KI und Staat

(Beitrag im Rahmen der Task Force „Künstliche Intelligenz“ des EVW
Von Dr. Lothar Weniger, August 2020)

Einleitung

Nachdem das Thema KI und Verwaltung schon an anderer Stelle behandelt wurde fokussiert dieser Beitrag auf Anwendungen von KI im Rahmen der Aufrechterhaltung der öffentlichen Ordnung. Dies ist in sich selbst ein weites Feld und reicht von der Optimierung des Verkehrsflusses durch intelligente Ampelschaltung, über Kriminalitätsbekämpfung, bis zur Nutzung von Überwachungssoftware zur Identifizierung und Unterdrückung von abweichenden Meinungen durch totalitäre Staaten.

Wie generell im Thema KI, sind die Problemstellungen nicht neu, der Verkehr wurde auch vor KI geregelt, Straftäter wurden gesucht und Dissidenten unterdrückt. Der Einsatz von KI ermöglicht allerdings eine gewaltige Effizienzsteigerung und erhebliche Kosteneinsparungen, besonders beim Personal. Potentiell verschiebt der Einsatz von KI damit das Machtverhältnis zwischen Bürger und Staat zu Gunsten des Staates.

In den westlichen Demokratien scheinen sich die Bemühungen zu verschärfen gesellschaftlich unerwünschte Beiträge und Meinungen im Internet aufzudecken und zu verfolgen. Wiederum ist nicht der Einsatz von KI der Grund hierfür, sondern ein gesteigerter Drang Abweichungen von „Political Correctness“ zu sanktionieren. KI liefert nur die technische Basis, die es ermöglicht riesige Volumina von Text nach diesen Abweichungen zu durchsuchen. Die westlichen Staaten bauen dafür keine eigenen Überwachungsapparate auf, sondern verpflichten die Plattformbetreiber dazu diese Recherchen durchzuführen.

Technik

Verschiedene Technologien werden erst durch KI in die Lage versetzt gewünschte Informationen einigermaßen zuverlässig aus riesigen Datenmengen herauszufiltern, bzw nach Training an großen Datenmengen, auffällige Muster oder Bedeutungen zu erkennen.

Beispiele sind:

- Gesichtserkennung (Computer Vision)

- Spracherkennung (NLP)

- Texterkennung (Filterung von Fake News, Hassreden)

- Automatische Übersetzung

- Upload Filter (Urheberrechte)

- Verhaltensmusteranalyse (Charaktererkennung)

Mit Hilfe von KI werden diese Techniken zwar flexibler und „intelligenter“ und sind nicht mehr nur auf diskret vorgegebene Muster und Protokolle angewiesen, dennoch bleiben sie vergangenheitsbezogen. Der Outcome des Programms wird im Wesentlichen auf Basis der erlernten Muster aus der Vergangenheit errechnet. Wenn es einen signifikanten Bruch in der Zeitreihe gibt, können die Ergebnisse wertlos werden. Ein Beispiel wäre die KI unterstützte Aufdeckung von Krankenversicherungsbetrug nach Ausbruch der Corona Pandemie. Diese sogenannte Modeldrift kennt man auch aus traditionellen Fachgebieten, wie der Ökonomie.

Ein interessantes Beispiel für die Unzulänglichkeiten heutiger KI – Algorithmen ist Facebook‘s Suche nach Hassreden in seinem sozialen Netzwerk. Das Unternehmen verfügt unbestritten über erstklassige KI Experten und Technologie. Dennoch reagiert man auf anhaltende Kritik mit der Ausweitung menschlicher Recherchen auf Kosten von KI Software.

Ein Problem ist die Mehrdeutigkeit von Sprache (Metaphern), die KI Algorithmen schwer erfassen können. Facebook’s Chief AI Scientist Yann LeCun sagt: „Current machines don‘t have common sense. They have very limited and narrow function”.

Ein weiteres Problem ist die Anpassungsfähigkeit der menschlichen Autoren. Die relative statischen KI Suchkriterien können erkannt und umgangen werden. Der Unterschied zwischen dem Erkennen und dem Verstehen von Inhalten kommt hier exemplarisch zum Ausdruck.

Anwendung

Im Bereich der öffentlichen Ordnung gibt es eine Vielzahl von KI-gestützten Anwendungen. Hier nur eine Auswahl:

1. Kriminalitätsbekämpfung.
    1.1. Gesuchte Personen finden (Kriminelle, Vermisste..)
    1.2. Mündungsfeuer orten (in USA werden nur 20% des Schusswaffengebrauchs gemeldet)
    1.3. Rudelbildung erkennen (Opernplatz, Plünderungen)
    1.4. Predictive Policing (Dynamische Identifizierung von Verbrechensschwerpunkten)
    1.5. Einschätzung individueller Straffällig-, Rückfallwahrscheinlichkeit (Entscheidung über Strafentlassung, Kaution, in China auch zur Beobachtung möglicher Straftäter)
    1.6. Zuschreibung von Persönlichkeitsmerkmalen durch Auswertung passiver Mobiltelefondaten. (Gerade von der LMU München erfolgreich zur Bestimmung der 5 wichtigsten Persönichkeitsmerkmale (big five) erprobt)
2. Predictive Public Maintenance (Straßenbau, Kanalisation…)
3. Zusagen von Förderkrediten und Subventionen
4. Genehmigung von Sozialleistungen
5. Erkennen von Steuerbetrug

6. Erkennen von Insiderhandel
7. Auffinden gestohlener Fahrzeuge
8. Sortierung von Notrufen nach Dringlichkeit (Cincinnati Fire Department)
9. Überwachung der öffentlichen Meinung. (Totalitäre Regime nutzen KI Tools zur Optimierung der Zensur und Überwachung von Dissidenten. Im Westen zielen ähnliche Verfahren eher auf die Identifizierung von Hassreden und Urheberrechtsverletzungen im Internet)

Anbieter

Besonders KI-basierte Gesichtserkennung wird schon von vielen Unternehmen angeboten.

Der wohl größte Anbieter ist der japanische **NEC** Konzern. Hier fokussiert man weniger auf Polizeianwendungen, als auf auf zivile Nutzungen, wie Zutritts- und Ausweiskontrolle und kontaktloses Bezahlen.

Auf Gesichtserkennung zur Unterstützung der Polizeiarbeit haben sich insbesondere **Clearview AI** (USA), **Ayonix** (Jp) und **Sense Time** (Chn) konzentriert. Sense Time ist einer der größten chinesischen Anbieter und hat einen geschätzten Marktwert von 8 Mrd USD. Die Firma steht im Westen in der Kritik, da sie eng mit der chinesischen Leon Technologies zusammenarbeitet, die wiederum mit ihren KI basierten Auswertungen ein wichtiger Akteur beim Aufbau des chinesischen Überwachungsstaates ist. **Microsoft** und **Amazon** sind weitere Anbieter, die aber Verkäufe an Polizeibehörden ausgesetzt haben, da in den USA das Thema Gesichtserkennung inzwischen sehr kontrovers diskutiert wird.

PredPol Inc bietet „predictive policing“ Software an.

Equivant Inc KI-basierte Entscheidungen über Haftentlassung, Untersuchungshaft, Kautionshöhe

Cloud Walk (Chn) bietet u. a. Software an, welche durch Beobachtung von Einzelpersonen deren Wahrscheinlichkeit eine Straftat zu begehen berechnet.

Hikvision (Chn) ist einer der größten Anbieter von Überwachungstechnologien generell. Die Firma liefert inzwischen Videoüberwachungskameras in denen die KI basierte Auswertung dezentral integriert ist. Die Firma hat sich in der Vergangenheit damit gebrüstet uighurische Personen identifizieren zu können. Sie ist mehrheitlich in Staatsbesitz.

Vorteile

Der wesentliche Vorteil des KI Einsatzes im öffentlichen Sektor ist die Möglichkeit große und diverse Datenmengen effizient auszuwerten. Damit können Entscheidungsprozesse optimiert und teilweise gänzlich automatisiert werden. Die potentiellen Kosteneinsparungen sind enorm. Eine Studie von Deloitte beziffert das jährliche Einsparpotential allein für die

USA auf bis zu 41 Mrd USD. Dabei sind die wirtschaftlichen Vorteile einer Verringerung von Kriminalität, Verkehrsstaus, Insider Handel etc. noch gar nicht berücksichtigt.

Im Rahmen einer Automatisierung von Prozessen kann KI Investitionskosten erheblich reduzieren. Oftmals erfordert eine Automatisierung umfangreiche Vorarbeiten und nach Einführung die Beachtung starrer Ablaufregeln. Zur automatischen Briefsortierung musste, z. B., ein flächendeckendes System von Postleitzahlen geschaffen werden und jeder Teilnehmer (Briefversender) muss diese Systematik kennen und verwenden. KI könnte ein solch aufwendiges System überflüssig machen.

Weitere technische Vorteile der KI-Algorithmen gegenüber der menschlichen Umsetzung sind die

-verringerte Häufigkeit von Input- und Übermittlungsfehlern,
-größere Prozessgeschwindigkeit,
-strikte Regelbeachtung,
-gleichzeitige Nutzung von mehr und diverseren Datenquellen.

Darüber hinaus erfolgt die Auswertung grundsätzlich objektiv und ohne menschliche Vorurteile oder Neigungen. Dieser Aspekt wird in der öffentlichen Debatte allerdings oft anders gesehen. Darauf komme ich im übernächsten Kapitel zurück.

Für totalitäre Regime ergibt sich ein Vorteil durch die Festigung ihrer Herrschaft mittels besserer Kontrolle und Lenkung der Gesellschaft. Mit der Auswertung einer Vielzahl von unterschiedlichen Datenquellen, wie z. B. Gesichtserkennung, Bewegungsprofile, Konsumgewohnheiten, emails, Telefonate, Kontakte, Lesematerial, Nachrichtenquellen, Einkünfte, Likes in sozialen Netzwerken etc, kann am Ende die Regimetreue jedes Einzelnen automatisiert eingeschätzt werden. Ein zentralisiertes Social Credit Score Modell, wie in China, wäre ohne KI-Algorithmen wohl kaum möglich. Wenn die Bürger dies akzeptieren, oder gar aktiv unterstützen, um ein gutes Standing zu erreichen, nähern wir uns dem perfekten Überwachungsstaat.

Die Wirkung eines öffentlichen Social Scoring Systems auch im privaten Leben, konnte vor kurzem in China beobachtet werden. Im April stürzte der Ölpreis ins Bodenlose. Das löste Nachschusspflichten bei chinesischen Privatanlegern aus, die auf Anraten mehrerer staatlicher Banken auf steigende Preise spekuliert hatten. Es wird berichtet, dass die Banken für den Fall einer Zahlungsverweigerung mit negativen Folgen für den individuellen Social Score drohten.

Nachteile

Die schon erwähnten Schwächen der KI-Algorithmen können zu einer falschen Auswertung der Inputdaten führen. Wenn der Entscheidungsbaum durchgängig automatisiert ist, resultieren daraus direkt entsprechend falsche Entscheidungen.

Auch wenn die KI Auswertungen nur als Entscheidungshilfe für den menschlichen Akteur gedacht sind, besteht die Gefahr, dass dieser sich unkritisch auf Vorschläge verlässt und sie 1 zu 1 umsetzt.

Eine weitere Schwäche von KI Algorithmen ist, dass sie mit Menschlicher Logik ausmanövriert werden können. Wenn, z. B., Kriminelle erkennen, nach welcher Systematik eine Predictive Policing Software Verbrechensschwerpunkte identifiziert, und somit, wo die Ordnungskräfte konzentriert sind, können sie das für ihre Zwecke nutzen. Das gilt für alle Mustererkennungsmodelle, wie beim Aufspüren von Sozialbetrug, Insiderhandel etc. In diesem Katz- und Mausspiel zwischen Ordnungskraft und Kriminellen kommt die, auf historischen Daten trainierte, KI schlecht mit.

Die verbesserten Überwachungsmöglichkeiten der Bürger mögen für totalitäre Staaten ein Vorteil sein, für die Gesellschaft aber wohl eher ein Nachteil. Wenn abweichende Meinungen sehr schnell erkannt und sanktioniert werden können, werden die Meinungsfreiheit und damit die dynamische Weiterentwicklung der Gesellschaft abgeschnürt.

Ethische Aspekte

Zwei moralisch, ethische Fragen mit KI-Bezug werden in diesem Zusammenhang besonders intensiv diskutiert, und zwar die Prävalenz von Rassismus in Polizeisoftware und von Hassreden im Internet. In beiden Fällen steht der Einsatz von KI in der Kritik.

Der Vorwurf des Rassismus

Besonders in den USA wird, im Zuge der „Black Lives Matter“ Bewegung, häufig unterstellt, dass KI-basierte Algorithmen Minderheiten, insbesondere Schwarze, benachteiligten. Einige Gemeinden haben daher schon die Nutzung von polizeiunterstützender Software ausgesetzt und Unternehmen wie Microsoft und Amazon bieten ihre Gesichtserkennungssoftware Polizeibehörden nicht mehr an.

Die Vorwürfe reichen von einer größeren Fehlerwahrscheinlichkeit bei der Gesichtserkennung von Minderheiten, über eine geringere Empfehlungsquote für Strafentlassungen schwarzer Gefängnisinsassen, bis zu vermeintlich diskriminierender Konzentration von Polizeikräften in „schwarzen“ Stadtteilen durch predictive policing Software.

Diese Vorwürfe sind zunächst überraschend, da Algorithmen natürlich nicht per se rassistisch sein können, sondern viel eher eine objektive und neutrale Analyse hervorbringen als - vielleicht auch nur unterbewusst - vorurteilsbehaftete Menschen.

Die Argumentation geht dann meist dahin, dass der Rassismus durch Programmierung und Training der KI von Menschen in die Software implementiert wird.

Das könnte sein. Aber vielleicht gibt es noch einen anderen Grund, der auch die Heftigkeit der Anschuldigungen verständlicher macht.

Die kalte Logik der KI-Algorithmen bringt Tatsachen an die Oberfläche, die die Gesellschaft möglicherweise nicht hören will. Schwarze sind in den USA proportional häufiger kriminell als z. B. Amerikaner mit asiatischem Hintergrund. Wenn eine KI Kriminalitätsschwerpunkte besonders häufig in „schwarzen“ Stadtvierteln anzeigt, ist das nicht unbedingt ein Anzeichen von Rassismus, sondern eine Folge der Auswertung der statistischen Datenbasis.

Die Firma Equivant hat nachgewiesen, dass Vorwürfen, die besagen, dass schwarze Straftäter bei den Empfehlungen zur Haftentlassung durch ihre KI Software benachteiligt würden, nicht haltbar sind, wenn man die höhere Kriminalitätsrate einbezieht.

Wäre es moralisch geboten schwarzen Kriminellen einen Bonus zu geben und asiatischen einen Malus damit sich der Verfolgungsdruck durch die Polizei ausgleicht?

Die Ungenügende Bekämpfung von Hassreden im Internet

Ein zunehmender gesellschaftlicher und gesetzgeberischer Druck, als ungebührlich erachtete Aussagen im Internet zu eliminieren und zu sanktionieren führt zu verstärktem Einsatz von KI-basierten Algorithmen. Man kann unterstellen, dass sich die Definition von inakzeptablen Inhalten in Demokratien in einem dynamischen, gesellschaftlichen Prozess bildet und durch einigermaßen transparente gesetzlichen Regeln formalisiert wird.

Da für die Umsetzung aber nur KI Algorithmen in Frage kommen erfolgt die faktische Identifizierung von verdächtigen Beiträgen durch diese Softwareprogramme. Wenn dazu berücksichtigt wird, dass die Plattformbetreiber, denen die Durchführung der Säuberung auferlegt wird, kein Interesse an geschäftsschädigenden Konflikten haben, ist zu befürchten, dass Inhalte eher großzügiger eliminiert werden als zu zaghaft.

Ist es ethisch vertretbar Mindermeinungen zu eliminieren (Fake News, Verschwörungstheorien, unwissenschaftliche Meinungen (Klimaleugner) etc) um zu erreichen, dass weniger Pöbeleien und Verunglimpfungen im Netz stehen?

In totalitären Staaten erreicht das Problem noch eine ganz andere Dimension. Hier zeigt sich, wie leicht Meinungen als schädlich definiert werden können, wenn sie dem herrschenden Regime nicht gefallen. Mit dem Vorgehen des Westens als Vorwand und denselben Tools werden hier Dissidenten aufgespürt und ausgeschaltet.

Das chinesische Vorgehen wird derzeit in Hongkong sehr transparent, aber auch vermeintlich demokratische Staaten wie die Türkei kündigen an die Zensur im Internet deutlich zu verschärfen, um „Beleidigungen“ und „Falschmeldungen“ auszumerzen.

Allgemeine Erwägungen

Auch im Verhältnis Bürger/Staat kommen also die Schwächen der KI zum Vorschein. Die Algorithmen sind am Ende nur menschengemachte Programme die innerhalb eines vorgegebenen Rahmens (Logik) ablaufen. Im Vergleich zum Menschen haben sie wenig Flexibilität und reagieren mechanistisch, ohne Empathie oder „Verständnis“ für die Situation des Einzelnen. Als Werkzeuge eines totalitären Staates sind sie erbarmungslos und kennen keine moralischen Hemmschwellen oder Gewissensbisse.

Natürlich fällt auch der Mensch seine Urteile auf Basis seiner Erfahrungen (Training, Datenbasis) und erlernter Verhaltensmuster (Algorithmus). Dabei werden persönliche Erfahrungen gegenüber „objektiven“ Fakten wahrscheinlich stark übergewichtet. Auch das kulturelle und soziale Umfeld hat einen großen, verzerrenden (?) Einfluss.

Neben einer notwendigen Demystifizierung der künstlichen Intelligenz, ist eine ebensolche der menschlichen Intelligenz vielleicht ganz angebracht.