

Chancen und Risiken von Künstlicher Intelligenz und Algorithmen aus antidiskriminierungsrechtlicher Perspektive

Dr. Duygu Damar, institut für finanzdienstleistungen e.V.

Struktur

Erster Teil: Zivilrechtliches Diskriminierungsverbot

- Europarechtliche Vorgaben
- § 19 AGG
- Unmittelbare / mittelbare Diskriminierung
- Präferenzbedingte / statistische Diskriminierung

Zweiter Teil: KI-spezifische Chancen und Risiken

- Technische Begriffe
- Potential: Beseitigung der präferenzbedingten, unmittelbaren Diskriminierung
- Herausforderung: Statistische (unmittelbare und mittelbare) Diskriminierung

Dritter Teil: Lösungsansätze

Zivilrechtliches Diskriminierungsverbot - I

■ EU-Recht

- Richtlinie 2000/43/EG des Rates vom 29. Juni 2000 zur Anwendung des Gleichbehandlungsgrundsatzes ohne Unterschied der Rasse oder der ethnischen Herkunft, ABl. 2000 L 180/22 → Anti-Rassismus-Richtlinie (AntiRasRL)
- Richtlinie 2004/113/EG des Rates vom 13. Dezember 2004 zur Verwirklichung des Grundsatzes der Gleichbehandlung von Männern und Frauen beim Zugang zu und bei der Versorgung mit Gütern und Dienstleistungen, ABl. 2004 L 373/37 → Gender-Richtlinie (GenderRL)

■ § 19 AGG

- Massengeschäfte und ähnliche Geschäfte (§ 19 Abs. 1 Nr. 1 AGG) → umfassendes Verbot
- Versicherungsverträge (§ 19 Abs. 1 Nr. 2 AGG) → umfassendes Verbot
- Gesundheitsdienste / Bildung / Zugang zu und Versorgung mit Gütern und Dienstleistungen, die der Öffentlichkeit zur Verfügung stehen (§ 19 Abs. 2 i.V.m. § 2 Abs. 1 Nr. 5-8 AGG) → partielles Verbot

Zivilrechtliches Diskriminierungsverbot - II

- § 19 Abs. 1 Nr. 1 AGG
 - Massengeschäfte: „typischerweise **ohne Ansehen der Person** zu vergleichbaren Bedingungen in einer Vielzahl von Fällen zustande kommen“
 - Ähnliche Geschäfte: „ bei denen das **Ansehen der Person** nach der Art des Schuldverhältnisses **eine nachrangige Bedeutung hat** und die zu vergleichbaren Bedingungen in einer Vielzahl von Fällen zustande kommen“
- § 19 Abs. 1 Nr. 2 AGG: Privatversicherungsverträge
- Umfassendes Verbot: rassistische Diskriminierungen und Diskriminierungen aufgrund der ethnischen Herkunft, des Geschlechts, der Religion, einer Behinderung, des Alters und der sexuellen Identität

Zivilrechtliches Diskriminierungsverbot - III

- § 19 Abs. 2 i.V.m. § 2 Abs. 1 Nr. 5-8 AGG
 - Sozialschutz, einschließlich der sozialen Sicherheit und der **Gesundheitsdienste**
 - soziale Vergünstigungen
 - die **Bildung**
 - **Zugang zu und Versorgung mit Gütern und Dienstleistungen, die der Öffentlichkeit zur Verfügung stehen, einschließlich von Wohnraum.**
- Partielles Verbot: rassistische Diskriminierungen und Diskriminierungen aufgrund der ethnischen Herkunft

Zivilrechtliches Diskriminierungsverbot - IV

☺ Verboten ☹ Nicht verboten	Rassistisch & ethnische Herkunft	Geschlecht	Religion	Behinderung	Alter	Sexuelle Identität
Massengeschäfte und ähnliche Geschäfte	☺	☺	☺	☺	☺	☺
Versicherungsverträge	☺	☺	☺	☺	☺	☺
Gesundheitsdienste	☺	EU ☺ AGG ☹	Vorschlag KOM(2008) 426 endgültig			
Bildung	☺	☹	Vorschlag KOM(2008) 426 endgültig			
SV bzgl. Zugang zu und Versorgung mit Gütern und DL	☺	EU ☺ AGG ☹	Vorschlag KOM(2008) 426 endgültig			

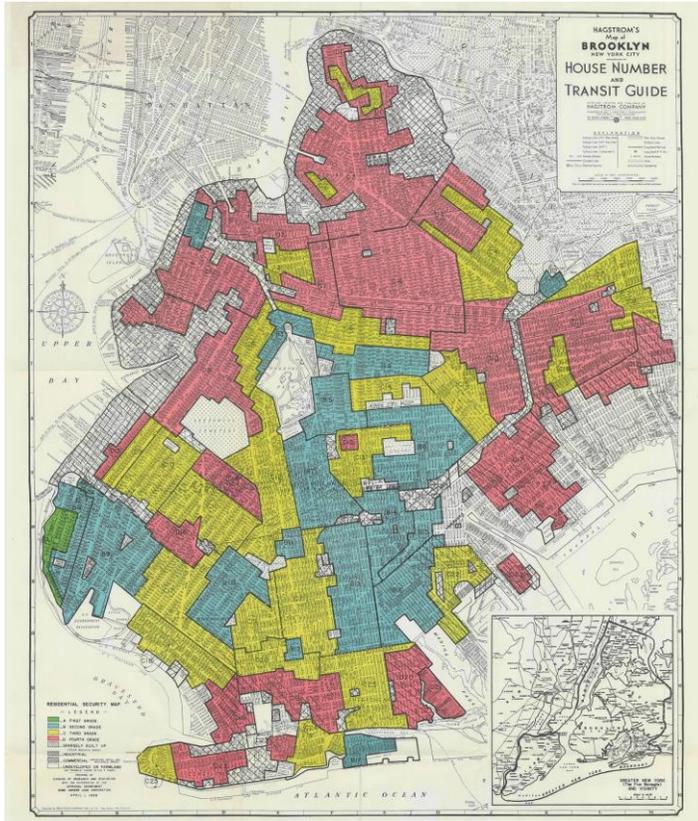
Zivilrechtliches Diskriminierungsverbot - V

- Versicherungsverträge → § 19 Abs. 1 Nr. 2 AGG
- Verbraucherdarlehensvertrag (VDV)
 - Allgemein-VDV (§ 491 Abs. 2 BGB) ./ Immobiliar-VDV (§ 491 Abs. 3 BGB)
 - § 505a BGB → Kreditwürdigkeitsprüfung (KWP); BT-Drs. 16/1780, S. 42
 - Immobiliar-VDV → stark personalisierte KWP → kein Massengeschäft → SV bzgl. Zugang zu und der Versorgung mit Gütern und Dienstleistungen, die der Öffentlichkeit zur Verfügung stehen (§ 19 Abs. 2 i.V.m. § 2 Abs. 1 Nr. 8 AGG)
 - Allgemein-VDV
 - Standardisierte Kreditprodukte, z.B. Sofortkredit, Kleinkredit → Massengeschäfte oder ähnliche Geschäfte i.S.d. § 19 Abs. 1 Nr. 1 AGG
 - Sonstige Allgemein-VDV → SV bzgl. Zugang zu und der Versorgung mit Gütern und Dienstleistungen, die der Öffentlichkeit zur Verfügung stehen (§ 19 Abs. 2 i.V.m. § 2 Abs. 1 Nr. 8 AGG)

Diskriminierung – Juristische Perspektive

- Unmittelbare (direkte) Diskriminierung - § 3 Abs. 1 AGG
 - Eine unmittelbare Benachteiligung liegt vor, wenn eine Person (i) wegen eines in [§ 19] genannten Grundes (ii) eine weniger günstige Behandlung erfährt, als eine andere Person in einer vergleichbaren Situation erfährt, erfahren hat oder erfahren würde.
- Mittelbare (indirekte) Diskriminierung - § 3 Abs. 2 AGG
 - Eine mittelbare Benachteiligung liegt vor, wenn dem Anschein nach neutrale Vorschriften, Kriterien oder Verfahren Personen wegen eines in [§ 19] genannten Grundes gegenüber anderen Personen in besonderer Weise benachteiligen können, [Rechtfertigungsmöglichkeit] es sei denn, die betreffenden Vorschriften, Kriterien oder Verfahren sind durch ein rechtmäßiges Ziel sachlich gerechtfertigt und die Mittel sind zur Erreichung dieses Ziels angemessen und erforderlich.

Mittelbare Diskriminierung – Beispiel: Redlining



Quelle: The New York Times, Self-Fulfilling Prophecies - How Redlining's Racist Effects Lasted for Decades, 24. August 2017

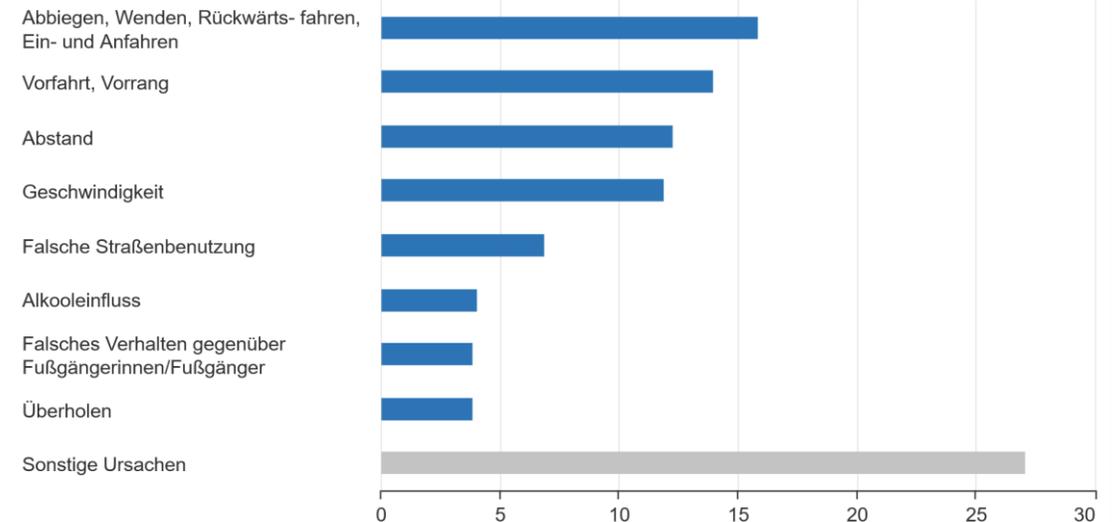
- Ausschluss vom Baufinanzierungsangebot
- Ohne Rücksicht auf die finanzielle Qualifikation der kreditbeantragenden Person

Mittelbare Diskriminierung – Beispiel: Kfz-Versicherung

- Prämienberechnung für die Kfz-Versicherung
- Motorleistung
- Statistisch: Autos mit stärkerer Motorleistung → Männer

Fehlverhalten der Fahrer bei Unfällen mit Personenschaden im Straßenverkehr 2020

in %



© Statistisches Bundesamt (Destatis), 2021

Diskriminierung – Ökonomische Perspektive

- Präferenzbedingte Diskriminierung → Abneigungen und Vorurteile gegenüber bestimmten Personengruppen
- Statistische Diskriminierung
 - Variable → Indikator für eine erwünschte oder unerwünschte Eigenschaft
 - Kriterium für das Rechtsgeschäft → erwünschte oder unerwünschte Eigenschaft (z.B. Fahrverhalten) → Durch die Variable ersetzt (z.B. Geschlecht) → Ersatzkriterium / Stellvertretermerkmal (*proxy criteria*)
 - Vermutung von persönlichen Eigenschaften aufgrund ihrer Zugehörigkeit zu einer bestimmten Gruppe → geschütztes Merkmal → klischeehafte Zuordnung → **unmittelbare Diskriminierung**
 - § 20 Abs. 2 S. 2 AGG: Statistische Diskriminierung wegen der Religion, einer Behinderung, des Alters oder der sexuellen Ausrichtung für Privatversicherungsverträge zulässig
 - § 20 Abs. 2 S. 1 AGG a.F. = Art. 5 Abs. 2 GenderRL a.F. → EuGH, Urt. v. 1.3.2011 – Rs. C-236/09 (*Test-Achats*): ungültig

Struktur

Erster Teil: Zivilrechtliches Diskriminierungsverbot

- Europarechtliche Vorgaben
- § 19 AGG
- Unmittelbare / mittelbare Diskriminierung
- Präferenzbedingte / statistische Diskriminierung

Zweiter Teil: KI-spezifische Chancen und Risiken

- Technische Begriffe
- Potential: Beseitigung der präferenzbedingten, unmittelbaren Diskriminierung
- Herausforderung: Statistische (unmittelbare und mittelbare) Diskriminierung

Dritter Teil: Lösungsansätze

Künstliche Intelligenz – Technische Begriffe

- Algorithmus: computertechnisch vordefinierte Berechnungsschritte
 - Algorithmen der direkten Programmierung
 - Lernende Algorithmen
- Maschinelles Lernen (*Machine Learning*)
 - Lernende Algorithmen
 - Trainingsdatensatz → bestimmte Muster und Korrelationen → Erzeugung eines Modells → Übertragung des Modells auf neue Kontexte

Künstliche Intelligenz - Potential

- Vorbeugung bzw. Beseitigung von präferenzbedingten, unmittelbaren Diskriminierungen
- Grundvoraussetzung: keine diskriminatorischen Qualitätsdefizite
 - Entwicklung lernender Algorithmen → persönliche Abneigungen oder Vorurteile
 - Trainingsdatensatz
 - *Historical bias*
 - Beispiel: KI-System zur Erkennung von Versicherungsbetrug
 - Überrepräsentation bestimmter Personengruppen

Consumer-Lending Discrimination in the FinTech Era*

Robert Bartlett	Adair Morse	Richard Stanton	Nancy Wallace
School of Law UC Berkeley	Haas School of Business UC Berkeley & NBER	Haas School of Business UC Berkeley	Haas School of Business UC Berkeley
rbartlett@berkeley.edu	adair@berkeley.edu	rhstanton@berkeley.edu	newallace@berkeley.edu

February 2019

Künstliche Intelligenz – Herausforderung I



Check out the [beta version](#) of the new UCI Machine Learning Repository we are currently testing! [Contact us](#) if you have any issues.

South German Credit Data Set

Download [Data Folder](#), [Data Set Description](#)

Abstract: 700 good and 300 bad credits with 20 predictor variables. Data from 1973 to 1975. Stratified sample from actual credits with bad credit.

Data Set Characteristics:	Multivariate	Number of Instances:	1000	Area:	Business
Attribute Characteristics:	Integer, Real	Number of Attributes:	21	Date Donated	2019-11-29
Associated Tasks:	Classification, Regression, Clustering	Missing Values?	N/A	Number of Web Hits:	22435

Source:

Statistische Diskriminierung

- Aufbau von statistischen Korrelationen im Trainingsdatensatz
- Kontoguthaben, Laufzeit des beantragten Kredits, Zahlungsmoral der kreditbeantragenden Person anhand vergangenen und parallel laufenden Kreditverhältnissen, Verwendungszweck des Kredits, Höhe des Kreditbetrags, Sparkontoguthaben, Beschäftigungszeit bei dem/der derzeitigen Arbeitgeber:in, Höhe der Kreditraten, **Familienstand und Geschlecht**, Bürg:in, Wohndauer an derzeitiger Adresse, Vermögen, Alter, Wohnungstyp, bisherige Kreditverhältnisse, Beruf, pflegebedürftige Personen, Festnetzanschluss, **Gastarbeiter:in**, Kreditrisiko

Künstliche Intelligenz – Herausforderung II

- *Trivedi, A Study on Credit Scoring Modeling with Different Feature Selection and Machine Learning Approaches, Technology in Society 63 (2020), 101413*
 - Sehr wichtig: Wohnungstyp, **Familienstand und Geschlecht**, Kontoguthaben, mitlaufende Kredite
 - Wichtig: Berufserfahrung, Stellung im Beruf
 - Keine Rolle: Gastarbeiter:in

- *Pedreschi/Ruggieri/Turini, The Discovery of Discrimination, in: Custers/Calders/Schermer (Hrsg.), Discrimination and Privacy in the Information Society, Data Mining and Profiling in Large Databases, Bd. 3, Berlin, Heidelberg 2013, S. 91 (100 f.)*
 - Ablehnung des Kreditantrags mit großer Wahrscheinlichkeit → **Gastarbeiterin** zum Zwecke eines Autokaufs

- Ergebnis: Geschlecht und ggf. ethnische Herkunft → Stellvertretermerkmal für Zahlungsmoral oder Zahlungsfähigkeit → **unmittelbare Diskriminierung**

Künstliche Intelligenz – Herausforderung III

- Verzicht auf geschützte Merkmale im Trainingsdatensatz?
- Ersatz: Vor- und Nachname, Geburtsort, Wohnort
- *Omitted variable bias* → keine Beseitigung von Verzerrungen, sondern → mittelbare Diskriminierung
- Verarbeitung bestimmter personenbezogenen Daten, wie z.B. ethnische Herkunft, überhaupt rechtmäßig?
 - Art. 9 DSGVO
 - Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz), KOM(2021) 206 endgültig vom 21.4.2021 (KI-Vorschlag)
 - Art. 10 Abs. 5 KI-Vorschlag → erlaubt für die Beobachtung, Erkennung und Korrektur von Verzerrungen im Zusammenhang mit Hochrisiko-KI-Systemen
 - KI-Systeme für **Kreditscoring** und **Kreditwürdigkeitsprüfung** → Hochrisiko-KI-Systeme (Art. 6 Abs. 2 i.V.m. Nr. 5 lit. b Anhang III KI-Vorschlag)

Struktur

Erster Teil: Zivilrechtliches Diskriminierungsverbot

- Europarechtliche Vorgaben
- § 19 AGG
- Unmittelbare / mittelbare Diskriminierung
- Präferenzbedingte / statistische Diskriminierung

Zweiter Teil: KI-spezifische Chancen und Risiken

- Technische Begriffe
- Potential: Beseitigung der präferenzbedingten, unmittelbaren Diskriminierung
- Herausforderung: Statistische (unmittelbare und mittelbare) Diskriminierung

Dritter Teil: Lösungsansätze

Lösungsansätze

- Anforderungen des KI-Vorschlags an Hochrisiko-KI-Systeme
- Menschliche Aufsicht
 - EIOPA, Consultative Expert Group on Digital Ethics in Insurance
 - Art. 14 KI-Vorschlag
 - KI-Systeme im Versicherungswesen → nach dem KI-Vorschlag keine Hochrisiko-KI-Systeme!
- Gezielte Erhebung von Daten zu geschützten Merkmalen als Kontrollvariable
 - Art. 10 Abs. 5 KI-Vorschlag

Ende der Präsentation
Vielen Dank für Ihre Aufmerksamkeit



institut für finanzdienstleistungen
Grindelallee 100
20146 Hamburg

www.iff-hamburg.de
institut@iff-hamburg.de

Fon: +49 (0)40 30 96 91 - 13

E-Mail: duygu.damar@iff-hamburg.de