



COMMERZBANK

Grundsätze der Nutzung von Künstlicher Intelligenz (KI) und Maschinellern Lernen (ML)

Prinzipien der Commerzbank AG¹ und inländischer Töchter in Bezug auf ihre Beschäftigten

Präambel

Künstliche Intelligenz (KI) und Maschinelles Lernen (ML) halten zunehmend Einzug in unser tägliches Leben und in die Arbeitswelt in der Commerzbank. KI und ML eröffnen vielfältige Chancen, unsere Kunden noch besser zu bedienen, Bankprozesse schneller, besser und effizienter zu gestalten sowie Verluste durch Risiken, Betrug und Finanzkriminalität zu verringern. Gleichzeitig verändert die Implementierung von KI und ML-Modellen die Arbeitsprozesse unserer Mitarbeitenden. Der Einsatz von KI soll die Mitarbeitenden bei ihrer Arbeit unterstützen und Prozesse effizienter gestalten. Ziel ist es auch, dass die Mitarbeitenden befähigt sind, mit den Veränderungen im Zusammenhang mit KI umgehen zu können.

Aus diesen Gründen ist es wichtig, dass Arbeitgeberseite und Konzernbetriebsrat ein gemeinsames Verständnis zum Einsatz von Künstlicher Intelligenz und Maschinellern Lernen in der Commerzbank etablieren. Gemeinsames Ziel ist es, Sorgen und Ängste in diesem sensiblen Themenumfeld ernst zu nehmen, die Mitarbeitenden und ihr Know-how bei neuen Prozessen und veränderten Arbeitsbedingungen zu integrieren und den Einsatz von KI und ML angemessen transparent zu machen und zu erklären. Auf dieser Basis können die Vorteile dieser Technologien ausgeschöpft und gleichzeitig Fehler und Missbrauch vermieden werden.

Die Begriffe „Künstliche Intelligenz“ und „Maschinelles Lernen“ werden im allgemeinen Sprachgebrauch nicht einheitlich verwendet. Traditionelle Programmierung basiert auf durch den Programmierer vorgegebenen Regeln. In Abgrenzung dazu wird in den hier vereinbarten Grundsätzen der Ausdruck „KI-System“ für die Implementierung von Modellen verwendet, die solche Regeln aus den ihnen vorliegenden Daten mit einem signifikanten Grad an Autonomie und Flexibilität selbst „lernen“.²

Grundsätze

Die Persönlichkeitsrechte unserer Mitarbeitenden werden im Kontext des Einsatzes von Künstlicher Intelligenz und Maschinellern Lernen in unserem Konzern geachtet und geschützt im Einklang mit den bestehenden Gesetzen. Darunter verstehen wir:

- Als Konzern sind wir uns der Tatsache bewusst, dass KI-Systeme die Möglichkeiten und Rechte von anderen betreffen können.
- KI-Systeme müssen die fundamentalen Rechte von Betroffenen, geltende Gesetze und demokratische Werte berücksichtigen.
- KI-Systeme müssen die Daten der Beschäftigten in einer nachvollziehbaren Weise verarbeiten.

¹ Dies betrifft sowohl die Commerzbank AG Inland sowie ausländische Niederlassungen und DTCs.

² Für Details s. Glossar

- Die strategische Ausrichtung des Konzerns im Themenumfeld KI und ML sowie die relevanten Konzepte und Begrifflichkeiten werden den Mitarbeitenden transparent dargelegt und erklärt.
- Bei der Wahl der Modelle und der Implementierung von KI-Systemen werden Chancen und Risiken adressiert und abgewogen.
- Der Einsatz von KI-Systemen sowie deren Qualitätsmaßstäbe werden so ausgerichtet, dass Diskriminierung vermieden wird.
 - Wir stellen sicher, dass KI-Systeme vertrauenswürdig und verantwortungsvoll eingesetzt werden.
 - Dies beinhaltet insbesondere Fairness, Transparenz, Robustheit, Sicherheit, Datenschutzkonformität und menschliche Kontrolle.
- Bei Interaktionen von Menschen mit Chatbots, Talkbots und Deep Fakes wird am Frontend rechtzeitig, klar und verständlich Transparenz darüber geschaffen, dass eine Interaktion mit einem KI-System vorliegt.
- In Prozessen mit Unterstützung von Künstlicher Intelligenz muss den Mitarbeitenden stets klar und verständlich sein, welche Rechte und Pflichten (z.B. Interventionsmöglichkeiten und Prüfpflichten) bestehen.
- Mitbestimmungsrechte des Betriebsrats und existierende Betriebsvereinbarungen gelten auch beim Einsatz von Künstlicher Intelligenz und werden durch die vorliegenden Grundsätze nicht eingeschränkt³. Es gelten insbesondere die Bestimmungen der Rahmen IT-KBV.
- Der Einsatz von KI-Systemen in personal-relevanten Prozessen („HR-Prozessen“) wird außerhalb der vorliegenden Grundsätze gesondert im Rahmen der bestehenden Mitbestimmungsrechte des Betriebsrats betrachtet.

Unsere Mitarbeitenden dürfen durch den Einsatz von KI-Systemen nicht gläsern werden. Das bedeutet:

- Etwaige manipulative, ausbeuterische und soziale Kontrollpraktiken⁴ werden in KI-Systemen nicht eingesetzt.
- KI-Systeme werden nicht eingesetzt, um Emotionen, Persönlichkeitseigenschaften oder die psychische Verfassung unserer Mitarbeitenden zu analysieren oder zu manipulieren. KI-Systeme, die das Wohlbefinden der Beschäftigten verbessern sollen, werden nur verwendet, wenn die Betriebsvereinbarung es zulässt. Es werden keine KI-Systeme eingesetzt, die dem Zweck dienen, Charaktereigenschaften zuzuschreiben.
- Auch für KI-Systeme gilt Artikel 9 Absatz 1 DSGVO: *„Die Verarbeitung personenbezogener Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person ist untersagt“*.⁵ Ausnahmen hiervon gelten gemäß Artikel 9 Absatz 2 DSGVO.

Die Commerzbank und die Arbeitnehmervertretung werden sich weiterhin regelmäßig zum Einsatz von Künstlicher Intelligenz und Maschinellen Lernen austauschen. Im Fokus dieses gemeinsamen Dialogs stehen insbesondere der sich fortwährend entwickelnde Einsatz von KI und ML in der Bank, die Überprüfung und Weiterentwicklung der vorliegenden Grundsätze sowie deren Operationalisierung.

³ Gemäß Datenschutz-Handbuch ([SEC-32002-DE – Gültig ab 01.06.2023 | Version 12.0 Kapitel 21](#)) gilt: „Werden personenbezogene Daten von Beschäftigten verarbeitet, so ist der zuständige Betriebsrat hinzuzuziehen.“ Darüber hinaus gelten alle Vereinbarungen der schriftlich fixierten Ordnung auch beim Einsatz von Künstlicher Intelligenz grundsätzlich weiterhin, insbesondere Grundsätze des Datenschutzes sowie der IT-Sicherheit.

⁴ Beispielfähig fällt „Social Scoring“ unter diese Kategorie. „Social Scoring“ bedeutet Bewertung oder die Klassifizierung natürlicher Personen auf der Grundlage ihres sozialen Verhaltens, sozioökonomisch Status, Bekanntschaften oder vorhergesagte Persönlichkeitsmerkmale.

⁵ Hiervon unberührt sind Verifikationssysteme, die lediglich die biometrischen Daten einer Person mit ihren zuvor bereitgestellten biometrischen Daten vergleichen. Ebenfalls ausgenommen sind Systeme, die der Identifizierung von natürlichen Personen in der Ferne dienen mit dem einzigen Zweck, zu bestätigen, ob eine bestimmte natürliche Person erlaubt ist, Zugang zu einem Dienst, einem Gerät oder einem Standort zu erlangen.

Glossar

| Begriff | Erläuterung |
|-----------------------------|---|
| Künstliche Intelligenz (KI) | <p>Künstliche Intelligenz bezieht sich auf die Erstellung eines Algorithmus, der Daten verwendet, um bestimmte Aspekte der Welt zu modellieren. Anschließend wird das Modell auf neue Daten angewendet, um Ergebnisse in Form von z.B. Inhalten, Vorhersagen, Klassifizierungen oder Empfehlungen zu generieren. Ziel ist dabei, menschliche Aktivitäten zu unterstützen oder letztlich auch bestimmte Entscheidungen zu treffen.</p> |
| KI-System | <p>„KI-System“ bezeichnet ein maschinenbasiertes System, das mit unterschiedlichen Graden an Autonomie Vorhersagen, Empfehlungen oder Entscheidungen generieren kann, die die physische oder virtuelle Umgebung beeinflussen. Im Rahmen dieser Grundsätze verwenden wir den Begriff KI-System für Software, die Maschinelles Lernen, genauer „nicht-parametrische“ Algorithmen mit einem vorhandenen Grad an Autonomie nutzt.</p> |
| Maschinelles Lernen (ML) | <p>Traditionelle Programmierung kodifiziert Regeln (z.B. Wenn-/Dann-Bedingungen). Sie friert die modellierte Realität ein. Demgegenüber zielt ML darauf ab, diese Regeln selbst zu identifizieren (zu „lernen“), die für die traditionelle Programmierung zu komplex sein können. ML-Modelle stellen statistische Input-Output-Beziehungen dar und eignen sich sehr gut zur Beschreibung komplexer und nicht-linearer Beziehungen, was es ihnen ermöglicht, auch auf neue Daten oder Veränderungen der Realität zu reagieren. Oft bietet das Ergebnis Wahrscheinlichkeiten für verschiedene mögliche Outputs, die Akzeptanzschwellen erfordern, um den jeweiligen Output zu nutzen (z.B. für eine „Ja oder Nein“-Entscheidung).</p> <p>Ist die mathematische Funktion, die erlernt werden soll, im Vorhinein bekannt oder festgelegt, spricht man von „parametrischen Algorithmen“. Hier werden nämlich „nur“ die Parameter der vorgegeben Funktion aus den Daten erlernt. Beispiele für „parametrische Algorithmen“ sind Lineare Regressionen oder Logistische Regressionen.</p> <p>Demgegenüber stehen „nicht-parametrische Algorithmen“. Diese machen keine starken Annahmen über die Form der Funktion und können so jede funktionale Form und Verteilung aus den Trainingsdaten lernen. Beispiele für „nicht-parametrische Algorithmen“ sind Neuronale Netze oder Ensemblemethoden (z. B. Gradient Boosting, Random Forest).</p> |
| Fundamentale Rechte | <p>International gilt die Allgemeine Erklärung der Menschenrechte der Vereinten Nationen. Darüber hinaus fallen unter den Begriff der fundamentalen Rechte die Werte gemäß Artikel 2 des Vertrags über die Europäische Union (EUV), auf die sich die Europäische Union gründet: die Achtung der Menschenwürde, Freiheit, Demokratie, Gleichheit, Rechtsstaatlichkeit und Wahrung der Menschenrechte einschließlich der Rechte von Personen, die Minderheiten angehören. Zudem gilt die Charta der Grundrechte der Europäischen Union. Diese Grundrechte werden ergänzt um nationale Regelungen wie beispielsweise das Grundgesetz in Deutschland. Da es sich um grundlegende Rechte von Individuen handelt, sind diese im besonderen Maße zu schützen.</p> |

| | |
|------------------------|---|
| Fairness | KI-Systeme sollen nicht zu ungerechtfertigter, nachteiliger Behandlung von Individuen, einer sogenannten Diskriminierung, führen. Diskriminierung bedeutet, dass jemand aufgrund einer sozial hervorstechenden Eigenschaft erniedrigt oder als moralisch minderwertiger behandelt wird. Fairness herrscht dort, wo eine Art von Gleichheit besteht, in der Diskriminierung weitestgehend verhindert, überwacht, und gemildert wird. Diskriminierung kann an unterschiedlichen Stellen entstehen sowohl im Modellierungsprozess selbst, während Modelllebenszyklus oder durch das eingesetzte Umfeld. |
| Transparenz | Transparenz fragt in erster Linie, ob die Grundfunktionalität eines KI-Systems für Anwender, Betroffene, Experten, das Management sowie alle externen Stakeholder ausreichend nachvollziehbar ist und ob Ergebnisse des Systems ggf. reproduziert und gerechtfertigt werden können. Es behandelt Aspekte der Rückverfolgbarkeit, Reproduzierbarkeit und Erklärbarkeit. |
| Robustheit | Robustheit befasst sich mit Risiken, die entstehen, wenn Eingabedaten beschädigt oder manipuliert werden, aber für die eine genaue Verarbeitung durch die ML-Komponente vorgesehen ist (z. B. falsche Zeichen in einem Wort). |
| Sicherheit | Sicherheit umfasst sowohl funktionale Sicherheitsmerkmale als auch die Absicherung gegen Angriffe und Manipulationen des KI-Systems. |
| Datenschutzkonformität | Diese Dimension bezieht sich auf den Schutz sensibler Daten im Rahmen der Entwicklung und des Betriebs eines KI-Systems. Dabei geht es sowohl um den Schutz personenbezogener Daten als auch um Geschäftsgeheimnisse. Hierbei verweisen wir auf die bestehende Datenschutzrichtlinie und das Datenschutzhandbuch sowie die SECAM-Klassifizierung eines IT-Systems, das auch datenschutzrelevante Aspekte umfasst. |
| Menschliche Kontrolle | KI-Systeme werden in der Regel in Situationen eingesetzt, in denen es eine Form der Interaktion mit der Außenwelt und typischerweise auch eine Interaktion mit dem Menschen gibt. Die mögliche Spannweite reicht von voller menschlicher Kontrolle („Human Control“) bis hin zu automatisierten Entscheidungen („Human-out-of-the-loop“). Dazwischen unterscheidet man „Human-in-the-loop“ und „human-on-the-loop“. Beim ersteren ist der Mensch bei den meisten Entscheidungen stark involviert und kann jederzeit den Prozess voll übernehmen, während bei zweiterem dem Menschen eine Überwachungsfunktion zukommt und das System die meisten Entscheidungen trifft. |
| Deep Fake | Deep Fake bedeutet im Allgemeinen manipulierte oder synthetische Audio-, Bild- oder Videoinhalte, die fälschlicherweise authentisch oder wahrheitsgemäß wirken. Sie enthalten Abbildungen von Personen, die Dinge zu sagen oder zu tun scheinen, die diese nicht sagten oder taten, sondern mit KI hergestellt wurden. |