

„Menschliche Intelligenz ist unschlagbar!“

Mit Künstlicher Intelligenz (KI) und Machine Learning zur Predictive Quality

Kann man mithilfe von KI Lottozahlen vorhersagen, und wo wird Machine Learning bereits heute in der Qualitätssicherung angewendet? Die QZ sprach mit Prof. Reiner Kurzhals und Daniel Philippi, beide Westphalia Datalab (Münster), über Chancen und Grenzen sowie über gängige Missverständnisse rund um KI.

QZ Mit welchen unrealistischen Kundenerwartungen werden Sie als Anbieter von KI-Dienstleistungen konfrontiert?

Reiner Kurzhals Wir hatten mal regionalen Vertretern einer deutschen Lotto-Gesellschaft erklärt, wie wir Machine Learning anwenden. Beim Zuhören wurde einer der Vertreter blass. Ich fragte ihn, ob alles in Ordnung sei. Er habe sich gerade gefragt, ob wir auch in der Lage sind, die Lottozahlen von morgen vorherzusagen.

QZ Und, sind Sie dazu in der Lage?

Kurzhals Lottozahlen werden einzig durch den Zufall bestimmt, also können wir sie nicht vorhersagen. Wir sind tatsächlich in der Lage, die Wahrheit in der Zukunft bestmöglich vorherzusagen – sofern sie vorhersagbar ist. Wenn es ein Muster gibt, dann finden wir das heraus. Der Zufall ist das natürliche weiße Rauschen.

QZ Welche Muster zum Beispiel?

Kurzhals Zum Beispiel haben wir im Moment ein großes Projekt in der vorausschauenden Augenheilkunde (Predictive Ophthalmology). Dort arbeiten wir mit dem Weltmarktführer in der Augenheilkunde zusammen, Novartis aus Basel (Schweiz). In diesem Bereich bezieht sich Künstliche Intelligenz auf die Bilderkennung. Wenn es ein Muster in den Daten, eine Wahrheit gibt, wenn in der Biologie ein Krankheitsverlauf angelegt ist, dann werden wir die abnehmende Sehfähigkeit zukünftig

mithilfe von KI früh genug erkennen und vorhersagen können.

QZ Warum ist der Hype um KI gerade in den letzten Jahren so stark geworden?

Kurzhals KI ist im Grunde genommen eine modische Umschreibung für Trial and Error. Hinter der Künstlichen Intelligenz stecken bis zu siebzig Jahre alte mathematische Verfahren: Matrizen. Diese müssen konvertiert oder invertiert werden. Früher hat der Rechner nach einer gewissen Anzahl von Versuchen bei entsprechend großer Datenmenge kapituliert. Jetzt haben wir eine extreme Stei-

gerung der Speicherkapazität, der Prozessoren, der Berechnungsmöglichkeit – paralleles Rechnen, Map-Reducing-Verfahren ... Deshalb sind wir heute in der Lage, das millionenfach schneller und effizienter durchzurechnen.

QZ Welche Ansätze haben sich im Bereich Predictive Quality konkret etabliert?

Daniel Philippi Wir hatten einen Automobilzulieferer als Kunden, der stichprobenartig Qualitätsmessungen durchführte. Die Prüfung der Qualität – auch stichprobenartig – ist aufwändig und teuer, da zum Teil auch zerstörende Prüfverfahren angewandt werden. Deshalb wollte der Kunde verstehen, wie Ausschüsse zustande kommen, ob es Muster und Zusammenhänge gibt zwischen den Voreinstellungen der Maschinen im Laufe des Produktionsschrittes oder in den Chargen. Zudem wurde untersucht, ob man mit Unterstützung einer Qualitätsvorhersage die Prüffrequenz verringern kann, sodass der Kunde weniger Prüfungen durchführen muss.

QZ Wie anspruchsvoll ist eine solche Anwendung?

Kurzhals Der Kunde hat seine Voreinstellungen und Zeitverläufe über den Walzprozess hinweg sowie die Werte der Sensoren aufgezeichnet, also lagen uns viele maschinelle Daten vor. Maschinelle Daten haben den Vorteil, dass sie sehr strukturiert hereinkommen, sodass man sie direkt analysieren kann. »»



Prof. Reiner Kurzhals

ist Gründer der Westphalia Datalab GmbH in Münster, Professor für Statistik an der Fachhochschule Münster sowie Gastprofessor an der EBS Universität für Wirtschaft und Recht.

KONTAKT

Reiner Kurzhals
kurzhals@westphalia-datalab.com

Ein anderes Extrem im Hinblick auf die Datenbeschaffenheit ist die Kundenbetreuung, da die Daten über ganz unterschiedliche Kanäle hereinkommen: E-Mail, Social Media und viele mehr. Die Herausforderung besteht darin, diese wenig strukturierten Daten in eine Form zu gießen, damit man sie überhaupt analysieren kann. Hier kommen hauptsächlich Text-Mining-Verfahren zum Einsatz, die die Masse abdecken. So kann

des hundertste Unternehmen es schafft, sich diese Kompetenz über die eigenen Mitarbeiter anzueignen. Es werden auch immer mehr Unternehmen Data Scientists nachfragen – und nicht jedes nur einen! Ein einzelner Data Scientist kann alleine nicht alle Aufgaben abdecken. Man bräuchte ein Team von sechs oder sieben Leuten, um alleine die Anwendungsmöglichkeiten im Bereich der Qualitätssicherung abzubilden. Es gibt

schen Methoden und relativ wenig Computer-Know-how solche Fragestellung zu beantworten. An dem Übergang zu Produktivsystemen kommt der Data Engineer ins Spiel, weil da mehr Elemente aus der Softwareentwicklung auftreten. Der Data Scientist kommt schwerpunktmäßig aus der Statistik und aus der Informatik. Der Data Engineer tendiert zum Softwareentwickler, er versteht, wie man ein System deployen kann, sprich in Produktivsysteme umwandeln kann. Er versteht dennoch auch die statistischen Zusammenhänge in einem Modell, aber weniger in der Tiefe als der Data Scientist, der sich in erster Linie mit der Generierung der Algorithmen beschäftigt.

Kurzahls Der Data Scientist sollte drei Kompetenzen haben. Er sollte statistisch modellieren können, die modernen Programmiersprachen beherrschen und ein Domänenwissen besitzen. Das muss nicht unbedingt ein Ingenieur sein! Wir haben hier Mathematiker, Physiker, BWLer, Biologen, Soziologen, Psychologen. Der Data Engineer bedient eher den Schwerpunkt Programmierung, und dort die Systemintegration.

QZ Ist ein Ende des Hypes in Sicht?

Kurzahls Vor dreißig Jahren haben sich die Industrien gefragt: Brauchen wir eigentlich diese komische IT? Diese Frage ist heute undenkbar. IT ist inzwischen das Herzstück einer jeden Firma. Allerdings hat die Führung der meisten Firmen in der Regel keinen Einblick darin, was genau dort vor sich geht. Jetzt kommt das Machine Learning, die Künstliche Intelligenz. Und wieder fragen sich alle: Brauchen wir das eigentlich? Versetzen Sie sich in die Lage der Firmenlenker: Sie verstehen nicht nur die IT seit Jahren nicht mehr, jetzt kommt noch etwas viel Anspruchsvolleres. Ich sage zu meinen Studenten immer: Im ersten Semester haben Sie die Mathe-Vorlesung gehabt. Mathe ist für die meisten schon ein hartes Brett. Jetzt kommen Sie zu mir in die Statistik-Vorlesungen und lernen das statistische Modellieren und Programmieren. Da ist Mathe ein Kinderspiel dagegen. Auch die Datenwissenschaften sind im Vergleich zur IT viel komplexer, viel anspruchsvoller.

„Die Welt spricht über Künstliche Intelligenz, aber kaum jemand weiß, was das genau ist.“

Prof. Reiner Kurzahls, Gründer Westphalia Datalab

man, ohne eine unendliche Anzahl an Mails und Posts durchlesen zu müssen, sich einen schnellen Überblick verschaffen, wo vermehrt Probleme anfallen. Wenn man das weiß, kann man dem Kunden gegenüber reagieren und das Produkt entsprechend anpassen.

QZ Lohnt es sich für Qualitätsmanager, sich bereits heute eine Expertise im Bereich KI oder Machine Learning anzueignen?

Kurzahls Der Bildungsmarkt hat sich in den letzten Jahren enorm verändert. Es gibt komplett neue Berufsbezeichnungen, zum Beispiel den Data Scientist. Am Standort Münster werden bis zu 150 Data Scientists im Jahr ausgebildet. Solche Standorte sind in Deutschland rar gesät, weil das deutsche Bildungssystem relativ langsam auf solche Veränderungen reagiert. Auf diesem Gebiet sind uns die Amerikaner voraus, auch Großbritannien und Skandinavien, Holland. Deutschland ist langsam, aber gründlich, und es gibt vereinzelt Hochschulen – Oldenburg, Magdeburg, München, Dortmund – die sich darauf spezialisiert haben, diese Nachfrage der Industrie zu decken.

QZ Wie hoch schätzen Sie die Nachfrage ein?

Kurzahls Die Nachfrage ist extrem, das Angebot gering. Ich schätze, dass nur je-

den Data Scientist, den Data Engineer, den UI Designer, den Entwickler mit Data Science Kenntnissen ...

RL Was ist der Unterschied zwischen Data Scientist und Data Engineer?

Philippi Data Scientist ist schwer greifbar, da es relativ interdisziplinär ist. Die Tätigkeit beginnt damit, solche Fragestellungen verstehen und identifizieren zu können, dann mit einfachen statisti-



Daniel Philippi

Daniel Philippi ist studierter Maschinenbauingenieur und seit 2017 als Data Scientist im Westphalia Datalab tätig. Zuvor war er als Unternehmensberater im Bereich Instandhaltung aktiv.

KONTAKT

philippi@westphalia-datalab.com

RL Im Rahmen von Data Science wird eine gewaltige Menge an Daten erhoben und verarbeitet. Wie stellt man die Sicherheit dieser Daten auf technischer und organisatorischer Ebene sicher?

Kurzhalts Sobald wir mit Kundendaten arbeiten, wenden wir Verschlüsselungstechnologie auf höchstem Grad an. Viele unserer Data Scientists haben sich auf Verschlüsselungstechnologien spezialisiert. Generell arbeiten wir sehr wenig mit Kundendaten aus dem B2C-Bereich. Zu 98 Prozent sind wir im B2B-Geschäft und eher mit Maschinendaten unterwegs als mit personenbezogenen Daten.

RL Wie wirkt sich die Datenschutz-Hysterie auf Ihr Geschäft aus?

Kurzhalts Der zweite oder dritte Mitarbeiter, den wir eingestellt haben, war ein Jurist, der auf Datenschutzbestimmungen spezialisiert war. Wie beim Thema Data Science herrscht auch beim Datenschutz ein Wissensdefizit. Daten sind für viele Unternehmen ein Existenz entscheidendes Asset. Da sehen wir uns in der Bring-schuld, dem Kunden zu erklären, was genau wir mit den Daten machen. Tatsächlich verlassen Kundendaten in den seltensten Fällen das Haus des Kunden.

QZ Ist die Cloud immer noch ein Thema?

Kurzhalts Derzeit verschieben viele Unternehmen, z. B. SAP, ihre Produkte und Services in die Cloud. Da tun wir uns in Deutschland noch ein bisschen schwer. Das Cloud-Geschäft ist das sicherste Geschäft, das es gibt – wenn man sich für ein möglichst großes Unternehmen entscheidet. Auf dem Markt gibt es nicht nur US-amerikanische Unternehmen, sondern auch zunehmend deutsche Unternehmen, die Cloud-Lösungen anbieten. Große Cloud-Anbieter wie die Telekom stellen rund um die Uhr eine enorme Manpower für die Sicherheit der Cloud-Landschaft zur Verfügung. Das können kleinere Unternehmen gar nicht leisten. Wenn sie anfangen, in eine eigene Cloud-Lösung zu investieren, ist es ungefähr so, als würde jedes Unternehmen sich einen eigenen Mietwagenverleih aufbauen.

RL Ein anderes Sicherheitsthema ist die Qualität einer KI. Wie stellt man sicher,

dass selbstlernende Systeme keinen Unsinn lernen?

Kurzhalts Man denkt oft, dass bei KI-Anwendungen irgendwas programmiert wird und dann rechnet die KI vor sich hin und keiner schaut mehr darauf. So ist es nicht. Der Mensch spielt immer noch eine große Rolle. Die menschliche Intelligenz ist nach wie vor unschlagbar. Künstliche Intelligenz kommt vor allem dann zum Einsatz, wenn der Mensch es nicht schafft, diese riesige Komplexität von Daten und die Unmenge an Versuchen nachzuvollziehen. Das kann die Maschine besser. Aber der Mensch ist in der Lage, Logik, Wissen und Gefühl miteinander zu vereinen. Das ist das, was uns Menschen ausmacht: das Einschätzen einer Situation, das Justieren, das Kontrollieren, das kann der Mensch gerade in diesem Bereich besser. Wenn unsere KI beim Kunden läuft (sog. On-Premise-Lösungen), dann müssen wir von hier aus immer wieder darauf gucken: Macht die Maschine noch das, was der Kunde will?

QZ Wie oft wird das überprüft?

Philippi Es wird immer dann geprüft, wenn es Veränderungen gibt. Bei manchen Kunden läuft ein Algorithmus unverändert den ganzen Monat, bei manchen ändert er sich einmal im Monat, bei wiederum anderen einmal am Tag. Bei manchen alle zehn Minuten! Es kommt auf die Datenmenge und vor allen Dingen auf die Fragestellung an: Wann verändert sich der Geschäftsprozess so, dass wir auch den Algorithmus anpassen müssen? Wir unterscheiden zwischen kurzlebigen und langlebigen Maschinenprozessen. Es gibt auch einen Unterschied zwischen kontinuierlicher Produktion und Batch-Prozessen, bei denen das Produkt vielleicht einmal am Tag gewechselt wird. Eine Erdölraffinerie zum Beispiel produziert das ganze Jahr über das Gleiche. Wenn man ein Modell hat, welches die Zusammenhänge gut erklärt, dann braucht man das eigentlich das ganze Jahr nicht anpassen, außer man möchte anhand von weiteren gesammelten Daten das Modell optimieren. Da muss man das Modell mit dem menschlichen Wissen verifizieren.

Das Interview führte Regina Levenshtein, QZ.

Agil arbeiten macht Spaß!



ISBN 978-3-446-45133-9 | € 34,-



ISBN 978-3-446-45682-2 | € 49,99



ISBN 978-3-446-45673-0 | € 22,-