

SEMANTISCHE KÜNSTLICHE INTELLIGENZ IN UNTERNEHMEN

von Börtecin Ege



Über die Buzzwords Künstliche Intelligenz und Machine Learning stößt man im Zusammenhang mit der Digitalisierung so gut wie immer, nur selten liest man von semantischen Netzwerken oder Ontologien. In seinem Beitrag verändert das Börtecin Ege, Autor des Buchs „Semantische Datenintelligenz im Einsatz“. Er beschreibt, wie semantische Künstliche Intelligenz neue Chancen für bekannte Technologien wie Machine Learning eröffnen kann und zeigt notwendige Entwicklungen auf, um den nächsten KI-Winter zu vermeiden.

Die Bedeutung der künstlichen Intelligenz für Unternehmen wächst und wächst. Doch was versteht man unter künstlicher Intelligenz wirklich? Ganz einfach ausgedrückt, verstehen wir unter künstlicher Intelligenz eine Software, die aufgrund der vorliegenden Eingaben und Daten gewisse Datenmuster selbstständig erkennen, erlernen und sich weiterentwickeln kann. Das bedeutet jedoch lange nicht, dass diese Technologie schon längst ausgereift ist. Aber auch wenn die künstliche Intelligenz noch eigene Probleme hat, die noch überwunden werden müssen, steht sie schon heute bereits auf eigenen Füßen. Dabei helfen nicht zuletzt semantische Technologien. In diesem Artikel möchten wir näher auf Einzelheiten eingehen.

Die Anfangszeiten

Sowohl die Wurzeln der künstlichen Intelligenz als auch von neuronalen Netzen reichen zurück bis in die 30er und 40er Jahre, obwohl man von den beiden erst seit einigen Jahren in den Medien hört. Drei Wissenschaftler spielten bei der Gründung dieser Disziplin eine wichtige Rolle: Der österreichische Logiker Kurt Gödel (1906-1978), der britische Logiker Alan Turing (1912-1954) und der US-amerikanische Computervissenschaftler John McCarthy (1927-2011). Die Erfindung von ersten neuronalen Netzen erfolgte dann erst in den 40er Jahren durch zwei US-amerikanische Wissenschaftler, Walter Pitts (1923-1969) und Warren McCulloch (1898-1969), die man oft auch künstliche neuronale Netze (KNN) nennt.

Der Durchbruch

Nach zwei sogenannten KI-Wintern in den 70er und 80er Jahren, in denen alle weiteren Entwicklungen schienen stillzustehen, kamen ab Ende der 90er Jahre wieder neue Erfolgsmeldungen wie IBM Deep Blues Sieg gegen den damals amtierenden Schachweltmeister Garry Kasparov (1997), IBM Watsons Sieg gegen seine zwei menschlichen Mitstreiter bei der US-Quizshow „Jeopardy!“ (2011) und nicht zuletzt die Siege von Google AlphaGo gegen zwei hochkaratige Go-Spieler, Fan Hui und Lee Sedol, in den Jahren 2015 und 2016.

Zum echten Durchbruch der künstlichen Intelligenz für die gewöhnlichen Anwender*innen kam es jedoch erst in den letzten Jahren mit der Entwicklung von Deep Learning und mit der immer steigenden Rechenleistung der heutigen Computergeneration. Eine weitere aktuelle und bahnbrechende Entwicklung in diesem Bereich ist das autonome Fahren. Unter autonomem Fahren verstehen wir Autos, die im Prinzip auch ohne menschliches Eingreifen selbst fahren können. Dank der aktuellen Entwicklungen im Bereich der Künstlichen Intelligenz, insbesondere durch Deep Learning, sind große technische Fortschritte in der Bildererkennung erzielt worden, und somit ist die Erkennungskraft von bestimmten Systemen massiv gestiegen. Auch wenn sie derzeit aus bestimmten Gründen wie schlechte Sicht, verschmutzte Sensoren oder Dunkelheit manchmal versagt und dadurch Fehler, Pannen oder gar Unfälle verursacht,

ist diese bahnbrechende Technologie doch auf Erfolgskurs. Nicht nur die größten Automobilhersteller wie Volkswagen, BMW und Mercedes, sondern auch Tech-Riesen wie Google und Uber befinden sich derzeit in einem Kopf-an-Kopf-Rennen, um als Erstes ein zuverlässiges, fahrerloses Auto mit unterschiedlichen Autonomiestufen - vom assistierten und teilautomatisierten Fahren, bis zum hoch- und vollautomatisierten sowie völlig autonomen Fahren, schon in absehbarer Zeit herzustellen. Um dieses Ziel zu erreichen, müssen jedoch zuerst noch viele technische Herausforderungen gemeistert werden. Daher gibt es weltweit viele Bestrebungen und Projekte in diesem Bereich, an denen sehr viele Wissenschaftler*innen intensiv mitarbeiten, wie zum Beispiel die Leitinitiative von VDA (Verband der Automobilindustrie), an der deutsche Autohersteller wie VW, BMW und Mercedes sowie viele andere Firmen gemeinsam arbeiten.

Machine Learning

Die künstliche Intelligenz (KI) besteht grundsätzlich aus drei verschiedenen Ansätzen, und zwar aus datenbasierten, semantischen und regelbasierten Ansätzen. Während die KI-Systeme wie neuronale Netze, die auf datenbasierten Ansätzen fungieren, aus vorliegenden Daten und Eingaben gewisse Datenmuster selbstständig erkennen und erlernen, werden semantische Ansätze, deren Bausteine man Ontologien nennt, derzeit vielmehr dafür eingesetzt, um eine semantische Plausibilität und Nachvollziehbarkeit bei der, von neuronalen Netzen getroffenen Entscheidungen, zu erreichen. Sowohl die datenbasierten als auch die semantischen Ansätze sind ein Teil von Machine Learning. Während die datenbasierten Ansätze wie neuronale Netze und Deep Learning auf Statistik und Wahrscheinlichkeitstheorie basieren, basieren die semantischen Ansätze rein auf wissensbasierte Systeme, die für die Implementierung von intelligenten Systemen unentbehrlich sind. Den dritten Ansatz, nämlich den klassischen regelbasierten Ansatz in der Form Wenn A dann B, der uns aus dem Alltag bekannt ist, findet man fast in jeder KI-Software. Folgendes Beispiel verdeutlicht dies: Um einem fahrerlosen Auto mitteilen zu kön-

nen, dass das Auto bei Rot anhalten soll, braucht man weder neuronale Netze noch einen semantischen Ansatz.

Blackbox-Effekt

Ein wichtiger Nachteil von datenbasierten Ansätzen ist, dass sie nicht nur mit richtigen und entsprechend bereinigten Daten für eine bestimmte Aufgabe oft stundenlang trainiert werden müssen, sondern man nach einer Entscheidung normalerweise auch nicht wissen kann, wie sie getroffen wurde. Dies nennt man Blackbox-Effekt.

Semantische Künstliche Intelligenz eröffnet neue Perspektiven

Wie oben bereits erwähnt, versucht man gegen bestimmte Probleme, die die datenbasierten Ansätze mitbringen und deren Bausteine aus Ontologien bestehen, mittels semantischer Technologien zu steuern. Unter einer Ontologie versteht man grundsätzlich ein logisches Regelwerk, das Wissen und Zusammenhänge über eine bestimmte Domain sehr präzise formuliert und darstellt.

Anders gesagt: Ohne Hilfe von semantischen Technologien und Ontologien fehlt einem neuronalen Netz sowie Deep Learning sogar ein gesundes Menschenverständnis. Schon jetzt kommt es oft vor, dass z.B. die Banken insbesondere bei der Feststellung der Kreditwürdigkeit ihrer Kund*innen Probleme bekommen bzw. die Kreditwürdigkeit ihrer Kund*innen falsch eingeschätzt wurde, weil sich solche rein datenbasierten neuronalen Netze nur auf konventionelle Machine Learning-Techniken stützen. Daher ist davon auszugehen, dass in absehbarer Zukunft solche rein datenbasierte neuronale Netze und auf Deep Learning aufbauende KI-Systeme, keine Akzeptanz finden werden.

Allein aus diesen Gründen muss mit Hilfe der semantischen Technologien unbedingt eine neue Generation von datenbasierten Ansätzen (neuronale Netze, Deep Learning) entwickelt werden, die nicht nur effektiver und transparenter arbeiten, sondern durch Fremdeingriffe wie z.B. einen Adversarial Attack (dank der semantischen Plausibilisierung) nicht so leicht verleitet werden können. Adversarial Attacks sind feindliche Angriffe

auf KI-Systeme wie Neuronale Netze, die Daten in böswilliger Absicht verändern und somit KI-Algorithmen täuschen, so dass diese am Ende falsche Ergebnisse liefern oder falsche Entscheidungen treffen. Ein einfaches Beispiel dazu wäre ein fahrerloses Auto, das dazu verleitet wird, in einer verkehrsberuhigten Zone Vollgas zu fahren.

Enterprise Knowledge Graphs

Ein anderer Anwendungsbereich, der erst durch semantische Technologien ermöglicht worden ist, sind Enterprise Knowledge Graphs bzw. Wissensgraphen. Im Gegensatz zu den heutigen relationalen Datenbanken ermöglichen Wissensgraphen eine effiziente Speicherung, Verarbeitung und Darstellung des Wissens durch einen großen Graphen, der aus zahlreichen Knoten und Kanten besteht. Sie sind mit ihrer auf Ontologien bauende Strukturen sowohl für die Verarbeitung von Wissen als auch Entdeckung von Zusammenhängen viel geeigneter als die relationalen Datenbanken. Sehr viele Firmen wie Facebook, Google aber auch staatliche Institutionen wie NASA haben bereits ihre eigenen Wissensgraphen im Einsatz. Während manche von solchen Wissensgraphen der Öffentlichkeit nicht zugänglich sind, sind bestimmte Wissensgraphen wie Wikidata, YAGO und DBpedia öffentlich.

Wo liegt der Schlüssel zum Erfolg bei Entwicklung von KI-Anwendungen?

Der Kern von heutigen Machine Learning basiert größtenteils auf datenbasierten Ansätzen, neuronalen Netzen und Deep Learning und somit auf Statistik und Wahrscheinlichkeitstheorie, jedoch nicht auf semantischen Ansätzen. Ein echtes KI-System sollte jedoch nicht auf Statistik und Wahrscheinlichkeitstheorie, sondern vielmehr auf semantischen Ansätzen aufbauen - zumindest größtenteils. Daher ist die semantische künstliche Intelligenz aktuell der einzige optimale Weg, damit die neue Generation von KI-Systemen nicht nur das Wissen effizient verarbeiten, sondern auch die vorhandenen Zusammenhänge besser erkennen und mögliche Blackbox-Effekte vermeiden kann. Sonst könnte es sein, dass bald ein dritter KI-Winter droht.