



SIND KI-SYSTEME NEUTRAL?

„Du hast zwar alles richtig beantwortet in der Abfrage, aber eine Eins bekommst Du nicht, weil mir die Farbe Deiner Hose nicht gefällt!“

Wenn Deine Lehrkraft so die Noten festlegen würde, fändest Du das sicher ziemlich unfair, oder? Leider gibt es in unserer Gesellschaft viele Situationen, bei denen wichtige Entscheidungen so getroffen werden: Jemand bekommt eine Wohnung aufgrund seiner Hautfarbe nicht, ein ausländischer Nachname oder das Geschlecht führen zur Absage bei der Bewerbung auf einen supercoolen Praktikumsplatz oder jemand wird wegen seiner Religion von anderen gemieden.

Warum verhalten sich Menschen so? Wir haben Vorurteile, die wir durch unser Umfeld aufgebaut haben. Dabei vertrauen wir nur auf unser Gefühl ohne überhaupt zu wissen, ob unsere Vermutungen in Wirklichkeit zutreffen. Viele haben auch Angst vor allem, was ihnen fremd ist oder was einfach anders ist als sie es gewohnt sind. Deswegen verhalten sie sich abweisend und beurteilen andere Menschen basierend auf Eigenschaften, die gar nichts über den Charakter, das Verhalten oder die Einstellung einer Person aussagen – vor allem nichts Schlechtes! Die Farbe Deiner Hose sagt ja auch nichts über Dich als Person aus!

Diskriminierung aufgrund von Herkunft, Hautfarbe, Geschlecht, Religion, sexueller Orientierung, körperlicher Beeinträchtigung oder vielen anderen Faktoren kommt in der Gesellschaft ständig vor. Wir können diese nur bekämpfen, wenn wir uns alle gemeinsam dagegen wehren, darüber aufklären und uns gegenseitig unterstützen. Zum Glück gibt es auch viele Organisationen, die solche Ungerechtigkeiten aufdecken und bekämpfen. Trotzdem haben wir unser Ziel, dass alle in der Gesellschaft fair und gerecht behandelt werden, noch lange nicht erreicht.



Frau Sarah Maier
Geb. 12.07.1985

Abitur: 1,0
Bachelor of Science in Wirtschaftsinformatik: 1,3
Master of Science in Wirtschaftsinformatik: 1,3

Mehrfähige Berufserfahrung als Softwareentwicklerin

Sehr gut geeignet. Beste Bewerberin. Einwandfreier und vielversprechender Lebenslauf.



Herr Martin Schmitt
Geb. 17.10.1986

Abitur: 3,5
Bachelor of Science in Mathematik in Harvard: 3,4
Master of Science in Mathematik in Harvard: 3,2

Keine Berufserfahrung. Häufiger Wechsel der Arbeitgeber.

Entspricht nicht allen Anforderungen. Qualifikation fraglich.

*Abgelehnt wegen:
Diskriminierung!*

04 (UN)FAIRE KI?!

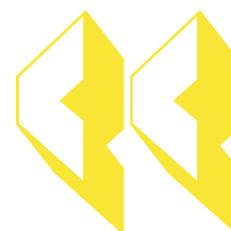


**FAIRE MASCHINEN?
- VON WEGEN!**

Wenn Menschen so voreingenommen sind, könnte es doch sinnvoller sein, KI-Systemen wichtige Entscheidungen zu überlassen, oder? Sie haben keine Gefühle und können daher ganz neutral und ohne Vorurteile urteilen – das wäre doch ziemlich super. Leider funktioniert das aber in Wirklichkeit nicht so: Wie ihr vielleicht schon an anderen Stationen gesehen habt, müssen KI-Systeme ihr Verhalten erst lernen. Und dafür brauchen sie ziemlich viele Daten. Wenn ein KI-System zum Beispiel entscheiden soll, welcher Bewerber eine Wohnung oder einen Ausbildungsplatz bekommt, dann muss das KI-System vorher anhand von Daten lernen, wer bisher eine Wohnung oder einen Ausbildungsplatz bekommen hat. Und woher kommen diese Daten? Richtig, aus den Entscheidungen von Menschen! Und da hast Du ja oben im Text schon gesehen, dass diese ziemlich unfair sein können.

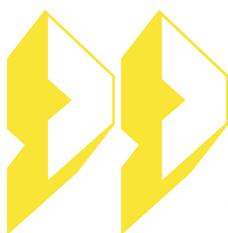
KI-Systeme, die aus von Menschen erzeugten und beeinflussten Daten lernen, wie sie sich verhalten sollen, lernen dieselben Verhaltensweisen, die auch wir Menschen zeigen. **Bei ihren Entscheidungen haben sie also auch Vorurteile und diskriminieren bestimmte Gruppen der Bevölkerung.** Man bezeichnet diese Voreingenommenheit als **Bias** (engl. Voreingenommenheit, Verzerrung). Das hast Du vielleicht beim Ausprobieren der Station bemerkt. Auch in der realen Welt gibt es entsprechende Beispiele von diskriminierenden KI-Systemen. Zum Beispiel gab es 2015 einen Rassismus-Skandal um die Bilderkennung von Google Fotos, weil das System Menschen mit dunklerer Hautfarbe auf Fotos als Gorillas markiert hat. Das ist nicht nur falsch, sondern darüber hinaus sehr beleidigend und lag daran, dass der Algorithmus von Google wohl das Aussehen von Menschen nur mit Bildern von eher hellhäutigen Personen gelernt hat. Die Trainingsdaten haben somit die Gesellschaft nicht richtig dargestellt und waren verzerrt.

Der Bias entsteht folglich meist aufgrund der benutzten Daten, kann aber zum Beispiel auch durch falsche statistische Messungen oder Fehler im Algorithmus verursacht werden. Ein KI-System mit Bias würdest Du sicher nicht gerne benutzen wollen, oder? Mit Menschen, die Entscheidungen treffen, kann man wenigstens über ihre Entscheidungen sprechen und ihre Vorurteile diskutieren oder widerlegen. Bei Maschinen klappt das nicht.





Wenn das alles so schlecht ist, warum gibt es dann überhaupt KI-Systeme, die irgendetwas entscheiden? Dadurch, dass wir wissen, dass viele Entscheidungen einen Bias haben, können wir auch versuchen, diesen bei KI-Systemen zu bekämpfen! Dazu wird versucht, aus den Daten alle Informationen zu löschen, die zu Diskriminierung führen könnten. Beispielsweise wird in KI-Systemen nicht mehr erlaubt, die Hautfarbe oder das Geschlecht in Entscheidungen einzubeziehen, damit ihr Urteil diesbezüglich neutral wird. Allerdings ist es ganz schön schwierig, wirklich alle Informationen in den Daten zu entdecken, die zu einem beeinflussten Ergebnis führen könnten. So kann eine Adresse zum Beispiel verraten, ob man eher arm oder eher reich ist, wenn in einem bestimmten Stadtteil eher ärmere oder eher reichere Menschen leben. Zudem können strukturierte Zusatzdaten (sogenannte Metadaten) in Datensätzen ebenfalls Informationen enthalten, die Rückschlüsse auf Personen zulassen – ganz ohne, dass ein Nutzer das bemerkt. Deswegen ist es besonders wichtig, bei der Erstellung von KI-Systemen gründlich zu prüfen, mit welchen Daten das System lernen soll. Auch die Ergebnisse des KI-Systems müssen immer wieder kritisch überprüft werden, ob sie auch neutral sind. Deswegen wird versucht, KI-Systeme so zu entwickeln, dass man alle Entscheidungen verstehen und zudem erklären kann, aufgrund welcher Information sie getroffen wurden. Das macht KI-Systeme transparent und fair. Zusätzlich braucht die Gesellschaft Regeln und Gesetze dafür, für welche Aufgaben KI-Systeme verwendet werden und welche Entscheidungen damit überhaupt getroffen werden dürfen, um Menschen zu schützen. Diese werden von den Politikern in Deutschland und Europa gerade entwickelt.



Und was sagt ein KI-System zur Frage: „Sind KI-Systeme rassistisch?“ ChatGPT hat Folgendes geantwortet:

„Künstliche Intelligenz (KI)-Systeme selbst sind nicht rassistisch, da sie keine Emotionen oder Vorurteile haben. Allerdings können KI-Systeme aufgrund von

Vorurteilen in den Daten, mit denen sie trainiert wurden, rassistische Ergebnisse produzieren. Wenn KI-Systeme auf Daten trainiert werden, die unbewusste oder sogar bewusste Vorurteile widerspiegeln, wie sie in der Gesellschaft existieren, können sie ungewollt rassistische oder diskriminierende Entscheidungen treffen. Zum Beispiel können KI-Systeme, die auf Daten von Bewerbungen trainiert wurden, automatisch männliche Kandidaten bevorzugen, wenn sie zuvor mehr männliche Bewerber ausgewählt haben, da dies die Grundlage ihres Trainings ist. Es ist daher wichtig, bei der Entwicklung von KI-Systemen auf eine sorgfältige Datenauswahl und -bereinigung zu achten, um Vorurteile und Diskriminierung zu minimieren. Zudem sollte eine angemessene Überwachung und Kontrolle von KI-Systemen durchgeführt werden, um sicherzustellen, dass sie fair und unvoreingenommen bleiben.“



QUELLEN

Graphik „Abgelehnt wegen: Diskriminierung!“:

erstellt von: Michaela Müller-Unterweger, Fotos: <https://pxhere.com/de/photo/706486>, <https://pxhere.com/de/photo/706040>, beide CCO-Lizenz, <https://pixabay.com/vectors/rubber-stamp-rejected-denied-vector-2054006/>, Fonts: Calibri, Really Free Font, CC-BY-ND Ermin Dwi Alita, Annisa Aulia Rasyidah, and Herbanu Tri Sasongko

Rassistischer Google-Algorithmus

<https://www.zeit.de/digital/internet/2015-07/google-fotos-algorithmus-rassismus>

ChatGPT

erstellt mit <https://chat.openai.com/> im April 2023

QUELLEN (UN)FAIRE KI?! - APP

Icons

Handschellen: https://www.flaticon.com/free-icon/handcuff_4435702?term=handcuff&page=1&position=2&origin=search&related_id=4435702

Tasche: https://www.flaticon.com/free-icon/portfolio_1063376?term=job&page=1&position=1&origin=search&related_id=1063376

Gebäude: https://www.flaticon.com/free-icon/buildings_558411?term=flat&page=1&position=3&origin=search&related_id=558411

Schule: https://www.flaticon.com/free-icon/classroom_3197831?term=school&page=1&position=20&origin=search&related_id=3197831

Datensatz: https://www.flaticon.com/free-icon/dataset_8003249?term=dataset&page=1&position=3&origin=search&related_id=8003249

Home-Button: https://www.flaticon.com/free-icon/home_25694?term=home&page=1&position=1&origin=search&related_id=25694

Portraits Testpersonen

Alan: <https://pixabay.com/de/photos/schwarzer-mann-afroamerikaner-3532117/>, Bild von jidaley

Maria: <https://pixabay.com/de/photos/frau-frauen-b%c3%bcro-arbeitsgesch%c3%a4ft-2773007/>, Bild von ernestoelava

Carmen: <https://pixabay.com/de/photos/frau-iranische-frau-alte-frau-iran-7262808/>, Bild von Javad_esmaeili

Fabian: <https://pixabay.com/de/photos/mann-rauchen-rauch-zigarette-4132758/>, Bild von 10634669



QUELLEN (UN)FAIRE KI?! - APP

Protraits Testpersonen

Stefanie: <https://pixabay.com/de/photos/frau-brille-business-frau-fachmann-1254454/>,
Bild von ukieiri

Almuth: <https://pixabay.com/de/photos/arbeiten-gesch%a4ftsfrau-weiblich-1219889/>, Bild von helpsg

Ramona: <https://pixabay.com/de/photos/marketing-gesch%a4ftsmann-3468598/>,
Bild von madartzgraphics

Artur: <https://pixabay.com/de/photos/m%a4nnliches-portr%a4t-mann-offiziell-1156365/>, Bild von akintans

Jürgen: <https://pixabay.com/de/photos/mensch-person-mann-nachbar-3429797/>,
Bild von 652234

Adam: <https://pixabay.com/de/photos/%a4lterer-herr-hunde-h%bcndchen-mann-1281284/>, Bild von brenkee

Familie Stretz: <https://pixabay.com/de/photos/familie-hund-portr%a4tk%bcche-liebe-7071153/>, Bild von Michelle_Raponi

Familie Nassar: <https://pixabay.com/de/photos/familie-spielen-spa%9f-1%a4chelt-kind-2901211/>, Bild von Chepearroyo

Paul: <https://pixabay.com/photos/boy-teenager-male-young-portrait-4433619/>,
Bild von djedj

Malaika: <https://pixabay.com/de/photos/m%a4dchen-kind-schwarze-afrikanisch-486950/>, Bild von Counselling

Melissa: <https://pixabay.com/de/photos/m%a4dchen-mensch-h%bcbsch-weiblich-707813/>, Bild von Axe77

Nikolaj: <https://pixabay.com/photos/homework-school-child-homeschooling-2803933/>,
Bild von anaterate

