

Daten verstehen lernen – Data Science Literacy als Schlüsselkompetenz der Zukunft

Ein falscher Umgang mit Daten kann gravierende Konsequenzen für wirtschaftliche und politische Entscheidungen haben. Professor Gero Szepannek lehrt Wirtschaftsinformatik an der Hochschule Stralsund und fordert in einem der derzeit weltweit meistgelesenen Fachartikel zum Thema Data Science zu einem bewussten Umgang mit Daten in der Hochschulbildung auf. Im Interview erklärt er, warum es nicht reicht, mathematisches Verständnis zu vermitteln, und warum der kritische Umgang mit Daten so wichtig ist.

Professor Szepannek, Ihr Artikel [„Why We Should Teach Causal Inference: Examples in Linear Regression with Simulated Data“](#) ist kürzlich im amerikanischen „*Journal of Statistics Education*“ veröffentlicht worden. Sie haben mit Ihren Mitautoren, Professor Dr. Jörg Horst (FH Bielefeld), Professor Dr. Karsten Lübke (FOM Hochschule Dortmund) und Professor Dr. Matthias Gehrke (FOM Hochschule Frankfurt) viel Aufmerksamkeit dafür bekommen.

Warum haben Sie als Autoren einen Bedarf gesehen, für die Ausbildung und Lehre in der Statistik einzuführen, dass Studierenden ein besseres Verständnis für Daten vermittelt werden sollte?

Gero Szepannek: Wir können derzeit in fast allen Aspekten des gesellschaftlichen Lebens eine Digitalisierung unserer Gesellschaft beobachten und leben mittlerweile in einer Welt von Big Data. Auf Grundlage der Daten, die uns umgeben, werden wichtige Entscheidungen getroffen.

Haben Sie ein Beispiel für Daten und daraus folgende Entscheidungen?

Szepannek: Ein anschauliches Beispiel liefert der aktuelle Umgang mit dem Coronavirus in den Medien. Man findet täglich Meinungen und Kommentare, deren Schlussfolgerungen sich zum Teil aus den vorhandenen Daten nicht begründen lassen. Einschätzen zu können, welche Schlüsse man aus Daten ziehen darf, aber insbesondere auch, welche nicht, wird heute immer wichtiger.

Haben Sie ein konkretes Beispiel, warum es so wichtig ist, den Unterschied genauer zu erfassen?

Szepannek: Bleiben wir bei der aktuellen Coronasituation: In einem vor wenigen Tagen im Telegraph erschienenen [Artikel](#) fanden Wissenschaftler heraus, dass glatzköpfige Menschen ein erhöhtes Risiko aufweisen. Hierbei wurde jedoch das Alter nicht in Betracht gezogen: so tritt eine Glatze vermehrt bei älteren Menschen auf, und die Ursache für das erhöhte Risiko ist entsprechend vermutlich nicht im Haarwuchs, sondern im Alter zu finden.

In Ihrem Artikel beziehen Sie sich auf das 2018 erschienene Buch „*The Book Of Why*“ von Professor Dr. Judea Pearl, einem der bekanntesten und einflussreichsten Informatiker und Philosophen aus den USA und Pionier im Bereich künstlicher Intelligenz. Haben Sie eine Rückmeldung von ihm erhalten?



Szepannek: Auf [Twitter](#) zeigt sich Pearl mit den Worten „This is incredible!“ begeistert von unserem Artikel. Darauf sind wir natürlich sehr stolz. Für mich ist es gleichzeitig auch ein Kompliment für unsere Lehre im Bereich Data Science am Studiengang Wirtschaftsinformatik der Hochschule Stralsund.

Warum ist der sensible Umgang mit Statistiken oder Definition von Zusammenhängen bisher nicht so im Fokus der Lehre gewesen?

Szepannek: Traditionell erfolgt die Ausbildung in Statistik in der Regel als Teilgebiet der Mathematik. Der Fokus liegt dabei eher im mathematischen Verständnis der Methoden, und deren Anwendung tritt dabei meist in den Hintergrund.

Mittlerweile bietet sich dank computergestützter Auswertungsmöglichkeiten, ein neues Lehrfeld, das die Daten und deren Auswertung stärker in den Blick nimmt. In meinen Vorlesungen nutzen wir zum Beispiel die Analyseumgebung R, um hier genauer auf Zahlen und Zusammenhänge zu schauen und zu lernen, was sie bedeuten.

Was genau ist Ihre Forderung?

Szepannek: Die Fähigkeit, mit Daten angemessen umgehen zu können, auch Data Literacy genannt, sollte im Sinne von „Data Science for all“ als Schlüsselkompetenz verstärkt in Curricula verankert werden. Dieses Grundverständnis wird zunehmend wichtiger, um den rasanten Entwicklungen im Bereich der künstlichen Intelligenz kritisch und unabhängig begegnen zu können.

Ebenfalls ist der Ausbau von Studienangeboten zur Advanced Data Science wegweisend, da die entsprechenden Fachkräfte in der Zukunft dringend benötigt werden. Wir sind gerade dabei unser Lehrangebot hierzu noch weiter auszubauen. Für die Studierenden ist Data Science Literacy die Schlüsselkompetenz der Zukunft.

Wie steigen Sie in der Lehre in dieses schwer verständliche Thema ein?

Szepannek: Wichtig finde ich, die Konzepte nicht abstrakt und theoretisch zu lehren, sondern anhand konkreter und realitätsnaher Beispiele. Die Beispiele vereinfache ich für die Studierenden zunächst, damit sie übersichtlich bleiben. Denn wenn wir diese vereinfachten Fälle nicht verstehen, wie sicher können wir dann sein, dass jeder von uns in großen Datensituationen die richtigen Fragen stellt und die richtigen Schlussfolgerungen ziehen wird?

Professor Szepannek, haben Sie vielen Dank für das Gespräch.

Das Gespräch führte Anke Weisbrich, Redakteurin der Hochschulkommunikation.

Bibliografische Angabe:

[Karsten Lübke](#) (FOM Hochschule Dortmund), [Matthias Gehrke](#) (FOM Hochschule Frankfurt), [Jörg Horst](#) (FH Bielefeld) und [Gero Szepannek](#) (Hochschule Stralsund). (2020) “Why We Should Teach Causal Inference: Examples in Linear Regression With Simulated Data”, [Journal of Statistics Education](#), DOI: 10.1080/10691898.2020.1752859.



Zur Person:

[Dr. Gero Szepannek](#) ist seit 2016 Professor für Statistik, Wirtschaftsmathematik und Machine Learning an der Hochschule Stralsund und Leiter des [Competence Center „Artificial Intelligence & Machine Learning“](#) des Institute for Applied Computer Science (IACS), sowie der Leiter der [Arbeitsgruppe Datenanalyse und numerische Klassifikation](#) (DANK) der Gesellschaft für Klassifikation. Nach seiner Promotion 2008 war er bis zum Wechsel in Forschung und Lehre Abteilungsleiter für Scoring & Ratingmodelle der Santander Consumer Gruppe zuständig. Seine Forschungsschwerpunkte sind Machine Learning, Computational Statistics (insbesondere mit der Programmiersprache R) und Data Literacy. In der Angewandten Forschung ist er in den Bereichen Kreditrisikoscoring, Customer Relationship Management sowie Music Information Retrieval aktiv.

Kontakt:

HOST - Hochschule Stralsund – University of Applied Sciences

Hochschulkommunikation

Fon +49 3831 456528

hochschulkommunikation@hochschule-stralsund.de

<http://www.hochschule-stralsund.de/>