

# Von Beobachtungsdaten zu Causal Discovery

Künstliche Intelligenz braucht ein Konzept für Kausalität

*„Big Data“ sind in aller Regel als sogenannte „Beobachtungsdaten“ zu interpretieren – also Daten, die nicht unter den kontrollierten Bedingungen eines Experimentes erhoben wurden, sondern rein beobachtend ohne eine Intervention. Die Analyse solcher Daten bringt viele Herausforderungen mit sich. Insbesondere darf man aus darin beobachteten Korrelationen nicht auf einen ursächlichen Zusammenhang zwischen zwei Größen schließen. Wissen über Ursache und Wirkung ist aber essenziell, wenn wir ein System intelligent steuern wollen, z.B. wenn wir gezielt in den Betrieb einer technischen Anlage eingreifen wollen, um Ursachen für Produktionsfehler zu eliminieren. „Künstliche Intelligenz“ ist in aller Munde – aber fehlt der Künstlichen Intelligenz nicht ein Konzept für Kausalität?*

*In diesem Beitrag erklären wir an plakativen Beispielen den Unterschied zwischen Beobachtungsdaten und Daten, die unter experimentellen Bedingungen erhoben werden. Gleichzeitig damit werden die Herausforderungen evident, die Beobachtungsdaten mit sich bringen. Wir erklären, was notwendig ist, um potenzielle Ursache-Wirkungs-Zusammenhänge aus solchen Daten zu ermitteln, und welchen Beitrag Xplain Data leistet, um große Mengen komplexer Beobachtungsdaten – sogenannte „Big Data“ – sinnvoll und intelligent zu nutzen.*

*In vielen Unternehmen entstehen zunehmend große Datenpools („Data Lakes“). Verantwortliche und Nutzer dieser Daten sollten sich über die hier dargestellten Zusammenhänge immer bewusst sein.*

„Korrelation impliziert nicht Kausalität“ – vielen ist diese Formulierung bekannt. Sie steht für die Aussage, dass sich anhand von sogenannten „Beobachtungsdaten“ Ursache-Wirkungs-Zusammenhänge nicht beweisen lassen. In den Daten einer Krankenversicherung mag sich z.B. zeigen, dass die Anzahl verschriebener Schmerzmittel mit der Häufigkeit einer Depression korreliert – je mehr Schmerzmittel einem Patienten verschrieben werden, desto häufiger finden wir später auch die Diagnose einer Depression. Trotzdem darf man aus diesem **beobachteten** Zusammenhang nicht auf eine Kausalität schließen – Schmerzmittel verursachen keine Depressionen.

Ein Beweis einer Kausalität bedarf eines Experimentes. Solche Experimente werden alltäglich durchgeführt. Der Ingenieur misst z.B. experimentell den Zusammenhang zwischen der Betriebstemperatur eines Bauteils und dessen Ausfallhäufigkeit und schließt daraus, ab welchen Temperaturen Ausfälle zu erwarten sind. Der Statistiker wertet die Daten einer kontrollierten Studie aus und darf dabei aus darin vorliegenden signifikanten Korrelationen auf Wirkungen und Nebenwirkungen eines Medikamentes schließen.

Warum darf der Ingenieur in einem Experiment und der Statistiker in einer kontrollierten Studie aus einer Korrelation auf eine Kausalität schließen, wir dürfen aber nicht aus der beobachteten Korrelation zwischen Schmerzmittel und Depression auf mögliche Nebenwirkungen der Schmerzmittel schließen?

Der Grund, warum das nicht erlaubt ist, ist allgemein bekannt (obwohl gleichzeitig auch häufig ignoriert oder jedenfalls oft nicht ausreichend im Bewusstsein). Der Unterschied liegt darin, dass in dem Schmerzmittel-Depression-Beispiel sogenannte „Beobachtungsdaten“ vorliegen, während in den anderen beiden Fällen Daten experimentell erhoben werden. Mit der experimentellen Erhebung wird sichergestellt, dass zwischen unterschiedlichen Gruppen eine Vergleichbarkeit besteht.

Der Ingenieur vermisst z.B. mehrere, exakt baugleiche Teile (oder sogar das identische Bauteil) bei unterschiedlichen Temperaturen. Der Statistiker stellt z.B. über eine Randomisierung sicher, dass in den Gruppen, die er vergleicht, zwar nicht exakt gleiche Patienten sind, diese aber jedoch im statistischen Mittel identisch sind.

Der Ingenieur würde nicht auf die Idee kommen, die Messung der Ausfallhäufigkeit bei einer niedrigen Temperatur an einem Fabrik-neuen Bauteil vorzunehmen, während die Messung der Ausfallhäufigkeit bei hoher Temperatur an einem Bauteil erfolgt, das schon Jahre in Betrieb war. Genau das würden wir aber implizit tun, wenn wir aus dem beobachteten Zusammenhang zwischen Schmerzmittel und Depression auf eine Kausalität schließen.

Wenn wir Patienten, die viele Schmerzmittel einnehmen, mit denen vergleichen, die selten oder noch nie Schmerzmittel benötigt haben, vergleichen wir implizit junge mit älteren Patienten. Bei älteren Patienten treten aber ohnehin öfter Depressionen auf. Für die häufige Einnahme von Schmerzmitteln und das häufige Auftreten einer Depression gibt es also eine gemeinsame Erklärung: das Alter. Je älter der Patient, desto häufiger benötigt er Schmerzmittel und desto häufiger finden wir auch Depressionen. Schmerzmittel und Depressionen sind also korreliert, obwohl es keinen kausalen Zusammenhang zwischen den beiden Faktoren gibt.

Unter „Beobachtungsdaten“ versteht man allgemein Daten, die unter rein beobachtenden Bedingungen (ohne eine Intervention) entstanden sind – im Unterscheid zu experimentell erhobenen Daten, in denen Parameter und Vergleichsgruppen aktiv kontrolliert werden. Dass Beobachtungsdaten keine Schlussfolgerungen über Ursache und Wirkung erlauben, ist bekannt ...

... aber sind wir uns dessen auch immer „bewusst“ bei vielen Schlussfolgerungen, die wir tagtäglich ziehen? Was bedeutet „Korrelation  $\neq$  Kausalität“ insbesondere im Zeitalter der „Big Data“, des „Machine Learning“ und der „Künstliche Intelligenz“?

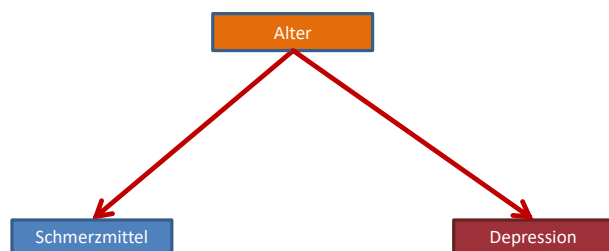
De facto alle Daten, die man unter „Big Data“ versteht, sind als Beobachtungsdaten zu bewerten. Und diese schöne, neue Welt der „Big Data“ soll uns nicht helfen „Ursache und Wirkung“ zu verstehen? Wie „intelligent“ ist eine Künstliche Intelligenz, wenn sie aus der Beobachtung eines Systems keine Schlüsse über Ursache und Wirkung ziehen kann, um basierend darauf ein System gezielt zu steuern?

Der Begriff der Kausalität spielt tatsächlich im Bereich der Künstlichen Intelligenz noch kaum eine Rolle! „Kausalität“ ist mathematisch sogar nicht definiert. Jeder Student kann die mathematische Definition eines Korrelationskoeffizienten anschreiben. Was aber ist „Kausalität“, und ist ein solcher Begriff basierend auf reinen Beobachtungsdaten überhaupt definierbar?

Man kann Kausalitäten anhand von Beobachtungsdaten nicht beweisen, sogar schwer mathematisch definieren. Unter den vielen schlechten Nachrichten nun aber zu der guten Nachricht: Man kann Hinweise auf kausale Zusammenhänge gewinnen – je nachdem wie geschickt man es macht mehr oder weniger starke Hinweise.

Wir beziehen uns hierbei auf eine Definition von Kausalität, wie sie von Kenney (1979) gegeben wurde. Ein Teil dieser Definition besagt, dass die Wirkung der Ursache zeitlich immer folgen muss. Ursache und Wirkung haben mit „vorher und nachher“ bzw. einer beobachteten Veränderung nach dem Auftreten einer Ursache zu tun.

Das zweite Element der Definition besagt, dass es „keine andere Erklärung“ für die beobachtete Veränderung danach geben darf. Im Falle „Schmerzmittel und Depression“ vermuten wir eine „andere Erklärung“ für den beobachteten Zusammenhang, eine andere als eine direkte Kausalität. Die andere Erklärung ist das Alter. Das Bild unten visualisiert den Zusammenhang in Form eines Graphen:



*Das „Alter“ erklärt indirekt den Zusammenhang zwischen Schmerzmittel und Depression – man bezeichnet Alter als einen „Confounder“, eine gemeinsame Ursache für beides.*

Um das Alter als indirekte Erklärung zu erkennen, muss das Alter der Patienten natürlich in den Daten bekannt sein. Fehlende Informationen zu wichtigen Confoundern führen zu falschen Schlüssen über Ursache und Wirkung.

Dabei kann es viele andere Faktoren geben, die als gemeinsame Erklärung in Frage kommen. Neben dem Alter in unserem Beispiel fallen uns schnell weitere plausible Faktoren ein, die sowohl die Einnahme eines Schmerzmittel nach sich ziehen als auch zu einer Depression führen könnten. Viele schwere Erkrankungen gehen mit einer Schmerztherapie einher und stimmen gleichzeitig den Patienten depressiv. Übergewicht könnte dazu führen, dass sich Schmerzen z.B. in Gelenken einstellen, und die übergewichtige Lebenssituation mag auch zu einer Depression führen. Übermäßiger Alkoholgenuss führt am nächsten Tag zu der Einnahme von Schmerzmitteln, und der langfristige Missbrauch zu einer Depression ...

... es könnte also sehr viele solcher „confounding“ Faktoren geben – alle müssen geprüft werden, und alle müssen natürlich in den Daten bekannt sein! Nur wenn keiner davon (oder keine Kombination dieser Faktoren) einen beobachteten Zusammenhang indirekt erklären kann, entsteht der Verdacht einer potenziellen, direkten Kausalität.

Daraus sind insbesondere zwei Schlussfolgerungen zu ziehen:

A) Um gute Hinweise auf potenziell kausale Zusammenhänge zu gewinnen, brauchen wir sehr umfangreiche Daten – idealerweise ein 360°-Bild des analysierten Objektes, z.B. ein vollständiges Bild des Patienten mit allen historischen Daten. Nur wenn uns alle relevanten Informationen zur Verfügung stehen, können wir Kausalitäten von banalen Korrelationen unterscheiden.

Eine Korrelation kann man anhand nur der beiden betreffenden Größen berechnen. „Kausalität“ ist in einem gewissen Sinne genau das Gegenteil. Um einen potenziell kausalen Zusammenhang zwischen zwei Größen zu bewerten, brauchen wir – neben den beiden Größen – auch „alle anderen Größen“. Der potenziell kausale Anteil einer beobachteten Korrelation ist der Teil, der sich nicht über andere Größen erklären lässt.

B) Umfangreiche Daten sind dabei aber wertlos, wenn nicht gleichzeitig auch sehr starke und schnelle Algorithmen zur Verfügung stehen, die in der Lage sind, in solchen umfangreichen und komplexen Daten „alle alternativen Erklärungen“ zu suchen und zu bewerten. Solche umfassenden Daten sind dabei notwendigerweise wesentlich komplexer als nur eine flache Tabelle, wie sie im Allgemeinen der Ausgangspunkt für klassische statistische Verfahren ist.

Xplain Data adressiert diese beiden Punkte. Unsere Object Analytics Database wurde mit dem spezifischen Ziel entwickelt, ein komplexes Objekt im analytischen Zugriff zu haben, übergreifend über alle Daten, die an dem Objekt hängen (rekursive Unterobjekte des Wurzel-Objektes). Im Bereich Gesundheit ist das Wurzel-Objekt in der Regel „Der Patient“, mit Millionen gespeicherter einzelner Patienten-Beispiele (Objekt-Instanzen) und daran gebunden mehreren Milliarden Ereignisse wie z.B. verschriebene Arzneimittel, gestellte Diagnosen, durchgeführte Prozeduren und Operationen etc.

Der Name „Object Analytics“ bringt zum Ausdruck, dass mit dieser Form der Datenhaltung ein 360°-Blick auf ein komplexes Objekt möglich wird. Die Notwendigkeit bei der Analyse von Ursache-Wirkungs-Beziehungen „alle alternativen Erklärungen“ zu evaluieren, war der Ausgangspunkt für die Entwicklung dieser Form der Datenhaltung. Die Causal-Discovery-Algorithmen entwickeln aus komplexen Daten automatisiert mehrere Millionen „alternative Erklärungen“, um für einen beobachteten Zusammenhang eine indirekte Erklärung zu finden. Nur wenn so keine indirekte Erklärung gelingen, wird ein Faktor als potenzielle, direkte Kausalität vorgeschlagen. Umfang der Daten und Stärke der Suchverfahren bestimmen dabei die Qualität des Ergebnisses.

Die Betonung liegt dabei noch immer auf „potenziell“ kausal – auch mit sehr umfangreichen Daten ist das noch immer kein Beweis. In der Praxis ist man in der Regel auch weit von „umfassenden“ Daten entfernt.

Das zieht die Notwendigkeit nach sich, dass von dem Algorithmus vorgeschlagene Hypothesen durch den Experten bewertet und interpretiert werden müssen. Um zu wertvollen Erkenntnissen zu kommen, muss der Experte z.B. einen Faktor zurückweisen können, um eine zweitbeste Erklärung anzufordern. Der Experte muss Folgefragen stellen können, z.B. warum ein Faktor, der als kausaler Einflussfaktor erwartet wurde, nicht als solcher identifiziert wurde ...

Wir kommen also zu einer weiteren Schlussfolgerung über ein notwendiges Element in einem Causal Discovery Prozess:

- C) Die algorithmische Intelligenz muss ergänzt werden durch eine Benutzeroberfläche, die Hypothesen für potenziell kausale Faktoten intuitiv verstehbar präsentiert und es dem Experten erlaubt, in einem Prozess von Fragen und Folgefragen zu einem Verständnis zu kommen.

Die Xplain Data GmbH stellt solche Oberflächen und Interaktionskonzepte mit dem Experten zur Verfügung. Diese werden ständig basierend auf den Erfahrungen aus Projekten in verschiedenen Industrien weiterentwickelt.