

ist da!



Der EU AI Act kommt! Wie betrifft das uns Aktuarinnen und Aktuare?

Dr. Fabian Bohnert, Dr. Mischa Pupashenko

Köln, 14. Dezember 2023

Kurzbeschreibung

Zum AI Act wurde am 8. Dezember 2023 eine vorläufige Einigung erzielt. Der AI Act wird nun in den Details finalisiert und ist dann nach der Verabschiedung innerhalb von zwei Jahren umzusetzen. Die Definition von künstlicher Intelligenz im Sinne des AI Acts ist sehr breit und umfasst viel mehr als nur KI-Modelle und -Systeme im engeren Sinne. Für Hochrisiko-KI-Systeme – wie die Preisbildung und Risikoprüfung in der Personenversicherung – ergeben sich umfangreiche Anforderungen hinsichtlich Governance, Risikomanagement/IKS, Daten, Sicherheit und Überwachung sowie Transparenz-, Informations- und Dokumentationspflichten. Darüber hinaus ist beispielsweise sicherzustellen, dass die Ergebnisse des KI-Systems erklärbar und nicht-diskriminierend sind.

In diesem Vortrag geben wir einen Überblick über die aktuellen **regulatorischen Entwicklungen** und die sich daraus ergebenden Anforderungen. Wir beleuchten exemplarisch **praktische Herausforderungen** für uns Aktuarinnen und Aktuare, die sich absehbar aus der Umsetzung des EU AI Acts ergeben, und stellen **Lösungsansätze** vor, z.B. um die **Erklärbarkeit** der oft als Black-Box wahrgenommenen Modelle zu verbessern.

Die in diesem Vortrag dargestellten Inhalte zum AI Act basieren auf den zum Zeitpunkt des Vortrags öffentlich verfügbaren Informationen und können aufgrund der weiteren Entwicklungen auch kurzfristig veraltet sein.

Heute bei Ihnen



**Dr. Fabian
Bohnert**

Director
KPMG FS Insurance



**Dr. Mischa
Pupashenko**

Senior Manager
KPMG FS Insurance



Agenda

- 01** Regulatorische Entwicklungen
- 02** Praktische Herausforderungen
- 03** Erklärbarkeit von KI-Modellen

Die EU reagiert mit dem AI-Act auf die Sorgen der Bevölkerung und aktuelle technische Entwicklungen

Das Stimmungsbild in der Gesellschaft...

KPMG – Studie



Drei aus vier Teilnehmern (75%) geben an, bereit zu sein, einem KI-System zu vertrauen, wenn Sicherheitsmechanismen bestehen, die ethischen und verantwortlichen Gebrauch sicherstellen



Nur zwei aus fünf Teilnehmern halten die derzeitigen Sicherheitsmaßnahmen rund um KI (Regeln, Regulierungen und Gesetze) für ausreichend



Die große Mehrheit der Teilnehmer (71%) hält KI-Regulierung für notwendig

Quelle: KPMG Studie – [Trust in Artificial Intelligence: A global study \(2023\)](#)



Jedes achte Unternehmen verwendet bereits KI – vornehmlich im Bereich der Spracherkennung und der Automatisierung

... spiegelt sich in den Plänen der EU wider



European Parliament
@Europarl_EN

Artificial Intelligence: new transparency and risk-management rules for AI systems have been endorsed by Parliament's internal market and civil liberties committees.

All MEPs are expected to vote on the mandate in June so that negotiations can start with the Council.

ARTIFICIAL INTELLIGENCE

- Once approved, new EU rules will be the world's first rules on Artificial Intelligence
- MEPs include bans on biometric surveillance, emotion recognition, predictive policing AI systems
- Tailor-made regimes for general-purpose AI and foundation models like GPT
- The right to make complaints about AI systems



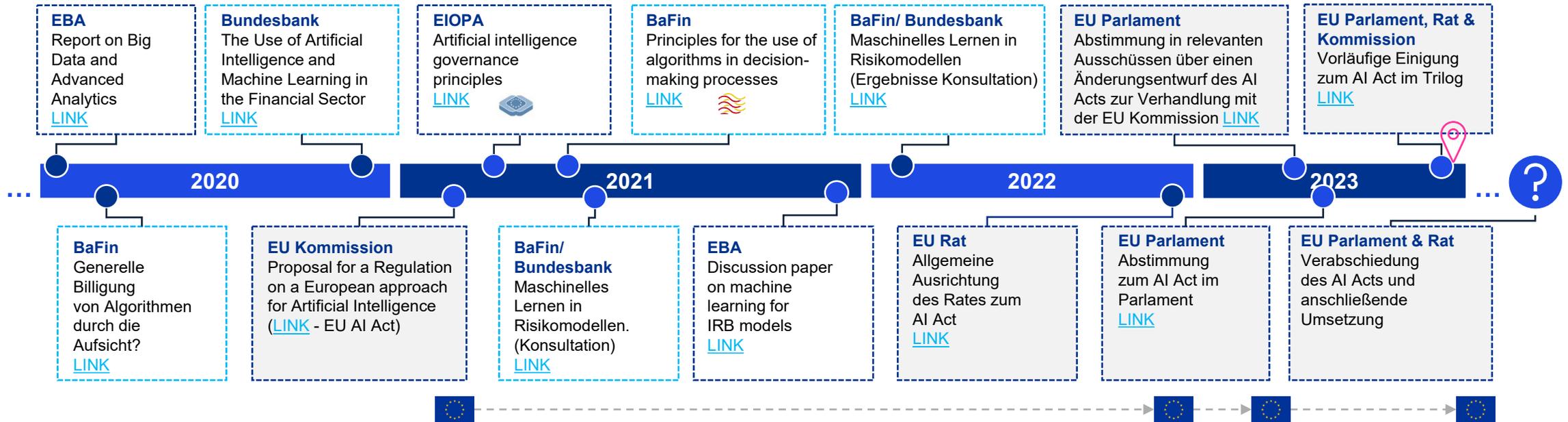
Quelle: Offizieller X-Account (ehem. Twitter) des Europäischen Parlaments, 11.05.23

01

Regulatorische Entwicklungen

Der EU AI Act ist da!
Wie betrifft das uns Aktuarinnen und Aktuare?

KI/ML & Regulatorik – Der Weg zur Gesetzgebung ist lang, die EU ist nach wie vor führend



Einordnung der Veröffentlichungen	Regulation/legislation			
	Europäische Union	European Commission (EU AI Act)	Principles	Diskussionspapiere
<ul style="list-style-type: none"> Auswirkungen der Veröffentlichungen unterschiedlich Regulatorische Vorgaben nur durch EU AI Act 	<ul style="list-style-type: none"> EIOPA 	<ul style="list-style-type: none"> BaFin 	<ul style="list-style-type: none"> EBA EIOPA BaFin Bundesbank 	

Regulierung in der EU: Der AI Act

Basierend auf aktuellem Stand des EU AI Acts

Übersicht: EU AI Act



- Der AI Act soll die **Sicherheit und Vertrauenswürdigkeit („Trustworthy AI“)** von **KI-Systemen** auf dem EU-Markt gewährleisten. Dazu werden **harmonisierte Standards** für deren Entwicklung, Inverkehrbringung und Einsatz entwickelt.
- Die Verordnung verfolgt einen **risikobasierten Ansatz** und legt Standards für verschiedene Risikostufen fest. Je nach Risikostufe gilt es entsprechende Maßnahmen umzusetzen. Der **Fokus liegt auf den „Hochrisikobereichen für KI“**. Im jüngsten Entwurf wurden **Foundation Models** noch einmal gesondert hervorgehoben und zusätzliche Regulatorik dafür ergänzt.
- Es wird Strafen für Verstöße geben, ähnlich zur DSGVO.
- Am AI Act wird seit 2021 gearbeitet. Am 8.12.2023 wurden die Verhandlungen erfolgreich abgeschlossen. Nun wird der AI Act ausformuliert und verabschiedet. Anschließende **Übergangsfrist von zwei Jahren**.

EU AI Act: Definition von Künstlicher Intelligenz (KI)

KI = "Ein maschinengestütztes System, das so konzipiert ist, dass es mit unterschiedlichen Autonomiegraden arbeitet und das für explizite oder implizite Ziele Ergebnisse wie Vorhersagen, Empfehlungen oder Entscheidungen generieren kann, die physische oder virtuelle Umgebungen beeinflussen".

- a) Konzepte des maschinellen Lernens, mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen unter Verwendung einer breiten Palette von Methoden, einschließlich des tiefen Lernens (Deep Learning)
- b) Logik- und wissensgestützte Konzepte, einschließlich Wissensrepräsentation, induktiver (logischer) Programmierung, Wissensgrundlagen, Inferenz- und Deduktionsmaschinen, (symbolischer) Schlussfolgerungs- und Expertensysteme
- c) Statistische Ansätze, Bayessche Schätz-, Such- und Optimierungsmethoden

Risikoklassen im EU AI Act



Risikoklassifizierung und Maßnahmen

Durchführung der Risikoklassifizierung und Maßnahmen



Wichtige Anforderungen des EU AI Acts



Risikomanagementsystem 

Qualitätsmanagementsystem 

Daten und Daten-Governance 

Genauigkeit, Robustheit und Cybersicherheit 

Fairness 

Technische Dokumentation 

Aufbewahrung der Protokolle 

 Offenlegung ggü. Nutzern

 Nutzer-Transparenz/Erklärbarkeit

 Human Oversight

 Konformitätsbewertung

 Folgenabschätzung für Grundrechte

 Bevollmächtigter Vertreter

 Meldung eines Verstoßes ...

Letzte Entwicklung: Grundsätzliche Einigung auf EU AI Act (1)

Basierend auf aktuellem Stand des EU AI Acts



Leitplanken für allgemeine KI-Systeme

- **Transparenzanforderungen für allgemeine KI-Systeme (GPAI) und GPAI-Modelle:** Dazu gehören die Erstellung technischer Dokumentationen, die Einhaltung des EU-Urheberrechts und die Bereitstellung von Zusammenfassungen der für das Training verwendeten Inhalte
- **Strengere Verpflichtungen für High-Impact GPAI-Modelle mit systemischem Risiko:** Dazu gehören Modellbewertungen, Bewertung und Minderung systemischer Risiken, kontradiktorische Tests („Adversarial testing“), Gewährleistung von Cybersicherheit, Meldung schwerwiegender Vorfälle sowie die Berichterstattung über Energieeffizienz



Pflichten für Hochrisikosysteme

- Klare Verpflichtungen für (aufgrund ihres erheblichen Schadenspotenzials für Gesundheit, Sicherheit, Grundrechte, Umwelt, Demokratie und Rechtsstaatlichkeit) als Hochrisiko eingestufte KI-Systeme: neu u.a. **verpflichtende Abschätzung der Folgen für Grundrechte**, die auch für den Versicherungs- und Bankensektor gilt
- Auch KI-Systeme, die zur **Beeinflussung von Wahlergebnissen und des Wählerverhaltens** eingesetzt werden, werden als hochriskant eingestuft
- Bürger haben das Recht, Beschwerden über KI-Systeme einzureichen und Erklärungen zu Entscheidungen zu erhalten, die auf KI-Systemen mit hohem Risiko basieren und sich auf ihre Rechte auswirken

Zusätzliche Auswirkungen auf die Finanzbranche vorhanden

- Breite Definition für KI / ML bleibt, umfasst grundsätzlich weiterhin bestehende statistische Modelle
- Preisbildung und Risikoprüfung bleiben High Risk
- Größte Änderung: Regulierung von generativer KI

Vorbereitung auf bestehende Fassung des AI Acts ist sinnvoll

Letzte Entwicklung: Grundsätzliche Einigung auf EU AI Act (2)

Basierend auf aktuellem Stand des EU AI Acts



Verbotene Anwendungen

- **Biométrische Kategorisierungssysteme**, die sensible Merkmale verwenden (z.B. politische, religiöse, philosophische Überzeugungen, sexuelle Orientierung, Rasse)
- **Ungezieltes Auslesen von Gesichtsbildern** aus dem Internet oder CCTV-Aufnahmen zur Erstellung von Gesichtserkennungsdatenbanken
- **Emotionserkennung** am Arbeitsplatz und in Bildungseinrichtungen
- **Social Scoring** basierend auf sozialem Verhalten oder persönlichen Merkmalen
- **KI-Systeme, die menschliches Verhalten manipulieren**, um deren freien Willen zu umgehen
- **KI-Systeme, die Schwächen von Menschen ausnutzen** (Alter, Behinderung, soziale oder wirtschaftliche Situation)



Ausnahmen für die Strafverfolgung

- **gezielte Suche nach Opfern** (Entführung, Menschenhandel, sexuelle Ausbeutung)
- **Abwehr einer konkreten und gegenwärtigen terroristischen Bedrohung**, oder
- **Lokalisierung oder Identifizierung einer Person**, die verdächtigt wird, eine der in der Verordnung genannten spezifischen Straftaten begangen zu haben (z.B. Terrorismus, Menschenhandel, sexuelle Ausbeutung, Mord, Entführung, Vergewaltigung, bewaffneter Raub, Beteiligung an einer kriminellen Vereinigung, Umweltkriminalität)

Zusätzliche Auswirkungen auf die Finanzbranche vorhanden

- Breite Definition für KI / ML bleibt, umfasst grundsätzlich weiterhin bestehende statistische Modelle
- Preisbildung und Risikoprüfung bleiben High Risk
- Größte Änderung: Regulierung von generativer KI

Vorbereitung auf bestehende Fassung des AI Acts ist sinnvoll

Letzte Entwicklung: Grundsätzliche Einigung auf EU AI Act (3)

Basierend auf aktuellem Stand des EU AI Acts



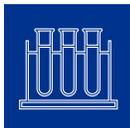
Maßnahmen zur Innovationsförderung und zur Förderung kleiner und mittelständischer Unternehmen (KMU)

- Sicherstellung, dass Unternehmen, insbesondere KMU, KI-Lösungen entwickeln können, ohne übermäßigen Druck von Branchenriesen, die die Wertschöpfungskette kontrollieren
- Zu diesem Zweck fördert die Vereinbarung sogenannte **Regulatory Sandboxes und Real-World-Tests**, die von nationalen Behörden eingerichtet werden, um innovative KI vor der Markteinführung zu entwickeln und zu trainieren



Sanktionen

- Die Nichteinhaltung der Regeln kann je nach Verstoß und Größe des Unternehmens zu **Bußgeldern** zwischen 35 Millionen Euro oder 7 % des weltweiten Umsatzes und 7,5 Millionen Euro oder 1,5 % des Umsatzes führen



Nächste Schritte und Inkrafttreten

- Der vereinbarte Text muss nun sowohl vom Parlament als auch vom Rat offiziell angenommen werden, um EU-Recht zu werden. Die Parlamentsausschüsse für Binnenmarkt und Bürgerrechte werden in einer der nächsten Sitzungen über das Abkommen abstimmen.

Zusätzliche Auswirkungen auf die Finanzbranche vorhanden

- Breite Definition für KI / ML bleibt, umfasst grundsätzlich weiterhin bestehende statistische Modelle
- Preisbildung und Risikoprüfung bleiben High Risk
- Größte Änderung: Regulierung von generativer KI

Vorbereitung auf bestehende Fassung des AI Acts ist sinnvoll

Stellungnahmen des GDV

Stellungnahme vom Juni 2023 zur damaligen Entwurfsfassung

Präzise Definition der Begrifflichkeiten:

- **Definition** sollte **präzise**, mit Fokus auf wesentliche Merkmale und im **Einklang mit international** anerkannten **Definitionen** sein
- **Abgrenzung zu klassischen Verfahren**, wie einfache Automatisierung, ist wünschenswert und wird im Vorschlag des EU-Rats gut abgebildet
- Klare **Abgrenzung zwischen Anbieter und Nutzer** nötig für die Definition von Rechten und Pflichten
- **Querverweise auf bestehende Governance-Regelungen** im Versicherungsbereich (z.B. Solvency II) sollten aufgenommen werden

Klassifizierung & Konformität von KI-Anwendungen:

- Grundsätzliche **Ablehnung der Aufnahme spezifischer Anwendungen** in Annex III und wenn unausweichlich Bevorzugung des Vorschlags des EP und damit Beschränkung auf Risikobewertung
- Begrüßt wird der Vorschlag (EP und Rat), dass Systeme aus Annex III nicht als hochriskant gelten, wenn eine **Folgenabschätzung** dies nahelegt
- Zustimmung zum Ansatz, das Anbieter **Konformitätsverfahren selbst ausführen**
- **Ablehnung zusätzlicher Nutzer-Folgenabschätzungen** (EP-Vorschlag) wegen Mehraufwands und bereits hoher HRS-Anforderungen

Verwendung von Daten

- Es ist sicherzustellen, dass die beim Abschluss von Versicherungsverträgen erforderliche **Risikobewertung** nicht unter das Verbot von **Social Scoring** fällt (EP-Vorschlag hier unzureichend)
- **Nutzung personenbezogener Daten** bei Hochrisikosystemen nur für Korrektur & Erkennung von Verzerrungen in Daten wie im EP-Vorschlag nicht ausreichend – **Erweiterung** um Datennutzung zu Beobachtungszwecken **notwendig**
- **Redundanz** im Kontext der **DSGVO** beachten

Aufsicht und Kontrolle:

- **BaFin als Überwachungsbehörde wird begrüßt**, solange sichergestellt ist, dass die **Zuständigkeit** geklärt ist und es keine zusätzliche KI-Behörde gibt bei Systemen ohne direkten Zusammenhang mit Finanzdienstleistungen
- **Kritik an Struktur** und Ausweitung der Befugnisse des **KI-Boards entsprechend Rats-Vorschlag** aufgrund fehlender gerichtlicher Kontrolle, quasi-verpflichtender Wirkung und möglicher praxisferner Entscheidungen
- **Vorschlag der Kommission** zur Struktur des **KI-Boards mit Kommission als Vorsitzende bevorzugt** gegenüber dem Vorschlag des Rats aufgrund der Rolle der Kommission innerhalb der EU

... und zur Einigung im Trilog

- GDV sieht die Festlegung von einheitlichen Regeln für alle grundsätzlich positiv
- Verordnung als „Maßstab für die Entwicklung von KI unter Beachtung ethischer Standards & europäischer Werte“
- Kritik daran, dass die Preisbildung und Risikoprüfung bei Lebens- und Krankenversicherungen weiterhin als Hochrisiko-KI-Systeme eingestuft werden
- Zusätzliche Verpflichtungen, obwohl es bereits strenge Vorgaben gibt, insbesondere bestimmte Transparenzpflichten, zum Beispiel zum Training des Modells und hohes Maß an Cybersicherheit
- Verbot von Social Scoring könnte ein Thema für Versicherer im Kontext „Pay as you live“ werden
- Es wird begrüßt, dass sich die Definition von KI an der OECD orientieren soll

02

Praktische Herausforderungen

Der EU AI Act ist da!
Wie betrifft das uns Aktuarinnen und Aktuare?

Exemplarische Herausforderungen im aktuariellen Umfeld

KI-Richtlinie / KI-Governance

Etablierung klarer, anpassungsfähiger KI-Richtlinien und Governance-Strukturen, die Datenschutz, Transparenz, kontinuierliche Modellaktualisierungen und verantwortungsbewusstes Risikomanagement gewährleisten.

OpRisk/ ORSA

Risiken im Zusammenhang mit KI-Systemen im Aktuariat zu identifizieren und zu managen: Bewertung der Risiken, Entwicklung der Risikomanagementstrategien und Umsetzung von Maßnahmen zur Risikominderung.

Explainable AI / Trusted AI

Die von KI generierten Ergebnisse für Menschen verständlich zu machen und dabei nach zuverlässigen, sicheren und ethisch vertretbaren KI-Systemen zu streben, um das Vertrauen der Nutzer und Beteiligten zu gewährleisten.

KI Compliance / AI ACT Readiness

Es muss sichergestellt werden, dass im aktuellen Umfeld eingesetzte KI-Systeme mit rechtlichen, aktuariellen und ethischen Standards konform sind.

Dokumentation

Neben der Beschreibung von technischen und fachlichen Details der Komponenten des KI-Systems muss die Dokumentation auch der Versionskontrolle dienen, so dass man in der Lage ist, auftretende Ereignisse zurückzuverfolgen und deren mögliche Ursache zu finden.

03

Erklärbarkeit von KI-Modellen

**Der EU AI Act ist da!
Wie betrifft das uns Aktuarinnen und Aktuare?**

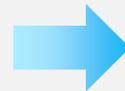
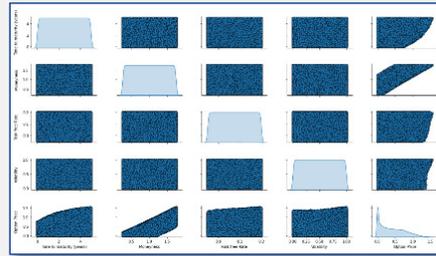
Was kann XAI adressieren?



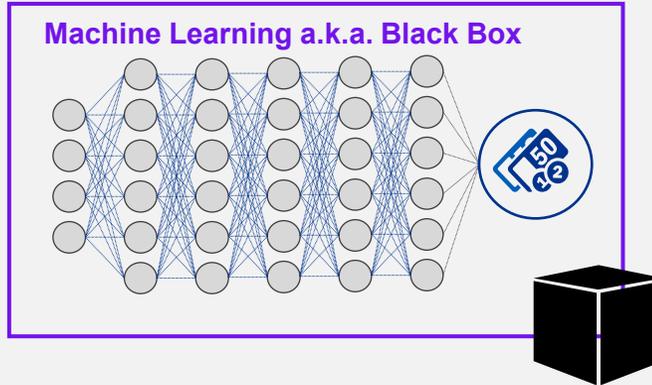
Black-Box-Dilemma

Ergebnisse von ML-Algorithmen sind häufig nur schwer zu interpretieren und zu erklären!

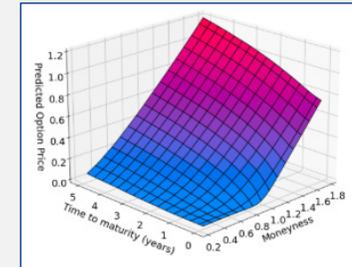
Trainingsdaten



Machine Learning a.k.a. Black Box



Ergebnisse / Schätzer



Darstellung / Quelle: KPMG

? Potenzielle Fragen, die bei der Nutzung beantwortet werden sollten:

Welches Feature / Merkmal hat den größten Einfluss auf das Endergebnis?

Wie ändert sich das Ergebnis bei Änderung eines Features / Merkmals?

Verwendet das Modell ungewollte, verbotene Praktiken (z.B. einzelne Merkmale ausgeschlossen sind – Diskriminierung) ?

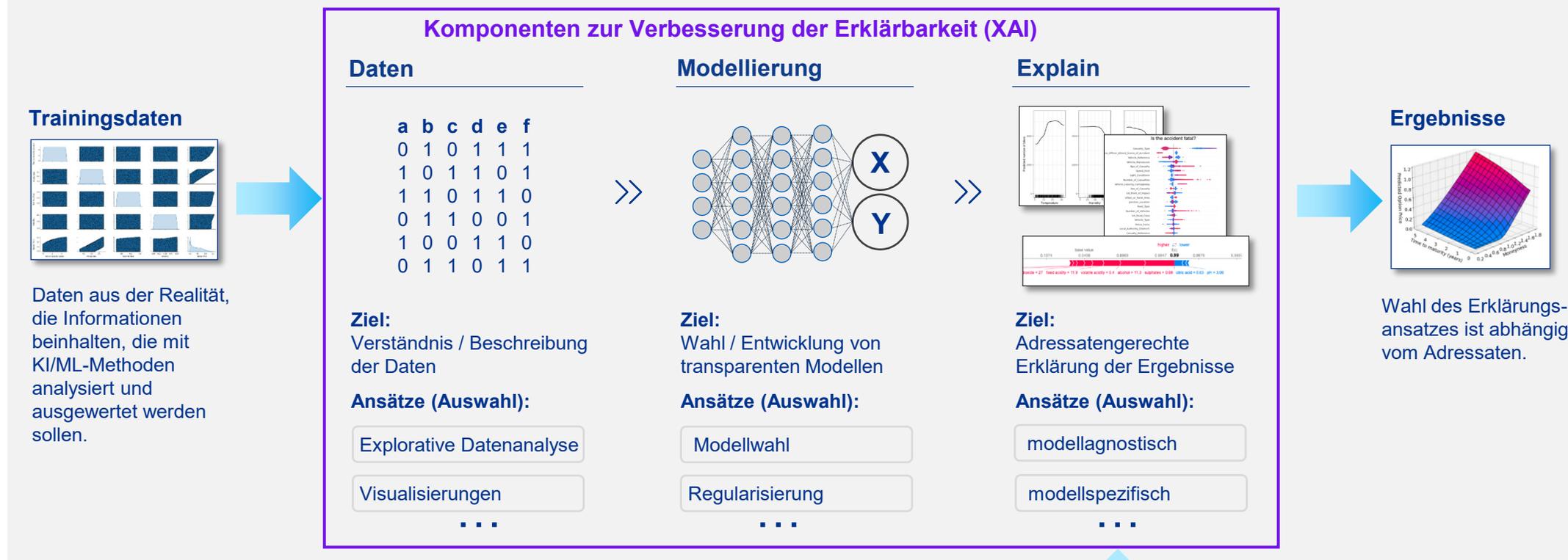
Welche Einflussfaktoren können vernachlässigt werden?

! Die Blackbox muss geöffnet werden, um die Ergebnisse zu erklären!

Das Verständnis für die Funktionsweise eines Machine Learning Ansatzes ist unerlässlich in der Anwendung.

Verbesserung der Erklärbarkeit

Welche Stellschrauben zur Erklärbarkeit der Black-Box gibt es?

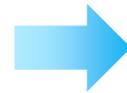
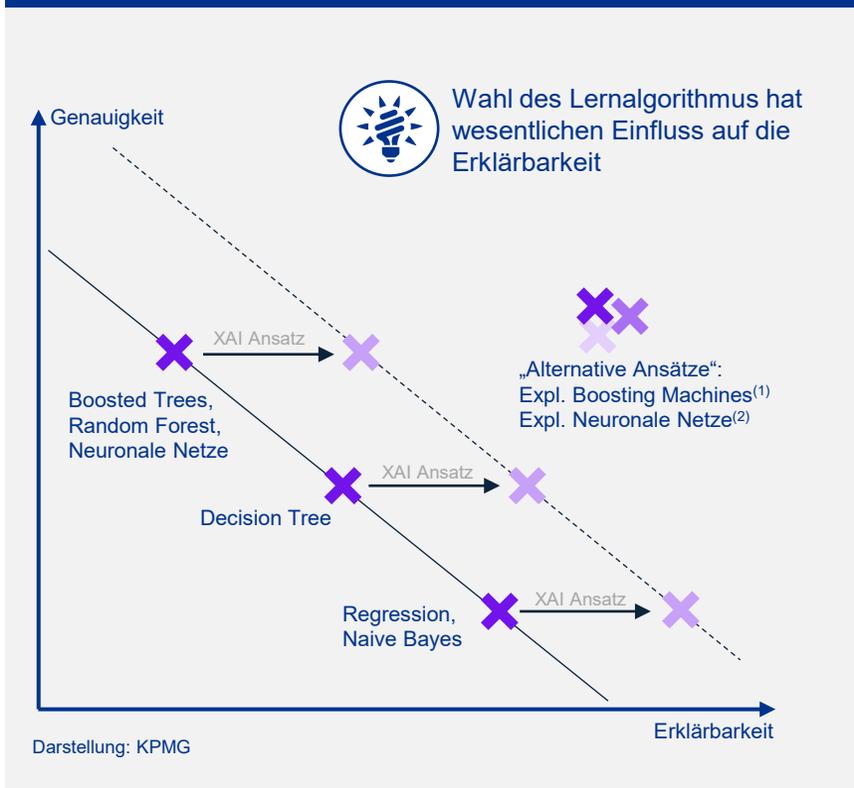


Fokus: Explain-Ansätze

Aber Achtung: Die Themen Modellierung und Explain hängen eng miteinander zusammen und können nicht losgelöst voneinander betrachtet werden.

Trade-off zwischen Genauigkeit und Erklärbarkeit

Trade-off für komplexe ML-Modelle



Maximierung: **Erklärbarkeit** + **Genauigkeit**

- 

Wahl des Modell

Methoden unterscheiden sich in Erklärbarkeit und Genauigkeit
Abzuwägen: „Schlechteres“ Model für verbesserte Erklärbarkeit wählen?
- 

Alternative Ansätze

Nutzung von Ansätzen, die Genauigkeit und Erklärbarkeit verbinden
Ansätze noch nicht weit verbreitet (Intrinsically interpretable models) und nicht immer geeignet
- 

Post-Hoc Analyse

Verbesserung der Erklärbarkeit bestehender Modelle durch Nutzung von XAI-Ansätzen
XAI-Ansätze: Unterscheidung zwischen modellagnostisch und modellspezifisch

Fokus: Post-Hoc Analyse
Die Anwendung von alternativen Ansätzen sind nicht Standard. Aus diesem Grund steht die Verwendung verbreiteter Modelle und deren Erklärbarkeit im Vordergrund.

⁽¹⁾ InterpretML: A Unified Framework for Machine Learning Interpretability (2019) // ⁽²⁾ Neural Additive Models: Interpretable Machine Learning with Neural Nets (2020)

Wrap-up

Regulatorische Entwicklungen

- Der EU AI Act zielt darauf ab, die **Sicherheit und Vertrauenswürdigkeit von KI-Systemen** auf dem EU-Markt sicherzustellen
- Es werden **harmonisierte Standards** für die Entwicklung, Inverkehrbringung und Einsatz von KI-Systemen entwickelt
- **Risikobasierter Ansatz** mit Standards für verschiedene Risikostufen; Fokus auf Hochrisikooanwendungen und besondere Betonung von Foundation Models
- Arbeit am EU AI Act begann 2021. Am 8.12.2023 wurden die Verhandlungen im Trilog erfolgreich abgeschlossen. Nun wird der EU AI Act ausformuliert und verabschiedet. Danach folgt eine **Übergangsfrist von zwei Jahren**
- **Strafen für Verstöße** ähnlich zur DSGVO

Praktische Herausforderungen

- **Breite Definition von KI / ML**, einschließlich statistischer Ansätze
- **Preisbildung & Risikoprüfung** sind voraussichtlich Hochrisikooanwendungen
- Identifikation und Management von **Risiken** im Zusammenhang mit KI-Systemen durch das Risikomanagement, inklusive Risikobewertung/-minderung
- Sicherstellung, dass eingesetzte KI-Systeme **rechtlichen, aktuariellen und ethischen Standards** entsprechen
- Streben nach verständlichen, sicheren und **ethisch vertretbaren KI-Systemen**, um Nutzer- und Stakeholder-Vertrauen zu gewährleisten
- **Dokumentation der technischen Details** von KI-Systemen inklusive Versionskontrolle

Erklärbarkeit von KI-Modellen

- **Transparenz** durch Erklärbarkeit der Funktionsweise von Algorithmen für alle Stakeholder, um Ergebnisse besser zu verstehen
- **Vermeidung von Bias und Diskriminierung** durch Analyse der Einflussfaktoren auf die Ergebnisse (**Fairness**)
- **Modellverständnis** zur Vermeidung unerwarteter Ergebnisse (**Robustheit**) und Validierung als Teil des Prozesses
- Anpassung **risikoreicher Modelle und Optimierung**, z.B. durch Ausschluss bestimmter Merkmale
- **Dokumentation** aller Analysen für Aufsichtsbehörden
- **Bewältigung des Black-Box-Dilemmas** und Abwägung zwischen Genauigkeit und Erklärbarkeit

Ansprechpartner



Dr. Fabian Bohnert
Director
FS Insurance

M +49 170 7016615
fbohnert@kpmg.com



Dr. Mischa Pupashenko
Senior Manager
FS Insurance

M +49 160 95242166
mpupashenko@kpmg.com



kpmg.de/socialmedia

kpmg.de

Die enthaltenen Informationen sind allgemeiner Natur und nicht auf die spezielle Situation einer Einzelperson oder einer juristischen Person ausgerichtet. Obwohl wir uns bemühen, zuverlässige und aktuelle Informationen zu liefern, können wir nicht garantieren, dass diese Informationen so zutreffend sind wie zum Zeitpunkt ihres Eingangs oder dass sie auch in Zukunft so zutreffend sein werden. Niemand sollte aufgrund dieser Informationen handeln ohne geeigneten fachlichen Rat und ohne gründliche Analyse der betreffenden Situation.

© 2023 KPMG AG Wirtschaftsprüfungsgesellschaft, eine Aktiengesellschaft nach deutschem Recht und ein Mitglied der globalen KPMG-Organisation unabhängiger Mitgliedsfirmen, die KPMG International Limited, einer Private English Company Limited by Guarantee, angeschlossen sind. Alle Rechte vorbehalten.

Document Classification: KPMG Public