

Autor: ggg ggg  
 Seite: 8  
 Ressort: K

Rubrik: KAPITALMARKTFORSCHUNG  
 Mediengattung: Tageszeitung  
 Nummer: 31

# Die KI braucht bei der Bankberatung immer noch menschliche Hilfe

Von **Kevin Bauer (Leibniz SAFE)**, **Oliver Hinz (Goethe-Universität Frankfurt)**, Cathy Liu (HEC Paris), Xitong Li (HEC Paris)

Börsen-Zeitung, 14.2.2023

Infolge der großen Fortschritte im Bereich des maschinellen Lernens übertrifft künstliche Intelligenz (KI) in vielen Fällen menschliche Prognosefähigkeiten und kann bereits mit minimalem oder zuweilen gänzlich ohne menschliches Zutun Hochleistungen erbringen. In der Finanzindustrie finden KI-Modelle unter anderem Anwendung in der Analyse von Kreditwürdigkeit, in der Anlageoptimierung sowie im Bereich Robo-Advisory.

Die große Mehrheit heutiger KI-Modelle fällt unter das sogenannte "Supervised Learning"-Paradigma - zu Deutsch: überwachtes Lernen. Supervised Learning ist ein Ansatz zur KI-Entwicklung, bei dem ein Modell auf Basis einer großen Menge historischer Eingabe-Ausgabe-Beispiele trainiert wird. Das Modell wird mehr oder weniger automatisiert so lange trainiert, bis es die zugrundeliegenden Muster und Beziehungen zwischen den Eingabeinformationen (Merkmalen) und den Ausgabeinformationen (Klassen) erkennen kann. Das zugrundeliegende Ziel eines trainierten Modells ist es, generalisierbare Muster zu identifizieren, sodass es auch für nie zuvor gesehene Observationen mit bestimmten Merkmalskombinationen genaue Vorhersagen über die unbekannte Klasse erzeugen kann. Beispielsweise können Banken basierend auf den historischen Anlageentscheidungen (Klassen) ihrer Kunden mit bestimmten soziodemografischen Eigenschaften (Merkmale) ein KI-Modell entwickeln, das attraktive Anlageprodukte für Neukunden identifiziert, indem es Vorhersagen darüber erzeugt, ob ein Neukunde mit spezifischen Eigenschaften in ein gegebenes Produkt investieren würde.

Noch nicht ganz einwandfrei  
 Obwohl einige KI-Modelle bereits

exzellente Ergebnisse im Hinblick auf die Vorhersagegenauigkeit erzielen können, sind diese Ergebnisse bei einer nuancierten Betrachtung für einzelne Subgruppen von Personen oder Objekten nicht immer zwangsläufig wünschenswert. Dies ist häufig darin begründet, dass moderne KI-Modelle typischerweise nicht in der Lage sind, mit seltenen und ungewöhnlichen Vorfällen umzugehen, geschweige denn Entscheidungen in moralischer und rechtlicher Hinsicht zu verantworten. Letztendlich lernen einzelne KI-Modelle nur die in historischen Daten vorhandenen Muster zu erkennen und diese "blind" auf neue Daten anzuwenden. Abhängig von der Qualität der Trainingsdaten und der verwendeten Trainingsmethoden bergen moderne KI-Systeme die potenzielle Gefahr, dass die Vorhersagegenauigkeit für bestimmte Merkmalskombinationen gravierend schlechter als für andere ausfällt. Für solch ein Versagen von KI-Modellen gibt es eine Vielzahl von Beispielen. Unter anderem konnten Forscher zeigen, dass Modelle zur Gesichtserkennung die Gesichter schwarzer Frauen viel schlechter erkennen als die weißer Männer, was nicht nur eine hohe Varianz der Vorhersagegenauigkeit zur Folge hat, sondern auch aus ethisch-moralischer Perspektive problematisch sein kann. Ein weiteres Beispiel, das für erhebliches Aufsehen sorgte, ist ein von Amazon entwickeltes KI-Modell, das die Lebensläufe von Bewerbern automatisiert bewerten sollte. Mitarbeiter hatten das Modell anhand von Lebensläufen programmiert, die über einen Zeitraum von zehn Jahren bei Amazon eingereicht worden waren und die mehrheitlich von männlichen Bewerbern stammten. Basierend auf diesen Daten lernte das Modell die Lebensläufe weiblicher Kandidaten ungenauer und unbegründet tendenziell schlechter zu bewerten,

sodass die automatisierte Verwendung des Modells im operativen Geschäft eine diskriminierende Einstellungspolitik zur Folge hätte haben können. Dieses (geschlechterspezifische) Versagen des Modells bewegte das Unternehmen letztlich dazu, das KI-Modell nicht für Anstellungsentscheidungen zum Einsatz kommen zu lassen.

Kombination als Lösung  
 Vor diesem Hintergrund ist es nicht weiter verwunderlich, dass sowohl politische Entscheidungsträger als auch private Unternehmen zunehmend Möglichkeiten diskutieren, wie das Problem potenziell gravierenden Versagens moderner KI-Modelle adressiert werden kann. Eine besonders faszinierende und möglicherweise vielversprechende Strategie besteht in der Kombination einer besonderen menschlichen Fähigkeit - der flexiblen und adaptiven Reaktion auf noch nie zuvor gesehene Szenarien - mit der von KI-Systemen, nämlich der Fähigkeit, hochdimensionale Zusammenhänge präzise zu identifizieren und skalierbar zu generalisieren. Im Idealfall führt die Kombination dieser komplementären Stärken in Form einer Mensch-KI-Kooperation zu einem hybriden System, das Aufgaben besser löst, als der Mensch oder die Maschine dies isoliert könnten. Im Ergebnis könnten hybride Mensch-KI-Systeme ein wirksames Mittel darstellen, um Probleme, die mit spärlichen, nicht repräsentativen oder strukturell verzerrten Daten einhergehen, zu überwinden. Und zwar indem menschliches Wissen bereits in den Modellierungsprozess oder zu einem späteren Zeitpunkt in den KI-gestützten Entscheidungsfindungsprozess einbezogen wird. Ein Beispiel für eine Mensch-KI-Kooperation ist im klinischen Bereich ein System zur Unterstützung intuitiver Entscheidungen in der Diagnose des Krebsstadiums,

das mit dem Arzt zusammenarbeitet, um die Diagnose zu vervollständigen.

Insbesondere in der Europäischen Union wird die Integration einer menschlichen Komponente in KI-Systeme zunehmend in neuen Regulierungen zum Einsatz digitaler Technologien aufgegriffen. So wird beispielsweise in Artikel 22 der Allgemeinen Datenschutzverordnung (DSGVO) der Begriff des "Eingreifens einer Person" eingeführt. Danach hat eine betroffene Person grundsätzlich das Recht, nicht einer ausschließlich auf einer automatisierten Datenverarbeitung beruhenden Entscheidung, die rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt, unterworfen zu werden. Stattdessen soll die Entscheidung effektiv im Rahmen einer Mensch-KI-Kooperation getroffen oder zumindest durch eine menschliche Überwachung abgesegnet werden. In jüngerer Vergangenheit hat auch die Europäische Kommission in ihrer KI-Verordnung menschliche Einflussnahme im Bereich der algorithmischen Entscheidungsfindung weiter gestärkt. Dies gilt insbesondere für KI-Dienste, die sich auf grundlegende Menschenrechte und soziale Gerechtigkeit auswirken könnten. In dem weitreichenden Verordnungsentwurf zur KI-Regulierung der Europäischen Kommission (Artificial Intelligence Act) wird explizit die Einführung einer verbindlichen menschlichen Aufsichtskomponente für hochriskante KI-Systeme vorgeschlagen, um Vertrauen ebenso wie ein einheitliches und hohes Niveau des Schutzes der Sicherheit und der Grundrechte zu gewährleisten.

Obwohl die Idee einer Mensch-KI-Kooperation bereits in Regulierungen und neuen Regulierungsentwürfen kodifiziert wird, gibt es gerade aus verhaltens- und wirtschaftswissenschaftlicher Perspektive derzeit allerdings noch auffällig wenige Erkenntnisse, die dieses Konzept des "human in the loop" durchleuchten. So ist insbesondere unklar, wie Dritte darauf reagieren, wenn sie mit einem hybriden Mensch-KI-System anstelle eines reinen KI-Systems oder eines Menschen allein interagieren. Genau an diesem Punkt setzt unsere Forschung an. Im Rahmen einer Feldstudie untersuchen wir, inwiefern ein hybrides Mensch-KI-System aus ökonomischer Perspektive einen positiven Beitrag in einer vertikalen Wertschöpfungskette leisten kann oder ob sich potenzielle Effizienzgewinne verwässern und letztlich nicht bei dem Endkun-

den ankommen. Wir untersuchen das Auftreten dieser Effekte in Zusammenarbeit mit einer deutschen Bank im Kontext der Anlageberatung. Genauer gesagt testen wir, wie Bankberater mit einem von uns entwickelten KI-basierten Beratungstool kooperieren und wie sich die hybride Beratung auf die Anlageentscheidungen von Kunden auswirkt.

#### Monetärer Anreiz

Um Erkenntnisse über die kausalen Effekte des Einsatzes einer hybriden Mensch-KI-Kooperation im Rahmen eines Anlageberatungsprozesses zu gewinnen, wählte unser interdisziplinäres, aus Wirtschaftswissenschaftlern, Informatikern und Verhaltensökonomien bestehendes Team einen experimentellen Ansatz. In enger Abstimmung und Zusammenarbeit mit einem Praxispartner aus der Finanzindustrie entwickelten wir ein experimentelles Protokoll, das im Rahmen einer mehrwöchigen Feldstudie implementiert wurde. In einem zweistufigen Prozess wurde das Experiment mit Kunden und Kundenberatern in zwei zentral gelegenen und gut besuchten Bankfilialen durchgeführt. Um die Anlageberatung so real wie möglich zu gestalten, haben wir das experimentalökonomische Paradigma der Incentivierung genutzt. Diesem Ansatz folgend haben alle Entscheidungen, die von den teilnehmenden Beratern und Kunden getroffen werden, reale monetäre Konsequenzen, d. h. dass sowohl Anlageberatungs- als auch Investitionsentscheidungen zu realen Gewinnen oder Verlusten aufseiten der Entscheider führen, die sich am Ende der Studie in einer Barauszahlung widerspiegeln. Die Verwendung monetärer Anreize erlaubt es uns, Rückschlüsse auf die wahren Präferenzen der Teilnehmer in Bezug auf ihre Interaktion mit der KI bzw. dem hybriden Mensch-KI-System zu ziehen.

#### Zweistufiger Ansatz

Unsere Studie betrachtet eine Anlage- und Investitionsentscheidung mit folgender Struktur: Ein Kunde verfügt über 100 Euro und möchte dieses Geld gerne in Form eines Mikrokredits zu einem festen Zinssatz an eine Privatperson verleihen. Der Bankberater des Kunden verfügt über eine Auswahl verschiedener Privatkreditanfragen. Für jede dieser Anfragen stehen dem Bankberater verschiedene Informationen zu dem potenziellen Kreditnehmer zur Verfügung (z. B. Alter, Kreditsumme, regel-

mäßiges Einkommen, Anlagedauer, Beruf des Kreditnehmers). Auf Basis der gegebenen Informationen, seiner eigenen Erfahrung und mithilfe eines von uns entwickelten KI-Modells, das anhand der Eigenschaften der Kreditnehmer das Kreditrisiko der verschiedenen Optionen vorhersagt, muss der Berater eine finale Risikoeinschätzung und eine Investitionsempfehlung an den Kunden aussprechen. Das von uns zuvor entwickelte KI-Modell wurde auf Basis einer großen Menge historischer Daten einer Privatkreditplattform trainiert. Das empirisch optimierte Modell ist ein neuronales Netzwerk, das eine gute Vorhersagequalität auf repräsentativen Testdaten erzielt. Die Bankberater wurden über die Vorhersagegenauigkeit informiert und konnten die von der KI erzeugten Vorhersagen in ihre Kundenberatung miteinbeziehen oder aber darauf verzichten. In Bezug auf den Entscheidungsfindungsprozess der Berater ist ex ante unklar, inwieweit das KI-Modell die Qualität der Beratung mit Blick auf profitable Anlageempfehlungen verbessern oder beeinträchtigen könnte.

Einerseits wird diese spezielle Mensch-KI-Kooperation als wünschenswert erachtet, da Experten ihr Fachwissen einbringen, potenzielle Fehler der KI überwachen und die Entscheidungsverantwortung gegenüber den Kunden übernehmen, wodurch die Beratungsqualität und das Kundenvertrauen gesteigert werden können. Andererseits könnte die menschliche Komponente in der hybriden Beratung ein Einfallstor für bekannte kognitive Verzerrungen menschlichen Denkens darstellen (z. B. Bestätigungsfehler oder Attributionsfehler), sodass sich die gute Vorhersagequalität der KI nicht ausreichend in die Empfehlungen des Bankberaters übersetzt, was letztlich zu suboptimalen Investitionsentscheidungen der Kunden führt. Vor diesem Hintergrund untersuchen wir zunächst die Auswirkungen des Human-in-the-Loop-Konzepts auf der Produktionsseite des Beratungsprozesses. Unsere Ergebnisse zeigen, dass die Qualität der Beratung nicht darunter leidet, wenn die Berater das letzte Wort bezüglich der Anlageempfehlungen haben, d. h. die Mensch-KI-Kooperation untergräbt die technische Vorhersagequalität des KI-Modells nicht. Wir stellen sogar fest, dass die Nutzung der KI den Bankberatern hilft, ihren Kunden bessere Anlageempfehlungen zu geben.

In einem zweiten Schritt konzentriert

sich unsere Studie auf die Kundenseite. Auch wenn die Kooperation der Berater grundsätzlich zu akkurateren Anlageempfehlungen führt, ist weiterhin unklar, inwiefern Kunden auf eine hybride Mensch-KI-Beratung reagieren. So ist es grundsätzlich denkbar, dass Kunden die Kompetenz infrage stellen, wenn ihr Berater Unterstützung durch eine Maschine erfährt oder sich eine generelle Aversion gegenüber maschinellen Beratungen auch auf hybride Systeme überträgt, sodass Kunden sich zu wenig auf die Beratung verlassen. Wenn dies der Fall wäre, könnten sich Effizienzgewinne auf der Produktionsseite nicht über die Wertschöpfungskette hinweg durchsetzen. Um die Effekte einer hybriden Mensch-KI-Beratung auf Seite der Kunden besser zu verstehen, vergleichen wir diese Form der Beratung mit reiner KI-Beratung und reiner Beratung durch einen menschlichen Berater.

#### Deutliche Präferenz

In unserer Feldstudie haben wir reale Bankkunden angesprochen, die die Bankfiliale in einem Zeitraum von etwa einem Monat besuchten. Die 100 Euro, die wir den Kunden zur Verfügung gestellt haben, konnten direkt behalten oder in einen Privatkredit investiert werden. Verliehen sie das Geld an einen privaten Kreditnehmer, wurde es tatsächlich angelegt und gemäß dem Auszahlungsverhalten des realen Kreditnehmers zurückgezahlt, beziehungsweise, wenn der reale Kreditnehmer zahlungsunfähig wurde, nicht zurückgezahlt. Um den Mehrwert einer hybriden Mensch-KI-Beratung zu isolieren, vergleichen wir die Investitionsentscheidungen der Kunden für drei verschiedene Szenarien: erstens eine reine KI-Beratung, zweitens eine reine Mensch-Beratung, und

drittens eine hybride Mensch-KI-Beratung.

Die Ergebnisse zeigen, dass die Kunden im Vergleich zu einer reinen KI-Beratung eine menschliche oder hybride Mensch-KI-Beratung in signifikant höherem Maße in ihre Investitionsentscheidungen miteinbeziehen. Obwohl die Qualität der reinen KI-Beratung vergleichbar hoch ist wie die der Mensch-KI-Beratung, lassen sich die Effizienzgewinne, die sich aus der hohen Vorhersagequalität der KI ergeben, nicht auf Kundenebene realisieren, da diese die Empfehlung einer KI nicht ausreichend berücksichtigen. Nur wenn die Beratung eine menschliche Komponente beinhaltet, also durch eine hybride Mensch-KI-Kooperation erfolgt, können sich die ökonomischen Vorteile des Einsatzes eines KI-Modells zur Anlageberatung über die Wertschöpfungskette hinweg durchsetzen und letztlich Kunden zugutekommen. Interessanterweise zeigen weiterführende Analysen, dass die positiven Effekte einer hybriden gegenüber einer reinen KI-Beratung insbesondere für Investitionsmöglichkeiten mit relativ hohem Risiko auftreten.

Um besser zu verstehen, welche Mechanismen der hohen Akzeptanz einer hybriden Mensch-KI-Beratung im Vergleich zur ausschließlichen KI-Beratung zugrundeliegen, haben wir zusätzlich zu der Feldstudie ein Online-Experiment durchgeführt, welches uns ermöglicht, dafür verschiedene externe Störeinflüsse zu kontrollieren. Das klassische Elaboration Likelihood Model (ELM) aus der Persuasionsliteratur legt nahe, dass sowohl ein zentraler Kanal (d. h. der Glaube an die Qualität des Ratschlags) als auch ein peripherer Kanal (d. h. einfache Hinweise im Persuasionskontext, die den Inhalt des Ratschlags nicht verändern) die Akzeptanz

der Anlageberatung antreiben könnten. Aufbauend auf dem ELM testen wir mehrere plausible Faktoren. Interessanterweise finden wir keine Hinweise dafür, dass der zentrale Kanal die hohe Akzeptanz der hybriden Beratung erklärt. Stattdessen deuten die Ergebnisse darauf hin, dass die Verbraucher wahrscheinlich durch den peripheren Weg beeinflusst werden, insbesondere durch die Beziehung von Mensch zu Mensch (d. h. sozialer Einfluss). So scheint es den Kunden beispielsweise schwerer zu fallen, Ratschläge aus einer Quelle, an der Menschen beteiligt sind, abzulehnen als reine KI-Ratschläge. Zusammenfassend zeigt unsere Studie aus einer angewandten Praxisperspektive die weiterhin hohe Relevanz einer menschlichen Komponente im Anlageberatungsprozess, insbesondere für relativ riskante Finanzprodukte. Obwohl die KI dem menschlichen Berater bezüglich der Bewertung von Investitionsmöglichkeiten individuell überlegen ist, lassen sich die technologischen Vorteile nur dann realisieren, wenn die Beratung in Form eines hybriden Mensch-KI-Systems erfolgt. Allgemeiner können unsere Ergebnisse auch als Evidenz dafür interpretiert werden, dass die häufig dokumentierte Abneigung von Verbrauchern gegen KI-Empfehlungen dadurch gemildert werden kann, dass Menschen in die Erstellung und Vermittlung von KI-Ratschlägen einbezogen werden.

----

Börsen-Zeitung, 14.2.2023

ggg ggg

----

Von Kevin Bauer (Leibniz SAFE), Oliver Hinz (Goethe Universität Frankfurt), Cathy Liu (HEC Paris), Xitong Li (HEC Paris)

**Wörter:**

2065