

Künstliche Intelligenz in der Neurologie

Anwendungen, Algorithmen, Anforderungen



Tim L. T. Wiegand^{1,2,3}, Laura I. Velezmore^{3,4}, Leonard B. Jung^{1,2,3}, Felix Wimbauer^{3,5}, Konstantinos Dimitriadis⁶ †, Inga K. Koerte^{1,2,7,8} †

- 1 cBRAIN, Klinik und Poliklinik für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie, LMU Klinikum, Ludwig-Maximilians-Universität München
- 2 Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston
- 3 LMU AIM, Ludwig-Maximilians-Universität München
- 4 Klinik und Poliklinik für Strahlentherapie und Radioonkologie, LMU Klinikum, Ludwig-Maximilians-Universität München
- 5 Computer Vision Group, School of Computation, Information and Technology, Technische Universität München
- 6 Neurologische Klinik und Poliklinik, LMU Klinikum, Ludwig-Maximilians-Universität München
- 7 Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München
- 8 Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston

† geteilte Letztautorenschaft

ZUSAMMENFASSUNG

Künstliche Intelligenz (KI) kann große Datenmengen verarbeiten und interpretieren und birgt daher großes Potenzial für die Neurologie. Viele neurologische Erkrankungen erfordern umfangreiche multimodale Diagnostik und personalisierte Therapiekonzepte, die durch KI automatisiert bzw. optimiert werden können. Hinter der KI verbergen sich komplexe Algorithmen. Ziel dieser Arbeit ist es, einen Überblick über Grundbegriffe, Algorithmen und Anwendungen der KI in der Neurologie zu vermitteln.

Es existiert eine Vielzahl an KI-Modellen. Besonders leistungsfähig sind neuronale Netze und Transformer. Neuronale Netze können durch die Verarbeitung statischer Eingabedaten, z. B. von radiologischen Bildern, Verdachtsdiagnosen oder Prognoseeinschätzungen abgeben. Ihre Funktionsweise basiert auf einer großen Anzahl kleiner Recheneinheiten, den Neuronen, die zu Netzen angeordnet sind. Transformer hingegen können sequenzielle Daten wie Text verarbeiten und somit, z. B. basierend auf Anamnesegesprächen, Verdachtsdiagnosen ausgeben oder Befundberichte formulieren. Sie berechnen hierzu die Beziehung der einzelnen Textsegmente zueinander, sodass diese bei der Verarbeitung berücksichtigt werden können. Ein grundlegendes Verständnis der Anwendungen und Funktionsweisen der KI, aber auch der Herausforderungen und Limitationen ist für den erfolgreichen Einsatz in der neurologischen Forschung und Praxis entscheidend.

Grundlagen

Durch ihre Fähigkeit, große Datenmengen zu verarbeiten und zu interpretieren, hat die Künstliche Intelligenz (KI) in fast alle Bereiche der Medizin Einzug gehalten. Auch für die neurologische Gesundheitsversorgung wird der KI ein großes Potenzial zugeschrieben [1]. Viele neurologische Erkrankungen erfordern umfangreiche multimodale Diagnostik und personalisierte Therapiekonzepte, die durch KI automatisiert bzw. optimiert werden können. Ein grundlegendes Verständnis der Anwendungen und Funktionsweisen der KI, aber auch der Herausforderungen und Limitationen ist für den erfolgreichen Einsatz in der neurologischen Forschung und Praxis entscheidend.

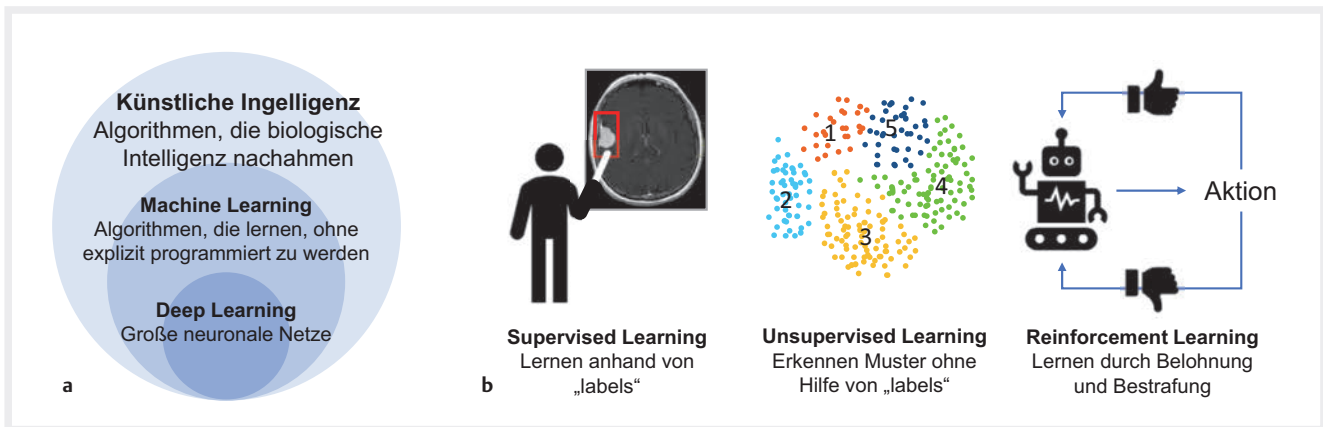
In diesem Übersichtsartikel geben wir zunächst eine allgemeine Einführung in die KI und definieren wichtige Grundbegriffe zu den Teilbereichen der KI und den zugrunde liegenden Lernprozessen. Anschließend wird erläutert, welche Daten sich für die Verarbeitung durch KI-Modelle eignen und wie sie in die Modelle eingespeist werden

können. Es folgen ausführlichere Erklärungen der Anwendungen von Funktionsweisen von neuronalen Netzen und Transformern als 2 besonders leistungsfähigen KI-Modellen. Abschließend werden einige Herausforderungen der Nutzung von KI in der Neurologie kurz beschrieben.

Künstliche Intelligenz und ihre Teilbereiche

Neben dem Begriff „Künstliche Intelligenz“ haben sich weitere Begriffe etabliert, die z. T. synonym für KI verwendet werden, jedoch Teilbereiche der KI darstellen (► **Abb. 1a**). Da sich die englische Terminologie in vielen Fällen auch im deutschen Sprachgebrauch durchgesetzt hat, werden wir auch in diesem Artikel häufig die etablierte englischsprachige Nomenklatur verwenden und lediglich auf die deutschen Übersetzungen verweisen.

- Künstliche Intelligenz: Unter diesem Begriff werden gemeinhin all jene Algorithmen subsummiert, die biologischer Intelligenz ähneln bzw. diese nachahmen. KI umfasst damit eine große Anzahl an Algorithmen von simpler Logik, über Regressionen, bis hin zu



► **Abb. 1** a Abgrenzung der Begriffe „künstliche Intelligenz“, „machine learning“ und „deep learning“. b Abgrenzung der Begriffe „supervised learning“, „unsupervised learning“ und „reinforcement learning“.

komplexen neuronalen Netzen. Die KI weist große Überschneidungen mit dem Feld der Statistik auf, die ebenfalls zur Entdeckung von Zusammenhängen und dem Treffen von Vorhersagen genutzt wird.

- Machine learning (maschinelles Lernen): Hierbei handelt es sich um einen Teilbereich der KI, der Algorithmen umfasst, die in der Lage sind, anhand von Daten selbstständig zu lernen und sich anzupassen. Das bedeutet, dem Algorithmus müssen keine ausdrücklichen Anweisungen gegeben werden, wie die gegebenen Daten verarbeitet werden sollen. Machine-learning-Algorithmen sind meist etwas komplexere KI-Algorithmen wie z. B. support vector machines, Random-forest-Modelle oder neuronale Netze. Sie lernen bspw. medizinische Eingabedaten wie radiologische Bilder zu verarbeiten und eine Diagnose auszugeben.
- Deep learning (tiefes Lernen): Hierbei handelt es sich wiederum um einen Teilbereich des machine learnings, der umfangreiche („tiefe“) neuronale Netze umfasst. Sie können besonders komplexe Aufgaben lösen und weisen z. B. eine höhere Präzision bei der Befundung radiologischer Bilder auf. Sie gehören damit zu den zukunftssträchtesten KI-Algorithmen.

Lernprozesse der KI

Es lassen sich verschiedene Lernprozesse der Algorithmen differenzieren. Sie beeinflussen wiederum, für welche neurologischen Anwendungsbereiche sich Algorithmen eignen (► **Abb. 1b**).

- Supervised learning (überwachtes Lernen): Hierbei handelt es sich um Algorithmen, die mittels exakt beschriebener Eingabedaten lernen, um diese Beschreibungen anschließend selbstständig generieren zu können. Die Beschreibungen der Eingabedaten werden „label“ genannt. Soll ein KI-Algorithmus aus dem Bereich supervised learning bspw. maligne Raumforderungen in Schädel-Magnetresonanztomografie (MRT)-Bildern erkennen, wird der Algorithmus anhand

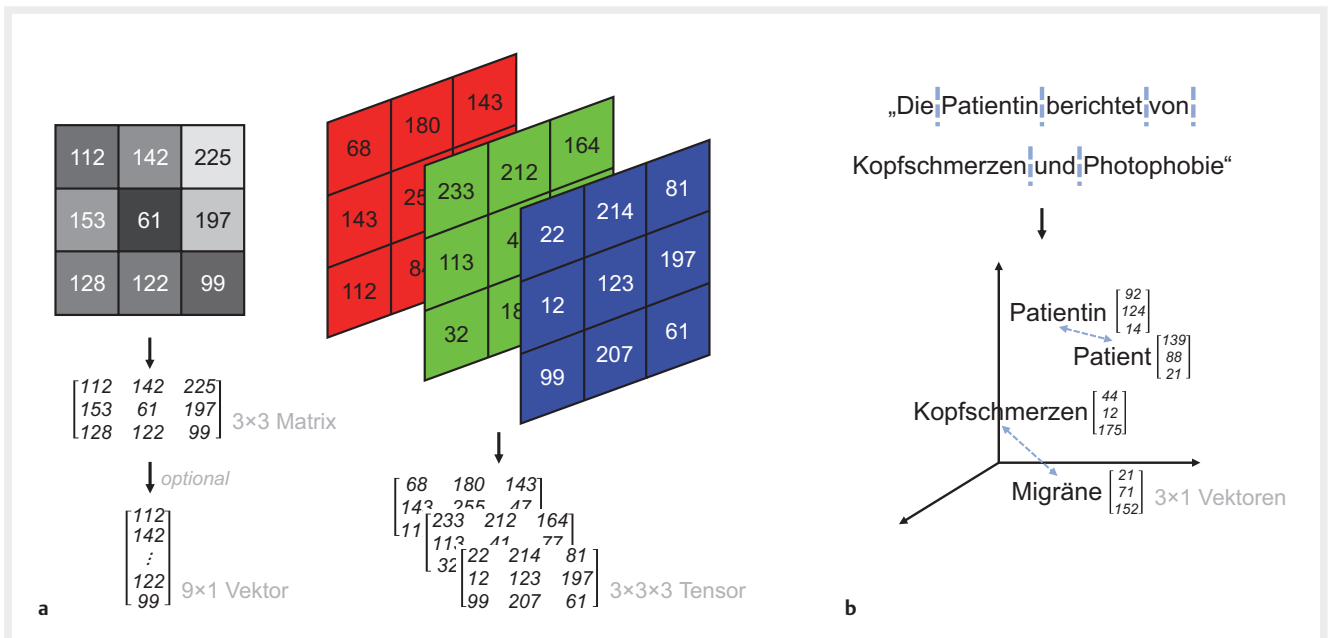
befundeter Bilddaten trainiert. Diese Trainingsdaten bestehen aus den MRT-Bildern sowie den labels, also den Beschreibungen, ob (und ggf. auch wo) jeweils eine maligne Raumforderung vorliegt oder nicht. Der Algorithmus lernt anhand dieser Beschreibungen zwischen MRT-Bildern mit und ohne Malignomen zu unterscheiden [2]. Die Qualität der label ist hierbei entscheidend, da der Algorithmus anhand dieser lernt.

- Unsupervised learning (unüberwachtes Lernen): Es umfasst Algorithmen, die ohne label lernen. Solche Algorithmen erkennen selbst Muster und Zusammenhänge in den Daten. Ein typisches Beispiel sind Cluster-Algorithmen, die Gruppen von ähnlichen Datenpunkten identifizieren und so bspw. anhand klinischer Merkmale oder genetischer Marker Subgruppen neuropsychiatrischer oder -genetischer Erkrankungen identifizieren [3, 4].
- Reinforcement learning (verstärkendes Lernen): Hierbei erlernt der Algorithmus Strategien, die Belohnungen nach sich ziehen bzw. Bestrafung vermeiden. Das Training erfolgt ebenfalls nicht anhand gelabelter Daten, sondern durch das Ausprobieren verschiedener Handlungsmöglichkeiten (z. B. Bewegungen der Gelenke eines chirurgischen Roboterarms), um positive Ergebnisse zu erzielen (z. B. eine chirurgische Prozedur zu meistern). Reinforcement learning ist besonders geeignet für die Verarbeitung von „Real-time“-Daten aus kontinuierlichen Messungen und Sensoren. Neurologische Anwendungen für reinforcement learning bestehen somit in Robotern wie bspw. für die (Neuro-)Chirurgie [5] oder in klinischen decision support systems wie etwa in der neurologischen Intensivmedizin [6, 7].

Daten

Statische und sequenzielle Daten

Die Wahl des geeigneten KI-Modells hängt maßgeblich davon ab, ob einzelne statische oder sequenzielle Daten



► **Abb. 2** **a** Einzelnes Schwarz-weiß-Bild (links) sowie Farbbild (rechts) als Beispiele statischer Eingabedaten. Die Zahlenwerte jedes Pixels können als Vektoren, Matrizen bzw. Tensoren dargestellt und verarbeitet werden. **b** Textabschnitt als Beispiel sequenzieller Eingabedaten. Der Text kann in tokens zerlegt werden. Diese können als Vektoren dargestellt werden, die zusätzlich die semantische Nähe der tokens zueinander abbilden (embeddings).

verarbeitet werden sollen. Statische Daten weisen keine zeitliche Abfolge auf und können somit einzeln in einen Algorithmus eingespeist werden. Beispiele für statische Daten sind klinische Scores, Laborwerte und radiologische oder histologische Bilder. Sequenzielle Daten weisen hingegen eine zeitliche Abfolge auf. Beispiele sind Texte, Sprachaufnahmen oder Elektroenzephalografie (EEG)-Aufzeichnungen.

Dateiformate und Einschleusen in Algorithmen

Für jede Art Daten existieren meist mehrere Dateiformate. Ein Bild kann z. B. als JPEG oder PNG gespeichert sein. In der Radiologie üblich ist das Dateiformat DICOM (Digital Imaging and Communications in Medicine). Je nachdem, ob etablierte Standards existieren oder Gerätehersteller eigene Dateiformate verwenden, sind unterschiedliche Verarbeitungsschritte notwendig, um die Daten in einen KI-Algorithmus einzuspeisen. In der klinischen Praxis übliche Dateiformate wie DICOM enthalten meist neben der eigentlichen Kerndatei (bspw. dem Bild) zusätzliche Informationen zur untersuchten Person, dem Aufnahmezeitpunkt, technischen Spezifikationen u. a. Häufig wird jedoch nur das eigentliche Bild in ein KI-Modell eingespeist. Im Kern handelt es sich bei Bildern um Tabellen mit Zahlenwerten, die die Farbwerte jedes Pixels im Bild repräsentieren. Schwarz-weiß-Bilder stellen 2-dimensionale Tabellen bzw. Matrizen dar, wobei die Zahlenwerte den Graustufen der Pixel entsprechen (► **Abb. 2a** links). Für Farbbilder existieren verschiedene Farbräume. Im RGB-(Rot-Grün-Blau)-Farbraum bestehen Bilder aus 3-dimensionalen Ta-

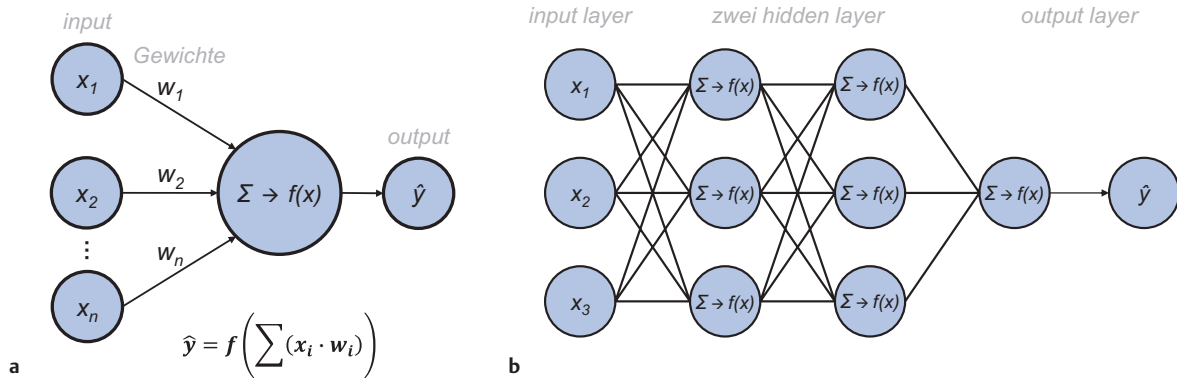
bellenn bzw. Tensoren mit 3 Schichten für jeden Farbanteil (► **Abb. 2a** rechts). Auf diese Weise lassen sich statische Dateien wie Bilder in Form von Zahlenwerten in KI-Algorithmen einspeisen.

Sequenzielle Daten müssen für viele KI-Algorithmen in vordefinierte Zeitfenster oder Segmente aufgeteilt werden. Ein Text wird dabei in einzelne Einheiten aufgeteilt, die als tokens bezeichnet werden. Tokens können Wörter, Zeichen oder Wortbestandteile sein. Ihnen werden spezifische Zahlenwerte zugewiesen, die ihre semantische Nähe zueinander abbilden (► **Abb. 2b**). Für diese Zahlenrepräsentation der tokens werden oft eigene Algorithmen verwendet. Das Ergebnis wird als embedding bezeichnet. Die embeddings stellen oft 1-dimensionale Tabellen bzw. Vektoren dar. Die Zahlendarstellungen können nun ebenfalls in KI-Algorithmen eingespeist werden.

Trainings-, Validierungs- und Testdaten

Um ein leistungsstarkes KI-Modell zu entwickeln, werden je nach Art des Modells und Komplexität der Anwendung häufig große Mengen an Daten benötigt. Der verfügbare Datensatz wird für viele Anwendungen in 3 Teile unterteilt.

- **Trainingsdaten:** Sie werden verwendet, um ein KI-Modell zu trainieren. Dies bedeutet, dass das mathematische Modell anhand der Trainingsdaten zunehmend verfeinert wird, bis es zu den Trainingsdaten passt. Mittels der Trainingsdaten passt ein Regressions-Algorithmus bspw. die Regressionsgerade so an, dass sie optimal durch die Punktwolke verläuft, also der Abstand zwischen Regressionsgerade und Daten-



► **Abb. 3** **a** Aufbau eines einzelnen künstlichen Neurons. Der input x wird zunächst gewichtet, d. h. mit den Gewichten w multipliziert. Die gewichteten Eingaben werden summiert (Σ) und in die Aktivierungsfunktion $f(x)$ eingesetzt (vergleichbar mit einem Aktionspotenzial). Hierdurch entsteht der output \hat{y} , der wiederum den input für nachfolgende Neurone darstellt. **b** Ein beispielhaftes neuronales Netz mit 2 hidden layer mit jeweils 3 Neuronen.

punkten minimal ist. Meist wird der größte Teil des Datensatzes, bspw. 70%, für das Training verwendet.

- Validierungsdaten: Sie dienen dazu, die Leistung verschiedener trainierter Modelle zu vergleichen. So können während des Entwicklungsprozesses bspw. mehrere Varianten neuronaler Netze, die alle mit denselben Trainingsdaten trainiert wurden, verglichen werden und das Modell mit der besten Leistung ausgewählt werden. Häufig werden ca. 10–20% des Datensatzes hierzu verwendet.
- Testdaten: Der verbleibende Teil des Datensatzes wird für die finale Bewertung der Leistung des Modells genutzt. Hierzu müssen Daten verwendet werden, die der Algorithmus nicht kennt, die also weder für Training noch Validierung verwendet wurden. Der Testdatensatz soll einen Eindruck der Modell-Leistung in der Realität vermitteln. Hierzu werden ca. 10–20% der Daten verwendet.

Modelle

Es existiert eine große Anzahl verschiedener KI-Algorithmen. Sie unterscheiden sich u. a. darin, ob sie statische oder sequenzielle Eingabedaten akzeptieren und ob sie statische oder sequenzielle Ausgaben produzieren. Im Folgenden werden die KI-Algorithmen neuronale Netze sowie Transformer als 2 moderne und besonders leistungsstarke Algorithmen näher beschrieben.

Neuronale Netze

Neuronale Netze eignen sich für die Verarbeitung statischer Eingabedaten, z. B. demografische Daten, klinische Befunde, medizinische Bilder und Laborwerte, und können basierend hierauf statische Ausgaben wie Verdachtsdiagnosen oder Prognoseeinschätzungen generieren [8]. Insbesondere große neuronale Netze aus dem Bereich deep

learning können sehr leistungsstark sein und sich für neurologische Fragestellungen eignen. Bei der Analyse neuroradiologischer Aufnahmen liefern neuronale Netze hilfreiche Informationen, z. B. Angaben zur 3-dimensionalen Ausdehnung eines Tumors oder Hinweise auf auffällige Areale in einem Bild [9]. Auch genetische Daten können von neuronalen Netzen verarbeitet werden, um etwa Risikofaktoren für neurologische Erkrankungen zu identifizieren oder Vorhersagen zur Krankheitsentwicklung zu treffen [10, 11].

Der Aufbau und die Funktionsweise künstlicher neuronaler Netze orientiert sich an biologischen neuronalen Netzen im Gehirn. Künstliche Neurone sind die funktionellen Einheiten, die zu großen Netzwerken verknüpft sind. Ein künstliches Neuron erhält mehrere Eingabewerte x (► **Abb. 3a**). Diese Eingabewerte repräsentieren Informationen oder Signale, die von anderen Neuronen oder externen Quellen stammen. Jeder Eingabewert weist ein Gewicht w auf, das seine Bedeutung oder Stärke in Bezug auf das Ergebnis des Neurons angibt. Die Eingabewerte x werden mit ihren Gewichten w multipliziert. Die Ergebnisse werden zu einem Zahlenwert addiert, der in eine Aktivierungsfunktion eingesetzt wird. Sie berechnet, wie stark ein nachfolgendes Neuron aktiviert wird. Es gibt eine Vielzahl unterschiedlicher Aktivierungsfunktionen. Sie ermöglichen es, komplexe, nicht lineare Zusammenhänge zu modellieren.

Mehrere Neurone können zu einem neuronalen Netz angeordnet werden (► **Abb. 3b**), das grundsätzlich beliebig viele Schichten und Neurone pro Schicht aufweisen kann. Die erste Schicht wird als input layer (Eingabeschicht) bezeichnet. Sie erhält die Eingabedaten (z. B. ein Röntgenbild). Es folgen die hidden layer (versteckte Schichten), in denen die Eingabe mit Gewichten und Aktivierungsfunk-

tionen weiterverarbeitet wird. Am Ende des Netzes folgt das output layer (Ausgabeschicht), das die finale Ausgabe (z. B. die Verdachtsdiagnose oder Klassifikation „gesund“ vs. „krank“) ausgibt.

Mithilfe der linearen Algebra können die beschriebenen Rechenschritte kompakt gebündelt durchgeführt werden. Die Eingabe entspricht bspw. einem Vektor, der die Zahlenwerte jedes Pixels des Röntgenbilds enthält (► **Abb. 2a** links). Die Gewichte können in Matrizen angeordnet werden. Sie werden initial bspw. zufällig festgelegt und während des Trainingsprozesses optimiert. Eingabe-Vektoren und Gewichts-Matrizen können multipliziert und die Ergebnisse in die Aktivierungsfunktionen eingesetzt werden. Diese Berechnungen lassen sich mit jeder zusätzlichen Schicht beliebig oft wiederholen. Zusätzliche Schichten erlauben komplexere Modellierungen und ggf. eine bessere Modell-Leistung, erfordern jedoch auch mehr Rechenleistung. Die finale Ausgabe entspricht ebenfalls einem Vektor. Soll das Netz zwischen einem pathologischen und einem unauffälligen Röntgenbild unterscheiden, kann die Klassifikation „krank“ z. B. durch einen Zahlenwert nahe 1 und „gesund“ durch einen Zahlenwert nahe 0 repräsentiert werden. Dieser gesamte Rechenprozess von Eingabe bis Ausgabe wird als forward pass bezeichnet.

Der Trainingsprozess eines neuronalen Netzes wird analog auch backward pass genannt. Hierzu wird zunächst eine Fehlerfunktion berechnet. Sie beschreibt den Fehler des Modells, also den Unterschied der aktuellen Ausgabe des Modells zur korrekten Ausgabe. Die korrekte Ausgabe entspricht dem label der Eingabedaten, also bspw. dem Zahlenwert 0, wenn es sich bei dem Trainings-Röntgenbild um ein unauffälliges Bild handelt. Wird die Fehlerfunktion minimiert, sinkt der Unterschied zwischen der Modell-Ausgabe und der korrekten Klassifikation. Die einzigen Variablen, die hierzu angepasst werden können, sind die Gewichte (Eingabe und Aktivierungsfunktionen sind vorbestimmt). Der backward pass nutzt demnach ein mathematisches Verfahren, das die Gewichte des Netzes schrittweise so anpasst, dass die Fehlerfunktion minimiert wird. Dieses Verfahren heißt gradient descent (Gradientenabstieg).

Eine spezielle Form der neuronalen Netze, die besonders für die Verarbeitung von Bildern geeignet ist, sind convolutional neural networks (CNNs) [12]. Sie verarbeiten mehrere benachbarte Pixel auf einmal. Hierdurch können Zusammenhänge zwischen benachbarten Pixeln berücksichtigt und Muster besser erkannt werden.

Transformer

Transformer eignen sich für die Verarbeitung sequenzieller Daten wie Text. Sie wurden 2017 erstbeschrieben [13] und bilden den Kern von Software wie ChatGPT (Chat Generative Pretrained Transformer) [14] der Firma OpenAI oder BERT (Bidirectional Encoder Representations from Trans-

formers) [15] und PaLM (Pathways Language Model) [16] der Firma Google.

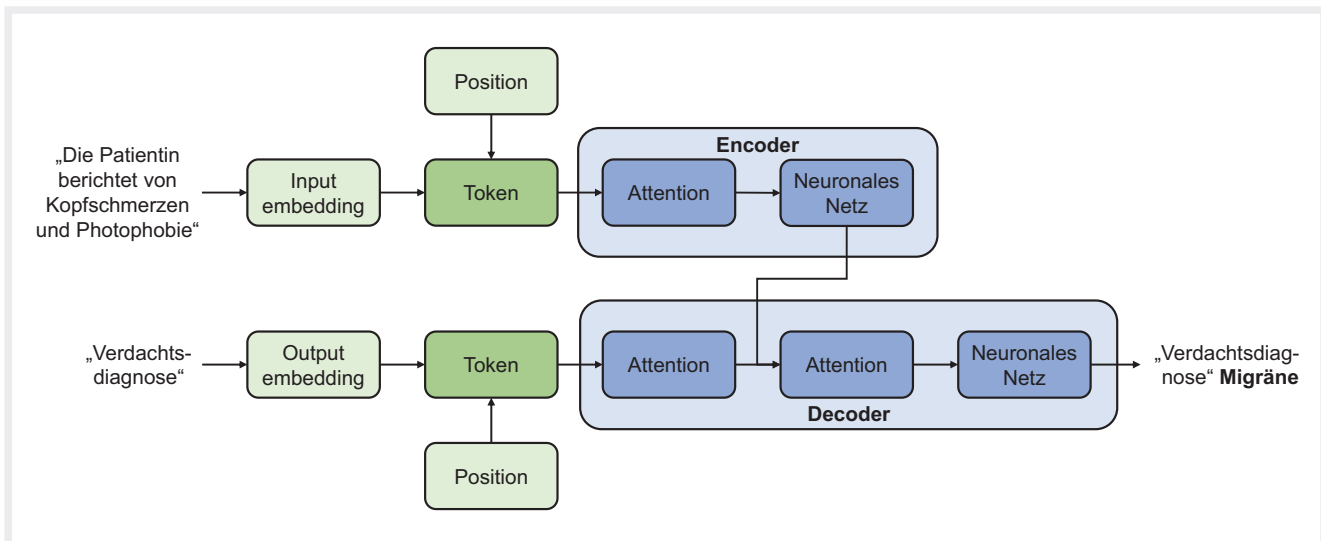
Eine Vielzahl medizinischer Anwendungen von Transformern scheint möglich. Jüngste Veröffentlichungen zeigen, dass ChatGPT die deutschen [17] und US-amerikanischen [18] medizinischen Staatsexamina bestehen und somit komplexe medizinische Fragestellungen beantworten kann. Transformerbasierte Sprachmodelle können neue medizinische Datenpunkte vor dem Hintergrund riesiger Mengen verfügbarer Literatur interpretieren. Somit können sie potenziell die Diagnosestellung unterstützen oder individualisierte Therapieempfehlungen abgeben [19]. Auch Texte können von Transformern verarbeitet und zusammengefasst werden, um z. B. bei Neuaufnahmen einen schnellen Überblick über die bisherige Krankengeschichte zu vermitteln [20]. Transformer können zudem strukturierte und eloquente Texte formulieren und damit u. a. die zeitaufwendige medizinische Dokumentation vereinfachen [21, 22].

Eine erste Besonderheit der Funktionsweise der Transformer ist die Art der Texteingabe. Während ältere Algorithmen wie etwa recurrent neural networks (RNNs, rekurrente neuronale Netze) lediglich ein token nach dem anderen verarbeiten können, verarbeiten Transformer die gesamte Texteingabe auf einmal. Der Textabschnitt wird in tokens unterteilt und diese zu input embeddings weiterverarbeitet (► **Abb. 2b**). Hinzu kommen Informationen zur Position der tokens im Eingabetext. Diese optimierten Vektor-Repräsentationen der tokens werden nun parallel in den ersten Abschnitt des Transformers eingespeist, den encoder (Kodierer; ► **Abb. 4**). Die Aufgabe des encoders besteht darin, den Eingabetext zu „verstehen“, bevor im decoder (Dekodierer) eine Ausgabe produziert wird. In der Praxis werden mehrere encoder- und mehrere decoder-Blöcke hintereinandergeschaltet, um komplexere Berechnungen zu ermöglichen.

Die zweite Besonderheit und der Kern der Transformer sind die Self-attention-Module [13]. Sie verwenden eine mathematische Formel, die für jeden token die Beziehung zu allen einzelnen token der Textsequenz (allen anderen und sich selbst) berechnet. Diese Beziehungen werden auch als attention bezeichnet und als Gewicht ausgedrückt. Ein höheres attention-Gewicht beschreibt eine stärkere Beziehung eines tokens zu einem anderen, z. B. bezieht sich der Artikel „Die“ stark auf „Patientin“ und wenig auf die anderen tokens (► **Abb. 4**). Die Funktionsweise der attention-Module ist komplex. Für besonders Interessierte folgt nun eine ausführlichere Erläuterung.

$$\text{attention} = \text{softmax} \left(\frac{\text{query} \cdot \text{key}^T}{\sqrt{\text{dim}_{\text{key}}}} \right) \cdot \text{value}$$

dim = Dimensionalität des key-Vektor



► **Abb. 4** Vereinfachte Darstellung einer Transformer-Architektur mit einem encoder- und einem decoder-Block. Das Modell erhält als Eingabetext eine kurze Fallvignette und erstellt schrittweise einen Ausgabertext mit einer Verdachtsdiagnose. Der Eingabetext wird zunächst in tokens umgewandelt, d. h. in Wortbestandteile unterteilt, die als Zahl bzw. Vektor ausgedrückt werden und Informationen zur Position der Wortbestandteile im Satz enthalten. In den attention-Blöcken wird die Beziehung der tokens zueinander berechnet. Diese optimierten Zahlenrepräsentationen von Wortbestandteilen können genutzt werden, um mittels neuronalen Netzen das nächste möglichst sinnvolle Wort auszugeben. Die Abbildung zeigt den Zustand des Modells, in dem es im vergangenen Zyklus das Wort „Verdachtsdiagnose“ ausgegeben hat und im jetzigen Zyklus das Wort „Migräne“ ausgibt.

Die Variablen-Bezeichnungen der Formel sind der Suchmaschinensprache entlehnt. Hierbei bezeichnet „query“ eine Suchmaschinenanfrage (z. B. nach einem Bild in einer Datenbank). „keys“ sind Kodierungen der einzelnen Elemente der Datenbank (vergleichbar mit den Namen oder Kurzbeschreibungen der Bilddateien). „Value“ sind die Elemente selbst (z. B. die Bilder). Es soll das Element (value) ausgegeben werden, für das die Anfrage (query) am besten zur Kodierung (key) passt. Hierzu werden die Anfrage (query) und alle Kodierungen (keys) hinsichtlich ihrer Ähnlichkeit verglichen. Je ähnlicher die query und ein bestimmter key sind, desto stärker wird das Element (value), das zu dem key gehört, gewichtet, und desto eher wird es letztlich ausgegeben (also z. B. ein spezifisches Bild nach einer Datenbankanfrage).

Im Kontext der Transformer bezeichnet „value“ das Textelement bzw. den token, für den die attention-Gewichte bzw. Beziehungen zu allen einzelnen Textelementen bzw. tokens berechnet werden soll. Die Gewichte berechnen sich durch den Vergleich der query- und key-Vektoren. Mathematisch gelingt dies über ein Skalarprodukt aus query und keys, also einer Multiplikation der Vektoren. „Query“ ist ein Anfrage-Vektor, der der Kodierung des keys eines values ähneln kann/soll. „Keys“ sind Kodierungen in Form von Vektoren für jeden einzelnen token. Query und die keys werden meist zufällig initialisiert und während des Trainings des Modells optimiert, um eine bestmögliche Gewichtung der values zu bewerkstelligen. Die Gewichte werden anschließend skaliert und auf Zahlenwerte zwischen 0 und 1 normalisiert. Abschließend wird jedes Gewicht mit den value-Vektoren multipliziert.

Die Ausgabe eines attention-Moduls sind nun noch weiter verbesserte Vektordarstellungen der tokens, die neben den initialen input embeddings und Positionskodierungen auch Informationen zur Beziehung zu anderen Wörtern im Text enthalten. Auf diese Weise werden Zusammenhänge zwischen selbst weit entfernten Textabschnitten berücksichtigt. Der attention-Mechanismus wird parallel für unterschiedliche Aspekte und Arten der Beziehungen durchgeführt, um unterschiedliche Facetten zu berücksichtigen. Im encoder durchwandern die token-Vektoren zunächst ein attention-Modul. Die resultierenden attention-Vektoren werden anschließend in einem klassischen neuronalen Netz-Modul weiter optimiert und danach in den decoder eingeschleust.

Transformer geben Schritt für Schritt weitere Wörter aus, bis ein ganzer Ausgabertext entsteht. In jedem Zyklus erhält der decoder hierzu die zuletzt erzeugte Modellausgabe als Eingabe. In unserem Beispiel entspräche dies dem zuletzt ausgegebenen Wort („Verdachtsdiagnose“; ► **Abb. 4**). Wie beim encoder, wird diese letzte Modellausgabe zunächst in embeddings und anschließend in einem attention-Modul in attention-Vektoren umgewandelt. Im nun folgenden attention-Modul laufen die attention-Vektoren des zuletzt ausgegebenen Textabschnitts sowie der initialen Texteingabe zusammen. Die Ausgabe dieses attention-Moduls sind schließlich attention-Vektoren, die die Beziehung jedes Worts der initialen Eingabe zur letzten Ausgabe darstellen. Es folgt ein weiteres neuronales Netz-Modul, das die Umwandlung der Vektoren in Text vorbereitet. Die finale Ausgabe des Modells eines Zyklus ist vereinfacht beschrieben eine Wahrscheinlichkeitsverteilung über alle Wörter im Wörterbuch. Das Wort mit der

höchsten Wahrscheinlichkeit wird ausgegeben (in unserem Fall „Migräne“). Auf diese Weise gibt der Transformer mit jedem Zyklus ein weiteres Wort aus, bis ein Ausgabertext als Antwort auf den Eingabetext entstanden ist.

Anforderungen für die KI in der Neurologie

Die Anwendungsmöglichkeiten von KI in der Neurologie sind breit. Es existieren leistungsstarke Modelle, die nahezu jede Form von Daten verarbeiten und auswerten können. Jedoch bestehen weiterhin einige Herausforderungen für die Nutzung von KI in Klinik und Forschung:

- **Menge und Qualität der Daten:** Für viele Anwendungen sind große Mengen Trainingsdaten notwendig mit Hunderten bis Tausenden Datenpunkten. Gerade für seltenere Erkrankungen ist dies oft nicht zu bewerkstelligen. Zudem ist die Qualität der Trainingsdaten entscheidend. Sie sollten eine breite Palette von Szenarien abdecken und repräsentativ für die Eingabedaten der realen Welt sein. Unterschiede hinsichtlich Alter, Ethnizität, Krankheitsmanifestation etc. können zu einer reduzierten Leistungsfähigkeit der Algorithmen in der Praxis führen, also z. B. zu fehlerhaften Klassifikationen radiologischer Bilder. Die Daten sollten frei von Artefakten, Ausreißern und Rauschen sein. Beim supervised learning müssen zudem die label zuverlässig sein, also die Beschreibungen der Daten anhand derer der Algorithmus lernt.
- **Infrastruktur:** Die notwendige technische Infrastruktur aus Software zur strukturierten Erfassung von Daten, Servern zur Speicherung der Daten, Rechenleistung zur Verarbeitung der Daten sowie entsprechenden Verknüpfungen existieren vielerorts noch nicht. Hinzu kommen Anforderungen an die Expertise der Anwender zur erfolgreichen Nutzung von KI und insbesondere der Interpretation der Ausgaben.
- **Erklärbarkeit:** Zwar können KI-Modelle „intelligente“ Ausgaben wie Verdachtsdiagnosen und Prognoseabschätzungen liefern, diese Ausgaben jedoch oft nicht näher erklären oder begründen. Ein wichtiges Forschungsfeld ist daher die explainable AI (xAI, erklärbare KI). Sie entwickelt Methoden, um die Ausgaben der KI-Algorithmen besser zu erklären, z. B. indem zusätzlich zur reinen Bildklassifikation die vermeintliche Auffälligkeit farbig markiert wird. Somit kann die Ausgabe bzw. Entscheidung des Modells besser nachvollzogen werden.
- **Ethische und juristische Anforderungen:** Der Einsatz von KI in der Medizin erfordert sorgfältige ethische und juristische Abwägung, z. B. hinsichtlich Datenschutz, Sicherheit und der sachgemäßen Anwendung von KI-Algorithmen, oder der Haftung bei Fehlern. Die Zulassungsverfahren sind daher oft langwierig und kostenintensiv. Haftungsfragen sind in vielen Fällen nicht hinreichend geklärt [23].

FAZIT FÜR DIE PRAXIS

KI weist großes Potenzial in der Neurologie auf, da sie sowohl Forschungsanwendungen als auch eine Vielzahl diagnostischer und therapeutischer Prozesse unterstützen bzw. automatisieren kann. Die Art der Eingabedaten und gewünschten Ausgaben beeinflusst die Auswahl des KI-Modells. Neuronale Netze eignen sich besonders für die Verarbeitung statischer Daten, Transformer für die Verarbeitung sequenzieller Daten. Die Anforderungen an die Datenmenge und -qualität, die IT-Infrastruktur, die Erklärbarkeit der Ergebnisse sowie die ethischen und rechtlichen Bestimmungen sind hoch und erfordern eine kontinuierliche Weiterentwicklung. Bis das volle Potenzial der KI in der Neurologie ausgeschöpft werden kann, wird es deshalb noch einige Zeit dauern.

GLOSSAR

- **Algorithmus:** Handlungsvorschrift bzw. Sammlung von Handlungsvorschriften zur schrittweisen (mathematischen) Lösung eines Problems
- **Attention:** Methode, um die Beziehung zwischen verschiedenen Teilen sequenzieller Eingabedaten zu berücksichtigen
- **Deep learning (tiefes Lernen):** Teilbereich der künstlichen Intelligenz und des machine learning mit umfangreichen neuronalen Netzen, die besonders komplexe Aufgaben lösen können
- **Embedding:** Zahlenrepräsentation von tokens, die zusätzliche Informationen wie z. B. zur semantischen Nähe der tokens enthalten kann
- **Gradient descent (Gradientenabstiegsverfahren):** Methode, die genutzt wird, um Algorithmen aus dem Bereich machine learning, z. B. neuronale Netze zu trainieren, indem der Fehler des Modells als Funktion ausgedrückt wird, die schrittweise minimiert wird
- **Künstliche Intelligenz:** Algorithmen, die biologischer Intelligenz ähneln bzw. diese nachahmen
- **Künstliche neuronale Netze:** Algorithmen aus den Bereichen künstliche Intelligenz bzw. machine learning, die biologischen neuronalen Netzen ähneln und zur Verarbeitung von Informationen und dem Lernen aus Daten verwendet werden
- **Label:** Exakte Beschreibungen bzw. „Beschriftungen“ der Daten, meist durch Menschen
- **Machine learning (machinelles Lernen):** Teilbereich der künstlichen Intelligenz mit Algorithmen, die anhand von Daten selbstständig lernen und sich anpassen, ohne dass explizit vorgegeben wird, wie die Daten verarbeitet werden sollen

- Modell: Repräsentation bzw. Abbildung der Realität, z.B. mithilfe einer mathematischen Formel
- Reinforcement learning (verstärkendes Lernen): Lernen durch Belohnung und Bestrafung
- Supervised learning (überwachtes Lernen): Lernen anhand exakt beschriebener Eingabedaten (labels), um diese Beschreibungen anschließend selbstständig generieren zu können
- Testdaten: Anteil der verfügbaren Daten, die genutzt werden, um die Leistung eines Modells abschließend zu beurteilen
- Token: Einteilung sequenzieller Daten wie Text in diskrete (endliche) Einheiten wie Wörter oder Wortteile, die meist durch Zahlenwerte bzw. Vektoren repräsentiert werden
- Trainingsdaten: Anteil der verfügbaren Daten, die genutzt werden, um ein Modell zu trainieren und dessen Leistung zu verbessern
- Transformer: Moderne Algorithmen aus dem Bereich künstliche Intelligenz bzw. machine learning, die u. a. einen attention-Mechanismus nutzen, um sequenzielle Daten zu verarbeiten und auszugeben
- Unsupervised learning (unüberwachtes Lernen): Lernen bzw. Mustererkennung ohne exakt beschriebene Eingabedaten (labels)
- Validierungsdaten: Anteil der verfügbaren Daten, die genutzt werden, um die Leistung verschiedener bereits trainierter Modelle zu vergleichen

Interessenkonflikt

Erklärung zu finanziellen Interessen

Forschungsförderung erhalten: nein; Honorar/geldwerten Vorteil für Referententätigkeit erhalten: nein; Bezahler Berater/interner Schulungsreferent/Gehaltsempfänger: nein; Patent/Geschäftsanteile/Aktien (Autor/Partner, Ehepartner, Kinder) an Firma (Nicht-Sponsor der Veranstaltung): ja; Patent/Geschäftsanteile/Aktien (Autor/Partner, Ehepartner, Kinder) an Firma (Sponsor der Veranstaltung): nein.

Erklärung zu nicht finanziellen Interessen

Dr. Koerte ist Professorin an der Ludwig-Maximilians-Universität München (bezahlte Stelle). Sie ist europäische Herausgeberin des Journal of Neurotrauma (unbezahlte Stelle) und Vizepräsidentin der European Neurotrauma Organization (unbezahlte Stelle). Sie erhält Forschungsgelder von den National Institutes of Health, dem Europäischen Forschungsrat und dem deutschen Bundesministerium für Forschung und Bildung. Sie erhält Mittel für eine Forschungsstudie über sportbedingte Gehirnerschütterungen von Abbott Inc. Das Klinikum der Ludwig-Maximilians-Universität erhielt Spenden für ihre Forschung von der Schatt-Stiftung und von Mary Ann Liebert Inc. Sie erhält Tantiemen für Buchkapitel, die von Thieme veröffentlicht werden. Ihr Ehepartner ist Angestellter bei Siemens und sie hält daher Aktienoptionen bei Siemens und Siemens Healthineers.

Wissenschaftlich verantwortlich

Wissenschaftlich verantwortlich gemäß CME-Zertifizierungsbedingungen für diesen Beitrag ist Prof. Dr. med. Inga K. Körte, München.

Korrespondenzadresse

Tim L. T. Wiegand

cBRAIN, Klinik und Poliklinik für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie Ludwig-Maximilians-Universität, St. Vinzenz-Haus Nußbaumstr. 5, 80336 München, Deutschland
tim.wiegand@med.uni-muenchen.de

Literatur

- [1] Patel UK, Anwar A, Saleem S et al. Artificial intelligence as an emerging technology in the current care of neurological disorders. *J Neurol* 2021; 268: 1623–1642. doi:10.1007/s00415-019-09518-3
- [2] Saeedi S, Rezayi S, Keshavarz H et al. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Medical Informatics and Decision Making* 2023; 23: 16. doi:10.1186/s12911-023-02114-6
- [3] Alashwal H, El Halaby M, Crouse JJ et al. The Application of Unsupervised Clustering Methods to Alzheimer's Disease. *Frontiers in Computational Neuroscience* 2019; 13. doi:10.3389/fncom.2019.00031
- [4] Koretsky MJ, Alvarado C, Makarios MB et al. Genetic risk factor clustering within and across neurodegenerative diseases. *Brain : a journal of neurology* 2023. doi:10.1093/brain/awad161. doi:10.1093/brain/awad161
- [5] Barnoy Y, O'Brien M, Wang W et al. Robotic surgery with lean reinforcement learning. *arXiv preprint arXiv:210501006* 2021
- [6] Liu S, See KC, Ngiam KY et al. Reinforcement Learning for Clinical Decision Support in Critical Care: Comprehensive Review. *J Med Internet Res* 2020; 22: e18477. doi:10.2196/18477
- [7] Liu Y, Qiao N, Altinel Y. Reinforcement Learning in Neurocritical and Neurosurgical Care: Principles and Possible Applications. *Computational and Mathematical Methods in Medicine* 2021; 2021: 6657119. doi:10.1155/2021/6657119
- [8] Valliani AA, Ranti D, Oermann EK. Deep Learning and Neurology: A Systematic Review. *Neurol Ther* 2019; 8: 351–365. doi:10.1007/s40120-019-00153-8
- [9] Yao AD, Cheng DL, Pan I et al. Deep Learning in Neuro-radiology: A Systematic Review of Current Algorithms and Approaches for the New Wave of Imaging Technology. *Radiol Artif Intell* 2020; 2: e190026. doi:10.1148/ryai.2020190026
- [10] Alharbi WS, Rashid M. A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics* 2022; 16: 26. doi:10.1186/s40246-022-00396-x
- [11] Gunavathi C, Sivasubramanian K, Keerthika P et al. A review on convolutional neural network based deep learning methods in gene expression data for disease diagnosis. *Materials Today: Proceedings* 2021; 45: 2282–2285. doi.org/10.1016/j.matpr.2020.10.263

- [12] Lecun Y, Bengio Y. Convolutional Networks for Images, Speech, and Time-Series; 1995
- [13] Vaswani A, Shazeer N, Parmar N et al. Attention is all you need. *Advances in neural information processing systems* 2017; 30
- [14] Brown T, Mann B, Ryder N et al. Language models are few-shot learners. *Advances in neural information processing systems* 2020; 33: 1877–1901
- [15] Devlin J, Chang M-W, Lee K et al. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* 2018.
- [16] Chowdhery A, Narang S, Devlin J et al. Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:2204.02311* 2022
- [17] Jung LB, Guder JA, Wiegand TLT et al. ChatGPT besteht schriftliche medizinische Staatsexamina nach Ausschluss der Bildfragen. *Dtsch Arztebl International* 2023; 120: 373–374. doi:10.3238/arztebl.m2023.0113
- [18] Kung TH, Cheatham M, Medenilla A et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health* 2023; 2: e0000198. doi:10.1371/journal.pdig.0000198
- [19] Singhal K, Azizi S, Tu T et al. Large Language Models Encode Clinical Knowledge. *arXiv preprint arXiv:2212.13138* 2022
- [20] Yang X, Chen A, PourNejatian N et al. A large language model for electronic health records. *NPJ digital medicine* 2022; 5: 194. doi:10.1038/s41746-022-00742-2
- [21] Patel SB, Lam K. ChatGPT: the future of discharge summaries? *The Lancet Digital Health* 2023; 5: e107-e108. doi:10.1016/S2589-7500(23)00021-3
- [22] Jeblick K, Schachtner B, Dext J et al. ChatGPT Makes Medicine Easy to Swallow: An Exploratory Case Study on Simplified Radiology Reports; 2022. doi:10.48550/arXiv.2212.14882
- [23] Hedderich DM, Weisstanner C, Van Cauter S et al. Artificial intelligence tools in clinical neuroradiology: essential medico-legal aspects. *Neuroradiology* 2023. doi:10.1007/s00234-023-03152-7. doi:10.1007/s00234-023-03152-7

Bibliografie

Nervenheilkunde 2023; 42: 591–599

DOI 10.1055/a-2050-0768

ISSN 0722-1541

© 2023. Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Punkte sammeln auf CME.thieme.de



Diese Fortbildungseinheit ist in der Regel 12 Monate online für die Teilnahme verfügbar. Unter <https://eref.thieme.de/CXLLM9V> oder über den QR-Code kommen Sie direkt zur Startseite des Wissenstests und zum Artikel. Sie finden dort auch den genauen Einsendeschluss. Sollten Sie Fragen zur Online-Teilnahme haben, finden Sie unter <https://cme.thieme.de/hilfe> eine ausführliche Anleitung.

Wir wünschen viel Erfolg beim Beantworten der Fragen!

VNR 2760512023163515836



Frage 1

Welche Aussage zu machine learning trifft zu?

- A Streng genommen handelt es sich hierbei um ein Synonym für „künstliche Intelligenz“.
- B Deep learning ist ein Teilbereich des machine learning.
- C Es umfasst Algorithmen wie simple Logik.
- D Neuronale Netze zählen nicht dazu.
- E Jeder Verarbeitungsschritt muss mathematisch exakt definiert werden.

Frage 2

Welche Zuordnung aus Anwendungsbeispiel und Lernprozess der KI passt am besten?

- A Pflegeroboter mit mehreren Laser-, Infrarot-, Kamera-Sensoren – unsupervised learning
- B Klassifikation von Schädel-CT-Aufnahmen ohne label – reinforcement learning
- C Klinisches decision support system auf der neurologischen Intensivstation – supervised learning
- D Cluster-Analysen basierend auf einigen genetischen Markern – unsupervised learning
- E Klassifikation von EEG in „gesund“ und „pathologisch“ anhand beschriebener Trainingsdaten – reinforcement learning

Frage 3

Welche Aussage trifft zu? Eine Matrix ...

- A ... entspricht einer 2-dimensionalen Tabelle.
- B ... ist nicht geeignet, um weitere mathematische Berechnungen mit ihr durchzuführen.
- C ... ist definiert als die Eingabe in ein neuronales Netz.
- D ... entspricht der algebraischen Repräsentation eines RGB-(Rot-Grün-Blau)-Farbbildes.
- E ... weist mehr Dimensionen auf als ein Tensor und weniger als ein Vektor.

Frage 4

Welche Aussage zu tokens und embeddings trifft zu?

- A Tokens und embeddings sind Synonyme.
- B Tokens enthalten einen höheren Informationsgehalt als embeddings.
- C EEG-Aufnahmen können nicht in tokens oder embeddings umgewandelt werden.
- D Tokens entsprechen ganzen Wörtern einer Textsequenz, aber nicht Wortbestandteilen.
- E Sowohl tokens als auch embeddings können Vektoren sein.

Frage 5

Welche Aussage zum Aufbau neuronaler Netze trifft zu?

- A Neuronale Netze weisen i. d. R. mehr Gewichte als Neurone auf.
- B Neuronale Netze bestehen aus einem encoder und einem decoder.
- C Vereinfacht beschrieben gilt, je mehr input layer, desto leistungsstärker ist das Modell.
- D Das output layer eines neuronalen Netzes besteht immer nur aus einem Neuron.
- E Der Aufbau, aber nicht die Funktionsweise künstlicher neuronaler Netze orientiert sich an biologischen neuronalen Netzen im Gehirn.

Frage 6

Welche Aussage zur Funktionsweise neuronaler Netze ist falsch?

- A Im Rahmen des forward pass werden u. a. die Eingabedaten x mit den Gewichten w multipliziert.
- B Aktivierungsfunktionen ermöglichen es, mit neuronalen Netzen auch komplexe, nicht lineare Zusammenhänge zu modellieren.
- C Für den backward pass kann das gradient-descent-Verfahren genutzt werden.
- D Klassische neuronale Netze sind besonders geeignet zur Verarbeitung sequenzieller Daten.
- E Insbesondere große neuronale Netze aus dem Bereich deep learning können sehr leistungsstark sein.

► Weitere Fragen auf der folgenden Seite ...

Frage 7

Welche Aussage zum Aufbau und der Funktionsweise von Transformer-Modellen ist **falsch**?

- A Ein Transformer kann mehrere encoder- und decoder-Blöcke enthalten.
- B Attention-Module finden sich lediglich im decoder eines Transformers.
- C In den attention-Modulen wird die Beziehung der tokens zueinander berechnet.
- D Transformer-Modelle können nicht nur Textsequenzen, sondern bspw. auch Bildsequenzen (z. B. aus einem Video) verarbeiten bzw. analysieren.
- E Im Rahmen des attention-Mechanismus bezeichnet „value“ das Sequenz-Element bzw. den token, für den die attention-Gewichte berechnet werden sollen.

Frage 8

Welche Aussage zu den Anforderungen für die KI in der Neurologie trifft zu?

- A Die Anforderungen an große Mengen von Trainingsdaten sind auch bei seltenen neurologischen Erkrankungen meist einfach zu erfüllen.
- B Algorithmen, die mit Daten erwachsener Patienten trainiert wurden, lassen sich ohne Weiteres auch für pädiatrische Patienten nutzen.
- C KI-Anwendungen für neurologische und neuroradiologische Fragestellungen benötigen i. d. R. weniger Rechenleistung als andere medizinische KI-Anwendungen.
- D KI-Anwendungen in der Neurologie eignen sich für diagnostische, weniger jedoch für therapeutische Fragestellungen.
- E Ein Beispiel für explainable AI (xAI) in der Neurologie wäre, dass ein Algorithmus pathologische Muster in einem EEG markiert, anstatt nur eine Verdachtsdiagnose auszugeben.

Frage 9

Sie möchten einen KI-Algorithmus entwickeln, der analysiert, ob in Schädel-CT-Aufnahmen ein ischämischer Schlaganfall erkennbar ist. Welchen Datensatz wählen Sie?

- A Einen großen, aber nicht anonymisierten Datensatz Ihres Krankenhauses.
- B Einen kleinen anonymisierten Datensatz Ihres Krankenhauses.
- C Mehrere große, frei verfügbare Datensätze wissenschaftlicher Konsortien.
- D Einen großen Datensatz mit Bildern verschiedener neurologischer Erkrankungen.
- E Einen großen Datensatz mit CT-Bildern ohne label.

Frage 10

Sie möchten einen KI-Algorithmus entwickeln, der analysiert, ob in Schädel-CT-Aufnahmen ein ischämischer Schlaganfall erkennbar ist. Welchen Algorithmus wählen Sie?

- A Das moderne und etablierte Transformer-Modell ChatGPT.
- B Ein neuronales Netz mit wenigen hidden layers, um beim Training Zeit zu sparen.
- C Ein neuronales Netz mit einer Architektur, die in einer Publikation gute Ergebnisse erzielte. Sie testen daher keine weiteren Varianten der Architektur.
- D Ein convolutional neural network (CNN). Sie vergleichen unterschiedliche Varianten mit dem Validierungsdatensatz.
- E Ein großes recurrent neural network (RNN) mit vielen hidden layers.