

Wenn Computer denken, sie denken

Die Neuroinformatik ist eine angewandte Wissenschaft, die viele Fachgebietsgrenzen überschreitet. Sie entwickelt mathematische Modelle und praktisch nutzbare Werkzeuge mit IT-Methoden, hat aber mit der Neurologie recht wenig gemeinsam.

Wer denkt, er könne das Denken eines Tages den Computern überlassen, weil sie größere Speicher haben, der irrt. Selbst Studenten der Neuroinformatik müssen Bio und Mathe büffeln, ehe sie eigene Geistesblitze entwickeln dürfen. Computer können zwar gewaltige Datenmengen speichern und diese auch um Größenordnungen schneller verrechnen als menschliche Nervenzellen, aber ihnen fehlt nahezu alles, was die Leistungsfähigkeit eines menschlichen Gehirns ausmacht: Erinnerung an bereits Erlebtes, emotionale Bewertung von Informationen und vor allem evolutionäre Anpassungsfähigkeit, die aus dem Willen zum Überleben resultiert.

Allerdings beschreibt das Institut für Neuroinformatik der Ruhr-Universität Bochum im Internet ganz klar, welche Träume die Forscherherzen höher schlagen lassen: „Ziel ist es, die Organisationsprinzipien des Nervensystems bis zur informationstechnischen Anwendbarkeit aufzuarbeiten.“ Das heißt nichts anderes, als dass die Wissenschaftler die Leistungen biologischer Nervensysteme in der künstlichen Siliziumwelt der Computer so naturgetreu wie möglich nachbilden wollen.

Wenn man als biologische Vorbilder einfache Strickleiter-Nervensysteme von Ringelwürmern und Stummelfüßlern auswählt, dann kommt man in der Tat mit zwei- oder dreischichtigen neuronalen Netzwerkmodellen aus und ist umso erstaunter, zu welchen Leistungen diese bereits fähig sind. Sie können zum Beispiel Gesichter in einer Menschenmenge erkennen oder verschiedene Krebsformen anhand von Biochipdaten über deren Genexpression unterscheiden.

Methodisch kommen bei genauerem Hinsehen allerdings kaum biologische als vielmehr mathematische Modelle der Nervennetze zum Tragen. Herbert Jaeger, Professor für Computational Science aus Bremen, gibt in der Zeitschrift *Technology Review* vom November 2006 offen zu: „Seit den Anfängen künstlicher Neuronen-

Netze haben Ingenieure weder Ahnung von der Biologie, noch richten Sie sich bewusst daran aus.“ Die Hauptaufgaben der Neuroinformatik sind derzeit vor allem technischer Art, z.B. Robotik, Bildverarbeitung oder Klassifizierung.

In den Urzeiten des Fachs, die gerade einmal 50 Jahre zurückliegen, stand allerdings die Erzeugung von *Künstlicher Intelligenz* ganz oben auf der Wunschliste und wurde mit gewaltigen Forschungsgeldern des amerikanischen Verteidigungsministeriums gefördert. Man wollte Maschinen entwickeln, die sich intelligent verhalten, indem sie zum Beispiel einen menschlichen Schachweltmeister besiegen. Dieses Problem lösen Computer seit 1996, als Garri Kasparow einen legendären Kampf gegen Deep Blue von IBM verlor, fast schon routinemäßig - zuletzt Anfang Dezember wieder, als Wladimir Kramnik gegen Deep Fritz aus Hamburg eine Niederlage kassierte.

Das hat jedoch mit Computer-Intelligenz nichts zu tun, sondern mit Rechenleistung, weil Computer bei der Berechnung von Millionen denkbarer Spielzüge dem Gehirn des Schachspielers so haushoch überlegen sind wie Baggergreifer der Hand eines Pianisten. Aber genau in dieser Stärke liegt auch ihre Schwäche: Computerprogramme denken immer nur das Eine, Gehirne können sich (fast) jedem Kontext anpassen.

Mit dem komplizierten Geflecht der rund zehn Milliarden Nervenzellen im Gehirn haben künstliche neuronale Netze so gut wie nichts gemeinsam. Jede einzelne Nervenzelle entspricht in ihrer Komplexität - wenn ein Vergleich von Biologie und Technik überhaupt zulässig ist - einem ganzen Computer. Insofern ist allenfalls das Internet mit seinen zahllosen angeschlossenen Computern ein künstliches neuronales Netz.

Deshalb versuchen Neuroinformatiker auch, Strategien zu entwickeln, um wenigstens ein paar Tausend parallele „Nanofestplatten“ auf einer Leiterplatte unterbringen zu können. Von zehn Milliarden Computern, die wie im Gehirn in einem kugelförmigen Gebilde von 2,5 kg Gewicht untergebracht werden könnten, träumen derzeit aber noch nicht einmal Utopisten.

Eine Nervenzelle enthält in der Tat Nanostrukturen (Nukleinsäuren, Prote-

ine und Lipide), die gewissermaßen den Strukturelementen auf den Computerplatinen entsprechen, sowie ionisierte Moleküle, die den Ladungszustand der Speicher-elemente im Computer repräsentieren.

Entscheidend aber für das Funktionieren biologischer Neuronennetze ist die interne Kommunikation mit Hilfe von Neurotransmittern, die an den Synapsen ausgeschüttet werden und somit eine Analogie zu Lötstellen, Verbindungskabeln oder drahtlosen Kommunikationselementen darstellen. Zu diesem elektrochemischen „Gedankenaustausch“ zwischen Nervenzellen gibt es in der Computerwelt keine wirkliche Analogie: Sie beschränkt sich auf elektrische Ladungszustände von Transistoren und Kondensatoren; Biochemie ist dabei nicht im Spiel.

So bleibt die Frage offen, ob die mathematische Modellierung biologischer

Sehende Roboter



Die Neuroinformatik erforscht u.a. autonome Roboter, die sich mit Hilfe von Videokameras und anderen Sensoren frei im Raum orientieren und bewegen können. Der klassische Ansatz benützt ausgeklügelte Programme, bei deren Design möglichst viele potenzielle Situationen

im Voraus berücksichtigt werden. Attraktiver sind Neuronale Netze, die aus Fehlversuchen lernen können. Ihre Implementierung benötigt zwar mehr Zeit, da die Roboter vor ihrem Einsatz erst trainiert werden müssen, aber dafür sind sie in unübersichtlichen Entscheidungssituationen weniger störanfällig. Ein Einsatzgebiet in der Medizin sind Geh- und Fahrhilfen für Behinderte und alte Menschen. Quelle: www.pr.uni-freiburg.de

Vorgänge dem forschenden Geist jemals Aufschlüsse über seine eigene Funktionsweise geben kann oder ob sie nur deshalb zu (scheinbar) logischen Resultaten führt, weil diese Logik genau diesem Geist entspringt. Dies ist allerdings kein Problem für Neuroinformatiker, sondern für Philosophen. ■

gh

Mustererkennung in der Krebsforschung

Am Beispiel von Biochipdaten wird gezeigt, wie man mit künstlichen Neuronalen Netzen Krankheitsstadien klassifizieren kann.

Die Identifizierung von Mustern in der Fülle ungeordneter sensorischer Reize ist eine der überlebenswichtigen Fähigkeiten des Gehirns. Einige Tricks der Natur lassen sich mit Künstlichen Neuronalen Netzen (KNN) für die maschinelle Mustererkennung nutzbar machen, was vor allem in der molekularen Medizin zunehmend Bedeutung gewinnt: Mit Biochips und anderen hochparallelen Techniken kann man nämlich gewaltige Datenmengen produzieren, aber nicht unbedingt immer sinnvoll deuten.

Die Abbildungen in der rechten Spalte zeigen eine komplette Datenauswertung am Beispiel synthetischer Genexpressionsdaten eines Krebsmodells¹⁾. Sie

und Abbau. Ausgehend von einem Datensatz für ein Normalgewebe wurden diese Konzentrationen mit einem pathophysiologisch begründeten Zufallsmechanismus verändert, so dass der Grad der Entartung von links nach rechts zunahm. Ein zeitlicher Sprung nach Experiment 4 sorgte dafür, dass zwei Klassen von frühen bzw. späten Krebsstadien entstanden (oberes Bild).

Anschließend wurden die Daten zur besseren Visualisierung spalten- und zeilenweise so umsortiert, dass Experimente und Gene mit ähnlichem Expressionsmuster nebeneinander zu liegen kommen (Clustering). Man erkennt jetzt deutlich die zwei Klassen.

Im letzten Schritt wurden die Daten für das Neuronale Netz digitalisiert, da KNN nur Werte von 0 und 1 als Eingangssignale empfangen können. Dies geschieht in der Diagnostik meist durch Bildung diskreter Wertebereiche wie z.B. normal oder

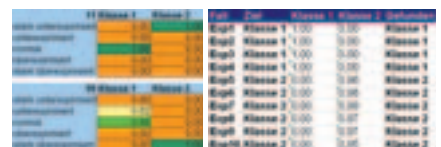
Komplette Datenanalyse für 10 Genexpressions-Experimente mit insgesamt 100 Sonden von der Transformation der Rohdaten über das Clustering bis zum Training eines künstlichen Neuronalen Netzes mit korrekter Reklassifizierung früher und später Tumorstadien.



Transformierte Originaldaten: Jede Zeile repräsentiert ein Gen, jede Spalte ein Experiment. Alle Signale wurden durch ein normales Referenzsignal dividiert. Ein Wert von 1,00 entspricht somit einer normalen Expression. Blaue und rote Farben markieren unter- und überexprimierte Werte.



Clustering der Daten: Man erkennt zwei Klassen von Experimenten (Spalten) und drei Klassen von Genen (Zeilen) mit unterschiedlichem Verhalten bei frühen und späten Tumorstadien.

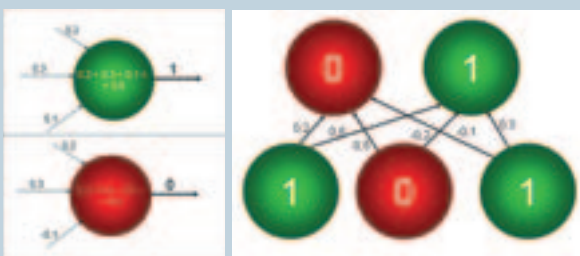


Trainiertes Neuronales Netz (Ausschnitt): Beispielhaft sind links zwei Gene dargestellt: Gen 11 ist mit einem Gewicht von jeweils 1,0 bei Klasse 1 normal und bei Klasse 2 stark unterexprimiert. Gen 59 (siehe auch Abb. 2) ist mit einem Gewicht von 0,62 bei Klasse 1 unterexprimiert (hellblau) und mit Gewicht 0,88 normal. Bei Klasse 2 ist es mit einem Gewicht von 1,00 stark überexprimiert (rot).

In der rechten Bildhälfte ist das Ergebnis der Reklassifizierung gezeigt: Alle 10 Experimente wurden mit Gewichten zwischen 0,95 und 1,00 der richtigen Zielklasse zugeordnet.

¹⁾ Gefördert von der Stiftung Pathobiochemie und Molekulare Diagnostik der DGKL e.V. (www.dgkl.de).

Biologische Logik



1958 publizierte der Psychologe Frank Rosenblatt das „Perzeptron“ (Abb. rechts), das Grundlage der künstlichen Neuronalen Netze werden sollte. Es besteht aus mehreren Lagen solcher Neuronen: Die untere Schicht sendet Signale von 0 bzw. 1 an die nächst höhere. Jedes Signal

wird mit einem Gewicht multipliziert, das die Synapsenstärke symbolisiert. Die Gewichte werden in einem „Lernprozess“ schrittweise so angepasst, dass sie jedes beliebige Eingangsmuster in ein definiertes Ausgabemuster überführen.

Auch Rosenblatt versuchte, mit seinem Modell Probleme der Logik zu lösen, was sich rückblickend als Irrweg herausstellte. Für die maschinelle Mustererkennung waren seine genialen Arbeiten jedoch eine wichtige Voraussetzung.

1943 führten die Mathematiker McCulloch und Pitts den Begriff des „Neurons“ in die Informatik ein (Abb. links). Wenn die Summe der Eingangssignale in diesem Computermodele einen Schwellenwert von beispielsweise 0,5 überschreitet, sollte es eine 1 als Äquivalent des Aktionspotenzials senden, sonst eine 0. Das Ziel bestand darin zu testen, ob ein Computer nach dem Vorbild des Gehirns „logisch denken“ kann.

wurden mit dem Programm SimChip (www.simchip.de) erzeugt. Der Vorteil solcher „in-silico“-Daten im Vergleich zu in-vivo-Ergebnissen liegt in ihrer Vorhersagbarkeit zur Plausibilisierung der Algorithmen.

Das Krebsmodell erzeugt Genexpressions-signale, also mRNA-Konzentrationen im Fließgleichgewicht zwischen Synthese

stark überexprimiert. Das Netz erhielt die beiden Klassen als Zielwerte vorgegeben (supervised learning) und setzte in der Tat die Gewichte so, dass anschließend alle gewünschten Zielklassen korrekt gefunden wurden (Bild unten). Mit dem so trainierten Netz können jetzt unbekannte Fälle klassifiziert werden. ■

gh