Wissen, erkennen, entscheiden

Blockchain, Internet of Things, Big Data, Artificial Intelligence (AI) sind Schlagwörter, wenn es um Computer und Digitalisierung geht. Über AI wird bereits seit Jahrzehnten geforscht und diskutiert.

Artificial Intelligence (Künstliche Intelligenz) ist eine Automatisierung intelligenten Verhaltens sowie maschinellen Lernens. Eine einzige Technologie, die künstliche Intelligenz ist, gibt es nicht, sondern es gibt unterschiedliche Technologien, die unter künstlicher Intelligenz zusammengefasst werden.

Die Abgrenzungen verschieben sich dabei laufend. Künstliche Intelligenz befasst sich heute mit den Funktionsbereichen Wissen, Erkennen, Entscheiden, sowie der Synthese menschlicher Ausdrucksformen, wie Sprache. Wenn heute diese Teilaspekte beispielsweise in digitalen Assistenten wie *Siri* oder *Alexa* zusammenkommen, spielen unterschiedliche Technologien mit, die auf viele Jahrzehnte Entwicklung und Forschung aufbauen.

Schwerpunkt Wissen. Sammlung, Repräsentation und Abruf von Wissen ist einer der ältesten Teilbereiche der künstlichen Intelligenz. Die Entwicklung von Expertensystemen begann in den 1960er-Jahren, seit den 1980er-Jahren gibt es kommerzielle Produkte als Expertensysteme. Das Ziel der Expertensysteme ist die Unterstützung bei der Lösung komplexer Probleme, die ein hohes Maß an Wissen über Inhalte und Zusammenhänge erfordert. Die Repräsentation von Wissen, eine strukturierte Wissensbasis, stellt dabei eine der großen Herausforderungen dar, denn es geht nicht nur um Informationen, sondern um deren Gewichtung und Beziehung und schließlich um

Roboter-Butler in Hotels führen Gäste zu ihren Zimmern.

Schlussfolgerungen, die ein Expertensystem anbieten soll. Expertensysteme können auf Einzelfällen (Beispielen) aufbauen, und damit über ein Ähnlichkeitsprinzip zu Ergebnissen führen, oder über Regeln aufgebaut werden, die auch als Entscheidungsbäume aufgebaut sein können. Die Regeln und ihre Hierarchien müssen direkt von menschlichen Experten in die Systeme eingepflegt werden. Während ursprünglich Expertensysteme typischerweise in einen Dialog mit dem Anwender verwendet wurden, der selbst meist Experte war, sind Expertensysteme heute oft wesentlicher Teil komplexer Entscheidungs- und Steuerungssysteme, sei es in medizinisch-diagnostischen Anwendungen, der Steuerung von technischen Anlagen oder der Kreditvergabe. Die kanadische Polizei setzt bereits seit den 1980er-Jahren ein Expertensystem ein, um einzelne Straftaten eventuell als Serie zu erkennen: *ViCLAS (Violent Crime Linkage Analysis System)*. Die Polizei in Österreich verwendet *ViCLAS* seit Mitte der 1990er-Jahre.

Schwerpunkt Erkennen. Das Erkennen von komplexen, unerwarteten, immer wieder neuen Situationen ist eine Herausforderung, die über die Möglichkeiten eines Expertensystems hinausgeht. Die Erkennung unvorhersehbarer Situationen ist die Voraussetzung für viele Entscheidungen, die bis vor Kurzem alleinige Domäne von Menschen waren, die bestenfalls von Expertensystemen unterstützt wurden.

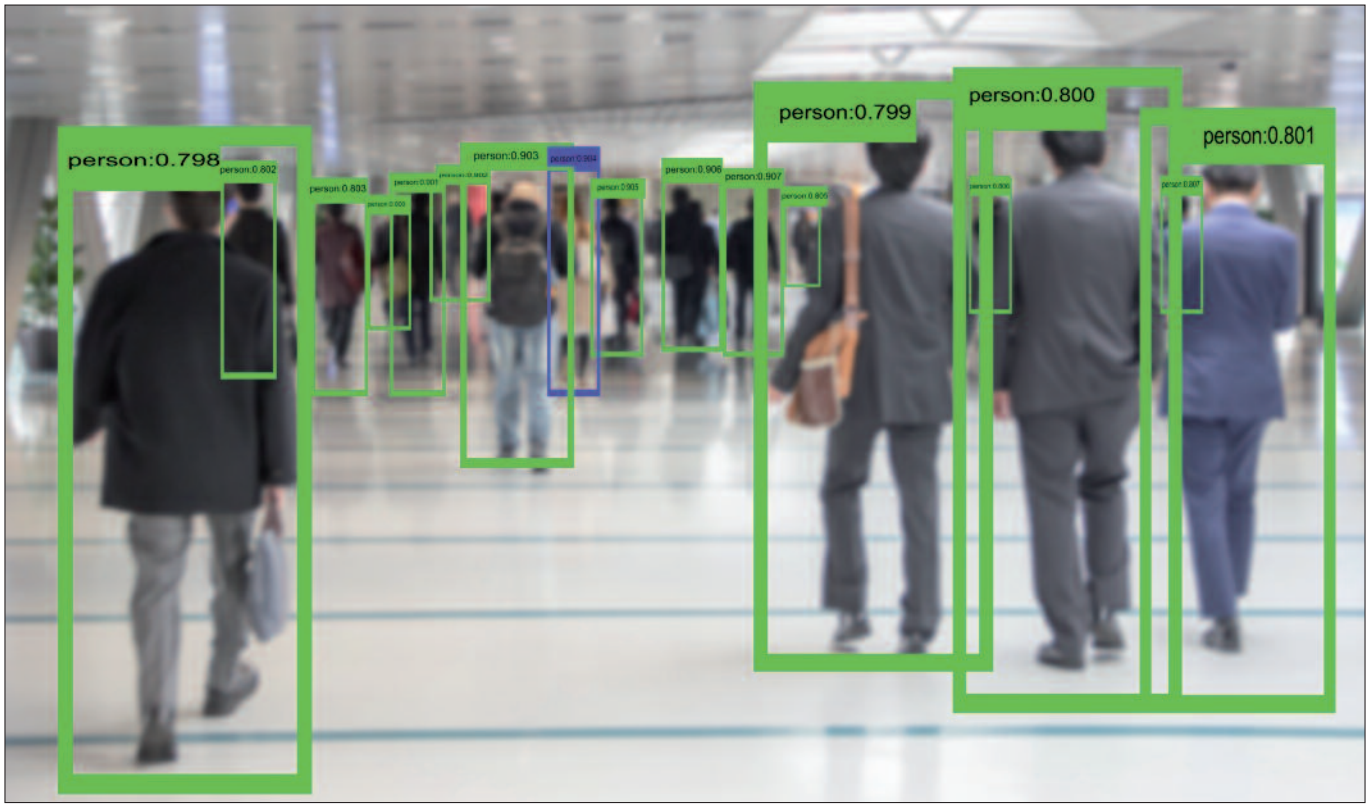
Neue Wege. Die klassischen Systeme der künstlichen Intelligenz (KI) bauen auf sogenannter „Symbolischer KI“ auf: Mit begrenztem, ausgewähltem Wissen werden Aussagen gefällt und Lösungen gefunden, um Entscheidungen zu fällen. Diese Form der KI funktioniert „Top-down“, indem sie aus Heuristik, Statistik, mathematischer Optimierung und Näherungsverfahren Aussagen trifft. Ein gegensätzlicher Ansatz ist aus der Biologie und Gehirnphysiologie abgeleitet: Die „neuronale KI“, die sich als „Bottom-up“ versteht. Dabei

EXPERTENSYSTEME

Toxische Algorithmen

Expertensysteme unterstützen bei Analyse und Entscheidung und haben das Potenzial zu einer Beschleunigung und Verbesserung von Entscheidungsprozessen. Sie haben auch problematische Aspekte: Sie fördern die Delegation von Entscheidungen und das blinde Vertrauen in Vorgaben durch ein System, dem hohe Autorität zugemessen wird. Anders als digitale Expertensysteme stehen menschliche Experten miteinander in Diskurs und können beispielsweise anhand von Sonderfällen die Bedeutung von Entscheidungsgrundlagen hinterfragen und anpassen. Da Expertensysteme für ihre Anwender meistens eine „Blackbox“ sind, also in ihrer Funktionsweise undurchschaubar, werden systematische Fehler, falsche Schlüsse oder Fehlinterpretationen nicht erkannt. Für dieses Problem hat sich auch der Begriff der „toxischen Algorithmen“ entwickelt.

Diesem Thema und seinen schwerwiegenden Folgen widmet sich beispielsweise Cathy O’Neil in ihrem Buch „Angriff der Algorithmen: Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden“ *(Carl Hanser Verlag, www.hanser.de)*.



Deep Learning kann nur in jenem Umfang sinnvolle Ergebnisse liefern, als ausreichend Übungsmuster geboten werden.

werden Strukturen, die aus dem Nervensystem und Gehirn bekannt sind, mit technologischen Mitteln nachgebaut, beziehungsweise simuliert. In den 1990er-Jahren wurde im Rahmen von sinnesphysiologischen Forschungen anhand der Nervenverbindungen des Auges die Grundlage von Mustererkennung entdeckt, die sich bereits in den Schaltplänen der Neuronen im Auge ergibt und im Gehirn spezifisch weiterverarbeitet wird.

Es war naheliegend, diese Schaltungen zu simulieren und einfache Mustererkennung durchzuführen. Doch die technischen Voraussetzungen waren damals nicht gegeben, komplexe Systeme mit ausreichender Breite und Tiefe zu bauen. In den letzten zehn Jahren wurden diese Grundlagen geschaffen und neuronale Netze in ausreichender Tiefe sind heute technisch möglich und auf breiter Basis außerhalb des Labors verwendbar.

Neuronale Netze haben eine besondere Eigenschaft: Sie sind lernfähig. Sie müssen erst lernen, um überhaupt verwendbar zu werden. Da dieses Lernen auf Mustern basiert, wird es „Deep Learning“ genannt. Das entspricht dem intuitiven Lernen, das beispielsweise Kinder beim Erlernen von Sprachen verwenden, bevor sie Konzepte und Regeln anwenden können. Wenn wir den Unterschied zwischen einem Dreieck und einem Viereck lernen, so ist das ein einmaliger konzeptueller Lernprozess, von dem wir letztlich auch Fünfecke oder andere Vielecke ableiten können.

Ein neuronales Netz lernt den Unterschied zwischen einem Dreieck und einem Viereck durch Beispiele und kann damit noch nichts über andere Vielecke sagen. Wenn beispielsweise ein neuronales Netz Kreise, Dreiecke und Rechtecke unterscheiden kann, werden andere Vielecke wahrscheinlich immer entweder als Kreis oder als Rechteck interpretiert werden, beziehungsweise reine Zufallsergebnisse herauskommen.

Neuronale Netze werden erfolgreich bei Mustererkennung verwendet – Funktionen wie Gesichtserkennung, sei es im *iPhone,* im öffentlichen Raum oder bei Zutrittssystemen, verwenden heute neuronale Netze, um die Übereinstimmung eines Gesichts mit einer Identität zu verifizieren. Im medizinischen Bereich werden für diagnostische Zwecke neuronale Netze verwendet, doch gab es diesbezüglich Fehlleistungen: Systeme, die in der Testphase verblüffend exakte Beurteilungen für ein bestimmtes Fachgebiet liefern konnten, versagten nach kurzer Zeit im medizinischen Betrieb und lieferten Fehldiagnosen, die lebensgefährdende Auswirkungen hätten haben können. Es stellte sich heraus, dass beim weitergehenden Anlernen in der Klinik, die das System verwendete, ab einem gewissen Zeitpunkt nicht mehr echte Fälle verwendet wurden, da die Eingabe zu aufwendig wurde, sondern hypothetische, unvollständige Fälle eingegeben wurden, und nur eine begrenzte Zahl von Ärzten ausschließlich ihre Fälle aus den jeweiligen Fachgebieten eingaben. Bereiche, die außerhalb der Fachgebiete dieser Ärzte lagen, lieferten dadurch Ergebnisse, die keine Grundlage hatten, letztlich also ungelernte, zufällig geratene Ergebnisse. Dadurch wurde das System bis zur Unverwendbarkeit verfälscht.

Deep Learning kann nur in dem Umfang sinnvolle Ergebnisse liefern, als ausreichend Übungsmuster geboten werden. Dort, wo diese Basis gegeben ist, kann „Deep Learning“ Ergebnisse liefern, die menschliche Entscheidungen an Geschwindigkeit und Zuverlässigkeit übertreffen – solange nichts Ungelerntes auftritt.

Schwerpunkt Entscheiden. Zuverlässige Entscheidung ist in verschiedenen

FOTO: POOBEST/STOCK.ADOBE.COM

Bereichen eine Anforderung an KI geworden, etwa beim autonomen Fahren oder der Steuerung von Verkehrsflugzeugen. Viele der Entscheidungsprozesse haben eine kritische zeitliche Komponente, die in ungetesteten Situationen, bei Fehlfunktionen oder Störungen, katastrophale Auswirkungen haben kann. Dabei können Nicht-Entscheidungen genauso schwerwiegend sein, wie falsche Entscheidungen.

Risikominimierung. Um das Risiko zu minimieren, werden heute meist mehrere voneinander unabhängige Systeme zugleich verwendet, die über Abstimmung zu einer Entscheidung gelangen – also etwa drei parallel arbeitende Kontrollsysteme, die in jeder Situation eine Mehrheitsentscheidung fällen, im Falle von Dissens eine Warnmeldung ausgeben. Auch wenn kritische Steuerungs-Systeme mit dieser Parallelität gebaut werden, kann es Gründe geben, diese Sicherung zu deaktivieren, wie das im Fall der beiden abgestürzten *Boeing 737 MAX* aus kommerziellen Gründen der Fall war.

Schwerpunkt Synthese. Mithilfe neuronaler Netze lassen sich Muster aus Sprache und Musik erlernen und für die Synthese von Sprache und Musik verwenden. Dabei entstehen beispielsweise verblüffend echt klingende Stimmen, die sogar Atemgeräusche simulieren. „Text-to-Speech“ gewinnt dadurch eine neue Ebene der „Natürlichkeit“. Ein interessantes Experiment in diesem Feld war, ein neuronales Netz ohne Vorgabe von Text „sprechen“ zu lassen – das, was dabei herausam, klang erstaunlich menschlich und war dem Gebrabbel eines Schlafenden nicht ganz unähnlich. Unter *https://deepmind.com/blog/article/wavenet-generative-model-raw-audio* sind Beispiele und Forschungsergebnisse zu diesem Themenbereich finden.

Im Rahmen der *Ars Electronica* 2019 wurde ein Projekt vorgestellt, in dem sich künstliche Intelligenz an die Vollendung von Mahlers 10. Symphonie, der Unvollendeten, heranmachte. Das Ergebnis ist verblüffend, es wird aber auch deutlich, dass musikalische Intention nicht einfach durch Mustererkennung ersetzbar ist. Mehr zu diesem Projekt unter *https://ars.electronica.art/aeblog/de/2019/09/02/mahler-unfinished/*.

Denkfähigkeit erfordert die Fähigkeit zur Reflexion, Erinnerung und Selbst-Identifikation. Wir können unser Denken außerdem nicht von unseren Wahrnehmungen, Vorstellungen und Gefühlen trennen. Denken ist in diesem Verständnis mit Bewusstsein verbunden, das wir als ein abgegrenztes, individuelles Ich verstehen, das eine Vergangenheit, Gegenwart und Zukunft hat, also eine zeitliche Kontinuität. Eine künstliche Intelligenz, die zum Denken befähigt ist, müsste Bewusstsein besitzen. Ob dies tatsächlich möglich ist, wird kontrovers diskutiert.

Alles, was derzeit den Eindruck der echten Menschlichkeit von künstlicher Intelligenz vermitteln soll, entstammt einer Simulation von Verhaltensweisen, die bisher nur bis zu einem gewissen Grad glaubwürdig menschlich ist, geschweige denn mit Bewusstsein oder Denkfähigkeit verbunden ist.

Michael Werzowa

Der Autor ist Experte für Netzwerk- und Datensicherheit, Vorstand der IoT Austria – The Austrian Internet of Things Network.