

Zeitschrift: Horizonte : Schweizer Forschungsmagazin
Herausgeber: Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung
Band: 29 (2017)
Heft: 113

Artikel: Augmented science
Autor: Fischer, Roland
DOI: <https://doi.org/10.5169/seals-821475>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 12.01.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Augmented Science

Künstliche Intelligenz wird ein immer stärkerer Teil des Forschungsprozesses. Noch verstehen die Forschenden nicht vollständig, was für einen Assistenten sie sich da ins Labor geholt haben. *Von Roland Fischer*

Intelligente Maschinen und selbstlernende Systeme beschäftigen die Forschung seit Jahrzehnten. Erste Berichte von Versuchen mit maschinellem Lernen bei der Identifizierung genetischer Muster wurden vor über 20 Jahren publiziert. Und in der Teilchenphysik wird schon so lange mit künstlicher Intelligenz (KI) experimentiert, dass manche Reviews um das Jahr 2000 sogar ein abflauendes Interesse feststellten. Sie drängten deshalb auf ein rasches Revival der Methode.

«Neuronale Netzwerke wurden schon in den 90er Jahren am Cern in diversen Experimenten studiert und eingesetzt», erinnert sich Sigve Haug vom Labor für Hochenergiephysik der Universität Bern. Man habe dies damals einfach nicht «Machine Learning» genannt.

Künstliche Intelligenz überall

Heute ist die Nutzung solcher KI-Methoden bei grossen Teilchenphysik-Experimenten schon fast der Normalfall, sei es in der Datenrekonstruktion oder der Datenanalyse. Und sie werden auch oft beim Distributed Computing genutzt, wo Programme lernen müssen, wann und wie Rechenprozesse am effizientesten verteilt werden. Aber KI ist nicht nur am Cern omnipräsent – ganz ähnlich präsentiert sich die Lage plötzlich so gut wie überall. Künstliche Intelligenz ist das Forschungs-Credo der Stunde. Physikalische Chemie, Molekularbiologie, medizinische Genetik, Astrophysik, sogar Digital Humanities: Wo mit grossen Datenmengen hantiert wird, ist auch KI nicht weit.

Ist die Entwicklung hin zur künstlichen Intelligenz als Laborassistent, hin zum Team von Mensch und Maschine in der Forschung also ein zwingender nächster Schritt? «Absolut», sagt Karsten Borgwardt vom Machine Learning and Computational Biology Lab der ETH Zürich. «In vielen Feldern der Lebenswissenschaften, wo mit Hochdurchsatzverfahren gearbeitet wird, geht es schon jetzt nicht mehr ohne.» Die Datenmengen seien schlicht zu gross, wenn man beispielsweise eine halbe Million Krankengeschichten mit den zugehörigen genetischen Daten verknüpfen

möchte. «Da kann kein Mensch mehr mit dem blossen Auge ein sinnfälliges, noch unbekanntes Muster erkennen.» Solchen Datenmengen sei nur noch mit effizienten statistischen Verfahren, wie Spezialisten wie Borgwardt sie derzeit entwickeln, beizukommen. Ohnehin sei die Grenze zwischen Statistik und maschinellem Lernen inzwischen fließend, merkt der junge ETH-Professor an.

Die künstliche Intelligenz als selbstverständlicher Partner im Forschungsprozess: Die Vision erinnert an Garry Kasparovs «Advanced Chess»-Idee, die er kurz nach seiner Niederlage gegen Deep Blue vor genau zwanzig Jahren vorbrachte. In Zukunft sollten nicht mehr Menschen gegeneinander oder gegen Maschinen antreten, sondern Mensch-Maschine-Teams. So könnte das Spiel auf eine ganz neue Ebene gehoben werden, glaubte Kasparov, ein Schach jenseits menschlicher strategischer Möglichkeiten.

Ein System kann
«überlernen»: Mit
zunehmendem Training
wird es schlechter.

«Machine Learning is the scientific method on steroids», schreibt auch der KI-Experte Pedro Domingos der Universität von Washington in seinem Buch «The Master Algorithm». Darin postuliert er so etwas wie eine Super-Machine-Learning-Methode, ein maschineller Allesverstehender. Durch intensiv eingesetzte KI würde die Forschung schneller, effizienter, tiefschürfender. Der Wissenschaftler könnte sich, von der statistischen Routine entlastet, ganz auf die kreativen Anteile seiner Arbeit konzentrieren. Domingo verspricht nichts weniger als eine neue goldene Ära der Wissenschaft.

Nicht alle Forscher, die sich mit KI beschäftigen, mögen in dieses Loblied einfach so einstimmen. Neven Caplar vom Institut für Astronomie der ETH Zürich ist durchaus ein Datennerd: Er betreibt den

Datenblog astrodataiscool.com und hat kürzlich mithilfe von maschinellem Lernen den Gender Bias in astronomischen Publikationen quantifiziert.

Der junge Forscher sieht zwar seit ein paar Jahren einen markanten Aufschwung bei Publikationen, die KI enthalten. Aber er hat doch einige Zweifel, ob die Methoden in seinem Feld einen grossen Durchbruch erleben werden. Astronomie sei «eine Wissenschaft von Biases»; es gehe nicht zuletzt darum, die Instrumente möglichst gut zu kontrollieren. Deshalb sollten sie so wenig wie möglich konzipiert werden wie eine «Black Box»: wie ein praktisches Tool, das zwar gute Resultate liefert, dessen genaue Funktionsweise man aber nicht nachvollziehen kann. Beim Umgang mit den Beobachtungsdaten sei die Interpretation durch den menschlichen Forscher nach wie vor zentral, sagt Caplar.

Gedopte Wissenschaft

«Ach, diese Black Box!» entfährt es seinem Kollegen Kevin Schawinski (siehe auch «Die Physik von allem», S. 30). Alle würden davon reden, dass KI eine Black Box sei – und wir also die Logik und Argumentation der Maschine gar nicht zu hinterfragen vermögen. Der Astronom Schawinski kann bei KI keine Black Box erkennen. Aus seiner Sicht handelt es sich schlicht um eine neue Methode der Forschung, die kalibriert und getestet werden muss, um sie richtig zu verstehen. Das sei nicht anders als bei jeder Methode, die sich die Wissenschaft neu zu eigen gemacht hat, sagt der Forscher. Schliesslich durchschaue kein Einzelner komplexe Versuchsanordnungen wie den Large Hadron Collider am Cern oder das Hubble-Teleskop als Ganzes. Schawinski vertraut da ebenso auf die Forschungs-Community: Diese wisse sicherzustellen, dass der wissenschaftliche Prozess robust funktioniert.

Mit Kollegen von den Computerwissenschaften hat Schawinski die Plattform space.ml lanciert, eine Sammlung von leicht zu bedienenden Tools zur Interpretation von astronomischen Daten. Er selber hat eine Methode entwickelt, um Bilder von Galaxien durch ein neuronales

«Maschinelles Lernen ist auf einer ewigen Gratwanderung zwischen Blindheit und Halluzination.»

Pedro Domingos

Netzwerk so zu verbessern, dass mehr Informationen extrahiert werden können, und zwar ohne dass der Computer dazu weitere Vorgaben braucht. Bei anderen Anwendungen kommt auch sogenanntes überwachtes Lernen zum Einsatz, bei dem auf ein Daten-Trainingsset zurückgegriffen wird: Von Menschen vorsortiert oder mit Metainformationen versehen, helfen die Trainingssets dem Computer, selber Regeln aufzustellen, die ihn befähigen, eine Aufgabe zu bewältigen.

Überangepasste Assistenten

Der Biostatistiker Karsten Borgwardt nutzt Überwachtes-Lernen-Methoden, um beispielsweise herauszufinden, ob Veränderungen im Genom eine schädliche Wirkung im Organismus haben. Dabei füttert er den Computer mit bereits erforschten Mustern und hofft, dass dieser in der Folge selbstständig weitere, noch unbekannte Zusammenhänge findet.

«Die Interpretation durch den menschlichen Forscher ist bei Beobachtungsdaten nach wie vor zentral.»

Neven Caplar

Doch es gibt einen Stolperstein: das Overfitting. Der Datenexperte muss im Auge behalten, ob der Computer im Trainingsset tatsächlich die wesentlichen Merkmale erkennt oder ob er sich zufällige Muster im Datenrauschen zur Regel macht. Auch das hat Pedro Domingos auf eine prägnante Formel gebracht: Er sieht maschinelles Lernen auf einer «ewigen Gratwanderung zwischen Blindheit und Halluzination». Auf der einen Seite erkennt der Algorithmus in den Datenbergen gar nichts. Auf der andern Seite des Overfitting beginnt er plötzlich Dinge zu sehen, die gar nicht da sind. Auf diese Weise kann man ein System tatsächlich «überlernen» - mit zunehmendem Training wird es dann wieder schlechter.

Ein Hauptgrund für Overfitting in der Genomik und Medizin ist laut Borgwardt, dass hinsichtlich des Trainingssatzes nicht immer die nötige Transparenz herrscht und deshalb nicht immer einschätzbar ist, wie stark die Trainingsdaten mit den

später auszuwertenden Daten überlappen. Wenn die Sets zu ähnlich sind, kann die Maschine im schlechtesten Fall nicht «generalisieren» und wendet einfach auswendig gelernte Fälle wieder an, falls sie eine Übereinstimmung findet. So entstehe kein echter Erkenntnisgewinn; die künstliche Intelligenz bleibt auf dem Stand einer gewöhnlichen Datenbank.

Auch wenn beim Training alles richtig gelaufen ist: Es bleibt das Problem, Zufallskorrelationen von echten, statistisch signifikanten Zusammenhängen zu unterscheiden. Je grösser die Datenmengen, desto grösser sei auch die Wahrscheinlichkeit, dass bloss aus Zufall gehäufte Genomvarianten vorkommen, sagt Borgwardt - und dass diese dann auch noch mit dem Auftreten einer Krankheit korrelieren. Ein wichtiger Teil seiner Arbeit besteht deshalb darin, Signifikanz in extrem hochdimensionalen Räumen zu bewerten. Das heisst: auch sehr komplexe und naturgemäss multikausale Situationen statistisch zu meistern.

Maschinelles Lernen für Quanten

Die KI kann nicht nur Zusammenhängen in komplexen Datensätzen nachspüren, sie kann auch ganz neue Materialien kreieren. Doch im Gegensatz zu den Lebenswissenschaften sei maschinelles Lernen in der physikalischen Chemie noch nicht sehr weit verbreitet, sagt der Chemiker und Materialforscher Anatole von Lilienfeld von der Universität Basel. Immerhin sieht auch er einen «rapiden Aufschwung» und glaubt, dass KI «unvermeidlicherweise» in zehn Jahren ein integraler Teil der meisten Studien sein wird.

Die Arbeit seiner Gruppe hat Pioniercharakter: Dank KI konnten die Forscher die Eigenschaften von Millionen theoretisch möglicher Kristalle durchrechnen, die aus vier bestimmten Elementen aufgebaut sind. Dabei identifizierte die KI neunzig bisher unbekannte Kristalle, die thermodynamisch stabil sind und als neuartige Werkstoffe in Betracht kommen. Die Effizienzsteigerung bei der Berechnung der Kristalleigenschaften - mehrere Größenordnungen schneller - verblüfft selbst den Fachmann: Sie ist so enorm, dass «nicht nur konventionelle Probleme gelöst werden können, sondern sich auch ganz neue Forschungsfragen auftun». Aber auch von Lilienfeld gibt ein paar Punkte zu bedenken: Maschinelles Lernen funktioniert

nur, wenn ein Ursache-Wirkung-Prinzip vorliege und genügend Daten zur Verfügung stünden. Entscheidend sei auch, dass der entsprechende Forscher «genügend Expertise besitzt, um effiziente Repräsentationen der zu untersuchenden Objekte und ihrer Eigenschaften zu ersinnen.»

Diese Expertise war auch für den theoretischen Physiker Giuseppe Carleo von der ETH Zürich der entscheidende Punkt. Er hat zusammen mit Kollegen einen Weg gefunden, die Wellenfunktion von Quantensystemen mit einem neuronalen Netz nachzubilden. Nachdem ihm dieser Schritt gelungen war, war die Optimierung der Wellenfunktion «eigentlich ein Kinderspiel». Der Algorithmus erledigte die Aufgabe schnell und ohne Probleme. Herkömmliche Methoden stossen bei solchen Rechenaufgaben rasch an ihre Grenzen: Die Simulation von komplexeren Quantensystemen galt bis vor kurzem noch als rechnerisches «Ding der Unmöglichkeit».

Der neue Ansatz setzt auf die Methode des unüberwachten Lernens (Unsupervised Learning), ohne Vorwissen für den Computer. Er sei für theoretische Physiker auch deshalb interessant, weil er ermögliche, «alte Probleme in neuen Perspektiven» zu sehen - im besten Fall profitierten die Ingenieurwissenschaften und die Grundlagenforschung von den Fortschritten.

Der Physiker hat sich vom letztjährigen Triumph des Algorithmus AlphaGo beim Spiel gegen einen Meister des Go-Spiels inspirieren lassen. Diese KI wurde immer stärker, indem sie unzählige Partien gegen sich selbst spielte. Carleo hat die Methode des bestärkten Lernens (Reinforced Learning), die die Spielintelligenz von AlphaGo in ganz neue strategische Sphären gehoben hat, für seine Zwecke adaptiert.

Er hat der Maschine im übertragenen Sinne beigebracht, die Suche nach der Lösung der Wellenfunktion wie ein Spiel zu sehen, bei dem das Ziel zwar klar ist, der Weg aber ganz offen - und dabei zu lernen, gute Lösungswege zu bevorzugen. Und tatsächlich hat die KI an dem Spiel sehr grossen Gefallen gefunden. So grossen Gefallen, dass sie es nun so gut beherrscht wie keine andere Intelligenz auf der Welt.

Roland Fischer ist freier Wissenschaftsjournalist in Bern.



Erhält der anthropomorphe Roboter einen Geburtstagskuchen? Zentraler ist wohl die Frage, ob er sich über das Geschenk freuen wird. Im Bild ist übrigens eine echte Frau, die als Roboter menschliches Verhalten imitiert.

Bild: Kevin Grennan, Android Birthday, 2011. Video.

Schauspielerin: Sylvi Kim