

W3_Ethik_Automatisierte Entscheidungen & Datendiskriminierung

Skript

Erarbeitet von
Marie von Lobenstein

Lernziele	1
Inhalt	2
Einstieg	2
Datendiskriminierung und automatisierte Entscheidungen	2
Das Amazon Beispiel	3
Vor- und Nachteilen von automatisierten Entscheidungen	3
Fun Fact: DSGVO	4
Take-Home Message	5
Quellen	5

Lernziele

- Automatisierte Entscheidungen als Begriff definieren können und verstehen.
- Vor- und Nachteile von automatisierten Entscheidungen nennen können.
- Ethische Fragen zum Thema automatisierte Entscheidungen benennen können.
- Hilfestellungen nennen können, bei der Verbesserung von automatisierten Entscheidungen.

Inhalt

Einstieg

KI hilft uns in vielen wichtigen Bereichen mit großen Datenmengen zurechtzukommen. Diese Datenmengen nutzen KI-Systeme als Grundlage, um uns bei Entscheidungen zu helfen. Das können Entscheidungen für banale Dinge sein, wie „welche Songs passen gut in meine Herzscherz-Playlist“ oder aber Entscheidungen, die uns nachhaltig betreffen, wie zum Beispiel „bekomme ich den Kredit“. Beispielsweise in der Medizin, im Recht oder auf dem Arbeitsmarkt werden KI's bereits genutzt oder getestet um uns in Entscheidungsfindungsprozessen zu unterstützen. Gleichzeitig werden aber leider auch immer mehr diskriminierende Effekte durch KI's bekannt, die uns bewusst machen, dass natürlich auch Datensammlungen und Algorithmen Teil der Gesellschaft sind und damit gesellschaftliche Diskriminierung reproduzieren. (Chun 2021) (Stalder 2016) (Noble 2018) (Finn 2017) Diese Diskriminierungseffekte werfen ethische oder moralische Fragen dazu auf, wie, ob und in welchem Umfang künstliche Intelligenz Entscheidungen für den Menschen treffen darf oder sollte. Besonders dann, wenn die Rede von automatisierten Entscheidungen ist.

Quelle [1] [2] [3] [4]

Datendiskriminierung und automatisierte Entscheidungen

In unserer traditionellen Wahrnehmung, haben vor allem andere Menschen oder wir selbst über bestimmte gesellschaftliche Prozesse entschieden. Andere Personen haben beispielsweise entschieden, ob wir den Job bekommen, für den wir uns beworben haben. Richter und Richterinnen haben darüber entschieden wie das Urteilsmaß in Gerichtsprozessen ausfällt und Ärzte und Ärztinnen waren dafür verantwortlich die passende Behandlung für unsere Erkrankungen zu empfehlen. Es werden bereits immer mehr KI-Systeme genutzt, um Menschen diese Entscheidungen zu vereinfachen. Ein spannendes und heiß diskutiertes Feld sind die sogenannten automatisierten Entscheidungen also „Entscheidungen ohne jeglichen menschlichen Einfluss.“ (Hoeren & Niehoff, 2018)

Hier mal ein Beispiel für solch einen Entscheidungsprozess durch ein KI-System.

Quelle [5]

Das Amazon Beispiel

Amazon hat 2018 Gebrauch von automatisierten Entscheidungen durch ein KI-System gemacht. Dieses KI-System hatte den ursprünglichen Zweck die Bewerbungsprozesse bei Amazon zu vereinfachen und eine Vorauswahl aus den eingegangenen Bewerbungen zu treffen. Dieses KI-System wurde durch eine Datenlage aus Lebensläufen vergangener Bewerberinnen und Bewerbern trainiert. Mit dem Wissen aus diesen Daten suchte das KI-System dann im Internet potenzielle Kandidaten und Kandidatinnen, die der Personalabteilung vorgestellt wurden. (Hamilton 2018) Das Entwicklungsteam des KI-Systems musste allerdings feststellen, dass das System Frauen in dieser Vorauswahl benachteiligt.

Einblendung Screenshot Business Insider Artikel "Amazon eine Künstliche Intelligenz für Bewerbungen gebaut, die Frauen diskriminiert".

Eine Erklärung dafür: die Daten, mit der das KI-System trainiert wurde, bestehen hauptsächlich aus den Bewerbungen von Männern. Die KI hat daraus „gelernt“, dass Männer Frauen im Bewerbungsprozess vorgezogen werden sollten, um vielversprechendere Mitarbeiter einzustellen. Die KI hat zum Beispiel Bewerbungen mit den Worten Frau herausgefiltert. Das Projekt wurde von Amazon daraufhin beendet. Das offizielle Statement dazu: das Vertrauen an die KI ist verloren gegangen.

Im schlechtesten Fall könnte diese KI also genutzt werden, um gezielt Menschen aus Bewerbungsprozessen auszuschließen. Was umgekehrt auch bedeuten würde, dass bestimmte Bevölkerungsgruppen einen Vorteil hätten Jobs zu bekommen. Eine ethische Frage könnte hier also sein, sollen wir KI benutzen für solche Bewerbungsprozesse, auch auf die Gefahr hin, dass sie bestimmte Menschen ausschließen könnte? Das war jetzt ein eher negatives Beispiel für automatisierte Entscheidungen, das zeigt, dass verzerrte Datensätze ein Problem sein können. Expertinnen und Experten erhoffen sich aber durchaus auch Vorteile durch automatisierte Entscheidungsprozesse.

Quelle [6]

Vor- und Nachteilen von automatisierten Entscheidungen

Erhoffte Vorteile sind beispielsweise eine schnellere Datenverarbeitung und Auswertung. Gerade in Verwaltungsprozessen könnten automatisierte Entscheidungen das System entlasten. Entscheidungen könnten in Echtzeit getroffen werden. Ein KI-System kann mit einem größeren Gesamtwissen ausgestattet werden, also mehr Quellen für Entscheidungen heranziehen. Auf der einen Seite erhofft man sich von diesen KI-Systemen einen rationalen Umgang mit den Daten, Maschinen sind nicht emotional involviert. Auf der anderen Seite hat die Vergangenheit gezeigt das schlechte Trainingsdaten und

einprogrammierter Bias, eben doch noch menschliches Wirken in KI aufweist. Und das meist zum Nachteil. (Puppe 2022) Außerdem ist ein KI-System nur so gut, wie die Daten mit der es trainiert wurde. In der Medizin zB arbeiten Ärzte und Ärztinnen oft mit vieldeutigen Symptomen, diese zu erkennen und eine Diagnose zu stellen ist sehr komplex. Die Frage ist also, kann eine KI überhaupt genug aus Daten lernen, menschliches Denken reproduzieren, um diese komplexen Diagnosen zu stellen? Oder im Fall von juristischer Entscheidungsfindung, können Dinge wie Gerechtigkeitsempfinden oder Resozialisierungschancen überhaupt algorithmisch erfasst werden? Aus dem Amazon Beispiel wird außerdem folgendes ersichtlich: das Vertrauen in KI-Systeme ist sehr fragil. Es heißt aber ja sehr schön „Aus Fehlern lernt man“. Das gilt auch für KI-Systeme. Erlauben wir also KI-Systemen in Zukunft aus Fehlern zu lernen, auch in Bereichen in denen Fehler hohe Kosten haben? Darf eine KI beispielsweise aus Fehlern lernen, die Menschenleben beeinträchtigen?

Quelle [7]

Fun Fact: DSGVO

In Deutschland müsste sich im Übrigen erst mal die Gesetzeslage ändern, wenn wir künstliche Intelligenzen voll automatisierte Entscheidungen treffen lassen wollen. Unter anderem, weil Artikel 22, Absatz 1 der DSGVO folgendes sagt:

Die betroffene Person hat das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung – einschließlich Profiling – beruhenden Entscheidung unterworfen zu werden, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt.

Das USA Beispiel

In den USA wird das sogenannte „predictive policing“ stark diskutiert. Predictive Policing ist die Praxis, unterstützt durch KI-Systeme, Voraussagen über das kriminelle Potential einer Person oder Voraussagen über mögliche Verbrechen zu tätigen. Größere USA Polizeibehörden experimentieren schon länger mit KI als Hilfe zur Verbrechensbekämpfung. Dabei haben verschiedene Studien gezeigt, dass dies bestehende systemische Rassismen reproduziert. (Heaven 2020) (O’Neil, 2016) (Fry, 2019) Dabei basiert „predictive policing“ meist entweder auf Standort basierten Auswertungen oder personenbezogenen Daten. Algorithmen werten Zusammenhänge zwischen Orten und Verbrechensraten aus oder Daten wie Herkunft, Alter, Vorstrafen, etc. Richter und Richterinnen können beispielsweise KI nutzen, um eine Angabe darüber zu erhalten wie wahrscheinlich die Rückfallquote für eine angeklagte Person ist und darauf basierend das Strafmaß aussprechen. US Amerikanische Polizist:innen erhoffen sich von KI-Systemen, dass sie es ihnen ermöglichen einzugreifen bevor eine Straftat begangen wird.

Wie in vielen Fällen der automatisierten Entscheidungen, sind auch hier die Trainingsdaten das große Problem. Denn diese speisen die KI mit verzogenen Verhaftungsstatistiken, die keine systemischen Faktoren mit einbeziehen. In den USA bedeutet das beispielsweise, dass

Schwarze Menschen eine mehr als zweimal so hohe Wahrscheinlichkeit haben, verhaftet zu werden. Es ist auch nicht unüblich, dass es eine überdurchschnittlich hohe Polizeipräsenz in Stadtteilen mit überwiegend schwarzer Bevölkerung gibt. Auch diesem Beispiel liegt die Annahme zu Grunde, dass KI's uns helfen, weil sie objektiver sind als Menschen. Wir trainieren aber die KI-Systeme oft mit verzerrten Datensätzen und den Biases, die in uns allen existieren. Das menschliche Wirken in den Prozessen von KI darf nicht unterschätzt werden. Deswegen ist es sehr wichtig, dass wir uns fragen WER, also der Mensch oder die Maschine, solche Entscheidungen trifft. Auch sollten wir uns mit der Frage beschäftigen, ob wir voll automatisierten Systemen, diese Entscheidungen überlassen wollen oder ob es den Mensch braucht, um diese Entscheidungen zu überwachen oder zu kontrollieren, also als Endentscheider*innen.

Quelle [8] [9] [10]

Take-Home Message

Sollten wir uns dazu entscheiden KI-Systemen in weiteren Bereichen unserer Gesellschaft Entscheidungen zu überlassen, dann müssen wir vorsorgen und sicher gehen, soweit das überhaupt möglich ist, dass sie dadurch niemanden diskriminieren. Dabei helfen können korrekte und inklusive Trainingsdaten, eine Leitung und Optimierung durch Fachexpert*innen und nachvollziehbare, also transparente Algorithmen. Und eine offene gesellschaftliche Debatte über die Potentiale und Risiken automatisierter Entscheidungsfindung.

Quellen

- Quelle [1] Chun, W. H. K., & Barnett, A. (2021). Discriminating data: correlation, neighborhoods, and the new politics of recognition. The MIT Press.
- Quelle [2] Stalder, F. (2016). Kultur der Digitalität (Erste Auflage, Originalausgabe). Suhrkamp.
- Quelle [3] Noble, S. U. (2018). Algorithms of oppression: how search engines reinforce racism. New York University Press.
- Quelle [4] Finn, E. (2017). What algorithms want: imagination in the age of computing. MIT Press.
- Quelle [5] Hoeren, T., & Niehoff, M. (2018). KI und Datenschutz-Begründungserfordernisse automatisierter Entscheidungen. Rechtswissenschaft, 9(1), 47–66. <https://doi.org/10.5771/1868-8098-2018-1-47>

- Quelle [6] Hamilton, I. A. (2018, October 11). Amazon hat eine Künstliche Intelligenz für Bewerbungen gebaut, die Frauen diskriminierte. Business Insider.
<https://www.businessinsider.de/wirtschaft/amazon-hat-eine-ki-gebaut-die-frauen-diskriminierte-2018-10/>
- Quelle [7] Puppe, F. (2022). Gesellschaftliche Perspektiven einer fachspezifischen KI für automatisierte Entscheidungen. Informatik Spektrum, 45(2), 88–95.
<https://doi.org/10.1007/s00287-022-01443-6>
- Quelle [8] Heaven, W. D. (2020). Predictive policing algorithms are racist. They need to be dismantled. MIT Technology Review.
<https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/>
- Quelle [9] O’Neil, C. (2016). Weapons of math destruction: how big data increases inequality and threatens democracy (First edition). Crown.
- Quelle [10] Fry, H. (2019). Hello world: being human in the age of algorithms. W.W. Norton & Company.