mit Anwendung von ML-Modellen im Notaufnahmesetting auf die Unterscheidung von zwei Klassen beschränken, so etwa, ob eine Sepsis vorliegt (ja/nein), konzentriert sich ENSURE auf die Klassifizierung der 20 häufigsten Diagnosen in der klinischen Notfallmedizin. Teil des Entwicklungsprozesses ist die Erprobung mehrerer ML-Modelle und deren retrospektive Überprüfung anhand von Testdaten. Ziel dieser Labortestung ist es, das Modell mit der besten Performance zu identifizieren und dieses dann prospektiv neben dem Expertensystem (regelbasierten System) in der sonstigen klinischen Prüfung nach MDR/MPDG (siehe hierzu den Artikel von Johner und Prütting in diesem Heft) zu erproben.

So wurden etwa Modelle der logistischen Regression, Random Forest Modelle, neuronale Netze sowie Support Vektor Maschinen auf Basis der Datenquellen trainiert. In Abb. 2 ist beispielhaft die normalisierte Mehrklassen-Konfusionsmatrix eines dieser trainierten Modelle (Random Forest Modell) abgebildet, welches retrospektiv mit einem Testdatenset erprobt wurde. Die Grafik beschreibt die Performance des Models für die 20 Diagnosen. So wird etwa S72 (Fraktur des Femurs) zu 97 % korrekt prädiktiert. Im gleichen Beispiel besteht eine 3,2 %ige Wahrscheinlichkeit, dass das Ergebnis als E87 (Hyperosmolarität und Hypernatriämie) prädiktiert wird.

Erklärbarkeit

Während die Performance entlang der Parameter Spezifität, Sensitivität, AUC-Werte etc. die Bewertung einer Klassifikationsgüte zulassen, besteht dennoch das Problem, dass ML-Modelle und deren Entwicklung nicht hinreichend im Kontext der Erklärbarkeit für die Endnutzer sind [11]. So besteht für Endanwender nicht nur der Anspruch einer möglichst hohen Klassifikationspräzision des Modells, sondern auch die Anforderung zur Transparenz der Entstehung bzw. Berechnung des Modells. Diese Anforderung ist in der technischen Praxis eine Trade-off Entscheidung. So können beispielsweise Deep Learning Netze sehr präzise in ihrer Aufgabe sein, allerdings gelten sie allgemeinhin als Blackboxes und können somit nur schwer zur Erklärung ihrer Ergebnisse genutzt werden. Das Gegenbeispiel sind etwa Regressionsmodelle, welche leichter eine Erklärung für das Rechenergebnis liefern können, aber in der Regel weniger präzise bei der Problemlösung sind.

Quellen

[1] Schöpke, T. et al. Statusbericht aus deutschen Notaufnahmen, Notfall Rettungsmed. 17 (2014) 660–670.
[2] http://bmg-projekt-ensure.uni-goettingen.de/home/ (zuletzt zugegriffen am 31.08.2022)
[3] https://medizininformatik.umg.eu/en/research/projects/ensure/ (zuletzt zugegriffen am 31.08.2022)
[4] Bernstein, S.L. et al. The Effect of Emergency Department Crowding on Clinically Oriented Outcomes, Academic Emergency Medicine. 16 (2009) 1–10.
[5] Richter J., Vogel S. Illustration of Clinical Decision Support System Development Complexity. Stud Health Technol Inform. 2020;272:261-264. doi:10.3233/SHTI200544
[6] Sutton, R.T., Pincock, D., Baumgart, D.C. et al. An overview of clinical decision support systems: benefits, risks, and strategies for success. npj Digit. Med. 3, 17 (2020). https://doi.org/10.1038/s41746-020-0221-y
[7] Vogel S., Krefting D. Towards a Generic Description Schema for Clinical Decision Support Systems. Stud Health Technol Inform. 2022;294:119-120. doi:10.3233/SHTI220409
[8] Blaschke S.; Walcher F.; Kulla M.; Wrede C. (Hrsg). SOP Handbuch Interdisziplinäre Notaufnahme. Medizinisch Wissenschaftliche Verlagsgesellschaft (MWV), Berlin, 2. Auflage 2022 (in press)
[9] https://www.aktin.org/de-de/ (zuletzt zugegriffen am 31.08.2022)
[10] D. Brammen, et al. Das AKTIN-Notaufnahmeregister – kontinuierlich aktuelle Daten aus der Akutmedizin. 117 (2020) 24-33. https://doi.org/10.1007/s00063-020-00764-2
[11] L.H. Gilpin, et al. Explaining Explanations: An Overview of Interpretability of Machine Learning, ArXiv:1806.00069 [Cs, Stat]. (2019).
[12] L.S. Shapley, 17. A Value for n-Person Games, Volume II, Princeton Univ. Press, 1953: pp. 307–318.

Erklärbare KI-Verfahren für ML als Brückenschlag zwischen hoher Performance und Erklärbarkeit teilen sich in zwei Bereiche: zum einen modellabhängige Ansätze, zum anderen modellunabhängige Ansätze. Im Projekt ENSURE wurde der Fokus auf modelunabhängige Ansätze gelegt, um eine Vielzahl unterschiedlicher ML-Modelle parallel mit gleichartigen Erklärbarkeitsansätzen zu entwickeln. Eine der häufigsten Erklärbarkeitsansätze ist Shap (Shapley Additive exPlanations) [12]. So wurde im Projekt beispielsweise bei der Diagnoseklassifikation auf Basis unterschiedlicher Merkmalsparameter ein Random Forest Model mit Shap kombiniert (siehe Abb. 3).

Shap beschreibt entlang der Klassifikationsberechnung, welches Gewicht und somit welchen Einfluss die jeweiligen Merkmalsparameter auf das Ergebnis hatten. Für den Anwender ist dies eine zusätzliche Information, welche bei der Interpretation und kritischen Betrachtung der Modelempfehlung unterstützen kann.

Ausblick des Projektes

Klinische Entscheidungsunterstützungssysteme, so auch ENSURE, stehen großen Herausforderungen auf unterschiedlichen Ebenen gegenüber, um in die Routineversorgung gebracht zu werden. Ein Erfolgsfaktor ist es, CDSS schrittweise im Laborsetting zu testen und schließlich im prospektiven Feldversuch zu erproben. Auf diese Weise soll die Robustheit sowie Generalisierbarkeit, aber auch die Handhabbarkeit der Systeme evaluiert werden. Insbesondere die regulatorischen Anforderungen des MPDG stellen hierbei für die Entwickler und Hersteller eine hohe Hürde dar, um ihre Systeme in die klinische Routineversorgung zu bringen. Im Kontext einer sog. sonstigen klinischen Studie soll im Projekt ENSURE das gleichnamige Entscheidungsunterstützungssystem erprobt werden, um die erste Hürde einer prospektiven Pilotstudie zu nehmen. ■

Bedeutung von AI-Literacy bei medizinischen Anwender:innen

- AI-Literacy wird im medizinischen Kontext immer wichtiger.
- Es gibt noch vergleichsweise wenig Bildungsangebote in der strukturierten Aus- und Weiterbildung von Mediziner:innen.
- Dafür entstehen immer mehr alternative Qualifizierungsmöglichkeiten, wie z.B. weiterführende Masterstudiengänge oder Onlinekurse.
- Ein Beispiel ist hier der BMBF-geförderte KI-Campus, der Onlinekurse im Schnittbereich von Medizin und AI anbietet.

Artificial Intelligence (AI) ist in den letzten Jahren zu einer Schlüsseltechnologie geworden, die in immer mehr Lebensbereiche einzieht. Ein Bereich, der sowohl gesellschaftlich als auch individuell von größter Bedeutung ist, ist die Medizin. Spezielle Computerprogramme werden geschrieben, um komplexe Probleme in der Medizin zu lösen, wie z.B. die Früherkennung von Abstoßungsreaktionen nach Organtransplantationen, die Diagnostik von unterschiedlichen Krebsarten anhand multimodaler Daten oder der Einsatz von Pflegerobotern. Während traditionelle AI-Verfahren noch darauf angewiesen waren, genaue Regeln und algorithmische Abfolgen von Expert:innen übermittelt zu bekommen, lernen neuere Programme anhand von Beispielen selbstständig Muster zu erkennen und auf neue Daten anzuwenden. Solche sogenannten Maschinellen Lernverfahren werden vermehrt eingesetzt, um große Datenmengen zu prozessieren und versteckte Dateneigenschaften zu finden, die ein einzelner Mensch nicht mehr überblicken kann. Immer mehr klinische Wissenschaftler:innen aber auch Ärzt:innen sind daran interessiert, innovative Computeralgorithmen zu verwenden, um ihre Daten zu analysieren und die klinische Versorgung zu verbessern.

Diese Computeralgorithmen haben jedoch eine hohe Komplexität und unterliegen immer strikteren gesetzlichen Rahmenbedingungen bezüglich ihrer Entwicklung und Nutzung. Es stellt sich die Frage: was müssen medizinische Anwender:innen eigentlich wissen, um mit AI in ihrem Arbeitsalltag verantwortungsbewusst arbeiten zu können? Welche Anforderungen sollten gestellt werden? Und was kann man von Ärzt:innen und medizinischen Anwender:innen erwarten, die meistens keine oder kaum relevante Vorbildung im Bereich der Mathematik und Informatik haben?

AI-Literacy

Die Bündelung der Kompetenzen, die notwendig sind, um mit AI umzugehen, werden allgemein »AI-Literacy« genannt und werden z.B. bei Long und Magerko [1] definiert als »the competencies necessary in a future in which AI transforms the way that we communicate, work, and live with each other and with machines«. Im Vorschlag der Europäischen Kommission für eine Verordnung zur Festlegung harmonisierter Vorschriften für AI werden Anforderungen an Personen, die die Aufsicht über ein Hochrisiko-AI-System übernehmen, definiert [2]. Die konkrete Ausgestaltung für Mediziner:innen ist jedoch noch unklar und wird kontrovers diskutiert. Wir schlagen vor, die erforderlichen Kompetenzen zu unterteilen in

- Basiskompetenzen (Grundlagen von AI & Maschinelles Lernen auf konzeptioneller Ebene, ethische Fragestellungen, praktische Anwendungen etc.),
- tiefergehende Kompetenzen (Programmierung, technisches Verständnis der verschiedenen AI-Algorithmen) und
- übergeordnete Kompetenzen (Translation, Verantwortlichkeit, Veränderung des Arzt-Patienten-Verhältnis etc.).

Während die Basiskompetenzen und die übergeordneten Kompetenzen wichtig für alle medizinischen Anwender:innen von AI sein sollten, sind die tiefergehenden Kompetenzen vor allem für diejenigen Mediziner:innen wichtig, die selbst AI-Algorithmen entwickeln und einsetzen möchten.

Es stellt sich hierbei die Frage, ob und wie diese Kompetenzen bereits heute in die medizinische Aus-, Weiter- und Fortbildung implementiert sind.

AI in der medizinischen Aus-, Weiter- und Fortbildung

Die Vermittlung von AI-Kompetenzen an (angehende) Mediziner:innen stellt sich in Deutschland heterogen dar und findet insbesondere in der Fort- und Weiterbildung kaum statt. Laut unserer Erhebung boten 71% der medizinischen Hochschulen in Deutschland im Jahr 2021 Lehrveranstaltungen zum Thema AI an [3]. Der Großteil dieser Angebote fand sich jedoch im Wahl(pflicht)bereich statt, nur drei medizinische Fakultäten gaben an, AI-Kurse im Pflichtcurriculum verankert zu haben. Das sollte sich aber bald ändern: Im überarbeiteten nationalen kompetenzbasierten Lernzielkatalog Medizin (NKLM), dem Rahmenlehrplan für das Medizinstudium in Deutschland, werden AI-Kom-

Lina Mosch
Institut für medizinische Informatik / Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Charité – Universitätsmedizin Berlin
lina.mosch@charite.de

Jun. Prof. Dr. rer. nat. Kerstin Ritter
Charité – Universitätsmedizin Berlin
kerstin.ritter@charite.de

Abb. 1
Dr. med. KI als Beispiel für ein Online-Angebot für die Entwicklung von AI-Kompetenzen in der Medizin

petenzen stärker abgebildet. Medizinabsolvent:innen müssen beispielsweise »den Begriff der personalisierten Medizin sowie die Grundlagen und medizinischen Anwendungen von maschinellen Lernverfahren und AI-Systemen erläutern können« [4].

In den Weiterbildungsordnungen für die 49 Fachärzt:innenausbildungen findet sich AI bisher in keinem Lernziel. Unter den 56 Zusatz-Weiterbildungen, die sich an eine abgeschlossene Facharztweiterbildung anschließen lassen, findet sich nur die Zusatz-Weiterbildung »Medizinische Informatik«, die in ihrem Lernzielkatalog das Thema AI aufgreift.

Fortbildungskurse für Ärzt:innen im Rahmen der sogenannten »Continuing Medical Education (CME)« werden durch die Landesärztekammern (LÄKs) zertifiziert. Zudem bietet jede LÄK ihren Mitgliedern auch eigene, durch die LÄK (mit)entwickelte CME-Kurse an. Den Gesamtbestand an CME-Fortbildungsangeboten in Deutschland zu untersuchen, ist aufgrund der Dezentralisierung und der Vielzahl von Anbieter:innen schwierig. In der CME-Datenbank der Bundesärztekammer, in der 2021 ca. 87.000 Kurse verzeichnet waren, thematisierten nur 30 Kurse (entsprechend 0,03%) AI in der Medizin explizit. 2021 boten laut der Erhebung unserer Arbeitsgruppe sieben von 17 LÄKs eigene AI-bezogene CME-Kurse an. Weitere sieben meldeten jedoch zurück, sie würden auch in Zukunft keine Angebote zu AI-Themen planen [3].

Sonstige Angebote

Dem Mangel an Bildungsangeboten in der strukturierten Aus- und Weiterbildung steht ein deutlicher Zuwachs an alternativen Qualifizierungsmöglichkeiten für Mediziner:innen im Bereich AI entgegen: Mittlerweile gibt es mindestens 14 Masterstudiengänge, die sich an Mediziner:innen richten und dem Themenbereich AI gewidmet sind, davon wurden 11 in den Jahren 2020 oder 2021 erstmalig angeboten [3].

Im Bereich Data Science und AI gibt es viele Online-Angebote, um sich Kompetenzen in diesem Bereich anzueignen. Wissensdatenbanken wie die Medizin-Lernplattform Amboss, Podcasts, Videos und Massive Open Online Courses (MOOCs) greifen das Thema AI und Data Science explizit auf die Medizin bezogen auf. MOOCs werden meist auf Plattformen wie coursera, Udemy, Futurelearn oder dem KI-Campus angeboten. Der KI-Campus ist eine auf Grundlage der KI-Strategie der Bundesregierung entwickelte, durch das Bundesministerium für Bildung und Forschung geförderte Open Source Lernplattform für AI – unter anderem mit zahlreichen Angeboten im Bereich Medizin und AI. Ein Beispiel ist hier der von uns entwickelte Podcast sowie das begleitende Online-Lernprogramm Dr. med. KI (s. Abb. 1) mit bislang drei Staffeln: Dr. med. KI – basics, Dr. med. KI – experts und Dr. med. KI – ethics. Geplant sind noch zwei weitere Staffeln (»- international« und »- coding«). Das Programm wird kontinuierlich weiterentwickelt und bietet neben dem Podcast- und Videomaterial viele weitere Möglichkeiten zur Wissensvertiefung wie z.B. Wissensabfragen, Quizze und Austausch in Foren. Das Angebot wird sehr gut angenommen und hat derzeit über 5.000 Hörende. Zusätzlich wird es an der Charité im Wahlpflichtbereich des Modellstudiengangs Medizin eingesetzt und der Einsatz an weiteren Universitätskliniken ist geplant.

Praxisbeispiel: Charité – Universitätsmedizin Berlin

An der Charité – Universitätsmedizin Berlin und dem Berlin Institute of Health (BIH) wird auf verschiedenen Ebenen versucht, AI-Kompetenzen zu vermitteln. Auf der Ebene der angehenden klinischen Wissenschaftler:innen gibt es seit 2019 das (Junior) Digital Clinician Scientist Program (DCSP), dessen Ziel es ist, die Teilnehmer:innen in die Lage zu versetzen, den digitalen Wandel aktiv mitzugestalten. Dafür erhalten sie eine geschützte Forschungszeit von 20% (Junior), bzw. 50%, die sie in ein selbstgewähltes Projekt mit einem starken technologischen und/oder AI-Fokus einbringen, sowie Möglichkeiten zur Weiterbildung im Bereich Digital Health und AI, und in einem begleitenden Mentoring sowohl von klinischer als auch technischer Seite. Dieses Programm ist explizit für angehende Fachärzt:innen mit Habilitationsvorhaben gedacht.

Auf der Ebene der interessierten Studierenden bietet die Charité seit dem Wintersemester 2019/2020

das 3-wöchige Wahlpflichtmodul »KI in der Medizin« mit insgesamt 60 Unterrichtseinheiten an. Ziel ist es, durch Vermittlung konkreter AI-Kompetenzen Ängste und Vorbehalte gegenüber AI in der Medizin abzubauen und Handlungsmöglichkeiten für den eigenen Berufsweg zu finden. Neben einer Einführung in Grundkonzepte von AI und Maschinellem Lernen sind eine Vertiefung in zielgruppenorientierten Programmier-Tutorials und Expertenvorträge zu Anwendungen von AI in der Medizin Teil des Moduls. Den Abschluss bilden die Betrachtung und Diskussion ethischer und rechtlicher Fragestellungen, z.B. zu den Themen Datenschutz, Fairness und Transparenz von AI-Anwendungen in der Medizin. In Kleingruppen stellen die Studierenden eigene (konzeptionelle) AI-Projekte im Bereich der Medizin vor. Die Themen werden in einem Kreativitätsworkshop erarbeitet und reichen von Optimierung von Fahrstühlen in Krankenhäusern, über KI-gesteuertes Monitoring in der Intensivmedizin zur multimodalen Datenauswertung in der Tumorsprechstunde. Das Angebot ist für Studierende der Berlin University Alliance (BUA) geöffnet und somit interdisziplinär.

Zur Vermittlung der Grundlagen von AI und Maschinellem Lernen setzen wir in dem Wahlpflichtmodul das Online-Lernprogramm »Dr. med. KI« ein. Die Studierenden erarbeiten sich so Grundkonzepte aber auch Anwendungsmöglichkeiten in der Medizin selbstständig. Die erlernten Inhalte werden dann in (Online-)Präsenzveranstaltungen vertieft ("inverted classroom«-Konzept), z.B. durch die Möglichkeit Fragen zu stellen, das gemeinsame Lösen von Übungsaufgaben und Gruppendiskussionen. Dies hat den Vorteil, dass die Studierenden sich mit den Themen zunächst individuell auseinandersetzen, sich Gedanken machen können und sich damit deutlich besser in die Lehrveranstaltung einbringen können. Zusätzlich wird dadurch der Austausch innerhalb der Gruppe erhöht und der Unterricht gewinnt an Lebendigkeit und Flexibilität. Die Lehrenden profitieren von der Möglichkeit, auf erlernte Inhalte zurückgreifen zu können und nicht die gesamte Zeit dozieren zu müssen, sondern auf Interessen und Fragen der Studierenden eingehen und Gruppenimpulse setzen zu können. Sie erleben dies als sehr befreiend. Das positive Feedback der Studierenden bestätigt diesen Eindruck. Zusätzlich ist aber auch der direkte Austausch mit medizinischen Anwender:innen wichtig, weshalb wir häufig unterschiedliche Wissenschaftler:innen und Ärzt:innen einladen, um über deren AI-Anwendung zu sprechen und Herausforderungen zu reflektieren. Die Programmiertutorials erscheinen uns als Präsenzworkshop am sinnvollsten, wir eruieren aber gleichzeitig, ob hier auch ein Onlinekurs (Dr. med. KI – coding) Sinn machen würde.

Eine Verankerung von AI-Kompetenzen im Pflichtcurriculum des Medizinstudiums ist an der Charité derzeit in Planung, wird aber an anderen Universitäten schon verbindlich umgesetzt, beispielsweise am Universitätsklinikum Heidelberg mit einer holistischen Verankerung von AI-Lernzielen im Kerncurriculum.

Ausblick

Zusammenfassend sehen wir einen derzeitigen Mangel an AI-Angeboten in der strukturierten medizinischen Aus-, Weiter- und Fortbildung. Im Medizinstudium finden sich zwar bereits zahlreiche innovative Angebote im Wahl(pflicht)bereich, zudem werden AI-Inhalte hier mit Überarbeitung des NKLM sowie der Planung weiterer Angebote innerhalb der Fakultäten präsenter. Es bleibt jedoch in der Verantwortung der Ärzteschaft, bereits fertig ausgebildete Ärztinnen und Ärzte für AI-Anwendungen im klinischen Alltag zu qualifizieren. Online-Angebote wie sie z.B. auf dem KI-Campus angeboten werden, sind als skalierbare Lückenfüller für diese Aufgabe geeignet. Jedoch benötigen klinisch tätige sowie niedergelassene Ärzt:innen freigestellte Zeit und konkrete Weiterbildungskonzepte, um sich im Bereich AI umfassend zu qualifizieren. ■

Quellen

[1] Long D, Magerko B. What is AI Literacy? Competencies and Design Considerations. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems [Internet]. New York, NY, USA: Association for Computing Machinery; 2020 [zitiert 23. August 2022]. S. 1–16. (CHI '20). Verfügbar unter: https://doi.org/10.1145/3313831.3376727

[2] Europäische Kommission: Vorschlag für eine VERORDNUNG DES EUROPÄISCHEN PARLAMENTS UND DES RATES ZUR FESTLEGUNG HARMONISIERTER VORSCHRIFTEN FÜR KÜNSTLICHE INTELLIGENZ (GESETZ ÜBER KÜNSTLICHE INTELLIGENZ) UND ZUR ÄNDERUNG BESTIMMTER RECHTSAKTE DER UNION [Internet]. 2021. Verfügbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206

[3] Mosch L, Back A, Balzer F, Bernd M, Brandt J, Erkens S, u. a. Lernangebote zu Künstlicher Intelligenz in der Medizin [Internet]. Zenodo; 2021 Sep [zitiert 23. August 2022]. Verfügbar unter: https://zenodo.org/record/5497668

[4] Charité – Universitätsmedizin Berlin, Medizinischer Fakultätentag: Nationaler Kompetenzbasierter Lernzielkatalog Medizin (Version 2.0), VII.2-13.1: Umgang mit Digitalisierung [Internet].2021 [zitiert 23. August 2022]. Verfügbar unter: https://preview.nklm.de/zend/objective/list/orderBy/@objectivePosition/lve/212188

Künstliche Intelligenz zur Erkennung von Zustandsänderungen in hochdimensionalen Patientendaten

Maximilian Kapsecker[1,2], maximilian.kapsecker@ukbonn.de

Leon Nissen, M.Sc.[1], Christoph Weinhuber[2], Prof. Dr. Stephan Jonas[1]

1) Institut für Digitale Medizin, Universitätsklinikum Bonn

2) Professur Digital Health, Technische Universität München

- Tragbare Geräte ermöglichen die zeit- und ortsunabhängige Aufnahme klinisch relevanter Gesundheitsdaten.
- Hochdimensionalen Daten können durch Representational Learning in ein zugängliches Format zur Analyse überführt werden.
- Cluster Analysen und visuelle Darstellung der niedrig-dimensionalen Daten lässt das Erkennen von Anomalien zu.
- KI-Analysen von Elektrokardiogrammen sind durch deren diagnostischen Bezug zu kardiovaskulären Erkrankungen motiviert.

Einleitung und Motivation

Mit der Gegenwärtigkeit digitaler Alltagsbegleiter wächst die Verfügbarkeit großer Datenmengen mit Bezug zur körperlichen und geistigen Gesundheit. Insbesondere körpernahe Geräte, das inkludiert exemplarisch Smartphones oder Smartwatches, ermöglichen das Messen von Vitalparametern, wie dem Energieumsatz, der zurückgelegten Strecke, dem Puls oder Elektrokardiogrammen (EKGs). Die Nutzer werden somit befähigt, sich selbst zu überwachen (Quantified Self), sportlich aktiver zu werden oder insgesamt einen gesünderen Lebensstil zu führen.

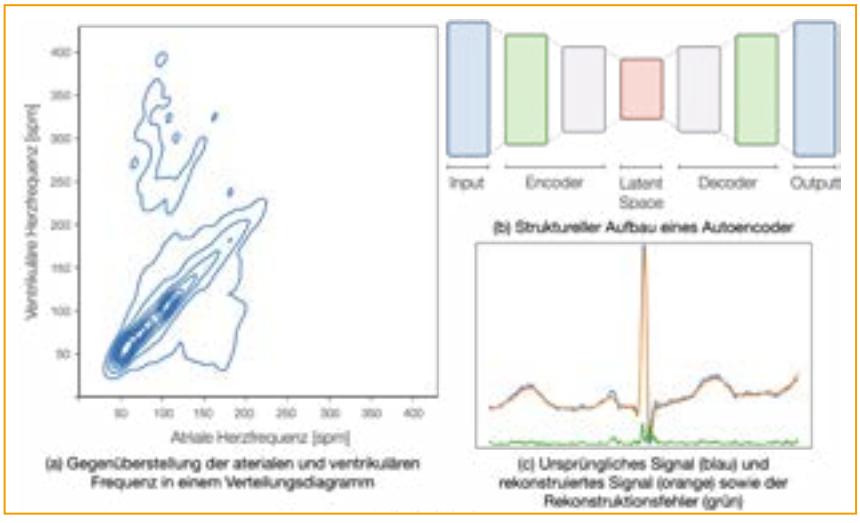
Diese Daten finden bis Dato selten ihren Weg in die klinische Routine. Dazu sind die Menge und Komplexität der aufgenommenen Daten, insbesondere die Zusammenhänge zwischen den Daten mehrerer Sensoren und damit die notwendige Zeit für eine manuelle Auswertung zu hoch. Zudem sind viele der Daten bisher nicht als klinisch relevant betrachtet worden (z.B. Bewegungsmuster oder Tippverhalten). Jene Parameter und gegebenenfalls deren Zusammenhänge können weitere Evidenz generieren. Allerdings ist oft bereits die Hypothesenbildung schwierig, da keine klinische Expertise vorliegt.

In diesem Artikel wird im Folgenden beschrieben, wie Methodik der künstlichen Intelligenz eingesetzt werden kann, um in komplexen Daten Zusammenhänge hervorzuheben, die unter anderem zur Generierung von Forschungshypothesen genutzt werden können.

Methoden: Repräsentative Darstellung hochdimensionaler Daten

Niedrig-dimensionale Daten, wie beispielsweise ein zweidimensionaler Vektor bestehend aus grundlegenden Herzparametern wie der atrialen und ventrikulären Herzfrequenz, können mit gängigen statistischen Methoden ausgewertet werden und sind mittels Visualisierung intuitiv zugänglich (s. Abb. 1a). Im Gegensatz dazu sind hochdimensionale Daten in der Regel ungeeignet für die direkte Auswertung durch Cluster- oder Visualisierungsmethoden. Daraus motiviert sich der Bereich des Representational Learning mit dem Ziel, die strukturell wichtigsten Eigenschaften der hochdimensionalen Daten in einen niedrig-dimensionalen Raum einzubetten. Einen Zugang stellt dabei die Klasse der Variational Autoencoder (VAE) dar [1]. Als Unsupervised Representational Learning Methoden komprimieren VAE mittels eines Encoders hochdimensionale Signale in eine latente, niedrig-dimensionale Repräsentation, sodass durch den Erhalt der strukturell wichtigsten Informationen aus dieser latenten Repräsentation ein Decoder das Ursprungssignal rekonstruieren kann (s. Struktur in Abb. 1b). Der Encoder und der Decoder folgen dabei einer neuronalen Netz-Architektur, welche auf das Ziel einer möglichst identischen Signalrekonstruktion optimiert werden (s. Abb. 1c). In diesem Prozess wird zusätzlich der latente Vektorraum (engl.: latent space) regularisiert, indem die komprimierten Datenpunkte in eine vordefinierte Verteilung (i.d.R. Normalverteilung) überführt werden. Durch den Erhalt der markantesten Informationen und der Annahme der Normalverteilung können potentielle Anomalien in den äußeren Quantilen erkannt werden. Durch die oftmals hohe Individualität der Gesundheitsdaten bieten sich Techniken wie das Transfer Learning oder Fine Tuning zur Personalisierung der Modelle an. Ein auf allgemeinen/diversen Daten konditionierter

Abb. 1



VAE kann auf den persönlichen Daten, beispielsweise auf mobilen Endgeräten, trainiert werden. Obgleich die Vergleichbarkeit mit einer breiteren Population teilweise verloren geht, ermöglicht das personalisierte Verfahren, Veränderungen in den eigenen Daten zu detektieren. Das sensitive und frühzeitige Erkennen von Änderungen in den Signalen, beispielsweise in der postoperativen Überwachung, kann von besonderer Relevanz sein. In der klinischen Routine kann zudem die Visualisierung der Abweichungen dem Arzt eine Vorauswahl an zu prüfenden Daten oder Zeitpunkte geben und die Möglichkeit eröffnen, die betroffenen bzw. kritischen Abschnitte zu interpretieren.

Im Folgenden verwenden wir EKGs als Beispiel zur Erkennung von Veränderungen in komplexen Signalen im Bereich der kardiovaskulären Erkrankungen. Die World Health Organization klassifiziert kardiovaskuläre Erkrankungen als häufigste Todesursache weltweit mit weitreichenden sozialen und wirtschaftlichen Folgen. EKGs zählen in diesem Kontext zu den diagnostischen Standardverfahren in der klinischen Versorgung. Mit der zusätzlichen Motivation von öffentlich zugänglichen EKG-Datenbanken stellt dieser Themenkomplex einen spannenden Anwendungsfall für die Erkennung von Zustandsänderungen in hochdimensionalen Daten dar. Zur deskriptiven Beschreibung der im Folgenden dargestellten Methodik wird zunächst der in 2020 veröffentlichte Datensatz von Zheng et al. verwendet [2]. Der Datensatz enthält über 10.000 zehn Sekunden 12-Kanal-EKGs mit Metadaten zu dem Vorkommen von Arrhythmien und weiteren kardiovaskulären Erkrankungen.

Ergebnisse: Anomalieerkennung in Elektrokardiogrammen

Zur Demonstration des beschriebenen Ansatzes wurden die EKGs (Kanal I) in einem ersten Schritt vorverarbeitet. Die einzelnen EKGs wurden mit Hilfe der Python Bibliothek NeuroKit2 [3] geglättet, normalisiert und eine Erkennung der R-Welle durchgeführt. Mit der Information zu der Lokalität der R-Welle wurden einzelne Herzschläge, je eine halbe Sekunde zur linken und rechten Seite der R-Welle, aus dem EKG segmentiert (entspricht 500 Samples / Dimensionen; s. Abb. 1c). Indem sich somit der R-Peak in jedem Segment an derselben Stelle befindet, wird die Komplexität, die das Machine Learning Modell abfangen muss, verringert. Die daraus resultierende geringere Datenvariabilität erfordert ein proportional kleineres Modell mit weniger freien Parametern und einem kleineren Basisdatensatz für das Training. Mit dem von Zheng et al. generierten Datensatz wurde ein VAE trainiert, wobei die Regularisierung zu einer Normalverteilung in dem latenten Vektorraum in der Optimierung durch einen Gewichtungsfaktor abgeschwächt wurde. Durch das Abschwächen der Regularisierung wird insbesondere eine bessere Rekonstruktion gewährleistet, welche im Prozess der Optimierung als Antagonist zu dem Regularisierungsterm gesehen werden kann. In dem experimentellen Rahmen hat sich eine Dimension von acht für den latenten Vektorraum als geeignete Größe zur strukturwahrenden Komprimierung der EKGs herausgestellt.

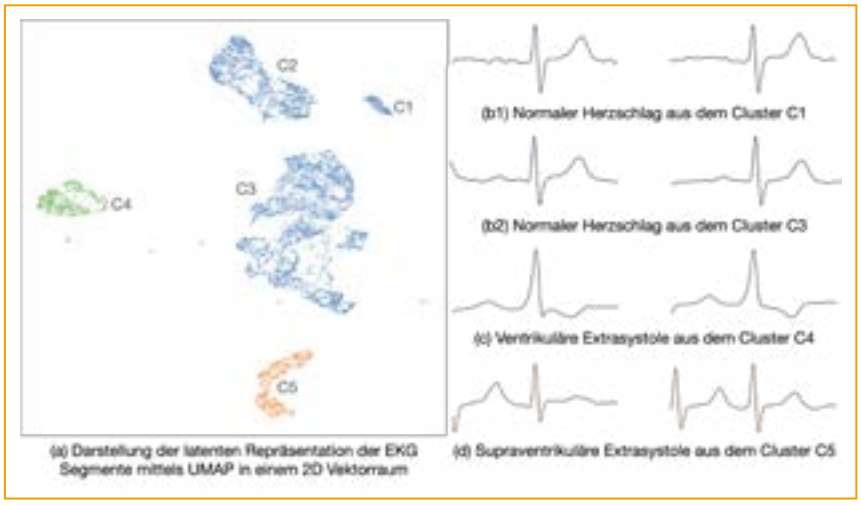


Abb. 2

Zur Validierung des optimierten VAE wurde eine zweite Datenquelle herangezogen, der Icentia11k Datensatz [4]. Hierbei handelt es sich um 1-Kanal EKGs unterschiedlicher Länge von 11.000 Patienten. Zu jedem einzelnen Herzschlag liegen Annotationen vor. Zur Demonstration wurde hier ein EKG eines convenience-gesampelten Subjekts auf dieselbe Weise verarbeitet und es erfolgte auf den Daten ein Feintuning des VAE mittels weniger Epochen (20). Anschließend erfolgte via dem Encoder für das Subjekt ein Mapping der Daten in den 8-dimensionalen Vektorraum. Um einen Eindruck der Verteilung der Punkte in dem niedrig- dimensionalen Vektorraum zu erhalten, wurde mittels des UMAP-Algorithmus ein weiteres Mapping der encodierten Datenpunkte in einen 2-dimensionalen Raum durchgeführt [5] (s. Abb. 2).

Diskussion: Anwendung und Limitation

Die Ergebnisse verdeutlichen die Generalisierbarkeit der gewählten Methodik. Mit wenigen Epochen kann der vorkonditionierte VAE auf die ungesehenen individuellen EKGs optimiert werden und bietet somit einen Zugang für eine ressourcenschonende Ausführbarkeit in mobilen Systemen. Die Validierung der Rekonstruktion auf Testdaten und die visuelle Untersuchung (s. Abb. 2a) des latenten Vektorraums hat gezeigt, dass VAE in der Lage sind, relevante strukturelle Eigenschaften der EKGs in einem niedrig-dimensionalen Raum darzustellen. Dies ermöglicht insbesondere das personalisierte Erkennen anomaler struktureller Änderungen.

Ein unauffälliges EKG kann dennoch verschiedene Ausprägungen aufweisen (s. Cluster C1-3 und zuge-

hörige EKGs in Abb. 2), sodass die Annahme eines normalverteilten Verhaltens der Punkte im latenten Vektorraum nur bedingt zutrifft und somit eine Limitation darstellt. Eine alternative Erklärung ist das Vorhandensein mehrerer verschiedener Zustände, die als »normal« klassifiziert wurden, sich jedoch in den Daten unterscheiden. Daraus kann man Hypothesen zur weiteren Überprüfung ableiten. Die Verallgemeinerung auf ein Gaussian Mixture Model als zugrundeliegende Verteilung ist Gegenstand weiterführender Arbeit und wurde unter anderem durch Guo et al. untersucht [6].

Da das Verarbeiten sensibler personenbezogener Gesundheitsdaten strengen Richtlinien unterliegt, sind konventionelle Machine Learning Applikationen häufig auf die Verfügbarkeit von Open Source Datenbanken limitiert. Aus dieser Problemstellung hat sich das Forschungsfeld Federated Learning entwickelt [7], das einen Zugang zu Machine Learning Verfahren unter Berücksichtigung einer dezentralen Datenhaltung und generellen Einhaltung der datenschutzkonformen Richtlinien bietet. Besonders die Supervised-Verfahren wurden bisher im Rahmen der dezentralen Architektur erforscht [8]. Autoencoder im Allgemeinen sind durch ihre Ähnlichkeit zu klassischen neuronalen Netzen ein naheliegender Zugang für Unsupervised Federated Learning. Dieses Konzept ist unter anderem von Singh et al. vorgestellt worden [9] und soll als Grundlage für Machbarkeitsanalysen im Bereich datenschutzkonformer Analysen hochdimensionaler Patientendaten zur Erkennung von Anomalien dienen, insbesondere auf mobilen Geräten.

Ausblick

Der vorgestellte Ansatz bietet die Möglichkeit, Anomalien oder bisher unbekanntes Verhalten in hochdimensionalen, nicht-annotierten Daten zu erkennen und graduelle Veränderungen zu quantifizieren. Selbst erhobene Gesundheitsdaten können somit semi-automatisch analysiert werden. In einem Anwendungsszenario stellt es somit einen Indikator für potentielle Veränderungen des Gesundheitszustandes dar und ermöglicht ein entsprechendes Warnsystem. Zudem können bisher unbekannte Aspekte durch neue Forschungshypothesen, ergänzt durch Expertenwissen und visuelle Inspektion, untersucht werden. Als Beispiel kann das von Porumb et al. demonstrierte Modell angeführt werden, welches auf Basis des EKGs hypoglykämische Zustände nicht-invasiv eingrenzen kann [10]. Zudem bietet der Ansatz die Möglichkeit, Veränderungen des Vitalzustandes in den postoperativen kontinuierlichen Überwachungen schnell zu erfassen und eine frühe Intervention zu ermöglichen. ■

Quellen

[1] Kingma DP, Welling M: Auto-Encoding Variational Bayes 2014.
[2] Zheng J, et al.: A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Sci Data 2020; 7(1): 48. (ISSN 2052-4463)
[3] Makowski D, et al.: NeuroKit2: A Python toolbox for neurophysiological signal processing. Behav Res 2021; 53(4): 1689-1696. (ISSN 1554-3528)
[4] Tan S, et al.: Icentia11K: An Unsupervised Representation Learning Dataset for Arrhythmia Subtype Discovery 2019.
[5] McInnes L, Healy J, Melville J: UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction 2020.
[6] Guo C, et al.: Variational Autoencoder With Optimizing Gaussian Mixture Model Priors. IEEE Access 2020; 8: 43992-44005. (ISSN 2169-3536)
[7] McMahan B, et al.: Communication-Efficient Learning of Deep Networks from Decentralized Data. Proc Int Conf Artif Intel Stat. PMLR 2017; 1273–1282.
[8] Kairouz P, et al.: Advances and Open Problems in Federated Learning. MAL 2021; 14(1-2): 1-210. (ISSN 1935-8237, 1935-8245)
[9] Singh S, Bhardwaj S, Pandey H, Beniwal G: Anomaly Detection Using Federated Learning. In: Bansal P, Tushir M, Balas VE, Srivastava R, editors. Proc Int Conf Artif Intel Appl, Singapore, 1164. Auflage 2021: 141-148. (ISBN 9789811549915 9789811549922)
[10] Porumb M, Stranges S, Pescapè A, Pecchia L: Precision Medicine and Artificial Intelligence: A Pilot Study on Deep Learning for Hypoglycemic Events Detection based on ECG. Sci Rep 2020; 10(1): 170. (ISSN 2045-2322)

Michael Gratius[1], michael.gratius@stud.hs-hannover.de

Sensorbasierte Klassifikation von Tanzfiguren mittels maschineller Lernverfahren

Diese Studie entstand im sechsten Semester des Bachelor-Studiengangs »Medizinisches Informationsmanagement« im Rahmen des Wahlpflichtprojektes »Sensorbasierte Klassifikation von Aktivitäten mittels maschineller Lernverfahren« der Hochschule Hannover in Zusammenarbeit mit dem Peter L. Reichertz Institut für Medizinische Informatik. Übergeordnete Aufgabenstellung des Projektes war die auf Sensordaten und maschinellen Lernverfahren basierende Analyse von Aktivitäten. Das konkrete Ziel der Projektarbeit war die Klassifikation von sechs unterschiedlichen Figuren des Tanzes »Rumba«. Dieses Ziel wurde gewählt, weil es sich durch die besondere Aktivität von den üblicherweise klassifizierten Aktivitäten, wie z.B. Gehen, Sitzen und Stehen, unterscheidet und nicht nur eine besondere Herausforderung, sondern auch ein wenig erforschter Bereich ist. Einige Publikationen befassten sich bereits mit den Themen »Gesellschaftstanz und Klassifikation« durch maschinelles Lernen [1, 2], behandelten aber andere Fragestellungen oder nutzten andere Methoden. Der Projektablauf wird in Abb. 1 dargestellt und nachfolgend erläutert.

Datenaufzeichnung

Die ausgewerteten Bewegungsdaten der Studie wurden von dem weiblichen Part eines Tanzpaares generiert. Die Daten wurden mittels eines Handyaccele-

rometers, der sich in der hinteren Hosentasche der Probandin befand, bei einer Abtastrate von 100 Hz aufgezeichnet. Jede Figur wurde zwanzigmal ausgeführt.

Datenvorverarbeitung

Der Datensatz hat eine Länge von 21 Minuten und es entstanden aufgrund der Abtastrate 126.450 Samples dreidimensionaler Beschleunigungsdaten. Nach Analyse der Daten in Python wurden die Samples der Tanzfiguren isoliert und in tabellarischen DataFrames gespeichert. Jede Tanzfigur wurde mit 40 Features beschrieben, die aus mehreren jeweils 164 Samples umfassenden Fenstern mit einer Überlagerung von je 50 Prozent extrahiert wurden.

Data Mining

Das Data Mining wurde mit Python PyCaret umgesetzt. Dazu wurde der Featuredatensatz mit Train-Test-Split stratifiziert in 70 zu 30 Prozent aufgeteilt. Für die Erstellung der Modelle wurde die tenfold Cross-Validation genutzt und die Baumtiefe auf 10 limitiert, um die Gefahr von Overfitting zu minimieren. Als primäre Messeinheit für die Güte des Modells wurde Cohen's Kappa gewählt. Das beste Ergebnis der verglichenen Modelle bot ein Extra Trees Classifier mit einem Kappa von 88,46 Prozent, auf den anschließend verschiedene Optimierungsstrategien angewandt wurden. Tuning, Bagging und Boosting konnten keine wesentliche Verbesserung erzielen. Als weiterer Ansatz wurde ein Soft Blending mit einem Random Forest Classifier getestet. Die geringe Verbesserung der Performance stand jedoch nicht im Verhältnis zum Anstieg der Komplexität als auch der erhöhten Gefahr von Overfitting. Abschließend wurde der unveränderte Extra Trees Classifier kalibriert und erreichte in der Validierung mit den Testdaten einen Kappa von 84,44 Prozent. Cohen's Kappa ist im Vergleich zu den Trainingsdaten etwas gesunken, der geringe Unterschied ist aber keine Indikation für ein Overfitting des Modells.

Diskussion

Das erstellte Modell klassifiziert die sechs Tanzfiguren mit einem Kappa von 84,44 Prozent, was für den Rahmen der Studie ein zufriedenstellendes Ergebnis ist. Dennoch gibt es einige Limitationen und mögliche erweiterte Anwendungsgebiete des Modells bzw. des Projekts. Zum einem könnte die Klassifizierung durch verschiedene Faktoren wie andere Figuren, Tänzer*innen oder Tänze ausgedehnt werden. Zum anderen könnte eine andere Art von Modell oder andere Features gewählt werden. Des Weiteren könnten weitere Sensoren an zusätzlichen Körperstellen die Klassifikationsrate verbessern. Mit einem erweiterten Modell ist ein Nutzen für den Tanzsport denkbar. Werden tänzerisch gesehen annähernd perfekte Bewegungsdaten aufgezeichnet, könnten darauf basierend Tänzer*innen während des Trainings korrigiert oder in Turnieren bewerten werden.

Bei Interesse kann das vollständige Studienprotokoll beim Projektteam angefordert werden. ■

Quellen

[1] Voss N, Nguyen PX. End–to–end Classification of Ballroom Dancing Music Using Machine Learning [Conference Paper]. Marseille, Frankreich: Computer Music Multidisciplinary Research; 14.–18. Oktober 2019.

[2] Matsuyama H, Hiroi K, Kaji K, Yonezawa T, Kawaguchi N. Ballroom Dance Step Type Recognition by Random Forest Using Video and Wearable Sensor [Conference Paper]. London, Vereinigtes Königreich: ACM International Joint Conference on Pervasive and Ubiquitous Computing & International Symposium on Wearable Computers; 9.–13. September 2019.

Carolin Cissée[1]
carolin.cissee@stud.hs-hannover.de

Luise Stach von Goltzheim[1]
luise.stach-von-goltzheim@stud.hs-hannover.de

Nektarios Ladas[2]
Dr. Matthias Gietzelt[2]

[1] Hochschule Hannover, Fakultät III, Bachelor Medizinisches Informationsmanagement

[2] Peter L. Reichertz Institut für Medizinische Informatik, TU Braunschweig und Medizinische Hochschule Hannover

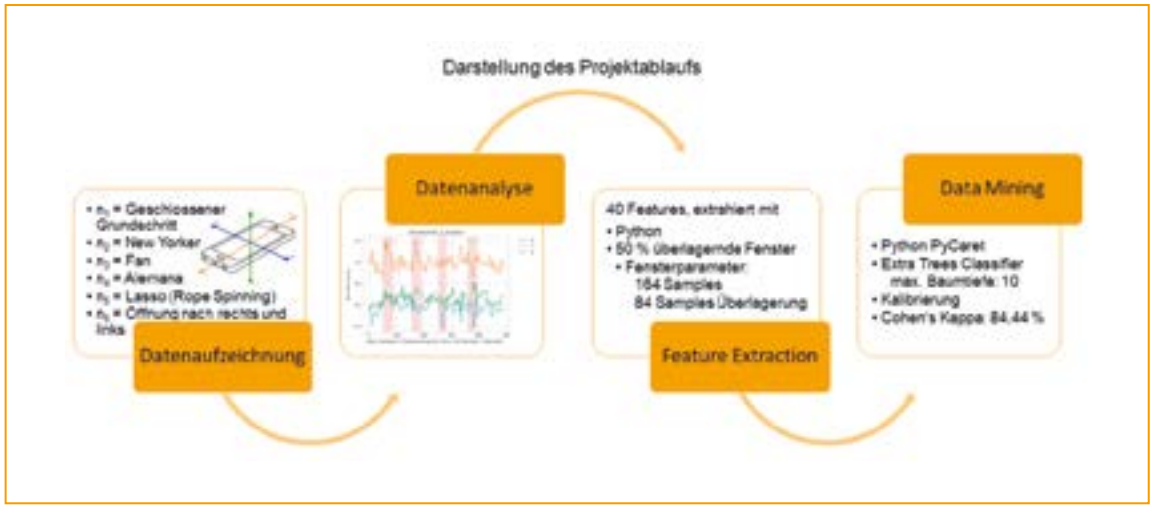


Abb. 1: Darstellung des Projektablaufs

Gedenksymposium zu Ehren von Dr. Carl Dujat

Prof. Dr. Paul Schmücker

Zur Erinnerung an das Wirken von Dr. Carl Dujat und zur Würdigung seiner Verdienste führen die GMDS-Arbeitsgruppe »Archivierung von Krankenunterlagen (AKU)«, der Berufsverband »Medizinischer Informatiker e.V. (BVMI)«, das Competence Center für die Elektronische Signatur im Gesundheitswesen e.V. (CCESigG) und die ENTSCHEIDERFABRIK am 01. Dezember 2022 in der Zeit von 10.00 bis 15.30 Uhr ein Gedenksymposium zu Ehren von Dr. Carl Dujat durch. Dieses findet in Berlin im Tagungszentrum des Hotels Aquino (Hannoversche Straße 5b, D-10115 Berlin-Mitte) statt. Wegen der Corona-Pandemie war das Gedenksymposium leider nicht früher als Präsenzveranstaltung möglich. Im Vorfeld des Symposiums mussten leider mehrere Termine wegen der Pandemie verschoben werden. Eine online-Veranstaltung wurde von den Veranstaltern nie als angemessen erachtet.

Während des Gedenksymposiums werden die beteiligten Verbände die Leistungen von Dr. Dujat aus ihrer Sicht würdigen. Wegbegleiter wie Prof. Dr. Reinhold Haux, Prof. Dr. Peter Haas und Dr. Andreas Beß werden sein Leben und Wirken beleuchten und reflektieren. Weiterhin werden Bernhard Calmer und Prof. Dr. Paul Schmücker über die Entwicklungen der klinischen Informationsverarbeitung in den letzten 60 Jahren referieren und die Arbeiten von Dr. Carl Dujat in diese einordnen. Abschließend wird Sebastian Semler die Konvergenz von Health-IT in Forschung und Versorgung für die letzten 20 Jahre aufzeigen.

Im Rahmen der Würdigung der Arbeiten von Dr. Carl Dujat wird die Entwicklung der rechnerunterstützten medizinischen Informationsverarbeitung von Beginn bis heute nicht nur dargestellt, sondern auch reflektiert.

Nähere Informationen zum Gedenksymposium einschließlich Programm finden Sie auf der Homepage der GMDS unter www.gmds.de. Wir bitten Sie, sich aus organisatorischen Gründen dort für das Gedenksymposium anzumelden. Die Teilnahme am Gedenksymposium ist kostenfrei. ■

Preisverleihung des Berufsverbandes Medizinischer Informatiker e.V. (BVMI) zu Ehren von Dr. Carl Dujat

Dr. Andreas Beß Präsident des BVMI

Dr. sc. hum. Carl Dujat ist als einer der in Deutschland engagiertesten Medizinischen Informatiker im Alter von 56 Jahren überraschend und viel zu früh verstorben. Während seiner Tätigkeiten im Archivbereich der Universitätsklinika Heidelberg und Aachen sowie als Mitbegründer und Vorstandsvorsitzender der promedtheus AG setzte er sich für die Medizinische Informatik in verschiedensten Funktionen und für sachgerechte Lösungen in der Patientenversorgung ein. Er war Mitbegründer der Arbeitsgruppe "Archivierung von Krankenunterlagen (AKU)" der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS), in der er sich in leitender Funktion bis zu seinem Tode engagierte. Von 2008 bis 2013 war er Präsident des Berufsverbandes Medizinischer Informatiker e. V. (BVMI). Weiterhin war er Mitbegründer der ENTSCHEIDERFABRIK und maßgeblich an der Gründung des Competence Centers für die Elektronische Signatur im Gesundheitswesen (CCESigG) beteiligt. Auch hat er wesentliche Beiträge bei der Erarbeitung von Stellungnahmen und Empfehlungen zur digitalen Archivierung, Elektronischen Signatur und Elektronischen Gesundheitskarte erbracht. Ferner organisierte er fortwährend mit Fachkolleginnen und -kollegen wertvolle Veranstaltungen für die Branche, u. a. die legendären Archivtage der AKU. Zusätzlich baute er auch den conhIT-Kongress als Vorläufer der DMEA mit auf. Seine exzellente Expertise und sein Engagement werden heute noch immer sehr geschätzt. Er war Experte für Elektronische Patientenakten und digitale Archive im Krankenhaus sowie für Elektronische Signaturen im Gesundheitswesen. Er hat sich intensiv mit den Themen der rechts- und revisionssicheren elektronischen Archivierung von Krankenunterlagen und der praktischen Umsetzung von elektronischen Sicherungsverfahren auseinandergesetzt. Neben der kontinuierlichen Bearbeitung fachlicher Themen hat er den Ausbau des Netzwerks der Gesundheits-IT umfangreich unterstützt und gefördert. Für seine Leistungen sind wir ihm zu großem Dank verpflichtet.

Die Bewerbungen für das Jahr 2022 sind per E-Mail oder postalisch einzureichen, postalisch jedoch in dreifacher Ausfertigung. Bewerbungen sind zu richten an

Berufsverband Medizinischer Informatiker e.V. (BVMI)
- Geschäftsstelle des BVMI -
Oberlinstraße 26, 41239 Mönchengladbach
eMail: hans-werner.ruebel@bvmi.de
Einsendeschluss für die Bewerbungen: 31. Dezember 2022
Dr. Andreas Bess, Präsident des BVMI

In Erinnerung an seine Verdienste schreibt der Berufsverband Medizinischer Informatiker e.V. (BVMI) den

Dr. Carl Dujat-Gedächtnispreis

für hervorragende Arbeiten im Bereich Medizinische Informatik aus. Dazu gehören insbesondere Arbeiten im Archivwesen, zum Dokumentenmanagement, zur digitalen Archivierung, zu Elektronischen Patientenakten, zur Rechtssicherheit von Patientenunterlagen und zum Informationsmanagement im Gesundheitswesen. Der Preis ist mit € 5.000,– dotiert.

Bewerben kann sich jede Person bzw. jede Gruppe mit exzellenten Arbeiten zu den vorgegebenen Themen. Auch kann eine Person bzw. eine Gruppe für den Preis vorgeschlagen werden. An eine Person bzw. eine Gruppe kann der Preis nur einmal vergeben werden. Die Bewerbung darf maximal 4.000 Zeichen umfassen. Diese sollte nach Möglichkeit strukturiert sein sowie präzise und nachvollziehbar insbesondere die verwendeten Methoden, Techniken und unterstützten Prozesse treffend beschreiben. Der Bewerbung sollte ein aussagekräftiger Lebenslauf inklusive Passfoto beigefügt werden, bei Gruppenbewerbungen aussagekräftige Lebensläufe inklusive Passfotos.

Der Preis wird in der Regel alle zwei Jahre verliehen. Über die Vergabe entscheidet ein Gutachter-Kollegium mit dem Präsidenten des BVMI als Vorsitzenden. Im Jahr 2023 erfolgt die erste Verleihung am 24. April im Rahmen der DMEA-Satellitenveranstaltung. ■

Der DVMD lädt zu seiner 53. Mitgliederversammlung ein.

Samstag, den 12. November 2022, 11:00 – 13:00 Uhr in Berlin.

Details zur Versammlung haben wir für Sie auf https://dvmd.de/events/mitgliederversammlung/ zusammengefasst. Bitte beachten Sie: eine Teilnahme ist nur nach Anmeldung bis spätestens 15. Oktober über das Mitgliederportal möglich.

Köpfe im DVMD: Bärbel Wellmann

Berufliches Tätigkeitsfeld
- Seit 2022 in Ruhestand.

Beruflicher Werdegang
- 1985 Abschluss Diplom-Biologie an der JWGoethe-Universität zu Frankfurt am Main
- 1985 – 1999 Elternzeit, währenddessen freiberufliche wissenschaftliche und medizinische Dokumentationstätigkeit bei Asta Medica und Senckenberg in Frankfurt am Main
- 1999 – 2021 Tumordokumentation Krebsregister Hessen, Agaplesion Markuskrankenhaus in Frankfurt am Main, Universitätsfrauenklinik Mainz und Krebsregister Rheinland-Pfalz in Mainz, OptiPath Frankfurt am Main, Nordwestkrankenhaus Frankfurt am Main

Weiterbildungen
- Tumordokumentation
- Datenbanken: ALBIS, GTDS, ODSeasy, ORBIS, SAP
- Qualitätsmanagement im Gesundheitswesen (EMBA-Zertifikat)
- Qualitätsmanagement im Medizinischen Informationsmanagement
- Komplementäre evidenzbasierte Onkologie/Integrative Onkologie (PRIO-Zertifikat)
- Psychoonkologie (DKG-Zertifikat)

DVMD-Mitgliedschaft
Mitglied seit 2003

Der DVMD ist mir wichtig, weil….
- Als Quereinsteigerin habe ich mich von meinem ursprünglichen Ziel, Wissenschaftliche Dokumentarin zu werden, systematisch in der Tumordokumentation weiterentwickelt und »hochgedient« – und dies auch mithilfe der vielen Fort- und Weiterbildungsangebote im DVMD.
- Diverse DVMD-Fachtagungen boten mir die Möglichkeit, sowohl mit Kollegen in Kontakt zu kommen, als auch meine persönliche Karriere zu fördern.
- Der DVMD bietet eine breite Plattform für die Medizinische Dokumentation – ist informativ, berücksichtigt die vielfältigen Themenfelder der Medizinischen Dokumentation und unterstützt versiert Fort- und Weiterbildung.
- Die Meddok-Liste ist ein unschätzbar wichtiges Tool für die Suche nach einem geeigneten Arbeitsplatz oder beruflicher Neuorientierung.
- Auch nach Beendigung meiner beruflich aktiven Phase fühle ich mich im DVMD e.V. gut aufgehoben und bleibe auch im Ruhestand gerne Mitglied – denn ich verfolge mit Freude die Unterstützung und Förderung des Nachwuchses in unserem Verein!

ID Information und Dokumentation im Gesundheitswesen

Klassifikationen - Terminologien - NLP

Semantische Analyse von Routine- und Forschungsdaten mit ID LOGIK®

Standardisierte Medikationsressourcen per FHIR®

Literatur:

König M, Sander A, Demuth I, Diekmann D, Steinhagen-Thiessen E. Knowledge-based best of breed approach for automated detection of clinical events based on German free text digital hospital discharge letters. PLOS ONE. 2019 Nov 27, https://doi.org/10.1371/journal.pone.0224916

Sander A, Wauer R. Integrating terminologies into standard SQL: a new approach for research on routine data. J Biomed Semantics. 2019 Apr 24;10(1):7

Sander A, Wauer R. From single-case analysis of neonatal deaths toward a further reduction of the neonatal mortality rate. J Perinat Med. 2018 Dec 19;47(1):125–133

Impressum

Charakteristik:
medizin://dokumentation/informatik/informationsmanagement/ (mdi) ist eine praxisorientierte Zeitschrift mit Fachartikeln zur Thematik der Medizinischen Dokumentation und des DV-Einsatzes im Gesundheitswesen und damit angrenzenden organisatorischen Fragen. Sie transportiert Erfahrungsberichte zu Top-Themen sowie aktuelle Entwicklungen direkt in die Praxis. Zielgruppe sind die ca. 2.600 tätigen Mitglieder der beteiligten Verbände, Entscheidungsträger im Management und DV-Management von Gesundheitsversorgungseinrichtungen und bei einschlägigen Industrie-Unternehmen wie Software-Häusern, Pharma-Firmen, CROs sowie leitende Mitarbeiter, Ärzte, Pflegekräfte und Therapeuten.
Verlag und Vertrieb:
Eigenverlag und Eigenvertrieb
ISSN: 1438-0900
Auflage: 1.400 Stück

Erscheinungsweise:
4-mal jährlich, jeweils zum Quartalsende
Herausgeber:
mdi GbR
c/o BVMI Berufsverband Medizinischer Informatiker e.V.
Oberlinstr. 26
41239 Mönchengladbach
Fon: 02166 2171148
Fax: 02166 134545
info@bvmi.de | www.bvmi.de
c/o DVMD Der Fachverband für Dokumentation und Informationsmanagement in der Medizin e.V.
Lobdengaustraße 13
69493 Hirschberg
Fon: 06201 4891884
Fax: 06201 4890459
dvmd@dvmd.de
www.dvmd.de
Nachdruck und Kopien:
Nur mit Genehmigung der Redaktion und unter Angabe der genauen Quelle

Manuskripte:
Zuschriften, die den Inhalt der Zeitschrift betreffen, sind direkt an die Redaktionsanschrift zu senden. Für unverlangte Manuskripte wird keine Haftung und keine Verpflichtung zur Veröffentlichung übernommen. Beiträge, die anderweitig parallel eingereicht wurden, werden nicht angenommen. Die Redaktion behält sich vor, aus technischen Gründen Kürzungen vorzunehmen. Namentlich gekennzeichnete Beiträge geben die Meinung des Verfassers wieder.
Redaktionsteam:
Prof. Dr.-Ing. Oliver J. Bott, Hannover | Prof. Dr. Andreas J. W. Goldschmidt, Frankfurt | Angelika Händel, Erlangen | Markus Stein, Berlin (Leitung) | Prof. Dr. Paul Schmücker, Mannheim
Redaktionsanschrift:
Siehe Verbandsanschrift des BVMI

Autorenrichtlinien:
unter https://www.bvmi.de/mdi/mdi-kontakt/autorenrichtlinien
Bestellungen:
Über die Verbandsanschrift des BVMI. Abbestellungen sechs Wochen zum Jahresende.
Bezugspreis:
Jährlich 49 Euro inkl. MwSt., inkl. Versandkosten. Ausland plus Versandkosten, für BVMI- und DVMD-Mitglieder frei.
Anzeigenpreisliste:
Nr. 22 vom Januar 2021
Anzeigenverwaltung:
DVMD e.V.
Katharina Mai
Lobdengaustraße 13
69493 Hirschberg
Fon: 06201 489-1884, Fax: -0459
dvmd@dvmd.de
Layout:
Fleck · Zimmermann, Berlin
Titelfoto:
shutterstock/HQuality
Druck:
Kössinger AG, Schierling

Archivar 4.0

Das Leistungsportfolio für die digitale Krankenhauszukunft. Powered by DMI.

Mit den zertifizierten Services und Tools der DMI Unternehmensgruppe professionalisieren Krankenhäuser ihr Datenmanagement und schaffen Sicherheit für die Bewältigung heutiger und künftiger Herausforderungen.

Nehmen Sie gern Kontakt mit unseren Berater*innen auf. Wir freuen uns auf Ihr Projekt.

www.dmi.de/leistungen

magrathea

Jobs
Praktika
Blöde Ideen

www.magrathea.eu