

## Humanistische Union

# Künstliche Intelligenz und Bürgerrechte

Ein Thema für die Humanistische Union!

In: Mitteilungen 240 (3/2019), S. 3-6

Am zweiten Tag der HU-Konferenz, dem Samstagnachmittag, befassten wir uns in drei Vorträgen mit Künstlicher Intelligenz. Aus Sicht der Humanistischen Union ein neues Thema, das bisher keinen Schwerpunkt unserer Arbeit bildete, aber aus Sicht der Bürgerrechte immer mehr an Bedeutung gewinnen wird: beispielsweise durch sein Potenzial zur Manipulation von Wahlen oder zur Überwachung und Klassifizierung von Menschen – Stichwort Programmierter Rassismus.

Den Anfang machte Stefan Hügel: Im Vortrag Künstliche Intelligenz und Politik gab er einen einführenden Überblick über die Anwendung von künstlicher Intelligenz und ihre Auswirkungen. Während in der Wissenschaft die Erwartungen an eine künstliche Intelligenz (KI) zuletzt deutlich reduziert wurden, wird die Technologie in der politischen Öffentlichkeit in den letzten Jahren deutlich aufmerksamer – und zum Teil kritischer – verfolgt. Stefan Hügel sortierte diese Debatte, indem er zunächst einmal die verschiedenen Begriffe von starker und schwacher KI, computergestützten Algorithmen, Lern- und Entscheidungssystemen voneinander abgrenzte. Basis der aktuellen Debatte um künstliche Intelligenz ist der Begriff des Algorithmus, der heute häufig irreführend verwendet wird. Ergebnisse des maschinellen Lernens werden hauptsächlich durch die Daten bestimmt, aus denen ein solches System „lernt“. Die Daten werden durch einen universellen Lernalgorithmus verarbeitet, der Parameter und Schwellwerte justiert, die jeweils statistische Wahrscheinlichkeiten ausdrücken. Unter Laborbedingungen ist diese Form der künstlichen Intelligenz zu erstaunlichen Leistungen fähig, wie sich zuletzt bei dem für das Go-Spiel entwickelten AlphaGo und seinen Weiterentwicklungen zeigte.

Im zweiten Teil seines Vortrags ging er ausführlich auf die verschiedenen Gefährdungspotenziale der KI für Demokratie und Öffentlichkeit ein. Maschinelles Lernen reproduziert die vorgefundene Welt und dies kann zu Diskriminierungen, z.B. sogenanntem programmiertem Rassismus führen, wenn solche Vorurteile in vorgefundene Daten eingeschrieben werden. Als Beispiel: Ergibt sich aus der Statistik, dass bevorzugt Männer in Führungspersonen gelangen, so werden beim maschinellen Lernen auch in der Zukunft vorzugsweise Männer für solche Positionen vorgeschlagen werden. Das hat zu einer politischen Debatte geführt; aufgrund der Komplexität und mangelnden Nachvollziehbarkeit gibt es aber keine einfache Lösung. Der grundsätzliche Ansatz wurde im Vortrag von Prof. Bast behandelt (s.u.).

Zusätzlich können die Ergebnisse für „Nudging“ genutzt werden, zur subtilen Beeinflussung von Menschen. Der Fall des Data-Science-Unternehmens Cambridge Analytica, wo Daten für gezielte Wahlwerbung genutzt wurden, deutet auf die Gefährdungen unserer Demokratie hin. Es ist nicht klar, ob die Verantwortlichen in der Politik die Problematik vollständig verstanden haben. Erforderlich ist letztendlich eine Einhegung solcher Konzepte, bei der ein wohlverstandener Datenschutz eine zentrale Rolle spielt. An den Vortrag von Stefan Hügel knüpfte Prof. Dr. Britta Schinzel, Professorin Emerita für Informatik und Gesellschaft an der Universität Freiburg vertiefend an. Sie stellte zunächst dar, was „künstliche Intelligenz“ eigentlich bedeutet, was möglich ist, und wie. Dazu ging sie auf die Geschichte der künstlichen Intelligenz ein: Erste Überlegungen hierzu gab es bereits in den 1930er Jahren; viele der heute genutzten Methoden sind nicht eigentlich neu – sie basieren vor allem auf künstlichen neuronalen Netzen (KNN) und Stochastik –, sondern können jetzt erst auf leistungsfähigen Maschinen, mit höherer Speicherdichte und einem umfangreichen Datenbestand („Big Data“) erfolgreich eingesetzt werden.

Künstliche Intelligenz ist keine Simulation des menschlichen Gehirns, sondern nutzt Standardverfahren der Mathematik und Informatik. Das gilt auch für das maschinelle Verstehen gesprochener Sprache durch Systeme wie Siri oder Alexa, die akustische Wellenformen mit Wörtern und Satzkonstruktionen in Beziehung setzen, mögliche Antworten durch statistische Inferenz ableiten und einen Antwortsatz auf Basis

der höchsten Wahrscheinlichkeit generieren. Die dabei gewonnenen Daten werden zur Optimierung, Individualisierung, Profilbildung und für weitere Nutzungen verwendet. Das gleiche Prinzip gilt bei der Bilderkennung. Es werden keine kausalen Zusammenhänge und genauen Ergebnisse ermittelt, sondern Näherungen und Korrelationen.

Den zweiten Abschnitt ihres Vortrags bildeten die vielfältigen Anwendungen des maschinellen Lernens, die heute eingesetzt werden für Sprach- und Informations-Medien, soziale Netzwerke, Bilderkennung, Geo-Systeme einschließlich automatischem Fahren, Geo-Aufklärung und für autonome Waffen.

Die Probleme und gesellschaftlichen Wirkungen stellte Prof. Schinzel im dritten Teil ihres Vortrags detailliert dar: Ressourcenverbrauch – Metalle und seltene Erden, die häufig unter menschenunwürdigen Bedingungen abgebaut werden, Arbeit 4.0 mit Crowd- und Clickworking, Überwachungskapitalismus, Bias und Ungleichgewichte und Diskriminierung durch Training an vorurteilsbehafteten Trainingsdaten oder durch methodologische Fehler. Diese Problematik ist in der Praxis häufig nicht durch Antidiskriminierungsrecht zu fassen.

Den Abschluss des Vortrags bildeten die Gegenmaßnahmen: Hier wurde in erster Linie die Anwendung von Datenschutzgesetzen – z.B. Zweckbindung und Datensparsamkeit – und Verweigerung der Deep-Packet-Inspektion genannt; dies könnte nach Ansicht einer Datenschützerin bereits rund 80% aller Probleme, Vorurteile, Diskriminierung, Missbrauch, Trollbots etc. vermeiden. Es bleibt die Frage nach ethischer Nutzung und den gesellschaftlichen Zielen der KI. Forderungen wie Transparenz, Datenschutz und Ethik by Design gehen an Deep Neural Networks in der Praxis einfach vorbei. Es gibt aber Initiativen: Die japanische Regierung hat eine Initiative für „AI 5.0“ aufgesetzt, es gibt ethische Standards des World Economic Forum und der EU-Kommission, und die deutsche Datenschutzkonferenz hat in ihrer Hambacher Erklärung Prinzipien aufgestellt.

Den Schlusspunkt der Vorträge zur Künstlichen Intelligenz setzte Prof. Dr. Hannah Bast, ebenfalls Professorin an der Universität Freiburg, wo sie in der Informatik den Bereich Algorithmen vertritt und in die Enquete-Kommission „Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale“ des Deutschen Bundestags berufen wurde. Bereits in den Sitzungen der Enquete-Kommission am 15. Oktober 2018 und am 14. Januar 2019 hat sie einige der behandelten Aspekte dargestellt.

Bei der Behandlung von Verfahren der Künstlichen Intelligenz ist es wichtig, so Professor Bast, zwischen klassischen Algorithmen und maschinellem Lernen zu unterscheiden. Klassische Algorithmen können bei „einfachen“ Problemen eingesetzt werden, bei denen die Anzahl der Parameter bei der Berechnung noch beherrschbar ist. Einfaches Beispiel ist die Auswahl einer Wohnung, beispielsweise mit den Parametern Wohnfläche und Preis. Die Bewertung kann in einem Algorithmus programmiert werden und wird immer gleich angewendet. Doch auch beim Schachspiel kann man noch in dieser Weise arbeiten: Spielstellungen können bewertet werden, und damit kann über den nächsten Spielzug entschieden werden. Offensichtlich benötigt man dafür erheblich mehr Rechenleistung als bei dem einfachen Wohnungsbeispiel; moderne Computer können dies aber leisten. 1997 wurde zum ersten Mal der amtierende Weltmeister – Garri Kasparow – von einem Computer – IBM Deep Blue – geschlagen.

Anders funktioniert das maschinelle Lernen: Um eine Klassifikation oder Bewertung von Daten zu erhalten, werden in einer Trainingsphase Spielstände von einem neuronalen Netz bewertet. Entspricht die Bewertung nicht dem gewünschten Ergebnis, werden Parameter des neuronalen Netzes so lange verändert, bis das gewünschte Ergebnis erreicht wird. Dabei müssen die anderen Eingabewerte ebenfalls berücksichtigt werden: Bereits richtig klassifizierte Eingaben können durch spätere Änderungen der Parameter wieder verfälscht werden, so dass weitere Veränderungen notwendig sind. Ziel ist es, Parameter so einzustellen, dass für alle Eingabewerte aus den Trainingsdaten das „richtige“ Ergebnis errechnet wird. An die Stelle der Bewertungsfunktion des klassischen Algorithmus tritt die „gelernte“ Bewertungsfunktion des neuronalen Netzes. Während dem ersten die Bewertungsfunktion noch explizit von „außen“ mitgegeben wurde, hat das neuronale Netz sie quasi „selbst“ ermittelt. Dies lässt sich anhand von Simulationsbeispielen anschaulich nachvollziehen. Die Komplexität von Go ist so hoch, dass nur noch Verfahren des maschinellen Lernens mit Erfolg angewendet werden können. Im Gegensatz zum Schach ist die Anzahl der möglichen Spielstellungen so groß, dass keine klassisch algorithmischen Bewertungsfunktionen mehr programmiert werden können. Das ist auch der Grund, warum hier der Computer den menschlichen Weltmeister erst ca. 20 Jahre später – 2016 – schlagen konnte.

Probleme des maschinellen Lernens, die Prof. Bast außerdem behandelte, waren Transparenz und „Bias“ – die Ursache für Diskriminierung, den bereits genannten programmierten Rassismus durch maschinelles Lernen. Die Parameter eines neuronalen Netzes haben (für den Menschen) keine nachvollziehbare Interpretation, und es besteht kein systematischer Zusammenhang zur den klassifizierten Objekten. Es ist nicht garantiert, dass für den Menschen gleiche Objekte vom neuronalen Netz gleich klassifiziert werden. Manchmal genügt die Änderung eines Pixels – für uns Menschen nicht erkennbar – für eine völlig andere Erkennung.

Das Problem Bias stellte Prof Bast anhand der Spracherkennung dar: Dabei können fast gleiche Wörter sehr unterschiedliche Bedeutung haben, wie „Hering“ und „Ehering“, die sich nur in einem Buchstaben unterscheiden. „Bundestag“ und „Parlament“ sind dagegen sehr unterschiedlich, bezeichnen aber das gleiche Konzept. Deswegen werden Wörter häufig als Punkte in einem Raum dargestellt, wobei Wörter mit ähnlicher Bedeutung nahe beieinander liegen. Damit ergeben sich Relationen wie z. B. „Deutschland verhält sich zu Berlin wie Schweden zu Stockholm“.

Dabei können ungewollte Effekte auftreten, wenn z.B. aus den Texten, die als Eingabedaten dienen, die – offensichtlich falsche – Relation „Mann verhält sich zu Chefarzt wie Frau zu Krankenschwester“ abgeleitet wird. Das wird dadurch gelöst, dass man zwischen definitiven (Mann, Frau) und stereotypen (Chefarzt, Krankenschwester) Wörtern unterscheidet und letztere aus der Klassifikation „herausrechnet“. Welche Wörter man dabei als definitiv und welche als stereotyp betrachtet, ist nicht zuletzt auch eine politische Frage.

---

<https://www.humanistische-union.de/publikationen/mitteilungen/240/publikation/kuenstliche-intelligenz-und-buergerrechte/>

Abgerufen am: 02.04.2023