COMMERZBANK

Ein risikobasierter Ansatz zu Künstlicher Intelligenz

15 Februar 2023

Die Position der Commerzbank zur europäischen KI-Verordnung (AI Act)

COMMERZBANK

Key Messages

- Ad-hoc-Analysen mit ML-Techniken liegen außerhalb des Geltungsbereichs des AI Act.
- Eine Symbiose zwischen Mensch und Maschine bei Labelling, Feedbackschleifen und Entscheidungsfindung erhöht die Effizienz und Effektivität eines Prozesses und mitigiert gleichzeitig Risiken.
- Die meisten Risiken im Zusammenhang mit dem Einsatz von KI-Modellen und -Systemen bei Finanzdienstleistungen sind nicht neu und bereits durch den erfolgreichen Umgang mit traditionellen Modellen in der Vergangenheit bekannt. Banken können hier auf bestehende Strukturen aufbauen.
- Nur wenn es sich konkret um eine „Ja oder Nein"- Kreditentscheidung handelt, sollte das System als "High-Risk" gemäß Anhang III, 5 (b) angesehen werden.
- Zertifikate bringen möglicherweise nicht den beabsichtigten Nutzen, können aber schnell kostspielig werden und redundante Belastungen für Banken verursachen. Bereits bestehende Aufsicht über Credit Scoring Modelle durch die zuständigen Behörden im Bankensektor muss als äquivalent zu Zertifizierungen betrachtet werden.
- Es muss festgelegt werden, wie Banken Zertifikate nutzen können. Dies ist insbesondere für die effiziente Nutzung von „General Purpose AI" wesentlich.
- ML-Modelle sind gegenüber Verbrauchern erklärungsbedürftig. Form und Umfang dieser Erklärungen sind allerdings nicht ohne Angabe von Adressaten und Kontext zu ermitteln. Verbraucher benötigen die richtige Menge an verständlichen Informationen, die es ihnen ermöglicht, getroffene Entscheidungen zu prüfen.
- Valide Bedenken sprechen gegen die Offenlegung detaillierter Informationen über KI-Systeme. Insbesondere im Zusammenhang mit der Betrugsbekämpfung könnte die Offenlegung von Details über die Methoden zur Identifizierung betrügerischer Aktivitäten dazu beitragen, diese zu umgehen.
- Aufgrund der Verflechtungen von Künstlicher Intelligenz mit verschiedenen anderen Regelungen wie der DSGVO fordern wir eine kohärente Harmonisierung der Vorgaben.

A. Einführung

Künstliche Intelligenz (KI) und Maschinelles Lernen (ML) sind für die Zukunft der Bankenbranche von entscheidender Bedeutung, um den vielfältigen Herausforderungen des digitalen Zeitalters und den daraus entstehenden zusätzlichen (Cyber-) Risiken zu begegnen. Die Implementierung von KI und ML in der stark regulierten Bankenbranche ist komplex. Deswegen braucht es einen risikobasierten Ansatz, der die Vorteile dieser Technologien mit den verbundenen regulatorischen Hürden abwägt.

In diesem Positionspapier sollen die wichtigsten Aspekte der europäischen KI-Verordnung aus Bankenperspektive zusammengefasst werden. Weitere Details finden Sie im White Paper der Commerzbank, welches den risikobasierten Ansatz der Commerzbank zur ML-Governance detailliert und es sich zum Ziel gesetzt hat, KI und ML zu entmystifizieren und als das zu begreifen, was sie wirklich sind: Kontrollierbar und nicht magisch!

B. Aktuelles regulatorisches Umfeld

Derzeit befindet sich die Europäische Union (EU) mitten im Gesetzgebungsprozess zur Festlegung harmonisierter Regeln für KI, im Folgenden „AI Act"[1] genannt. Der AI Act zielt darauf ab, die KI-Komponenten von IT-Systemen unabhängig vom Wirtschaftszweig, in dem sie eingesetzt werden, zu regulieren. Die Verordnung ist weltweit wegweisend und wird ein Gütesiegel für vertrauenswürdige KI liefern, die in Europa hergestellt und verwendet wird.

C. Definitionen

Unsere Definition von KI steht im Einklang mit der eines KI-Systems gemäß dem AI Act (vgl. Artikel 3 und S. 6 (6,

[1] Alle weiteren Referenzen basieren auf dem „General Approach" des Rates der Europäischen Union vom 6. Dezember 2022 "Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts".

Anfang Dezember 2022 adaptierte der Rat der Europäischen Union seinen „General Approach". Sobald das Europäische Parlament seine eigene Position eingenommen hat, kann der „Trilog" zwischen dem Europäischen Rat, dem Parlament und der Kommission aufgenommen werden. Dies wird Stand heute für das zweite Quartal 2023 erwartet.

6a-b)). Basierend auf dieser Grundlage leiten wir ab, dass Ad-hoc-Analysen mit ML-Techniken außerhalb des Geltungsbereichs des AI Acts liegen.

D. Risikobasierter Ansatz

KI-Innovationen führen zu Verschiebungen bei Kosten- und Nutzen-Abwägungen und erfolgreiche KI-Implementierungen müssen immer die Kosten der Risikominderung berücksichtigen. Die Wahl des Models ist dabei situationsbedingt:

- Eine hohe Qualität und ein großes Spektrum von Daten sind wesentliche Voraussetzungen und verbessern sowohl traditionelle als auch ML-Modelle.
- In einigen Fällen funktionieren einfachere datengetriebene Ansätze oder „klassische“ statistische Methoden ähnlich gut wie fortgeschrittene KI-Modelle, jedoch ohne einige der damit verbundenen Risiken. Nichtsdestotrotz erfordert die Beschreibung nicht-linearer, komplexer Zusammenhänge üblicherweise ML Methoden.
- Durch die Symbiose von Mensch und Maschine bei Labelling, Feedbackschleifen und Entscheidungsfindung können Effizienz und Effektivität eines Prozesses erhöht und gleichzeitig Risiken mitigiert werden.

Daher verfolgt die Commerzbank einen risikobasierten Ansatz. Systeme, die KI enthalten, werden in KI-Risikoklassen eingeteilt, abhängig von einer Vielzahl von Faktoren wie Modellkomplexität, Wirkung und geschäftlicher Kritikalität.

Kreditwürdigkeitsprüfung

Die folgende Bankenfunktion ist gemäß Absatz 5 (b) des Anhangs III des Vorschlags des Europäischen Rates ein Hochrisiko-KI-System:

“KI-Systeme, die bestimmungsgemäß für die Kreditwürdigkeitsprüfung oder Kreditpunktebewertung natürlicher Personen verwendet werden sollen (…)”

Obwohl die Absicht klar ist, nämlich Bürger zu schützen, lässt die Formulierung insbesondere im englischsprachigen Text („evaluate the creditworthiness“) Raum für Interpretationen.

Es ist nicht die Absicht dieses Positionspapiers, die Standards zu senken, sondern für eine realistische, effektive und angemessene Umsetzung zu werben. Es ist eines der Ziele des AI Acts, die Fähigkeit der Kunden, einen Kredit zu erhalten, zu wahren. Wir halten es daher für selbstverständlich, dass es hier ausschließlich um Situationen geht, in denen es sich um eine „Ja oder Nein“-Kreditentscheidung handelt. Daher kommen wir zu dem Schluss, dass Lifecycle-Anwendungen[2] genau das sind, was mit der Klausel zu „völlig unwesentlichen“ KI-Komponenten erreicht werden soll (Artikel 6).

Machine Learning Governance

Das Risikomanagement der Commerzbank folgt dem Prinzip der “Three Lines of Defence”. Die meisten Risiken im Zusammenhang mit dem Einsatz von KI-Modellen bei Finanzdienstleistungen sind nicht neu und sind aus der erfolgreichen Nutzung traditioneller Modelle in der Vergangenheit bekannt. Banken haben vorhandenes Wissen und können auf bestehenden Strukturen aufbauen.

E. Zertifikate

Zertifikate und CE-Kennzeichnungen sollen das Vertrauen in die angebotenen Lösungen stärken und damit Innovationen und Investitionen insbesondere für kleine und mittelständische Unternehmen fördern. Zwar ist die Absicht klar, Innovation zu unterstützen, doch das Konzept hat Grenzen:

[2] Während der gesamten Dauer einer Kundenbeziehung kann es dazu kommen, dass möglicherweise Informationen verarbeitet werden, die mit der Bonität einer Person zusammenhängen (sogenannte Lifecycle-Anwendungen z. B. Tools zur Informationsbereitstellung oder Klassifizierung).

COMMERZBANK

- Was *genau* wird zertifiziert? Es kann immer nur ein Point-in-Time-Snapshot oder vergangenheitsorientiert sein. Es wird möglicherweise nicht über einen bestimmten Zeithorizont hinaus gelten.
- Umstände wie Datenänderungen könnten die Aussagekraft und Aktualität einer Zertifizierung grundlegend verändern. Daher wäre es sinnvoll, den gesamten Lebenszyklus des Modells zu zertifizieren. Doch ein komplexes Zertifikat kann von Menschen nicht leicht verstanden werden und nur ein verständliches Zertifikat ist wirksam und kann Vertrauen wie beabsichtigt herstellen.
- Die Zertifizierung durch Dritte belegt nur, wie gut das Modell in der entworfenen Umgebung funktioniert, während es in anderen Umgebungen weniger effizient oder unbeabsichtigt funktionieren könnte. Es könnte somit als „Carte Blanche“ gesehen werden, um das Modell überall zu verwenden.
- Es ist fraglich, inwieweit Banken sich auf ein Zertifikat stützen können oder diese Art von Bewertungen selbst durchführen müssen, da die Verantwortung hier niemals ausgelagert werden kann. Es muss festgelegt werden, wie Banken Zertifikate nutzen können. Dies ist insbesondere für die effiziente Nutzung von „General Purpose AI“ wesentlich.

Zertifikate bringen möglicherweise nicht den beabsichtigten Nutzen, können aber schnell kostspielig werden und redundante Belastungen für Banken verursachen. Die Modelle zur Bonitätsbewertung werden bereits von den zuständigen Behörden geprüft. Dies sollte bereits als ein Zertifikat von hoher Qualität angesehen werden. Eine harmonisierte Standardisierung muss daher aufzeigen, dass diese laufende Überwachung den Prüfungserfordernissen vollumfänglich entspricht.

F. Vertrauenswürdige und verantwortungsvolle KI

Hier wollen wir uns auf die drei Aspekte mit der lebhaftesten öffentlichen Debatte und dem größten Interpretationsspielraum konzentrieren: Transparenz, Erklärbarkeit und Fairness.

Transparenz

Transparenz bedeutet, klar, offen und ehrlich zu kommunizieren, wie und warum die Daten einer Person verwendet werden (vgl. Artikel 13). Die Informationspflichten zur Datenerfassung und -nutzung sowie zur automatisierten Verarbeitung sind in den Artikeln 13, 14 und 22 DSGVO[3] festgelegt.

Erklärbarkeit / Interpretierbarkeit

Der richtige Umfang und die richtige Form der Erklärung können nicht ohne Angabe von Adressaten und Kontext bestimmt werden. Die Verbraucher müssen die entsprechende Menge an verständlichen Informationen erhalten, damit sie getroffene Entscheidungen überprüfen können. Zum anderen muss das Material beispielsweise für externe Prüfungen (Audits) vollständig sein und über festgelegte Zeiträume aufbewahrt werden.

Andere Aspekte sprechen allerdings gegen die Offenlegung detaillierter Informationen zu KI-Systemen. Insbesondere im Zusammenhang mit der Betrugsbekämpfung könnte die Offenlegung von Details über die Instrumente zur Ermittlung betrügerischer Aktivitäten dazu beitragen, diese zu umgehen. Das ist nicht wünschenswert. Darüber hinaus wurde betriebswirtschaftliches Know-how oft intern entwickelt und mit großen Investitionen in Zeit und Geld ausgestattet. Die Veröffentlichung dieser Informationen kann daher zu einem erheblichen Verlust an

[3] „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“ (Artikel 22 §1 DSGVO).

geistigem Eigentum, vertraulichen Geschäftslogiken und zur Entstehung von Nachteilen gegenüber Wettbewerbern führen.

Fairness

Diskriminierung kann vielfältig sein und unabhängig von den Absichten der Designer kann sie an verschiedenen Stellen in der Modellierung in den Algorithmus kommen. Fairness muss in allen Phasen des Modelllebenszyklus sorgfältig überwacht werden, wobei eine entsprechende interne Governance zu gewährleisten ist.

Wenn ein potenziell unterscheidendes Merkmal nicht aufgezeichnet oder später aus dem Datensatz gelöscht wird, ist es mathematisch immer noch möglich, dass das Modell dieses Feature weiterhin unterscheidet. Dies ist der Fall, wenn das unterscheidende Merkmal mit dem Ergebnis korreliert[4].

Die Erstellung von AI- und ML-Modellen ist ein iterativer Prozess. Natürlich sind Transparenz und Fairness wichtige Voraussetzung. Aber viele entsprechende Validierungsschritte können nur während oder nach dem Modellierungsprozess abgeschlossen werden. Von Anfang an eine endgültige Fairness-Validierung zu fordern, ohne Innovation damit zu blockieren, ist daher weder machbar noch sinnvoll. Robuste und vertrauenswürdige KI-Systeme können dazu beitragen, unbewusste menschliche Vorurteile zu überwinden, denn KI liefert wiederholbare und nachverfolgbare Ergebnisse.

G. Beziehung zu anderen Gesetzen und Gerichtsbarkeiten

Aus regulatorischer Sicht müssen die Entwicklung, Training, Evaluation und der Einsatz von KI-Systemen neben dem AI Act verschiedene weitere Anforderungen[5] erfüllen.

Andere Jurisdiktionen sind dabei, KI-bezogene Regelungen mit unterschiedlichen Blickwinkeln und Prioritäten umzusetzen. Gleichzeitig argumentieren andere, dass die meisten KI-Themen bereits ausreichend durch bestehende Vorschriften abgedeckt sind und möglicherweise nur entsprechend detailliert sein müssen. Diese Entwicklungen bergen das Risiko einer Fragmentierung der Märkte und stellen multinationale Organisationen vor große Herausforderungen.

H. Schlussfolgerung

Wir schätzen die allgemeine Ausrichtung des AI Acts – einen Ansatz für ein Gütesiegel von KI, die in Europa hergestellt oder verwendet wird. Die Definitionen sind auf dem neuesten Stand der Technik und ein risikobasierter Ansatz ist geeignet. Da das Gesetz für alle Branchen gelten soll, werden einige Bankspezifika jedoch nicht hinreichend berücksichtigt. Dies ist besonders relevant bei der Betrachtung der feinen Grenze zwischen Lifecycle-Kundenmanagement und Bonitätsbewertung sowie der Nachteile, die zusätzliche Zertifizierungen in einem bereits extern auditierten Fachgebiet wie der Bonitätsbewertung haben. Hier ist es wichtig, die bereits laufende Aufsicht durch die zuständigen Behörden als gleichwertig mit Zertifizierungen zu betrachten.

Es ist nicht die Absicht dieses Positionspapiers, die Standards zu senken, sondern auf die bereits bestehende und effektive Regulierung in der Bankenbranche hinzuweisen und für eine realistische, effektive und angemessene Umsetzung der neuen Anforderungen zu werben. Aufgrund der Verflechtungen von Künstlicher Intelligenz mit verschiedenen anderen Regulierungen, wie oben beschrieben, fordern wir eine kohärente Harmonisierung der Vorgaben.

[4] Vgl. Amazon Hiring Tool: IIF Bias and ethical implications in machine learning (S. 10).

[5] Darunter sind: Directive 2013/36/EU, Datenschutzgrundverordnung, European Data Act, European Data Governance Act, BAIT; MaRisk, EBA Guidelines on ICT, DORA, MiFiD and zumindest indirekt durch viele weitere Regulierungen wie Gesetze zu Verbraucherschutz, Gleichbehandlung und Nicht-Diskriminierung etc.

Commerzbank AG
Zentrale
Kaiserplatz
Frankfurt am Main
www.commerzbank.de

Postanschrift
60261 Frankfurt am Main
Tel. + 49 69 136-20
info@commerzbank.com

Big Data & Advanced Analytics

Julia Sterling
Julia.sterling@commerzbank.com

Thomas Stadje
Thomas.stadje@commerzbank.com

Der Geschäftsbereich „Big Data & Advanced Analytics“ der Commerzbank ist das Kompetenzzentrum für alle KI- und ML-bezogenen Themen, von der internen KI-Modellierung und -Implementierung bis hin zu Beratungsaufgaben.