



Redaktionsgespräch mit Gery Zollinger

„Eine KI kann keine moralischen Urteile fällen“

Herr Zollinger, der Einsatz künstlicher Intelligenz setzt sich immer mehr durch. Wie intelligent und selbstständig denkend sind denn die heute gebräuchlichen Systeme überhaupt?

Grundsätzlich unterscheidet man bei einem Machine-Learning-Algorithmus zwei Phasen: Learning und Running. In der ersten, der Learning-Phase geht es darum, den geeignetsten Algorithmus auszuwählen und anzulernen. Dies ist momentan noch ein manueller Prozess. Dabei werden das Design des Algorithmus und die Input-Parameter aufgrund

mit einiger Mühe und Aufwand erledigen könnten. KI-Systeme entlasten, indem sie die Analyse von Daten automatisieren und standardisieren. Diese automatisierten Analysen können in verschiedensten Bereichen relevant sein. KI kann beispielsweise HR-Verantwortliche bei der Vorauswahl von Bewerbern unterstützen, für Anlageberater relevante Marktnachrichten analysieren oder für Kundenberater Kontaktpunkte liefern. Als Bestandteil einer Beratungsplattform kann eine KI den Relationship Managern auch personalisierte Portfolioempfehlungen vorschlagen, die sich an den Zielen, der

über soziale Kanäle und Plattformen wie WhatsApp oder den Facebook Messenger interagieren möchten, wächst. Im Grunde wird solch ein Conversational Banking überhaupt erst mithilfe der Automatisierung durch KI-Systeme möglich. Beispielsweise sorgt die Intelligenz von Systemen für das Natural Language Processing dafür, die Intention von Kundenanfragen automatisiert zu verstehen. Dadurch ist eine Beratungsplattform in der Lage, den zuständigen Beratern gleich passende Antworten vorzuschlagen. Erst die KI schafft die Basis für eine dialogische Kommunikation zwischen Beratern und Kunden. So wird die Kommunikation im Conversational Banking individuell, persönlich und schnell – ohne die Berater zu überlasten.

„Der Mensch muss den Algorithmus anlernen, von selbst wird er nicht intelligent.“

eines „Training Set“ definiert. Das bedeutet also: Der Mensch muss den Algorithmus anlernen. Von selbst wird solch ein komplexer Algorithmus nicht intelligent – und er wird auch nicht eigenständig intelligenter. Man bezeichnet die folgende Running-Phase zwar gern als „Self-Learning“, aber das ist in diesem Kontext ein falsch verwendeter und irreführender Begriff. Denn in der zweiten Phase des Machine Learning wird der „angelernte“ Algorithmus nur automatisch eingesetzt und auf neue Input-Daten angewandt. Denken im menschlichen Sinn ist das natürlich nicht.

Welche Vorteile bringt KI für die Institute?

KI kann Menschen bei vielen repetitiven Aufgaben unterstützen, die sie selbst nur

Risikoaffinität und den Vorlieben der jeweiligen Kunden orientieren. Auch beim Kreditrating von Kunden kann die KI dazu dienen, menschliche Entscheidungen zu unterstützen. Für automatisierte Systeme mit künstlicher Intelligenz entstehen immer neue Einsatzfelder, nicht nur in Finanzinstituten.

In welchen Geschäftsbereichen von Finanzinstituten könnte diese Zukunftstechnologie noch eine stärkere Rolle spielen als beispielsweise nur in der Datenaufbereitung?

Ein aktuell sehr interessantes Einsatzfeld ist das Conversational Banking, die dialogische Kundenkommunikation über unterschiedlichste digitale Kanäle. Denn die Zahl der Kunden, die mit ihrer Vermögensberatung oder ihrer Privatbank auch

Es gibt jedoch nicht nur Vorteile, sondern auch Risiken von KI. Menschen müssen Vertrauen in die Technik setzen. Das ist aber nicht immer der Fall. Es gibt auch viele Vorbehalte, etwa wegen einer potenziell unethischen KI. Zu Recht?

Das Problem ist, dass menschliche Vorurteile und auch geschlechtsspezifische Ungleichheiten unbewusst ihren Weg in die Entscheidungen einer KI finden können. Das hat schon damit zu tun, dass sich die KI notwendigerweise auf Daten und Entscheidungen aus der realen Welt stützt – in der Learning-Phase wird der Algorithmus schließlich damit trainiert. Und schon das Design des Algorithmus und die Auswahl der relevanten Parameter können unethischen Entscheidungen Vorschub leisten – wenn die KI sich beispielsweise auf problematische Attribute wie das Geschlecht eines Kunden stützt oder auf sei-

nen Wohnort. Es ist einfach so, dass die Resultate einer KI letztlich immer auf menschlichem Urteilsvermögen beruhen. Darum kann eine menschliche Voreingenommenheit – die ganz unbewusst sein mag – am Ende fest im KI-System verankert sein. Es besteht die Gefahr, dass ein Algorithmus sich dann systematisch und auf unethische Weise irrt.

Welche konkreten Risiken können sich für Finanzinstitute dadurch ergeben?

Eine unethische und voreingenommene KI kann ein beträchtliches Reputationsrisiko bedeuten. Stellen Sie sich vor, die HR-Abteilung arbeitet mit einem KI-Sys-

tem unbedingt vermeiden sollten, ist ein Wildwuchs bei der Entwicklung der KI-Modelle, die sie für ihre Datenanalyse nutzen wollen. Deswegen ist es sinnvoll, ein technisches Framework zu definieren, das die Datenwissenschaftler dann bei allen Systemen für KI, Maschinelles Lernen und Prädiktion anwenden. Algorithmen einfach nach Gutdünken zu erstellen, wäre für das Ziel einer ethischen KI fatal. Das einheitliche technische Framework stellt dagegen sicher, dass schon bei der Entwicklung ethische Aspekte beachtet werden. So kann es zum Beispiel sinnvoll sein, mit einem standardisierten Fragebogen zu arbeiten, den die Datenwissenschaftler bei der Entwicklung zu beantworten haben. Ein ganz zentraler Aspekt

„Am Ende müssen immer Menschen beurteilen, ob das KI-Modell moralischen Ansprüchen genügt.“

tem zur Vorauswahl der Bewerber, und kein Bewerber einer bestimmten Ethnie schafft es je bis zu den persönlichen Vorstellungsgesprächen. Käme dies an die Öffentlichkeit, wäre das skandalös. Oder denken Sie an ein System für das Kreditrating, das Bewerber nur aufgrund ihres Wohnorts systematisch schlechter bewertet, als es sachlich und fachlich geboten ist. All dies bedroht nicht nur die Reputation eines Finanzinstituts, es kann ihm auch wirtschaftlich schaden. Denn durch Algorithmen, deren Resultate von irrationalen Vorurteilen geprägt sind, entgehen einem Institut vielleicht wertvolle Geschäftschancen, vielversprechende neue Kunden – oder kompetente Mitarbeiter und Mitarbeiterinnen.

Wie können die Institute die Vorentscheidfreiheit von KI sicherstellen, wie können sie dafür sorgen, dass ihre KI gut und ethisch funktioniert?

Wie gesagt: Ob die KI-Lösung später ethisch agiert und rationale Entscheidungen trifft, hat wesentlich damit zu tun, wie sie entwickelt worden ist. Die Weichen werden schon in der Entwicklungsphase gestellt. Was Finanzinstitute da-

um ist es, dass die Möglichkeit, die Ergebnisse der KI später zu prüfen und zu monitoren, schon in der Entwicklungsphase verankert ist. Es geht darum, Key Performance Indicators zu definieren. Anhand dieser KPIs müssen sich die Ergebnisse der KI-Lösung später kontinuierlich überprüfen lassen.

Wird durch die Kennzahlen sozusagen eine Sicherheitsstufe eingezogen, indem bei zu großen Abweichungen vorher eingestellter Werte Alarm geschlagen wird?

Im Idealfall ist das so. Es gibt dann für die Datenwissenschaftler ein Monitoring-Dashboard, das ihnen stets den Überblick

über die ethisch relevanten Kennzahlen des konkreten Modells liefert. So lässt sich beispielsweise zur Vorhersagegenauigkeit des Modells eine Qualitätskennzahl erheben. Schon in der Entwicklungsphase legt man dafür eine erwartbare

Genauigkeit fest und definiert eine tolerierbare Abweichung. Wenn nun im laufenden Betrieb die Qualität um mehr als 15 Prozent abweicht, kann das beispielsweise auf eine Drift, eine Veränderung in den Daten hinweisen. Das ist auch der Moment, an dem das Monitoring-Werkzeug Alarm schlägt. Denn jetzt heißt es für die Datenwissenschaftler nachzusteuern. Unter Umständen wird es sogar erforderlich, ein ganz neues Modell zu entwickeln.

Eine Neuentwicklung des KI Systems ist auch nötig, wenn der Algorithmus Entscheidungen für Datenbereiche treffen muss, die völlig neu sind und in der Learning-Phase gar nicht vorhanden waren. Etwa wenn ein Kreditrating-System, das ursprünglich nur für Personen über 65 Jahre entwickelt wurde, nun auf Personen ab 18 Jahren angewendet werden soll.

Die EU-Kommission hat in Ihrem Regulierungsvorschlag klare Vorstellungen zu einer rechtmäßigen, ethisch vertretbaren und robusten KI gemacht. Was genau verbirgt sich hinter dem Vorschlag? Und wie ist dieser aus Ihrer Sicht zu beurteilen?

Die EU-Kommission hat ja schon im April 2021 einen „Artificial Intelligence Act“ (AI Act) vorgeschlagen. Das zeigt, dass sie zu den fortschrittlichsten Einrichtungen gehört, was die Regulierung im Sinne einer ethischen KI angeht. Das ist absolut zu begrüßen. Zudem unterscheidet der Regulierungsvorschlag KI-Lösungen nach ihrem Risikopotenzial. Da gibt es drei Kategorien: KI-Systeme mit minimalem Risiko, mit starkem Risiko und mit unan-

„Kleinere Institute ohne umfassendes Data-Science-Team stehen bei KI-Systemen vor Problemen.“

nehmbarem Risiko. Dieses völlig inakzeptable Risiko sieht die EU-Kommission bei Algorithmen, die menschliche Grundrechte beeinträchtigen können. Diese Kategorie von KI-Systemen will der AI Act komplett verbieten. Die meisten

Details des Regulierungsvorschlags beziehen sich damit auf KI-Systeme mit starkem Risiko.

Welche Konsequenzen ergeben sich daraus für die verantwortungsbewusste Entwicklung und Anwendung von KI-Systemen?

Es gibt weltweit schon etliche Organisationen, die sich damit beschäftigen, Best Practices für den KI-Einsatz zu definieren. Mit dem AI Act ist die EU-Kommission da weiter als andere. Es lohnt darum, sich dessen Vorgaben genau anzusehen und sich bereits jetzt daran zu orientieren. So ist es für ein Finanzinstitut sinnvoll, seine KI-Modelle hinsichtlich ihres Risikoniveaus auf den Prüfstand zu stellen: Wo ist das ethische Risiko hoch – wie etwa bei Kreditrating und Bewerberauswahl? Und wo ist es eher niedrig – wie etwa bei NLP-Funktionen im Conversational Banking? KI-Systeme mit hohem Risiko brauchen natürlich besonderes Augenmerk und ein konsequentes Monitoring.

Auch für die EU-Kommission ist das ein wesentlicher Punkt: die Kontrolle der Systeme und ihr Monitoring. Die Kommission verlangt von KI-Systemen nicht nur grundsätzlich Fairness und Nicht-Diskriminierung, sondern auch Transparenz bei Daten, Systemen und KI-Geschäftsmodellen. Zudem sollen die Systeme rechenpflichtig und überprüfbar sein. Ein konsequentes Monitoring der KI-Modelle – und ein verantwortungsvoller Entwicklungsprozess – sind aber nicht nur aus regulatorischer Perspektive und aus Compliance-Gründen sinnvoll. Für ein Finanzinstitut ist es bei anderen KI-Systemen mit moderatem Risiko oft ähnlich wichtig, die Performance zu überwachen. Auch bei den virtuellen Assistenten, die die Vermögensberater bei ihrer Arbeit und im Conversational Banking unterstützen, braucht es eine Qualitätskontrolle.

Denn wenn der virtuelle Assistent regelmäßig die Intention von Kundenanfragen missversteht oder geschlechtsspezifische Antwortvorschläge macht, stellt auch das ein Qualitätsproblem dar. Es wird Kunden verärgern. Und wenn ein

KI-Modell häufig individuelle Anlagevorschläge generiert, die auf die Kunden gar nicht passen, ist auch das ein gravierendes Qualitätsproblem. Das Institut beeinträchtigt so seine Kundenbeziehung und letztlich entgeht im Geschäft.

Schon vor dem Hintergrund der immer weiter zunehmenden Regulierung: Droht das Ganze nicht insbesondere kleinere Häuser zu überfordern?

Kleinere Institute ohne umfassendes Data-Science-Team stehen bei KI-Systemen, die sie in der eigenen Umgebung betreiben, natürlich vor einem gewissen Problem. Für sie ist es mitunter schwierig, die Kontrolle zu realisieren, die fachlich oder regulatorisch nötig wäre. Wenn zum Beispiel die KI des virtuellen Assistenten Markt- und Wirtschaftsnachrichten kundenindividuell zusammenstellt, um Kontaktanlässe zu liefern und Anlageoptionen aufzuzeigen, dann steht der isolierte Betrieb vor Ort der Qualitätskontrolle im Weg. Gerade bei solchen KI-Systemen mit moderatem Risiko – wie dem virtuellen Assistenten oder dem Conversational-Banking-Tool – kann es darum besser sein, sie im Software-as-a-Service-Modell zu beziehen. Denn der SaaS-Anbieter hat selbst viel bessere Möglichkeiten zur Qualitätskontrolle. Allerdings ist solch ein SaaS-Ansatz bei KI-Systemen, die sehr nah am Kerngeschäft eines Finanzinstituts sind, vermutlich nicht optimal. Bei Recommendation Engines oder Systemen für das Kreditrating, bei denen ja auch das ethische Risiko höher ist, kann es für das Institut sinnvoller sein, sie in eigener Regie zu entwickeln und die Qualität der Ergebnisse selbst zu monitorieren. In diesem Zusammenhang sind natürlich die erwähnten KPI-Dashboards und Alarmfunktionen sehr hilfreich.

Was würden Sie sagen, was darf KI und was nicht?

Was KI nicht darf und auch nicht kann, ist den Menschen ihre Entscheidungen abzunehmen. Der Zweck von KI-Systemen ist es, Menschen zuzuarbeiten und ihnen zu helfen, fundierte Entscheidungen zu treffen. Auch das obligatorische Monito-



Gery Zollinger



Head of Data Science & Analytics,
Avaloq Group AG, Zürich

Innerhalb weniger Jahre wird künstliche Intelligenz, kurz KI, die Finanzbranche massiv verändern: Neuronale Netze übernehmen die Beurteilung und Abwicklung von Kreditanfragen, riesige Datenbestände werten Unternehmen mit Hilfe von Deep Learning aus. Das hilft, Betrug vorzubeugen und ressourcenintensive, repetitive Prozesse und Kundenservices ohne Qualitätseinbußen automatisiert zu bearbeiten. Die Banken haben das Potenzial von KI in Zeiten der Digitalisierung also erkannt, da sie viele Vorteile bringt, aber der Umgang mit KI birgt auch gewisse Risiken, weshalb zur Vorsicht zu raten ist. Schließlich lernen die Systeme schon bei der Entwicklung von ihren menschlichen Vorbildern. Das kann beispielsweise dazu führen, dass unbewusst menschliche Vorurteile oder geschlechtsspezifische Ungleichheiten ihren Weg in die Entscheidungen einer KI finden können. Welche Konsequenzen das für die Institute haben kann, erklärt der Autor im Redaktionsgespräch. (Red.)

ring eines KI-Systems kann ja keine KI übernehmen. Am Ende müssen immer Menschen beurteilen, ob das KI-Modell moralischen Ansprüchen genügt und die Anforderungen an eine ethische KI erfüllt. Es gibt zwar schon KI-Systeme, die andere KIs trainieren können, aber als technisches Werkzeug kann eine KI keine moralischen Urteile fällen. Das können nach wie vor nur Menschen. Wir müssen dafür sorgen, die Idee der ethischen KI umzusetzen, und die Qualität unserer KI-Systeme überwachen.