

Künstliche Intelligenz und maschinelles Lernen in der intensivmedizinischen Forschung und klinischen Anwendung

Artificial intelligence and machine learning in intensive care research and clinical application

A. Peine¹ · C. Lütge² · F. Poszler² · L. Celi^{3,4} · O. Schöffski⁵ · G. Marx¹ · L. Martin¹

► **Zitierweise:** Peine A, Lütge C, Poszler F, Celi L, Schöffski O, Marx G et al: Künstliche Intelligenz und maschinelles Lernen in der intensivmedizinischen Forschung und klinischen Anwendung. *Anästh Intensivmed* 2020;61:372–384. DOI: 10.19224/ai2020.372

Zusammenfassung

Kaum einer Entwicklung wird ein größerer Einfluss auf unser tägliches Arbeitsleben vorhergesagt, als der künstlichen Intelligenz (KI). Ein populäres Anwendungsfeld der künstlichen Intelligenz ist das sogenannte maschinelle Lernen, die Disziplin, die sich mit der Generierung von computerisiertem Wissen aus Erfahrung durch einen selbstadaptiven Algorithmus beschäftigt. Insbesondere der hohe Anwendungsbezug, zum Beispiel im Bereich der Mustererkennung und Prädiktion, macht das „Machine Learning“ heute zu einem bevorzugten Anwendungsfeld.

Insbesondere in der Intensivmedizin, charakterisiert durch außergewöhnlich hohe Datendichte und weitverbreitete computergestützte Datenerfassungsroutinen, hat das maschinelle Lernen in jüngster Zeit auch im wissenschaftlichen Kontext relevanten Einfluss gewonnen. Die Datendichte der Intensivmedizin, resultierend aus der mittlerweile stetig steigenden Anzahl konnektierter Geräte und Datenströme, macht die Anwendung von KI zu einem bevorzugten Anwendungsfeld in der Forschung und Entwicklung. Dies eröffnet neue Horizonte für die Praxis. So werden KI-basierte Algorithmen in Zukunft nach Validierung nicht nur das Verhalten der beteiligten Professionen, sondern auch direkt die Behandlung der Patienten beeinflussen können.

In diesem praktischen Leitfaden soll ein „Hands-on“-Einstieg in das maschinelle Lernen für die Anwendung in der Inten-

sivmedizin mittels Praxisbeispielen aus dem wissenschaftlichen Alltag gegeben werden. Weiterhin werden ethische und regulatorische Aspekte der neuen Technologien adressiert.

Summary

Hardly any other development is predicted to have a greater impact on our daily working life than artificial intelligence (AI). A popular field of application of artificial intelligence is the so-called “machine learning”, the discipline that deals with the generation of computerised knowledge from experience through self-adaptive algorithms. Especially the high practical relevance, for example, in the field of pattern recognition and prediction, makes machine learning a preferred field of application in the medical domain.

Especially in intensive care medicine, characterised by an exceptionally high data density and widespread computer-based data acquisition routines, machine learning has recently gained relevant influence in a scientific context as well. The data density of intensive care medicine, resulting from the steadily increasing number of connected devices and data streams, makes the application of AI a preferred field of application in research and development. This opens up new horizons for practice. Thus, after validation, AI-based algorithms in future will not only be able to influence the behaviour of the professions involved, but also directly influence the treatment of patients.

- 1 Universitätsklinikum der RWTH Aachen, Klinik für Operative Intensivmedizin und Intermediate Care (Direktor: Univ.-Prof. Dr. med. Gernot Marx, FRCA)
- 2 Institute for Ethics in Artificial Intelligence, Technische Universität München (Direktor: Prof. Dr. phil. Christoph Lütge)
- 3 Laboratory of Computational Physiology, Harvard Medical School – MIT Division of Health Sciences & Technology, Cambridge, MA, USA (Direktor: Roger G. Mark, M.D., Ph.D.)
- 4 Beth Israel Deaconess Medical Center, Boston, MA, USA
- 5 Lehrstuhl für Gesundheitsmanagement, Universität Erlangen-Nürnberg (Direktor: Prof. Dr. Oliver Schöffski)

Interessenkonflikt

A.P., G.M. und L.M. sind Mitgründer der Clinomic GmbH. A.P. ist Geschäftsführer der Clinomic GmbH. G.M. erhielt Forschungsunterstützung und Beraterhonorare von B. Braun Melsungen, Biotest und Adrenomed außerhalb der eingereichten Arbeit.

L.M. erhielt Forschungsunterstützung, Vortragshonorare und Beraterhonorare der Sphingotec GmbH außerhalb der eingereichten Arbeit.

Alle anderen Autoren erklären, dass sie keinen Interessenkonflikt haben.

Schlüsselwörter

Künstliche Intelligenz –
Maschinelles Lernen –
Algorithmen – Big Data

Keywords

Artificial Intelligence –
Machine Learning – Algo-
rithms – Big Data

This practical guide is intended to provide a “hands-on” introduction to machine learning for use in intensive care medicine, based on practical examples from everyday scientific life. Furthermore, ethical and regulatory aspects of the new technologies are addressed.

Einführung

Selbständig arbeitende Narkosegeräte, intelligente Algorithmen, die eine gefährliche Krankheit schon Jahre vor dem Ausbruch vorhersagen, Roboter, die Routineaufgaben übernehmen – kaum einer Technik wird ein größerer Einfluss auf unser tägliches Leben vorhergesagt als der künstlichen Intelligenz (KI). Diese entspringt der Kombination verschiedener Naturwissenschaften, insbesondere der Informatik, Mathematik und Physik, und wurde in ihrer Entwicklung wesentlich auch durch Erkenntnisse und Beobachtungen aus der Biologie, Medizin und Philosophie beeinflusst. Heute wird der Bereich der künstlichen Intelligenz

überwiegend als Teilgebiet der Informatik beschrieben. Ein sehr populäres Anwendungsfeld der künstlichen Intelligenz ist das sogenannte maschinelle Lernen (meist als „Machine Learning“ bezeichnet), die Disziplin, die sich mit der Generierung von computerisiertem Wissen aus Erfahrung durch einen selbstadaptiven Algorithmus beschäftigt [1]. Insbesondere der hohe Anwendungsbezug, zum Beispiel im Bereich der Mustererkennung und Prädiktion, macht das „Machine Learning“ heute zu einem bevorzugten Anwendungsfeld [2].

Als Teilgebiet des maschinellen Lernens lässt sich das sogenannte tiefgehende Lernen bezeichnen („Deep Learning“). Unter der Verwendung hierarchischer Klassifikationen, dem Aufbau eines biologischen Nervensystems ähnelnd, können Algorithmen mit Hilfe von „Deep Learning“ Aufgaben erfüllen, deren Komplexität noch vor wenigen Jahren unerreichbar war. So werden diese mittlerweile bevorzugt im Bereich

der visuellen Klassifikation (z.B. der Erkennung von Hautkrebs oder der videogestützten Chirurgie [3,4]) sowie in komplexen Optimierungsprozessen (z.B. der Verarbeitung von Vitaldaten in Echtzeit auf Intensivstationen [5]) eingesetzt.

Insbesondere in der Intensivmedizin, charakterisiert durch außergewöhnlich hohe Datendichte und weitverbreitete computergestützte Datenerfassungsroutinen, hat das maschinelle Lernen in jüngster Zeit auch im wissenschaftlichen Kontext relevanten Einfluss gewonnen, z.B. bei der Optimierung der Vasopressordosierung [6] oder bei der Vorhersage ungeplanter Wiederaufnahmen von Patienten [7,9] sowie bei telemedizinischen Anwendungen [8]. Insbesondere die Veröffentlichung zweier medizinischer Datenbanken, der „Medical Information Mart for Intensive Care III database (MIMIC-III)“ [10], bestehend aus Daten von 61.532 Intensivpatienten des Beth Israel Deaconess Medical Center (USA), und der „eICU Collaborative Research

Database v2.0 (eICU)" [11], bestehend aus Daten von 200.859 Intensivpatienten von über 300 Intensivstationen der USA, hat zu einer Demokratisierung der Forschung in diesem Bereich geführt. Jüngst ist auch ein europäisches Äquivalent, die „AmsterdamUMCdb“, mit assoziierten Daten von 20.181 Intensivpatienten veröffentlicht worden [12]. Mit diesen retrospektiven Daten können Wissenschaftler nun KI-Systeme auch ohne Zugang zu proprietären Krankenhausdaten und damit einhergehenden eventuellen Datenschutzbedenken trainieren.

Neben der technischen Entwicklung bestimmt das Vertrauen in das KI-System die Umsetzung am Krankenbett. Vertrauen kann definiert werden als die Erwartung eines Individuums, bei eigener Unsicherheit und Verletzlichkeit von einem externen Akteur, bei der Zielerreichung unterstützt zu werden. Im Bereich der Medizin wäre dies entsprechend die Erwartung eines Patienten gegenüber einem behandelnden Arzt, die Ursache seiner körperlichen Beschwerden zu diagnostizieren und erfolgreich zu behandeln. Ein solcher Akteur kann ein KI-System sein, wobei zum heutigen Zeitpunkt Patienten dem Arzt im Kontrast zu einer Maschine insgesamt deutlich stärker vertrauen [13].

Beeinflusst wird dieses Vertrauen in das System durch eine Vielzahl heterogener Faktoren, zum Beispiel Zuverlässigkeit, Fehlalarmrate, Komplexität der Aufgabe und vor allem die Erklärbarkeit der Entscheidung. Während die Zuverlässigkeit und Fehlerrate durch den stetigen technologischen Fortschritt zunehmend optimiert werden können, sinkt indessen die Erklärbarkeit entsprechender Systeme. Durch die Einführung des probabilistischen Lernens und des „Deep Learning“ wurden Anwendungen beispielsweise immer erfolgreicher, jedoch zugleich immer undurchsichtiger. Funktionsweise und Anwendung von KI-Systemen gleichen daher oft einer mysteriösen „Black-Box“, deren Inneres (der Grund für die jeweilige Entscheidung) unverständlich bleibt. Wie kann also überhaupt sichergestellt werden, dass ein System die ethischen Grundsätze

für KI wie Nicht-Nachteiligkeit, Vorteilhaftigkeit, Autonomie und Gerechtigkeit tatsächlich erfüllt, wenn nicht einmal erklärt werden kann, welche Faktoren und Verarbeitungsprozesse dem Ergebnis eines KI-Systems zugrunde liegen? Die Einhaltung ethischer Grundsätze könnte demnach für undurchschaubare KI-Systeme nie erwiesen werden. Problematisch ist zudem, dass Individuen, die bestimmte Daten (mithilfe von KI-Systemen) verarbeiten, auch hierfür verantwortlich sein und Rechenschaft ablegen sollten [14]. Erklärbarkeit bedeutet also sowohl Verständlichkeit als auch Verantwortlichkeit. Wenn Entscheidungen von Medizinerinnen durch hoch entwickelte KI-Systeme ergänzt und in einigen Fällen sogar übersteuert werden, sollten die menschlichen Experten demnach auf Anfrage immer noch die Möglichkeit und Fähigkeit besitzen, den Prozess der Maschinenentscheidung nachzuvollziehen und zu erläutern [15]. Ein solcher Experte muss nicht zwingendermaßen der behandelnde Arzt sein, auch der Programmierer des verwendeten KI-Systems käme dafür infrage, wobei von diesem nur die technische, nicht aber die medizinische Expertise erwartet werden könnte. Wichtig ist nur, dass die Funktionsweise bzw. die Entscheidungsrationale des Systems gegenüber Dritten (z.B. Richtern, Jurys) erläutert werden kann. Vor allem in moralisch belasteten Situationen, die einer gerichtlichen Begutachtung standhalten müssen, ist dies essenziell. Im medizinischen Bereich können solche Situationen häufig eintreten, wenn beispielsweise ein Behandlungsfehler zum Tod des Patienten führt. Ein wesentliches Kriterium erklärbarer künstlicher Intelligenz – insbesondere in der Medizin – bleibt somit Kausalität sowie die Messung der Qualität der Erklärbarkeit. Aufgrund dieser Prämissen existiert inzwischen Forschung zu „erklärbarer KI“, die darüber Aufschluss liefern soll, warum neuronale Netze ihre Entscheidungen treffen [16] und wie von Menschen interpretierbare Modelle entwickelt und optimiert werden können [13]. Gut möglich ist also, dass in Zukunft adäquate Erklärungen für die getroffenen Entscheidungen von KI-

Systemen erzeugt werden können [17]. Auch die Europäische Kommission hat sich jüngst in einem Whitepaper zu dem Thema positioniert [18].

In diesem praktischen Leitfaden soll ein „Hands-on“-Einstieg in das maschinelle Lernen für die Anwendung in der Intensivmedizin mittels Praxisbeispielen aus dem wissenschaftlichen Alltag gegeben werden. Dieser richtet sich explizit an Einsteiger in diese Disziplin und soll den Ausgangspunkt für die Vertiefung mittels weiterer Literatur bilden.

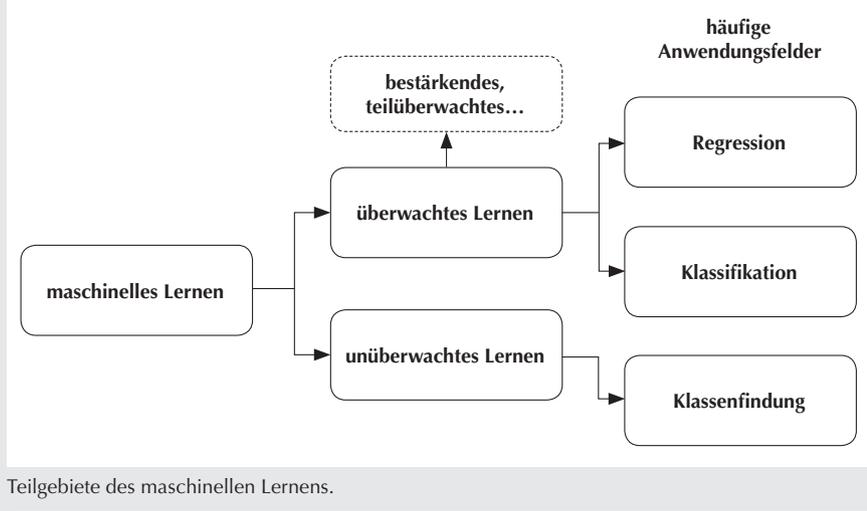
Maschinelles Lernen und Anwendungsfelder

Das maschinelle Lernen bezeichnet die Subdisziplin der Computerwissenschaften, die sich auf Basis vorhandener Datenbestände mit der automatisierten Entwicklung von Algorithmen durch Computer zur eigenständigen Problemlösung beschäftigt.

Arthur L. Samuel, einer der Pioniere des maschinellen Lernens, definierte den Begriff 1959 wie folgt: „(Machine Learning is the) field of study that gives computers the ability to learn without being explicitly programmed“ [19]. Er beschreibt damit die wesentliche Idee und Grundannahme des maschinellen Lernens in seiner heute angewandten Form: Durch Training an Datensätzen generiert der Computer Erfahrung, welche er in Form eines trainierten Algorithmus als Wissen manifestiert. Bei der Durchführung von Aufgaben kann er das antrainierte Wissen abrufen und an bisher unbekanntes Daten anwenden.

Im Wesentlichen werden zwei unterschiedliche Ansätze unterschieden: Das überwachte Lernen („supervised learning“) und das unüberwachte Lernen („unsupervised learning“), wobei das überwachte Lernen noch in spezialisierte Bereiche wie bestärkendes Lernen („reinforcement“) und teilüberwachtes („semisupervised“) Lernen subsegmentiert werden kann (Abb. 1). Oft wird das sogenannte bestärkende Lernen („reinforcement“), bei dem ein Agent selbständig eine Strategie erlernt mit dem Ziel, erhaltene Belohnungen zu

Abbildung 1



Teilgebiete des maschinellen Lernens.

maximieren, als eigenständige Kategorie angesehen.

Das Grundprinzip des überwachten Lernens ist die Generierung einer mathematischen Funktion auf Basis eines Trainings mittels bekannter Eingabe- und Ausgabedatensätze. Diese Funktion

kann dann genutzt werden, um Aussagen über die der Beobachtung zugrunde liegende Gesetzmäßigkeit zu treffen, also etwa Assoziationen oder Prädiktionen.

Der Begriff des „überwachten“ Lernens ergibt sich aus der Tatsache, dass der Algorithmus während des Trainings den

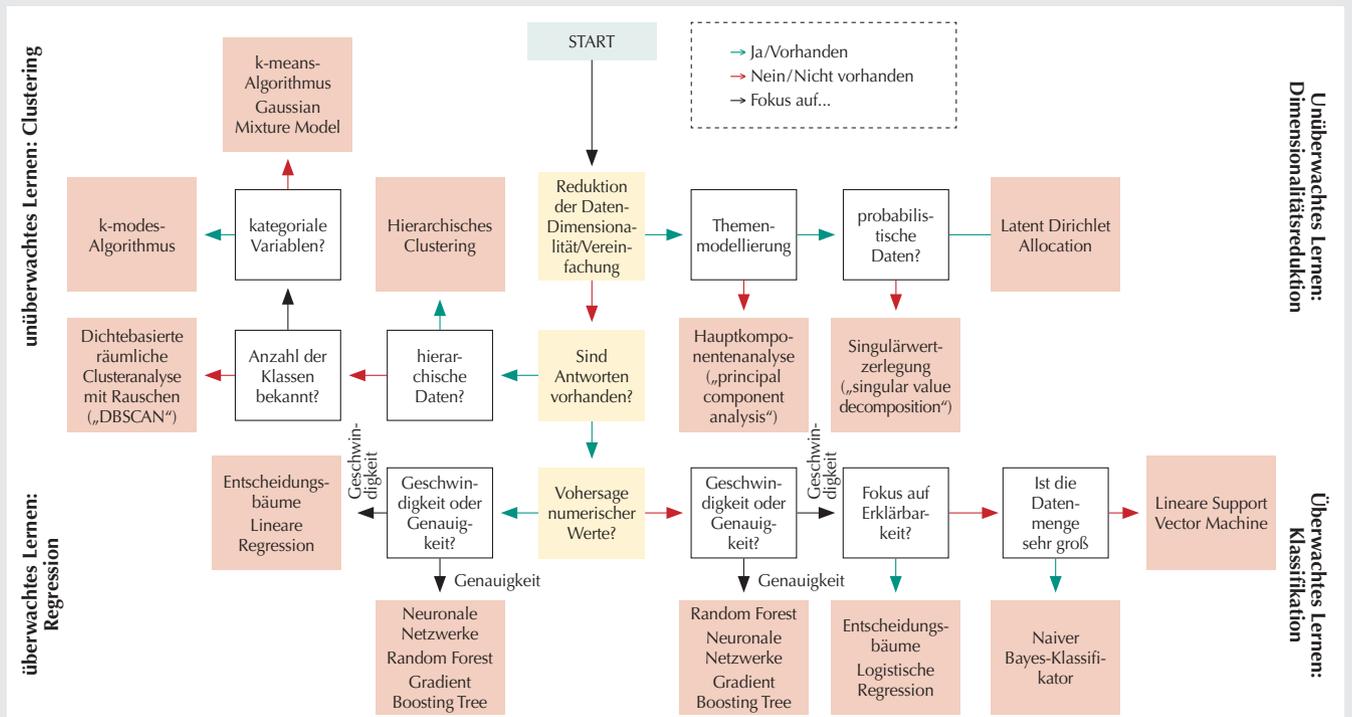
korrekten Funktionswert zur Verfügung hat, um die Güte seiner Prädiktion zu überprüfen – sinnbildlich wie ein Lehrer, der die Aufgaben seines Schülers überwacht, die richtigen Antworten kennt und gegebenenfalls korrigiert [20]. So vergleicht die KI bei jedem Trainingsdurchlauf die errechnete Antwort mit der richtigen Antwort und passt darauf basierend ihren Rechenweg so an, dass sich die Qualität der errechneten Antworten ständig verbessert und den richtigen Antworten immer näherkommt. Diese Algorithmen werden oft in Regressions- oder Klassifikationsaufgaben genutzt. Das in Abbildung 2 dargestellte Flussdiagramm gibt einen fragestellungsorientierten Überblick über etablierte Algorithmen.

Überwachtes Lernen

Lineare Regression

Beispiel: Ein Forscher möchte die optimale Dosierung für ein neues Medikament finden. Die Hypothese ist, dass sie eng mit anderen Laborbefunden (z.B. Leberenzymen, Kreatinin) des Patienten

Abbildung 2



Flussdiagramm der etablierten Algorithmen in der medizinischen Forschung, modifiziert nach [21].

Tabelle 1

Beispieldaten für eine Regressionsfragestellung im intensivmedizinischen Kontext.

Patient	Kreatinin mg/dl	GPT U/l	PCT ng/ml	Medikamenten- spiegel ng/ml
A	0,8	30	1	1.000
B	1,1	50	0,5	6.000
C	1,9	70	5	2.000
D	2,7	90	8	?

assoziiert ist. Die Fragestellung ist nun, ob es möglich ist, den Medikamentenspiegel im Blut allein durch die Verwendung der anderen Surrogate vorherzusagen (Tab. 1).

Zur Vereinfachung der Fragestellung wird hier zunächst nur eine Zielvariable betrachtet: Der Zusammenhang zwischen Kreatinin und dem resultierenden Medikamentenspiegel. Dieser ist in Abbildung 3 exemplarisch visualisiert.

Intuitiv lässt sich ein linearer Zusammenhang erahnen, in Abbildung 3 als gestrichelte Linie dargestellt. Dieser Zusammenhang lässt sich mit der Gleichung

$$f(x) = mx + b$$

beschreiben, wobei m die Steigung der Gerade und b den Schnitt mit der y -Achse angibt. Die sogenannte Hypothesenfunktion soll im Kontext des gestellten Problems die Frage beantworten, wie der Zusammenhang zwischen Kreatinin und Zielvariable zu beschreiben ist. Ziel des angewendeten Algorithmus soll es nun sein, die Variablen m und

b möglichst genau zu bestimmen mit dem Ziel, den mathematischen Zusammenhang zu beschreiben. Ein solcher Zusammenhang lässt sich z.B. durch den Algorithmus der „linearen Regression“ des Maschinenlernens näherungsweise lösen. Der Vollständigkeit halber sei erwähnt, dass dies nur einer der Ansätze für die Lösung eines Regressionsproblems ist. Die Vorgehensweise dieses Algorithmus lässt sich prototypisch auf andere Algorithmen des maschinellen Lernens übertragen, sodass hier eine genauere Erläuterung der grundsätzlichen Herangehensweise erfolgt.

Definition einer Fehlerfunktion: Sie soll angeben, wie gut oder schlecht die Abschätzung des Algorithmus gewesen ist. Für ein Variablenpaar (m, b) im genannten Beispiel soll ein „Fehlerwert“ ausgegeben werden, anhand dessen sich die Güte der aktuellen Prädiktion abschätzen lässt. Die Fehlerfunktion der linearen Regression kalkuliert für jedes Wertepaar (x, y) die quadrierte, summierte Distanz der vorhergesagten Funktion $(mx+b)$ zu den tatsächlichen Werten („mean squared error“). Die Quadrierung wird vorgenommen, damit sich einerseits negative und positive Abweichungen nicht in der Summe nivellieren und andererseits Ausreißer stärker ins Gewicht fallen und verbessert werden.

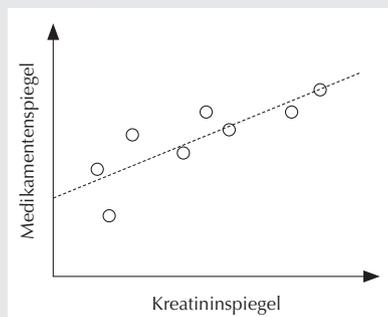
$$\text{Error}_{(m,b)} = \frac{1}{N} \sum_{i=1}^N (y_i - (mx_i + b))^2$$

Damit ergibt sich der Zusammenhang, dass eine bessere Vorhersage einen geringeren Wert der Fehlerfunktion ergibt. Visualisiert man nun die Fehlerfunktion in Abhängigkeit der Vorhersage, so lässt

sich diese als Oberfläche darstellen (Abb. 4a). Eine schlechtere Vorhersage (in der Abbildung rötlich dargestellt) befindet sich weiter entfernt vom „Tal“ (bläulich dargestellt), in dem der geringste Fehler besteht.

Optimierung der Fehlerfunktion: Ziel des Algorithmus ist es nun, eine Optimierung der Fehlerfunktion durchzuführen mit dem Ziel, den Ort des geringsten Fehlers und damit der höchsten Prädiktionsgüte zu finden. In diesem Fall wird das sogenannte Gradientenverfahren („gradient descent“) genutzt [22]. Dieser iterative Optimierungsalgorithmus kalkuliert für eine Position auf der Ebene den Ort des „größten Abstiegs“ mittels partieller Ableitung und bewegt sich nun in Schritten auf der Ebene in Richtung Tiefpunkt der Ebene (visualisiert im zweidimensionalen Raum in Abbildung 4b). Die Weite der Schritte (im Bereich des maschinellen Lernens als „Learning Rate“ bezeichnet) gibt an, wie weit sich der Algorithmus in Richtung Tal der Kurve bewegt. Diese Variable sollte ebenfalls optimiert werden, sodass dem Ziel mit möglichst großen Schritten entgegengefahren wird und keine Überschreitung des Ziels entsteht. Die vorhergehenden Erläuterungen beziehen sich jedoch nur auf eine univariate lineare Regression, d.h. den Zusammenhang zwischen einer Ursprungs- und Zielvariable. Im gewählten Beispiel ist der Medikamentenspiegel also nicht nur vom Kreatinin, sondern auch von anderen klinischen Surrogaten abhängig. Ein einfacher linearer Zusammenhang kann also nicht mehr angenommen werden. Vielmehr handelt es sich eher um einen multidimensionalen Vektor, welcher den Zusammenhang beschreibt. Die Visualisierung der Fehlerfunktion stellt sich analog nicht mehr als 2-dimensionale Ebene, sondern als komplexer Raum dar, wobei θ die zu optimierenden Parameter repräsentiert.

Mittels eines solchen iterativen Optimierungsverfahrens lassen sich Fragestellungen der Regression gut bearbeiten. Da es sich hierbei um komplexe Verfahren handelt, können Techniken des maschinellen Lernens hierbei bevorzugt ange-

Abbildung 3

Exemplarischer Zusammenhang zwischen Kreatininspiegel und Zielvariable.

Abbildung 4

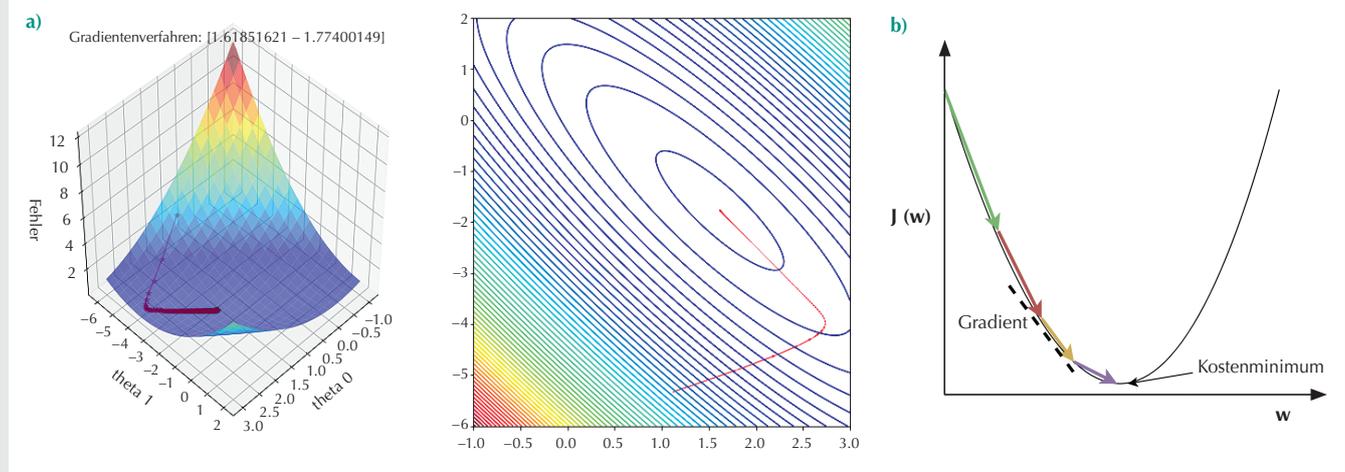


Abbildung 4a: Visualisierung eines exemplarischen Gradientenverfahrens im dreidimensionalen Raum.
Abbildung 4b: Visualisierung eines exemplarischen Gradientenverfahrens im zweidimensionalen Raum.

wendet werden. Ein besonderer Vorteil ist, dass ein einmal trainiertes Modell mit nur geringem Rechenbedarf erneut abgefragt werden kann.

Klassifikation mittels logistischer Regression

Beispiel: Ein Krankenhaus möchte die Wiederaufnahme-Quoten auf die Intensivstation für Patienten mit einer bestimmten Krankheit vorhersagen. Das Krankenhaus extrahiert die in Tabelle 2 genannten retrospektiven Daten aus der Patientendatenbank des Krankenhauses.

Der Zusammenhang zwischen Flüssigkeitsbilanz und Wiederaufnahme ist exemplarisch visualisiert in Abbildung 5.

Auch hier erfolgt zunächst eine Vereinfachung des Problems auf einen bivariaten Zusammenhang: der Flüssigkeitsbilanz des Patienten und der dichotomen Variable "Wiederaufnahme des Patienten ja/nein".

Eine mögliche Herangehensweise wäre hier die vorab beschriebene Methode der Generierung einer Hypothesenfunktion, ergänzt durch eine Schwellenfunktion, welche die beiden Gruppen nach ihrer Wahrscheinlichkeit, z.B. bei dem Erreichen des Schwellenwerts $x > 0,5$ in Gruppen von $y=1$ respektive $y=0$, klassifiziert. Dies ist in Abbildung 6a visualisiert, in der die Prädiktionen der Gruppen jeweils grün (Wiederaufnahme) oder rot (keine Wiederaufnahme) markiert sind. Ergänzt man nun jedoch weitere Werte außerhalb dieses Wertebereichs (wie in Abbildung 6b demonstriert), lässt sich die Problematik der linearen Regression für Klassifikationsprobleme erkennen: Durch die Veränderung der Parameter der Hypothesenfunktion (gestrichelt dargestellt) werden nun zunehmend Werte falsch der Kategorie „keine Wiederaufnahme“ zugeordnet. Ein zweites Problem ist, dass

die Hypothesenfunktion auch negative Werte annehmen könnte, was nicht dem vorhergesehenen Wertebereich $[0; 1]$ und somit nicht der Realität entspricht.

Die mathematische Lösung dieses Problems besteht in der Nutzung einer logistischen Hypothesenfunktion, welche folgende Eigenschaften zeigt: i) Die Wahrscheinlichkeit liegt im Intervall $[0, 1]$, ii) sie liegt symmetrisch um den Wendepunkt $y=0,5$, iii) Sie verläuft sigmoid asymptotisch zu ihren Intervallgrenzen $\{0, 1\}$ (Abb. 7).

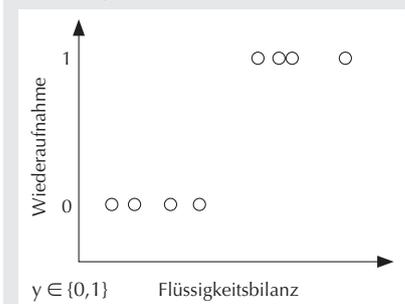
Eine direkte Berechnung der optimalen Regressionskurve ist nicht möglich, daher erfolgt auch diese einer iterativen Herangehensweise, meistens mittels der Methode der höchsten Wahrscheinlichkeit („Maximum likelihood“) [23]. Die

Tabelle 2

Beispieldaten für ein Klassifikationsproblem.

Patient	Liegedauer	Beatmungstage	Flüssigkeitsbilanz	Wiederaufnahme?
A	45	10	+5.500	1 (ja)
B	14	5	+500	0 (nein)
C	12	12	+2.300	0 (nein)
D	31	6	+1.500	?

Abbildung 5



Visualisierung des Klassifikationsproblems.

Abbildung 6

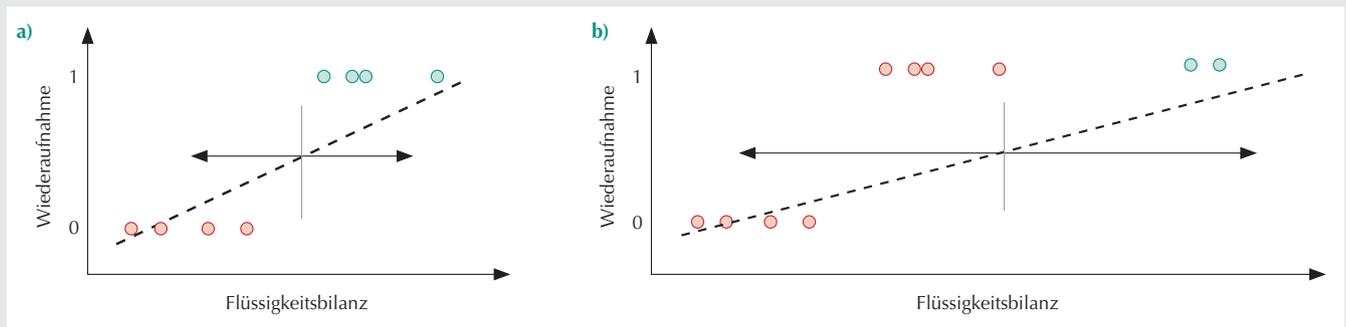
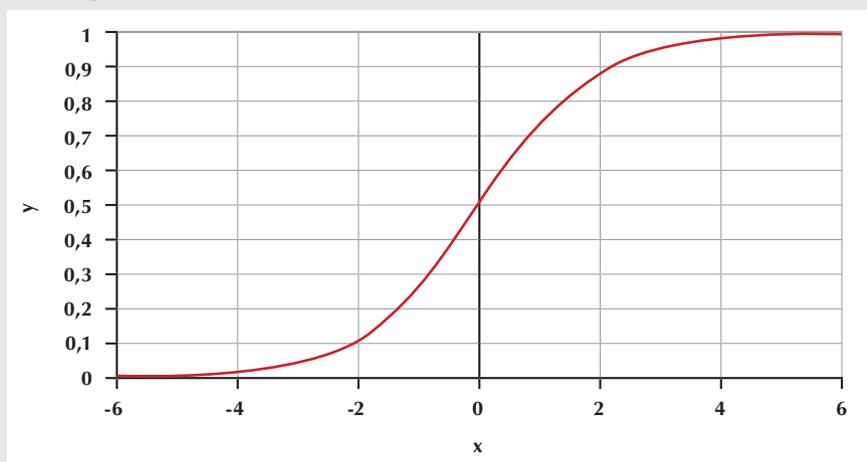


Abbildung 6a/b: Lineare Regression des dargestellten Problems: a) Klassifikation mittels Schwellenwert, b) Klassifikation bei vergrößertem Wertebereich.

Abbildung 7



Visualisierung der logistischen Funktion.

Hypothesenfunktion gibt hier nicht den erwarteten Wert aus, sondern die Wahrscheinlichkeit des Eintretens der dichotomen Variable – im Kontext des Beispiels also die Wahrscheinlichkeit, dass ein Patient auf die Station wieder aufgenommen wird. Die Optimierung des Modells erfolgt analog zur linearen Regression mit der logarithmischen Likelihood-Funktion („log loss“), welche mittels eines Gradientenverfahrens optimiert wird. Auch hier ist eine multivariate Berechnung (zum Beispiel unter Einbeziehung der anderen verfügbaren Variablen) möglich, auf die in diesem Kontext jedoch nicht eingegangen wird.

Unüberwachtes Lernen

Gruppierungsverfahren

Oftmals sind die grundlegenden Gesetzmäßigkeiten innerhalb von Datensätzen

nicht klar erkennbar. Gruppierungsverfahren, meist als „Clustering-Algorithmen“ bezeichnet, können genutzt werden, um zusammenhängende Gruppen innerhalb von Datensätzen zu klassifizieren. Da die „richtige“ oder „falsche“ Lösung nicht bekannt ist und der Algorithmus auch kein Feedback erhalten kann, wird diese Gruppe der Algorithmen als unüberwachte Lernalgorithmen bezeichnet.

Beispiel: In einem murinen Sepsis-Modell reagieren bestimmte Mausstämme unterschiedlich auf ein neues Medikament. Es wird angenommen, dass dies auf einen bestimmten Genpolymorphismus zurückzuführen ist. Die Forschungsfrage ist, wie man Mäuse in Kategorien gruppiert und möglicherweise ihr Reaktionsmuster vorhersagen kann.

Beispieldaten sind in Tabelle 3 dargestellt und in einer Punktwolke visualisiert.

Bereits bei der Betrachtung der Punktwolke in Abbildung 8 lassen sich drei Gruppen erahnen. Der grundsätzliche Ansatz wird hier anhand des sogenannten „k-means-Algorithmus“ erläutert, der bereits 1957 von dem Informatiker S. P. Lloyd's publiziert wurde und für Gruppierungsverfahren breite Anwendung findet [24].

Der Algorithmus geht nun wie folgt vor (Abb. 9):

- Zufällige Wahl der Clusterzentren (1): Der Algorithmus platziert die Zentren c (in Abbildung 9 rot, gelb und blau markiert) der Cluster zufällig innerhalb des Koordinatensystems der Daten.

Tabelle 3

Beispieldaten für ein Gruppierungsproblem.

Maus	Expressionslevel Gen A	Expressionslevel Gen B	Tumornekrosefaktor- α Level (pg/ml)
1	23	0,6	14
2	14	0,5	12
3	4	0	2

- Zuordnung zum nächsten Zentrum (2): Für jeden Punkt der Daten sucht der Algorithmus nun das nächste Zentrum, in dem er die (z.B. euklidische) Distanz zu allen Zentren ermittelt. Nun wird der Punkt dem Cluster des nächsten Zentrums zugeordnet bzw. in einer Abwandlung des Algorithmus, bei dem sich die Varianz des Clusters am wenigsten erhöht.
- Für jedes Cluster wird nun ein neuer Schwerpunkt berechnet, in dem alle Datenpunkte des Clusters j einbezogen werden und die Zentren dorthin verschoben werden [3].
- Es erfolgt erneut eine Zuordnung zum nächstgelegenen Zentrum. Diese Schritte werden wiederholt, bis keine Veränderung der Zuordnung mehr stattfindet, und damit die Gruppen gefunden sind (Schritt 4).

Nachteil dieses Algorithmus ist, dass die Klassifizierung stark von der initialen, zufälligen Wahl der Clusterzentren abhängig ist. Der Algorithmus findet damit nicht zwangsläufig die besten Lösungen. Weiterer Nachteil ist, dass die Anzahl der Clusterzentren (also der möglichen Gruppen) vorab gewählt werden muss. Bei der Wahl einer zu kleinen oder zu großen Anzahl Zentren können möglicherweise falsche Schlüsse gezogen werden.

Eine Abschätzung der möglichen Gruppen ist jedoch durch Visualisierung nicht immer möglich, etwa bei mehrdimensionalen (multivariaten) Fragestellungen, sodass möglicherweise andere Algorithmen, etwa die sogenannte „dichte-basierte räumliche Clusteranalyse mit Rauschen“ („DBSCAN“), von Vorteil sein können [26]. Weiterhin sei an dieser Stelle das sogenannte hierarchische Clustering erwähnt, welches die zu gruppierenden Elemente zu stetig größeren Clustern zusammenfasst und welchem gerade im Bereich der biologisch-medizinischen Fragestellungen eine wichtige Rolle zukommt [27].

Abbildung 8

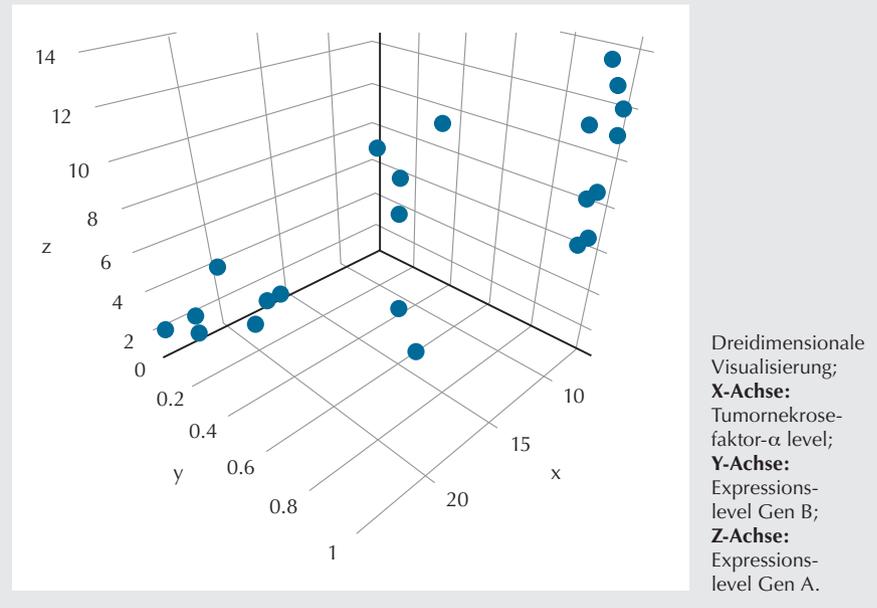
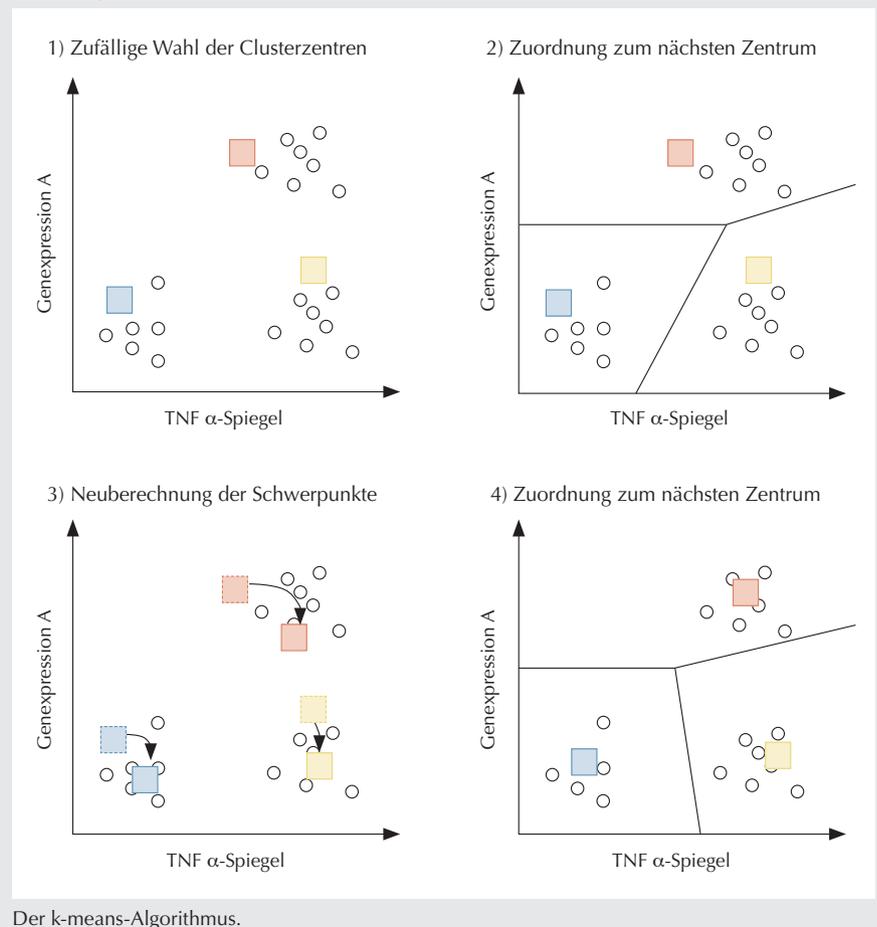


Abbildung 9



Der k-means-Algorithmus.

Mustererkennung mittels künstlicher neuronaler Netze

Bereits mit dem Aufkommen programmierbarer Computer in den 1940er Jahren beschrieben die Forscher Warren S. McCulloch und Walter Pitts in ihrem Werk „A logical calculus of the ideas immanent in nervous activity“ [28] Verknüpfungen von mathematischen Elementen in Form eines Netzes, welches in seiner Struktur dem Aufbau des Nervensystems ähnelt, heute als künstliche neuronale Netze („artificial neural networks“) bezeichnet.

Obwohl mittels künstlicher neuronaler Netze bereits wenige Jahre später beachtliche Erfolge erzielt werden konnten – etwa die Erkennung einfacher Schriftzeichen – konnten diese jedoch zunächst die in sie projizierten Erwartungen nicht bestätigen. Dies lag unter anderem an der damals vorliegenden mangelnden Rechenleistung. In neuester Zeit erleben diese Netzwerke eine Renaissance insbesondere im medizinischen Umfeld durch ihre herausragenden Leistungen im Anwendungsfeld der komplexen Mustererkennung. Insbesondere zeigen diese Netzwerke ihre Stärke in Aufgaben, die für Menschen vergleichsweise einfach zu lösen sind, einen Computeralgorithmus aber oft vor unlösbare Aufgaben stellte. So ist es beispielsweise vergleichsweise einfach, einem Computer auf höchstem Niveau Schach beizubringen, jedoch ausgesprochen schwierig, zwei Gesichter voneinander zu unterscheiden. Dieses Phänomen wird heute als Moravec's Paradoxon bezeichnet. Aus Gründen der Vereinfachung wird hier nur auf die Anwendung im überwachten Lernen eingegangen.

Beispiel: Ein Forscher möchte den Zusammenhang zwischen Volumentherapie und Rekapillarierungszeit bestimmen. Zu diesem Zweck wurden Bilder der Finger des Patienten zusammen mit den Informationen gesammelt, ob die Volumentherapie zu hämodynamischer Stabilität führte. Die Hypothese ist, dass ein Algorithmus auf der Grundlage dieser Informationen die Reaktion des Patienten auf die Volumentherapie vorhersagen kann (Tab. 4).

Die Eingabewerte für das neuronale Netz (im Folgenden als „Inputfeature“ bezeichnet) führen über einen Verarbeitungsschritt zu einer Ausgabe („Output“). Diese Struktur wird analog zur biologischen Struktur als Neuron bezeichnet. Je nach Eingabe kommt es zu einer Aktivierung oder Nichtaktivierung des Neurons. Kombiniert man nun mehrere Neuronen miteinander, so ergibt sich die in Abbildung 10 dargestellte Struktur.

Je nach Kombination der Inputfeatures entsteht nun also hier eine spezifische Aktivierung der Neurone innerhalb der zweiten Schicht. Da diese Zwischenschichten nicht sichtbar sind, werden sie als so genannte „hidden layer“ bezeichnet. Je mehr Zwischenschichten existieren, desto tiefer ist das neuronale Netz.

Betrachtet man nun ein solches Neuron mathematisch, so sorgen mehrere Funktionen innerhalb dessen für die Entscheidung zur Aktivierung oder Nichtaktivie-

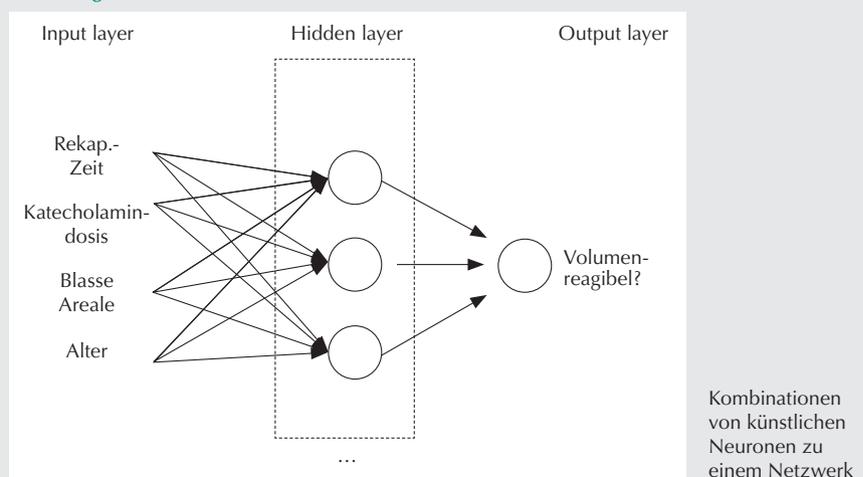
rung. Die Inputfeatures (hier Alter, Rekapillarierungszeit, etc...) werden je nach Gewichtung innerhalb des Modells mit einem Faktor multipliziert und innerhalb einer Übertragungsfunktion summiert. Die Aktivierungsfunktion ϕ , oft eine logistische Funktion, entscheidet nun über den Schaltzustand des Neurons. Sie ist je nach Art des neuronalen Netzwerks spezifisch. So ergeben sich, je nach Art des Inputfeatures, jeweils charakteristische Aktivierungsmuster. Das Netzwerk wird nun analog der Vorgehensweise des überwachten Lernens mittels eines Datensatzes trainiert. Im Falle einer falschen Ausgabe wird der Fehler iterativ Schicht für Schicht in das Netzwerk zurückgeführt, dies wird als sogenannte „backpropagation of error“ bezeichnet. Hierbei werden die Gewichte und der Schwellenwert (oft als „bias“ bezeichnet) entsprechend angepasst. Im Ausgangsproblem sind die Inputfeatures des neuronalen Netzes jedoch nicht definierte

Tabelle 4

Beispieldaten für das neuronale Netzwerk.

Video	Rekap.zeit s	Blasse Areale	Katecholamin dosierung mcg/kg/min	Alter des Patienten	Reagibel auf Volumen- therapie?
1	4	5	0,05	66	1
2	2	1	0,1	44	0
3	5	7	0,2	72	1
4	2	2	0,05	23	0

Abbildung 10



Werte innerhalb eines Datensatzes, sondern einzelne Standbilder („frames“) aus dem Video des Patientendaumens. Dies ist eine Aufgabe, die aufgrund der Masse an Inputvariablen für einen Computer ungleich schwerer zu lösen ist und mehr Rechenleistung erfordert. Für ein solches Problem haben sich gefaltete neuronale Netzwerke („convolutional neural networks“) etabliert [32].

Inputfeatures sind hier keine einfachen numerischen Werte, sondern große Matrizen, in denen jeder Wert z.B. für eine Graustufe eines Pixels des Bildes steht. Über ein Filterverfahren wird die Komplexität des Bildes in sogenannten „convolutional layers“ reduziert. So könnten zum Beispiel in den oberen Schichten des Netzwerkes Kanten, Formen oder Gradienten erkannt und extrahiert werden, in den tieferen Schichten zum Beispiel einzelne Bildkomponenten, etwa Gegenstände oder Personen. Somit kann die Anzahl der Inputwerte und folglich die benötigte Rechenleistung verringert werden. In „pooled feature maps“ werden die relevanten Informationen zusammengefasst. Analog zum biologischen Nervensystem eines Kleinkindes erlernt das Netzwerk zunächst die Erkennung von Formen und Farben, später Gegenstände und deren Verbindung miteinander, bis schließlich eine Klassifikation stattfinden kann. Im Kontext der Fragestellung könnte zum Beispiel ein Aktivierungsmuster zunächst spezifisch die Kanten, ein tieferes Muster beispielsweise scharfe von unscharfen Begrenzungen bei der Rekapillarisation, unterscheiden.

Werkzeuge und Herangehensweise

Mittlerweile ist für die Anwendung des maschinellen Lernens keine umfangreiche Programmierfähigkeit mehr vonnöten. Software-Werkzeuge bringen die Implementierung der gängigen Algorithmen mit und vereinfachen deren Nutzung in grafischen oder vereinfachten Oberflächen.

Ein idealer Einstieg ist die kostenfreie verfügbare Software „Waikato Environ-

ment for Knowledge Analysis“ (WEKA), welche von der Universität Waikato in Neuseeland entwickelt wird [33,34]. Die Software bringt eine grafische Oberfläche mit, welche die Arbeit mit den Datenbanken stark vereinfacht. Die Software verfügt über den „Explorer“, welcher ein unkompliziertes Testen verschiedener Algorithmen ermöglicht, den „Experimenter“ für kontrollierte Experimente und den „Knowledge Flow“, welcher eine grafische Erstellung der Datenextraktions-Pipeline ermöglicht.

Einen umfangreicheren und vertieften Einstieg bietet die Beschäftigung mit der Programmiersprache Python, welche eine der schnellwachsenden Plattformen für angewandtes maschinelles Lernen darstellt. Eine Vielzahl von umfangreichen Toolboxes, etwa Scikit-Learn [35], Pandas [36] oder NumPy [37], bieten Implementationen nahezu jedes bestehenden Algorithmus, ermöglichen umfangreiches Datenmanagement sowie komplexe mathematische Operationen. Die von der Firma Google in offener Lizenz bereitgestellte Software Tensorflow [38,39] ermöglicht einen deutlich vereinfachten Einstieg in das Deep Learning mittels neuronaler Netze, etwa zur Bildverarbeitung [39].

Eine weitere, gerade im Bereich der Statistik und Mathematik eingesetzte Programmiersprache ist R [40]. Diese wird insbesondere aufgrund der großen Anzahl verfügbarer Implementierungen (z.B. dem leistungsstarken Caret [41], welches insbesondere für prädiktive Modelle eingesetzt wird) und der hervorragenden Schnittstellen zu Datenbanken und verknüpfter Software von vielen professionellen Datenwissenschaftlern bevorzugt genutzt. Die Extraktion der Daten aus den jeweiligen Datenbanken wird meist mittels der sogenannten Standardized Query Language (SQL) [42] vorgenommen, welche eine spezifische Abfrage der genauen Datensätze für die Weiterverarbeitung ermöglicht. Eine besondere Herausforderung ist hier die Verknüpfung von medizinischem Wissen mit technischer Expertise, erforderlich für die wissenschaftliche Genauigkeit für die getroffenen Prognosen. Eine große wissenschaftliche Gemeinschaft

aus Ärzten, Informatikern und Datenwissenschaftlern trägt diese Skripte für die wesentlichen intensivmedizinischen Fragestellungen zur gemeinschaftlichen Nutzung zusammen [43,44].

Literatur

1. Alpaydin E: Introduction to Machine Learning. Boston, USA: MIT Press 2009
2. Cleophas TJ, Zwinderman AH: Machine Learning in Medicine – a Complete Overview. ISBN 978-3-319-15195-3 Springer 2015
3. Esteva A, Kuprel B, Novoa RA, et al: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115–118
4. Tai Y, Wei L, Zhou H, et al: Machine learning-based haptic-enabled surgical navigation with security awareness. *Concurr Comput Pract Exp* 2019;31: e4908
5. Davoudi A, Malhotra KR, Shickel B, et al: Intelligent ICU for Autonomous Patient Monitoring Using Pervasive Sensing and Deep Learning. *Sci Rep* 2019;9:1–13
6. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA: The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018;24:1716
7. Lin Y-W, Zhou Y, Faghri F, Shaw MJ, Campbell RH: Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. *PLOS ONE* 2019;14:e0218942
8. Marx G, Dusch M, Czaplik M, Balzer F, Brokmann JC, Deisz R et al: Telemedizin für die vier Säulen der Anästhesiologie. Positionspapier der Deutschen Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI) und der Deutschen Gesellschaft für Telemedizin (DG Telemed). *Anästh Intensivmed* 2019;60: 191–207. DOI: 10.19224/ai2019.191
9. Kagerbauer S, Blobner M, Ulm B, Jungwirth B: Die Zukunft hat schon begonnen. Wie maschinelles Lernen Anästhesie und Intensivmedizin prägt. *Anästh Intensivmed* 2020;61:85–96. DOI: 10.19224/ai2020.85
10. Johnson AEW, Pollard TJ, Shen L, et al: MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035
11. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O: The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data* 2018;5:180178

12. Amsterdam Medical Data Science. <https://amsterdammedicaldatascience.nl/> (Zugriffsdatum: 14.12.2019)
13. Stewart J, Sprivilis P, Dwivedi G: Artificial intelligence and machine learning in emergency medicine. *Emerg Med Australas* EMA 2018;30:870–874
14. Gutachten der Datenethikkommission. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6
15. Holzinger A, Biemann C, Pattichis CS, Kell DB: What do we need to build explainable AI systems for the medical domain? *ArXiv171209923* *Cs Stat* 2017; published online Dec 28. <http://arxiv.org/abs/1712.09923> (Zugriffsdatum: 06.12.2019)
16. Wachter S, Mittelstadt B, Floridi L: Transparent, explainable, and accountable AI for robotics. *Sci Robot* 2017;2. <https://ora.ox.ac.uk/objects/uid:fd0a0307-dc89-488e-bdec-4a100d-1f073e> (Zugriffsdatum: 06.12.2019)
17. Health C for D and R: Artificial Intelligence and Machine Learning in Software as a Medical Device. FDA 2020; published online Jan 28. <http://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> (Zugriffsdatum: 29.03.2020)
18. White Paper on Artificial Intelligence: a European approach to excellence and trust. European Commission https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (Zugriffsdatum: 30.03.2020)
19. Samuel AL: Some studies in machine learning using the game of Checkers. *Ibm J Res Dev* 1959;DOI: 10.1147/rd.33.0210:71–105
20. Cunningham P, Cord M, Delany SJ: Supervised Learning. In: Cord M, Cunningham P (Hrsg.): *Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval*. Berlin, Heidelberg: Springer 2008;21–49
21. Li H: Which machine learning algorithm should I use? *SAS Data Sci Blog*. <https://blogs.sas.com/content/subconsciousmusings/2017/04/12/machine-learning-algorithm-use/> (Zugriffsdatum: 14.12.2019)
22. Bottou L: Large-Scale Machine Learning with Stochastic Gradient Descent. In: Lechevallier Y, Saporta G (Hrsg.): *Proceedings of COMPSTAT'2010*. Paris, Frankreich: Physica-Verlag HD 2010;177–186
23. Murshudov GN, Vagin AA, Dodson EJ: Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr D Biol Crystallogr* 1997;53:240–255
24. Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY: An Efficient k-Means Clustering Algorithm: Analysis and Implementation. *IEEE Trans Pattern Anal Mach Intell* 2002; 24:881–892
25. Rousseeuw PJ: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987;20:53–65
26. Birant D, Kut A: ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data Knowl Eng* 2007;60:208–221
27. Johnson SC: Hierarchical clustering schemes. *Psychometrika* 1967;32:241–254
28. McCulloch WS, Pitts W: A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943;5:115–133
29. Moravec H: *Mind Children: The Future of Robot and Human Intelligence*. Harvard University Press 1988
30. Grossberg S: Adaptive Resonance Theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Netw.* 2013; 37:1–47. doi:10.1016/j.neunet.2012.09.017
31. Kohonen T: Exploration of very large databases by self-organizing maps. In: *Proceedings of International Conference on Neural Networks (ICNN'97)* 1997:PL1-PL6 vol.1
32. Ciresan DC, Meier U, Masci J, Gambardella LM, Schmidhuber J: Flexible, High Performance Convolutional Neural Networks for Image Classification. 2011. Machine learning-based haptic-enabled surgical navigation with security awareness. *Concurr. Comput. Pract. Exp.* 31(19) (2019)
33. Weka 3 - Data Mining with Open Source Machine Learning Software in Java. <https://www.cs.waikato.ac.nz/ml/weka/> (Zugriffsdatum: 25.10.2019)
34. Witten IH, Frank E, Hall MA, Pal CJ: *Data Mining: Practical Machine Learning Tools and Techniques*, 4th edn. Burlington, USA: Morgan Kaufmann 2016
35. scikit-learn: machine learning in Python – scikit-learn 0.21.3 documentation. <https://scikit-learn.org/stable/> (Zugriffsdatum: 25.10.2019)
36. Python Data Analysis Library – pandas: Python Data Analysis Library. <https://pandas.pydata.org/> (Zugriffsdatum: 25.10.2019)
37. NumPy – NumPy. <https://numpy.org/> (Zugriffsdatum: 25.10.2019)
38. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J & Kudlur M: (2016) Tensorflow: A system for large-scale machine learning. In 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16) (pp. 265–283)
39. Géron A: *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques for Building Intelligent Systems*. Beijing Boston Farnham: O'Reilly UK Ltd. 2017
40. R: The R Project for Statistical Computing. <https://www.r-project.org/> (Zugriffsdatum: 25.10.2019)
41. Wing MKC from J, Weston S, Williams A, et al: caret: Classification and Regression Training. 2019 <https://CRAN.R-project.org/package=caret> (Zugriffsdatum: 25.10.2019)
42. Date CJ: *A guide to the SQL standard: a user's guide to the standard relational language SQL*. Reading, Mass.: Addison-Wesley Pub. Co. 1989
43. MIT-LCP/eicu-code. MIT Laboratory for Computational Physiology 2019. <https://github.com/MIT-LCP/eicu-code> (Zugriffsdatum: 25.10.2019)
44. MIT-LCP/mimic-cookbook. MIT Laboratory for Computational Physiology 2019. <https://github.com/MIT-LCP/mimic-cookbook> (Zugriffsdatum: 25.10.2019).

Korrespondenz- adresse

**Dr. med. Arne Peine,
MHBA**



**Priv.-Doz. Dr. med.
Lukas Martin, MHBA**



Klinik für Operative Intensivmedizin
und Intermediate Care
Universitätsklinikum Aachen
Pauwelsstraße 30
52074 Aachen, Deutschland

Tel.: 0241 80-37606
Fax: 0241 80-80444

E-Mail: apeine@ukaachen.de
lmartin@ukaachen.de

ORCID-ID: 0000-0003-4163-2402
(A. Peine)

ORCID-ID: 0000-0001-8650-5090
(L. Martin)