

KI-basiert ermittelte Kreditausfallrisiken sind mit Vorsicht zu genießen

Katja Langenbacher: Die EU ringt noch um eine zeitgemäße Regulierung. Diskriminierungsverbote allein reichen nicht aus



Künstliche Intelligenz (KI) erlaubt die Vorhersage künftiger Ereignisse, über die bislang „nur Gott Bescheid wusste“, so ließ sich Mikio Okumura, Präsident des japanischen Versicherungskonzerns Sampo, vor kurzem zitieren. Das Unternehmen hatte Daten über die Lebensgewohnheiten der Bewohner:innen von Altenheimen gesammelt und auf ihre Korrelation mit dem Zeitpunkt des Eintritts von Demenzercheinungen überprüft. Man erreiche auf diese Weise nicht nur Fortschritte in der medizinischen Prävention, sondern verbessere auch die risikoadjustierte Kalkulation von Versicherungsprämien.

In der von Fintechs eroberten Welt des Konsumentenkredits werden derartige Vorhersagen längst verwendet. Weniger überraschend ist dabei, dass Daten über Kontobewegungen, Einkaufs- und Bezahverhalten im Internet Auskunft über die Cashflow-Entwicklung eines potenziellen Kreditnehmers geben. Zur Einschätzung des Kreditausfallrisikos können aber auch Alter, Geschlecht oder Wohnort relevante Informationen liefern. Dasselbe gilt für den Bildungs- oder Migrationshintergrund, die Religionsgemeinschaft oder eine Parteizugehörigkeit. Plattformen, die sich mit der Einschätzung von Kreditausfallrisiken beschäftigen, bleiben hierbei nicht stehen. Einschlägige Korrelationen mögen etwa auch Freunde auf sozialen Medien, Musikgeschmack oder Freizeitverhalten liefern. Dasselbe gilt für die Marke des verwendeten Mobiltelefons, die Anzahl von Tippfehlern in Textnachrichten oder die Zeit, die für das Ausfüllen eines Online-Formulars benötigt wird.

Beide Seiten profitieren

Von dieser Kombination aus maschinellem Lernen und dem Zugriff auf „Big Data“ profitieren naturgemäß Kreditgeber. Aber auch Kreditnehmer:innen kann ein Vertrag angeboten werden, der mit einer nach herkömmlichem Muster durchgeführten Bewertung nicht möglich gewesen wäre. Im vergangenen Jahr hegte bekanntlich die Schufa mit dem Plan des kontodatenbasierten „Check Now“-Verfahrens derartige Pläne, um Kund:innen mit mangelnder Bonität den Abschluss von Mobilfunkverträgen zu vereinfachen. Die darin enthaltene, sehr weit gefasste Einwilligung zur Datenverwertung irritierte Verbraucherschutzorganisationen und führte umgehend zur Einstellung des Verfahrens.

Der europäische Gesetzgeber hat derartige Entwicklungen im Blick. Von besonderer Relevanz ist dabei der Entwurf einer [EU-Verordnung](#) über künstliche Intelligenz. Aber auch im Rahmen der Reform der Verbraucherkreditrichtlinie finden sich wichtige Bausteine einer künftigen Regulierung.

„Besondere Aufmerksamkeit“, so liest man in Erwägungsgrund (37) der geplanten KI-Verordnung, verdiene der Einsatz von KI, wenn es um den „Zugang zu grundlegenden privaten und öffentlichen Leistungen“ gehe. Die „Kreditpunktebewertung“ (Scoring) und die „Bewertung der Kreditwürdigkeit natürlicher Personen“ hat der EU-Gesetzgeber als besonders relevant identifiziert, weil diese nicht nur den Zugang zu Kredit beeinflussen. Auch auf die Bereitstellung wesentlicher „Dienstleistungen wie Wohnraum, Elektrizität und Telekommunikationsdienstleistungen“ mögen sich die Ergebnisse der algorithmenbasierten Prozesse auswirken. Wegen der zentralen Bedeutung für die Teilhabe von Bürger:innen an den Gewährleistungen der modernen Zivilgesellschaft wurden derartige KI-Systeme als sogenannte „Hochrisikosysteme“ eingestuft.

Dabei beschränkt sich der Regelungsgegenstand der KI-Verordnung auf Entwickler und Verwender des KI-Systems. Die Situation des Endverbrauchers liegt zwar außerhalb dieses Anwendungsbereichs, gibt aber doch Anlass zur Etikettierung als Hochrisikosystem. Diese führt unter anderem zu besonderen Compliance-Anforderungen: Nach Art einer Produktregulierung werden Qualitätssicherungsmaßnahmen formuliert, Zertifizierungsverfahren eröffnet und die Zuständigkeiten von Aufsichtsbehörden geordnet.

Gefahr der Diskriminierung

Im bereits zitierten Erwägungsgrund ist davon die Rede, dass KI-Systeme, die für die Einschätzung der Kreditwürdigkeit eingesetzt werden, die Gefahr von Diskriminierung bergen. In Anlehnung an Forschungsergebnisse der Informatik wird von „historischen Diskriminierungsmustern“ („historical bias“) gesprochen, die dazu führen, dass erlebte Diskriminierung „aufgrund der rassischen oder ethnischen Herkunft, einer Behinderung, des Alters oder der sexuellen Ausrichtung“ fortgeschrieben wird. Der Grund hierfür hat mit den Besonderheiten des maschinellen Lernens zu tun. Ein KI-System „lernt“ auf der Basis derjenigen empirischen Daten, die ihm sein Entwickler zur Verfügung stellt. Es erstellt folglich das Profil des erfolgreichen Kreditnehmers anhand von Personen, die in der Vergangenheit in der Lage waren, ausgereichten Kredit zu bedienen. Umgekehrt formuliert: Personen, die zu einer Gruppe von Menschen gehören, die in der Vergangenheit Schwierigkeiten hatten, einen Kredit zu erhalten, wird auch das KI-System zunächst einmal als risikoreiche Kandidat:innen einordnen.

Für bestimmte Mitglieder genau dieser Gruppe kann sich die KI-Beurteilung gleichwohl positiv auswirken („invisible primes“). Je ähnlicher nämlich die nichttraditionellen Parameter eines potenziellen Kreditnehmers denjenigen der klassisch erfolgreichen Kandidat:innen sind, desto höher die Chance, aufgrund dieser Übereinstimmung trotz des untypischen Profils doch Kredit zu erhalten. Am Beispiel eines Einwanderers oder eines jungen Menschen lässt sich das verdeutlichen. Diese(r) mag (noch) keine Kredithistorie vorweisen können, die ein(e) nach herkömmlichen Mustern beurteilte(r) Kandidat:in zur Verfügung hat. Möglicherweise stimmen aber andere Variablen, welche die KI als relevant erkannt hat, mit einem bislang erfolgreichen Profil überein. Sie oder er mag das Mobiltelefon eines hochpreisigen Herstellers verwenden, eine angesehene Hochschulausbildung vorweisen oder besonders zügig mit Online-Formularen umgehen können. Genau in diese Marktlücke sind einige Fintech-Unternehmen gestoßen und konnten zur Ausweitung des Kreditvolumens an bislang benachteiligte Gruppen beitragen.

Von der auf diese Weise erreichten Inklusion profitieren allerdings nicht alle Menschen in gleichem Maße. Eine ganze Reihe empirischer Studien konnte für die USA zeigen, dass die Ausweitung des Kreditvolumens historisch benachteiligten Gruppen in wesentlich geringerem Umfang zugutekommt. Das kann mit dem beschriebenen Lerneffekt der KI zu tun haben, wenn diese Gruppen keine der Parameter vorweisen können, welche die traditionell erfolgreiche Gruppe auszeichnen. Mitunter gewichtet die KI auch diejenigen Merkmale besonders stark, welche sich bei der Mehrheit der erfolgreichen Gruppe finden. Dann kann es vorkommen, dass zwar andere Merkmale für die Minderheit eigentlich aussagekräftiger wären, diese aber nicht hinreichend gewichtet werden („majority bias“). Wiederum an einem Beispiel verdeutlicht: In der Vergangenheit erfolgreich zurück gezahlter Kredit wirkt bonitätssteigernd. Ob aber das korrekte Bedienen einer „buy now pay later“-Vereinbarung oder gar einer privaten Kreditvereinbarung denselben Effekt hat, hängt davon ab, ob diese als Kredit eingeordnet werden. Geschieht das nicht, können keine positiven Punkte gesammelt werden. Das mag jüngere oder unerfahrene Kreditnehmer benachteiligen.

Recht auf menschliche Aufsicht

Für die rechtliche Einordnung diskriminierender Kreditpraxis verweist die KI-Verordnung noch im Wesentlichen auf die Rechtsordnungen der Mitgliedstaaten. In der Reform der Verbraucherkreditrichtlinie wird dieses Risiko hingegen erstmalig adressiert. „Verbraucher, die ihren rechtmäßigen Aufenthalt in der Union haben“, so ordnet Art. 6 an, „dürfen bei der Kreditvergabe nicht aufgrund von Staatsangehörigkeit, Wohnsitz oder eines in Art. 21 der Grundrechtecharta der EU genannten Grundes benachteiligt werden.“ Die Richtlinie sieht außerdem vor, dass Verbraucher:innen das Recht haben, menschliches Eingreifen zu erwirken, wenn die Kreditwürdigkeitsprüfung eine automatisierte Verarbeitung beinhaltet. Auch eine „aussagekräftige Erläuterung der Prüfung und Funktionsweise der verwendeten automatisierten Verarbeitung“ wird verlangt (Art. 18).

Mit dem Verbot diskriminierender Kreditpraxis sind allerdings zunächst einmal mehr Probleme geschaffen als gelöst. Bislang kennt das europäische Recht zwei Formen der Diskriminierung: die direkte und die indirekte Diskriminierung. Direkt diskriminiert, wer eine Person aufgrund eines geschützten Merkmals in einer vergleichbaren Situation weniger günstig behandelt als eine Person ohne dieses Merkmal. Stellt ein Kreditgeber etwa fest, dass Frauen bei statistischer Betrachtung ein höheres Kreditrisiko aufweisen als Männer, ist es ihm gleichwohl verwehrt, sein KI-Modell so zu trainieren, dass der Einfachheit halber alle Frauen mit einem Risikozuschlag belegt werden.

Frage der Vergleichbarkeit

Praktisch viel relevanter ist die mittelbare Diskriminierung. Sie greift, wenn ein dem Anschein nach neutrales Kriterium, etwa eine Teilzeitbeschäftigung, dazu führt, dass Personen des geschützten Kreises, etwa Frauen, schlechter behandelt werden als Personen ohne das geschützte Merkmal. Um das festzustellen, vollzieht man die Entscheidungsparameter nach. Führt das Weglassen des betroffenen Kriteriums dazu, dass das Ungleichgewicht verschwindet, kommt eine mittelbare Diskriminierung in Betracht. Ob es sich tatsächlich um vergleichbare Umstände handelt oder ein sachlicher Grund für die Ungleichbehandlung besteht, ist erst eine Folgefrage.

Voraussetzung einer mittelbaren Diskriminierung ist somit, dass an einzelne Kriterien angeknüpft werden kann. Genau aus diesem Grund lässt sich diese Rechtsfigur für KI-Systeme nicht ohne weiteres passgenau anwenden. Je größer nämlich die Datenmenge, die ein KI-System verwertet, desto eher wird der Algorithmus ein weiteres scheinbar neutrales Kriterium finden, welches dieselbe Vorhersage mit vergleichbarer Präzision erlaubt („redundant encoding“). Mit dem Geschlecht kann beispielsweise nicht nur das Innehaben eines Teilzeitjobs korrelieren, sondern auch eine bestimmte Körpergröße, Vornamen, Musikgeschmack oder Freizeitverhalten. Ein anspruchsvolles KI-System wird solche Parameter finden und Korrelationen zwischen ganzen Bündeln von Variablen herstellen.

Beschränkt der Entwickler des Systems den Zugriff auf bestimmte Variablen ist die Wahrscheinlichkeit hoch, dass Ersatzvariablen („proxies“) gefunden werden. Eine denkbare Lösung wäre die Durchführung vieler Runden, in welchen immer weitere Variablen aus der Menge erfasster Daten gestrichen werden. Die Folge wäre im Regelfall, dass die Qualität der Vorhersage des Modells leidet. Das – wenn auch begrenzte – inklusive Potenzial des KI-Scoring würde abnehmen und Modelle, die auf breiterer Datenbasis präzisere Vorhersagen erlauben, würden pönalisiert.

Ob es sich bei Diskriminierungsverboten tatsächlich um das beste juristische Werkzeug für den Umgang mit algorithmensbasierter Kreditvergabe handelt, ist nicht nur wegen des Aufsetzens auf konkretisierbaren Entscheidungsparametern zweifelhaft. Unsicher ist auch, ob die heutige Formulierung geschützter Gruppenmerkmale den spezifischen Herausforderungen der KI-Entscheidungsfindung gerecht wird. Ungleichgewichte mögen künftig zwischen ganz unerwarteten Gruppen entstehen, etwa Personen, die häufig oder selten Softwareupdates durchführen, die auf sozialen Medien präsent sind oder hierauf verzichten, die ihr Mobiltelefon sorgfältig aufladen oder deren Akku häufig leer ist. Von Diskriminierungsverboten wird eine derartige Gruppenbildung nur dann – gleichsam zufällig – erfasst, wenn die Gruppe mit einem geschützten Merkmal korreliert.

Wenige effiziente Reaktionsmöglichkeiten

Für Verbraucher:innen entsteht eine kafkaesk anmutende Situation: Sie wissen nicht, welche Daten für ihre Beurteilung relevant sind, beispielsweise das Liegenlassen unbezahlter Rechnungen, und können folglich nicht durch eine gezielte Verhaltensänderung, etwa das pünktliche Bezahlen, reagieren. Der Verwender des KI-Systems hat keinen Anreiz zur Offenlegung relevanter Variablen. Zum einen handelt es sich im Regelfall um geschützte Geschäftsgeheimnisse, zum anderen könnte eine Verhaltensänderung des Kreditnehmers die Aussagekraft der betreffenden Variablen schmälern. Erfährt diese(r) beispielsweise, dass die Installation einer Dating-App dem Creditscore schadet, die Verwendung einer Trading-App sich hingegen positiv auswirkt, mag sie oder er erstere löschen und letztere herunterladen („gaming the system“). Beschränkt sich ihre oder seine Verhaltensänderung hierauf, ändert sich seine Beurteilung zu Unrecht.

Der Entwurf eines Regelungsrahmens für KI-Systeme, die bei Scoringverfahren und bei der Kreditwürdigkeitsprüfung zur Verwendung kommen können, dürfte deshalb über Diskriminierungsverbote hinaus auszugreifen haben. Zu dessen zentralen Elementen wird man zunächst eine Qualitätskontrolle zählen dürfen, wie sie die KI-Verordnung vorzeichnet. Das betrifft die Qualität des KI-Systems selbst, aber auch die Zuverlässigkeit der verwendeten Daten. Werden diese beispielsweise auf sozialen Medien gesammelt, kommen Missverständnisse und Fehler häufig vor. Die Möglichkeiten der Datenschutz-Grundverordnung (DSGVO), Zugang und gegebenenfalls Berichtigung unzutreffender Daten zu verlangen, setzen die klar verständliche Aufklärung des Verbrauchers ebenso voraus wie effiziente Verfahren der Rechtsdurchsetzung.

Reichlich Diskussionsstoff

Die wissenschaftliche und die öffentliche Diskussion über diesen Fragenkreis steckt noch in den Kinderschuhen. Zweifelhaft ist beispielsweise, ob die Verbraucherkreditrichtlinie einen überzeugenden Weg gefunden hat, wenn sie vorgibt: „personenbezogene Daten, wie Daten, die auf Social-Media-Plattformen gefunden werden, oder Gesundheitsdaten, einschließlich Daten zu Krebserkrankungen, sollten bei der Kreditwürdigkeitsprüfung nicht verwendet werden“ (Erwägungsgrund 47). Soweit ausgereifte KI-Systeme derartige Informationen ohne weiteres über den Umweg der Ersatzvariablen, beispielsweise häufiger Google-Suchen nach bestimmten Krankheiten oder Medikamenten, nachvollziehen können, ist Verbraucher:innen hiermit kaum geholfen.

Illustrieren lässt sich die Debatte um einem zeitgemäßen Regelungsrahmen auch mit Blick auf KI-basierte Preisdiskriminierung. Insofern verwundert die Verbraucherkreditrichtlinie erneut, wenn sie sich gegenüber personalisierter Preisbildung offen zeigt (Erwägungsgrund 40). Zwar mögen hierdurch in manchen Situationen gesamtgesellschaftliche Effizienzgewinne entstehen. Darüber darf jedoch nicht vergessen werden, dass KI-Systeme nicht nur „invisible prime“-Kandidaten, sondern auch besonders vulnerable potenzielle Kreditnehmer:innen finden können. Unerfahrene, in Finanzdingen schlecht ausgebildete Kreditnehmer:innen oder Personen, die besonders dringend auf einen Kredit angewiesen sind, können sich dem moralisch agnostischen KI-System als attraktive Marktchance für einen hochpreisigen Kredit präsentieren. Beruhigen mag insoweit, dass die Richtlinie künftig eine Obergrenze für Zinssätze, effektiven Jahreszins und Gesamtkosten des Kredits vorsieht (Art. 31). Jedenfalls dem Import des US-amerikanischen „predatory lending“ dürfte damit ein Riegel vorgeschoben werden.

Katja Langenbucher ist Professorin für Bürgerliches Recht, Wirtschaftsrecht und Bankrecht im House of Finance der Goethe-Universität Frankfurt und koordiniert das LawLab – Fintech & AI als SAFE-Brückenprofessorin.

Dieser Beitrag ist zuerst in der [Börsen-Zeitung](#) erschienen.

Blogbeiträge repräsentieren die persönlichen Ansichten der Autorinnen und Autoren und nicht notwendigerweise die von SAFE oder seiner Beschäftigten.



Prof. Dr. Katja Langenbucher
Kordinatorin LawLab – Fintech & AI



BACK

AKTUELLES

Alle News

SAFE Finance Blog

Veranstaltungen

Videos

SAFE in den Medien

Newsletter

Forschungsmagazin Points

In this Section:

FOLLOW US



KONTAKT

Leibniz-Institut für
Finanzmarktforschung SAFE

Theodor-W.-Adorno-Platz 3
60323 Frankfurt am Main

Telefon: +49 69 798 30080

Fax: +49 69 798 30077

E-Mail: info@safe-frankfurt.de

SAFE

Über SAFE

Offene Stellen

So finden Sie uns

POLICY CENTER

SAFE Senior Fellows

Policy Publikationen

SAFE Finance Blog

Policy Center Team

RESEARCH

Forschungsabteilungen

Forschungsteam

Data Center

NEWS & EVENTS

Alle News

Presse

Expertendatenbank