

ÖFIT-Trendschau

Öffentliche Informationstechnologie in der digitalisierten Gesellschaft

Trendthema 45:

Neuronale Netze

Stand: Juli 2017



Herausgeber:

Mike Weber

Kompetenzzentrum Öffentliche IT

Fraunhofer-Institut FOKUS

Kaiserin-Augusta-Allee 31, D-10589 Berlin

Telefon: +49 30 3463 - 7173

Telefax: + 49 30 3463 - 99 - 7173

info@oeffentliche-it.de

www.oeffentliche-it.de

www.fokus.fraunhofer.de

Autorinnen und Autoren der Gesamtausgabe:

Mike Weber, Stephan Gauch, Faruch Amini, Tristan Kaiser, Jens Tiemann, Carsten Schmoll, Lutz Henckel, Gabriele Goldacker, Petra Hoepner, Nadja Menz, Maximilian Schmidt, Michael Stemmer, Florian Weigand, Christian Welzel, Jonas Pattberg, Nicole Opiela, Florian Friederici, Jan Gottschick, Jan Dennis Gumz, Jens Fromm, Dorian Grosch

Autorinnen und Autoren einzelner Trendthemen:

Michael Rothe, Oliver Schmidt

ISBN: 978-3-9816025-2-4

Juli 2017

Autorinnen/Autoren:

Dorian Grosch

Bibliographische Angabe:

Dorian Grosch 2017, Neuronale Netze, In: Mike Weber, Hg., 2016: ÖFIT-Trendschau: Öffentliche Informationstechnologie in der digitalisierten Gesellschaft. Berlin: Kompetenzzentrum Öffentliche IT, <http://www.oeffentliche-it.de/-/neuronale-netze>

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 3.0 Deutschland Lizenz (CC BY 3.0 DE) <http://creativecommons.org/licenses/by/3.0/de/legalcode>. Bedingung für die Nutzung des Werkes ist die Angabe der Namen der Autoren und Herausgeber.

Neuronale Netze

Mit künstlichen neuronalen Netzen wird Computern das Denken beigebracht – durch einen Lernprozess, der an den eines Kleinkindes erinnert. Anstatt Lösungen von Hand zu programmieren, lernt die Software durch Nachahmung der biologischen Prozesse im menschlichen Gehirn. Als der hellste Stern am Himmel der künstlichen Intelligenz verzeichnet diese Technologie aktuell große Fortschritte in der Forschung und Anwendung. Doch für welche Einsatzbereiche ist die Technik geeignet – und wo liegen die Grenzen einer lernenden Maschine?

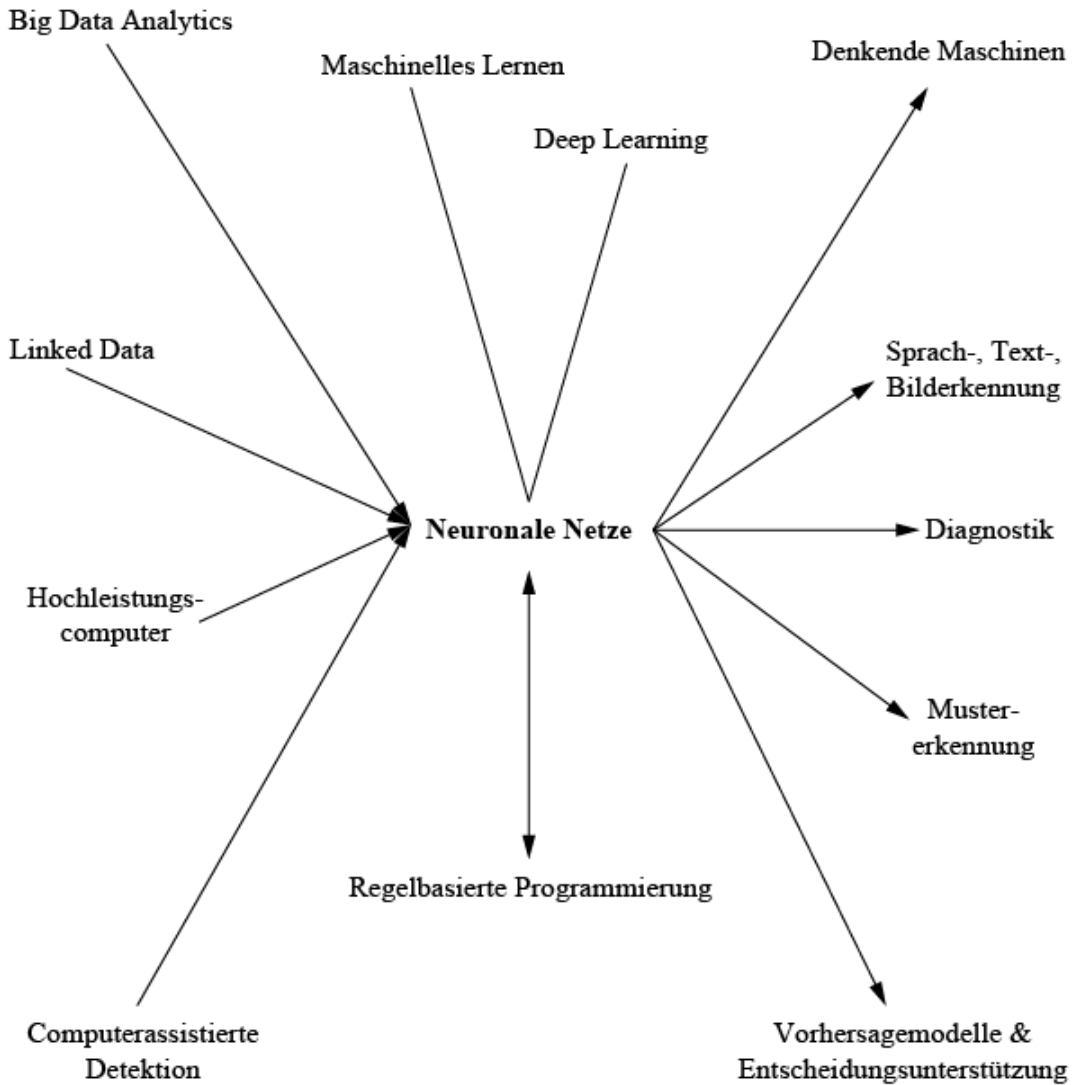
Rechensysteme mit Lernfaktor

Schon in den 60er Jahren gab es erste Versuche, künstliche neuronale Netze zu entwickeln, aber erst mit der steigenden Rechenkapazität und der höheren Verfügbarkeit der dafür notwendigen Datensätze ist die Technologie zur Marktreife gelangt.

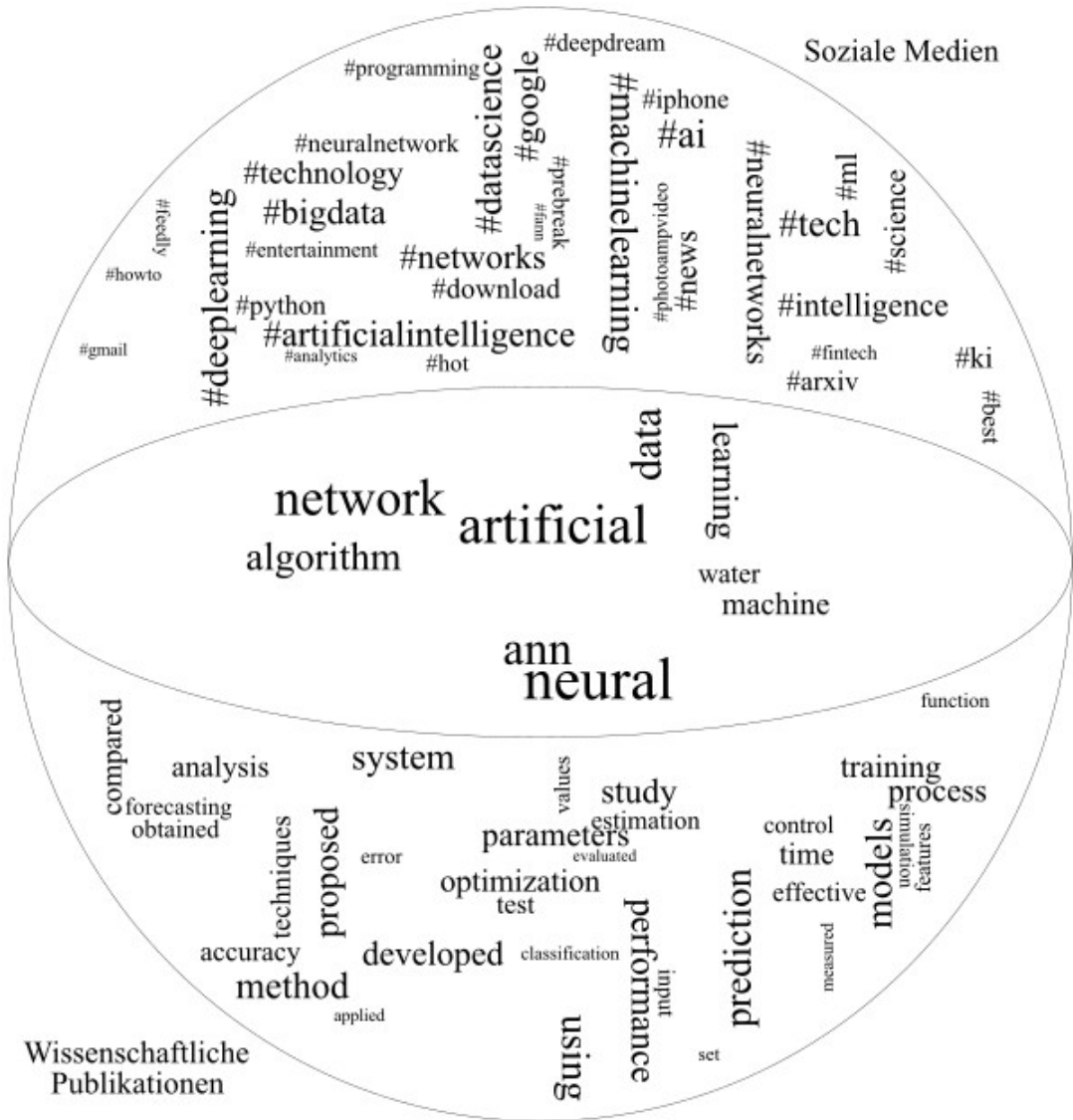
Neuronale Netze unterscheiden sich in ihrer Programmierform fundamental vom klassischen, regelbasierten Ansatz. Bei klassischen Computerprogrammen werden festgelegte Regeln Schritt für Schritt angewendet. Die Problemlösung ist also bereits in der Programmierung angelegt. Neuronale Netze werden demgegenüber so trainiert, dass sie die Lösungsstrategie auch dann entwickeln können, wenn diese den Entwicklern vorher nicht bekannt waren. Besonders bei der Sprach- oder Bilderkennung ist der klassische Ansatz problematisch, weil kein allgemein gültiges Prinzip zur Lösung des Problems existiert.

Handschrifterkennung bietet hier ein anschauliches Beispiel. Ein Mensch kann Handschriften prinzipiell genauso gut lesen wie gedruckte Schrift, weil die Konzepte von Buchstaben und Wörtern identisch sind. Für einen Computer gestaltet es sich ungleich schwerer, die als Pixel mehr oder weniger ausgeprägt erfassten Muster zu interpretieren. Die Erkennung wird zusätzlich durch individuelle Unterschiede in der Handschrift erschwert. Mit einem klassischen Ansatz ergäbe sich daraus ein immenses Regelwerk, dessen programmatische Umsetzung sehr aufwändig und fehleranfällig wäre. Ein neuronales Netz hingegen lernt selbstständig, den zusammenhängenden Pixeln Buchstaben zuzuordnen und daraus Worte zu konstruieren. Entsprechend wurde die Technologie schon früh etwa für den einfachen Anwendungsfall der Digitalisierung von Banküberweisungsformularen eingesetzt.

Begriffliche Verortung



Netzwerkartige Verortung des Themenfeldes



Gesellschaftliche und wissenschaftliche Verortung

Ein neuronales Netz besteht aus mehreren Schichten von miteinander verbundenen künstlichen Neuronen, die die Funktionsweise ihrer biologischen Gegenstücke simulieren sollen. Über Eingangsneuronen werden die zu analysierenden Daten in das neuronale Netz übertragen, das Ergebnis lässt sich an den Ausgangsneuronen ablesen. Zwischen Eingabe- und Ausgabeschicht können unterschiedlich viele innere Schichten aus verknüpften Neuronen zum Einsatz kommen. Für jedes künstliche Neuron wird ein Schwellwert definiert. Nur wenn dieser Wert überschritten wird, werden Signale an die anderen Neuronen weitergegeben. Zusätzlich werden die Verbindungen zwischen den Neuronen gewichtet. Abhängig von Art und Schwierigkeit der Aufgabenstellung werden verschiedene Arten von

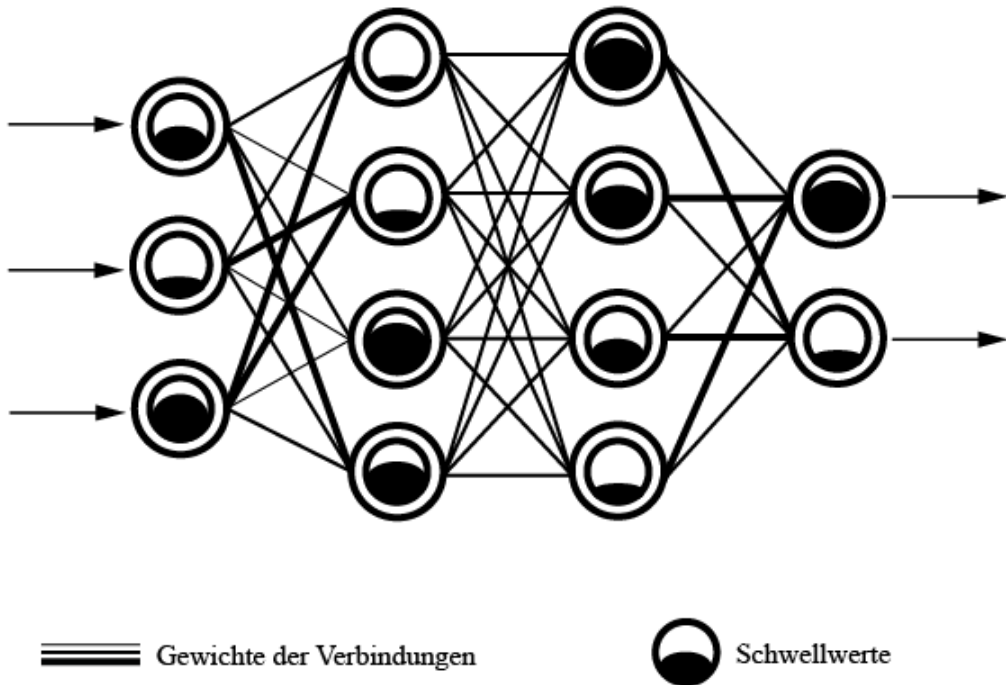
neuronalen Netzen eingesetzt, die sich in ihrem Aufbau und Anzahl der Neuronenschichten voneinander unterscheiden. Um ihre Aufgabe erfüllen zu können, müssen neuronale Netzwerke trainiert werden. Beispielsweise wird die Fähigkeit, einen Baum zu erkennen, anhand einer Vielzahl von Trainingsdaten in Form von beschriebenen Bildern mit und ohne Baum geübt.

Der Trainingsprozess läuft schrittweise ab. Am Anfang wird ein neuronales Netz mit einer Startkonfiguration aus zufälligen Belegungen für die Gewichte der Verbindungen initialisiert. Dann werden so lange der komplette Trainingsdatensatz durchlaufen, das Verhalten des Netzes analysiert und die Gewichte von einem Trainingsalgorithmus angepasst, bis das neuronale Netz den gestellten Anforderungen – etwa eine hinreichend hohe Trefferquote – gerecht wird. Dieses Trainingsverfahren benötigt in Abhängigkeit von der Schwierigkeit der Aufgabe von wenigen hundert bis zu mehreren hunderttausend Iterationen. Die aufwändige Prozedur kann jedoch teilweise oder komplett automatisch ablaufen und produziert in vergleichsweise kurzer Zeit flexible, anpassungsfähige und robuste Systeme, die vielfältig eingesetzt werden können. Die Entwicklung von neuronalen Netzen ist zudem leicht zugänglich, da viele kostenlose und sogar quelloffene Programmbibliotheken und Entwicklungsplattformen existieren, die von Universitäten und großen Softwarekonzernen entwickelt werden.

Eingabeschicht

Innere Schichten

Ausgabeschicht



Grundlegender Aufbau eines neuronalen Netzes

Blick in die Black Box

Eine Herausforderung bei der Anwendung neuronaler Netze liegt in der Auswahl der Trainingsdaten. Mit unvollständigen oder für die Anforderungen unpassenden Daten werden dem neuronalen Netz möglicherweise unerwünschte Problemlösungen antrainiert. Statt des Fotomotivs können beispielsweise Aspekte des Fotomaterials in die Bewertung eingehen, bspw. wenn Bäume häufiger auf Normal- statt auf Fotopapier gedruckt wurden. Die Trainingsdaten bestimmen das Verhalten des Netzes, deswegen besteht das Risiko, dass Verzerrungen und Vorurteile in den menschengemachten Daten angelernt werden. Außerdem kann ein neuronales Netz durch zu langes Training seine Generalisierungsfähigkeit verlieren und dann nur noch mit dem Trainingsdatensatz richtige Ergebnisse liefern. Generell

gilt zu berücksichtigen, dass die Trainingsdaten und Anforderungen gemeinsam immer nur ein Modell der realen Welt darstellen, das notwendig stark vereinfacht und möglicherweise fehlerhaft ist.

Mit Beginn der Trainingsphase spielen menschliche Faktoren nur noch eine geringe Rolle, da das Training automatisiert durchgeführt wird. Das Training erstreckt sich je nach Datensatz und Konzeption über wenige Tage bis zu mehreren Wochen. Dies stellt eine signifikante Beschleunigung in der Entwicklung von spezialisierten Anwendungen dar, bietet aber weniger Erfolgssicherheit als ein konventionelles Entwicklungskonzept.

Die bedeutendste Eigenschaft dieser Technologie ist gleichzeitig ihre größte Schwachstelle: Das neuronale Netz optimiert sich selbst in einem Prozess, der automatisch von Algorithmen gesteuert wird und für die menschlichen Entwickler nur konzeptuell transparent ist. Die konkrete Arbeitsweise bleibt undurchschaubar, eine sogenannte Black Box, die im Optimalfall genau das tut, was sie soll, aber ihr Arbeitsprinzip nicht preisgibt. Eine Funktionsanalyse wird durch die sich ändernden Systemzustände erschwert oder gar unmöglich (siehe [Denkende Maschinen](#)), was neuronale Netze für den Einsatz in kritischen Systemen disqualifiziert. Medizintechnische Software oder Steuersysteme von Flugzeugen und Kraftwerken erfordern eine Zertifizierung der Software, für die der Einblick in die Funktionsweise essenziell ist. Die Nachvollziehbarkeit der Ergebnisse neuronaler Netze ist ein offenes Forschungsfeld, wobei erste Analyseansätze sich beispielsweise auf die Visualisierung aktivierter Neuronen, für die die Schwellwerte überschritten wurde, beschränken.

Allzweckmittel zur Mustererkennung

Generell eignet sich die Technologie besonders für den Bereich der Mustererkennung, also der Bestimmung von Regelmäßigkeiten in Daten. Prägnante Beispiele sind Sprach-, Text- und Bilderkennung, aber auch bei spezialisierten Anwendungen wie Frühwarnsystemen, medizinischer Diagnostik, Aktienhandel und Wettervorhersage kann der Einsatz neuronaler Netze einen signifikanten Nutzen darstellen. Dabei können auch in Echtzeit eingehende Daten bearbeitet und ausgewertet werden. Kombiniert mit Big Data ergeben sich somit neue Möglichkeiten der Analyse und Verarbeitung großer Datenmengen. Neue Möglichkeiten entstehen insbesondere durch die Fähigkeit, auch in unstrukturierten Datenmengen Regelmäßigkeiten zu entdecken und so Erkenntnisse zu gewinnen, ohne dass diese vorher bekannt und als Erkennungsziel definiert waren. Das neuronale Netz sucht in diesem Fall nach gemeinsamen Mustern in den Daten und kann damit mögliche statistische Zusammenhänge erschließen.

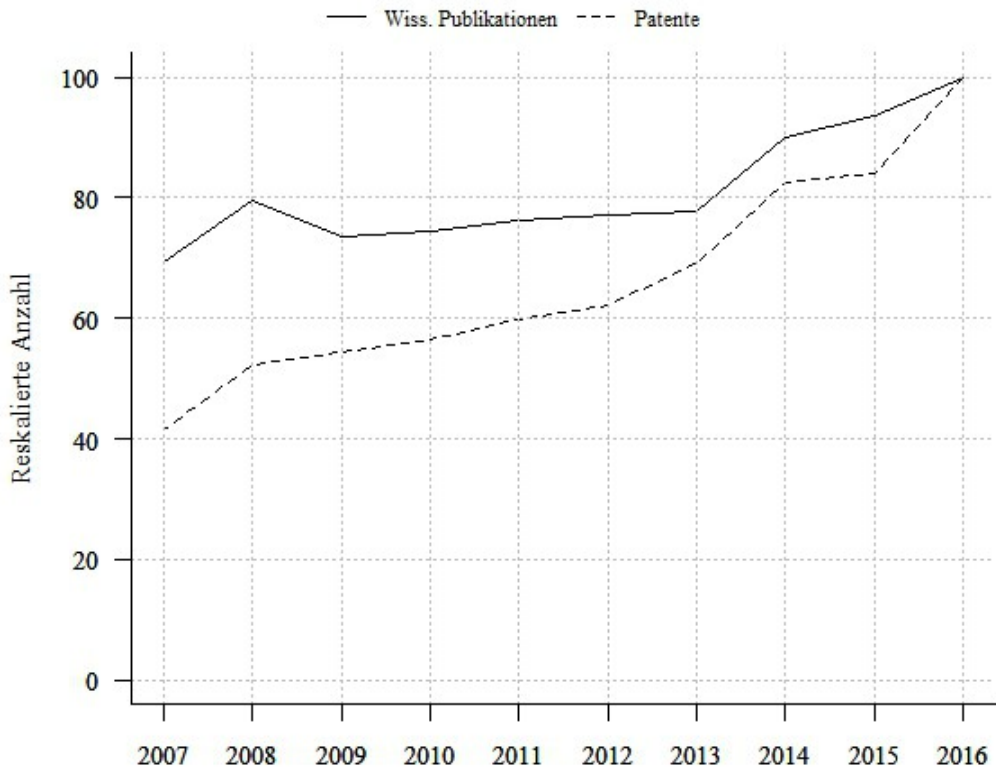
Noch ist das Potenzial von neuronalen Netzen nicht ausgeschöpft. Schon heute stecken sie in jedem Smartphone, bei der Stimm- und Handschrifterkennung oder zur Kategorisierung von Fotos auf der Speicherkarte. In Zukunft wird sich diese

Technologie durch immer größere Testdatenmengen und leistungsfähigere Rechner verbessern. Künstliche neuronale Netze können Entscheidungsunterstützung für Unternehmen und Verwaltungen (siehe [Verwaltung x.0](#)) bieten, aber auch für Wissenschaftler und Ärzte ein wertvolles Werkzeug darstellen sowie existierende technische Systeme intelligenter und leistungsfähiger machen.

Themenkonjunkturen



Suchanfragen und Zugriffe auf Wikipedia-Artikel



Wissenschaftliche Publikationen und Patentanmeldungen

Folgenabschätzung

Möglichkeiten

- Effiziente und anpassungsfähige Methode zur Mustererkennung
- Werkzeug zur effektiven Verarbeitung und Nutzung von Daten
- Fähigkeit zur selbstständigen Optimierung (s. [Selbstorganisation](#)) etwa bei Mustererkennung und Vorhersagemodellen (s. [Vorhersagende Polizeiarbeit](#))
- Schnellere Entwicklungszyklen möglich
- Erkennung von komplexen und neuen Zusammenhängen auch in unstrukturierten Daten

Wagnisse

- Überschätzung der Technologie aufgrund der Breite des theoretisch möglichen Anwendungsspektrums
- Fehlende Nachvollziehbarkeit und stark eingeschränkte Funktionsanalyse insbesondere bei Fehlfunktionen
- Technikgläubigkeit durch notwendiges Vertrauen in eine Black-Box
- Datengläubigkeit bei korrelierenden, aber nicht kausalen Mustern
- Subjektive Beeinflussung der Ergebnisse durch Realitätsmodell der Entwickler mit entsprechenden Diskriminierungsrisiken
- Hoher Bedarf an Trainingsdaten, die aufwendig vorbereitet werden müssen

Handlungsräume

Forschung fördern

In der Forschung beeinflussen sich Neurobiologie und Neuroinformatik gegenseitig. In beiden Forschungsfeldern besteht intensiver Forschungsbedarf. Besonders relevant für künstliche neuronale Netze ist dabei der Aspekt der Nachvollziehbarkeit der Abläufe innerhalb des Netzes.

Umsichtige Anwendung gewährleisten

Neuronale Netze können immer nur mit einer gewissen Wahrscheinlichkeit korrekte Ergebnisse erzielen. Ihre Einsatzmöglichkeiten sind daher insbesondere dann beschränkt, wenn Einzelpersonen direkt betroffen sind. Im Zweifel muss auf den Einsatz der Technologie verzichtet werden.