

Data Science: Aus der Datenflut Sinn machen. Daten spielen für KMU eine zunehmende Rolle als Entscheidungsgrundlage, Dienstleistung oder Wettbewerbsvorteil. Was steckt hinter den Trendbegriffen wie «Machine Learning», «Data Mining» oder «Big Data Analytics» und wie können die Methoden gewinnbringend angewendet werden?

VON ANDRÉ CSILLAGHY*

Wer kennt die Hinweise beim Onlineshopping nicht? «Personen, die diesen Artikel gekauft haben, interessierten sich auch für folgende Produkte.» Ein Beispiel für das Nutzen von Data Mining: Die systematische Anwendung statistischer Methoden auf grosse Datenbestände mit dem Ziel, neue Querverbindungen und Trends zu erkennen. Maschinelles Lernen – oder Machine Learning – kann zusätzlich helfen, aus den zahlreichen Käufen der Kunden personalisierte Empfehlungen zu generieren.

Aus grossen Datenmengen lernen. Was früher nur für Grossunternehmen möglich war, ist jetzt auch für kleinere Firmen anwendbar. Heute kann jedes KMU ein Datenarchiv aufbauen. Die Schwierigkeit besteht darin, aus grossen Datenmengen nutzbare Informationen zu gewinnen. Am Institut für 4D-Technologien FHNW entwickelt ein Team von Big-Data-Profis zusammen mit internationalen Partnern eine Software zur Vorhersage von Sonnenstürmen. Dabei nutzen sie komplexe Datenströme wie die Magnetfelder-Messungen des NASA-Satelliten Solar Dynamics Observatory SDO. Für die Kommunikations- und die Flugzeugindustrie eine nützliche Dienstleistung, denn heftige Sonnenstürme könnten die elektronischen Geräte beispielsweise in Satelliten beschädigen. Der Umgang mit Big Data ist anspruchsvoll: Die FHNW-Forschenden müssen täglich zwei Terabyte an Daten auswerten. Solch grossen, komplexen, schnelllebigen und schwach strukturierten Daten können nicht mit den herkömmlichen

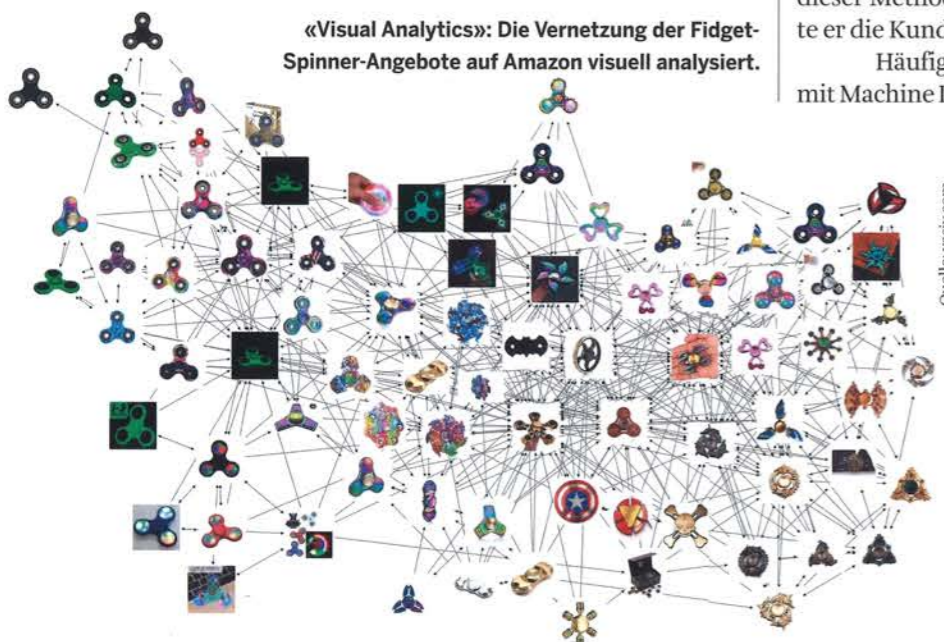
Methoden der Datenverarbeitung ausgewertet werden. Sie brauchen dafür speziell optimierte Algorithmen und sehr leistungsfähige Computer.

Big Data für kleine Unternehmen. Mit Cloud-Computing kann auch ein KMU Kapazität und Rechenpower einkaufen. Zusammen mit der entsprechenden Software können kleinere Unternehmen «Big Data»-Analysen durchführen und nützliche Informationen aus ihren eigenen Daten gewinnen. Als Quellen eignen sich beispielsweise Auswertungen aus den Verkäufen, Informationen der Website-Nutzung oder Messwerte aus Produktionsanlagen. Aktuelles Beispiel: Für ein Unternehmen, das sich auf die Schätzung von Immobilien spezialisiert hat, entwickelt das Institut für 4D-Technologien FHNW eine Software, die auf Basis von Vergleichsdaten direkt einen Schätzwert für den Baupreis ausgibt. Die Datenbasis liegt hier bei rund 150 000 Einträgen, also eine vergleichbar «kleine» Datenmenge, die aber mit Big-Data-Analysemethoden nützliche Informationen generieren kann.

Data Mining und Machine Learning. Sonnenstürme oder Immobilien-Schätzung: Die Anwendungsbereiche mögen unterschiedlich sein, die Herausforderung bleibt die gleiche: die Aufbereitung von Rohdaten in brauchbare Informationen. Ein Werkzeug dazu ist Data Mining, die systematische Anwendung von statistischen Methoden. Der Begriff umfasst den ganzen Prozess der «Wissensentdeckung» von der Datenbereinigung bis zur eigentlichen Analyse. Der Online-Händler Amazon war ein Pionier bei der erfolgreichen Anwendung dieser Methode. Mit persönlichen Empfehlungen ermunterte er die Kunden zum Kauf.

Häufig werden Daten-Analysen wie jene von Amazon mit Machine Learning gekoppelt: Das künstliche System lernt

«Visual Analytics»: Die Vernetzung der Fidget-Spinner-Angebote auf Amazon visuell analysiert.



Quelle: yastv.com

CAS IN DATA SCIENCE

Start: 20. Oktober 2017

Dauer: 12 Kurstage

Voraussetzungen:

Erfahrung bei der Anwendung von mindestens einer Programmiersprache

Weitere Informationen:

www.fhnw.ch/cas-data-science

aus Beispielen und kann nach einer Lernphase die Erkenntnisse selbstständig verallgemeinern. Im Gegensatz zum Data Mining, wo Gesetzmässigkeiten aus bestehenden Daten entdeckt werden, kann Machine Learning neue Muster erkennen und antizipieren.

Text-Mining und Natural Language Processing. Besonders bei Anwendungen im Web handelt es sich bei der Datenform oft um Textbausteine. Um aus den Textdaten Bedeutungsstrukturen zu gewinnen, gesellen sich neben den statistischen auch noch linguistische Methoden zur Analyse. Eine interessante Erweiterung von Text-Analyse ist Natural Language Processing (NLP): Dabei wird natürliche Sprache in Form von Text- oder Sprachdateien algorithmisch verarbeitet. Poppt auf einer Website eine virtuelle Person – ein Chat-Bot – auf und bietet Hilfe an, so steckt meist NLP dahinter. «Clippy» in Microsofts Windows war eine frühe – und eher unbeliebte – Anwendung dieser Technologie, heute haben sich Siri, Cortana oder Alexa im Alltag etabliert.

Visual Analytics. Der Mensch kann Muster oder Trends relativ schnell erkennen, wenn Daten visualisiert werden. Visuelle Methoden der Datenanalyse können daher sehr effizient sein. Zum Beispiel können Beziehungen zwischen Websites,

Tweets, Facebook-Freunden oder Amazon-Produkten als Graphiken visualisiert werden. Aus diesen lassen sich sehr leicht Gruppierungen – auch Clusters genannt – identifizieren.

Diverse Einsatzmöglichkeiten. Die Anwendung von Datenanalyse in der Praxis ist äusserst vielfältig: Von der Marketing-Strategie, der Immobilien-Software, über die Jobplattform, das Medikamenten-Kontrollsystem bis zur Vorhersage von Sonnenstürmen. Überall dort, wo eine grosse Anzahl Daten generiert wird, können «Big Data»-Methoden gewinnbringend eingesetzt werden. Dazu braucht es Data Scientists, die sich mit Statistik, Algorithmen und Datenstrukturen auskennen. Da es sich bei «Big Data»-Analysen oft um entdeckungsartige Vorgehen handelt, werden sogenannte interpretierte Sprachen wie Python benutzt. Diese sind im Allgemeinen einfacher zu lernen als Programmiersprachen wie Java, kommen jedoch auch mit ihren eigenen Herausforderungen.



PROF. DR. ANDRÉ CSILLAGHY ist Leiter des Instituts für 4D-Technologien FHNW. Der Informatiker hat sich auf Methoden der Integration, Selektion, Speicherung und Reduktion von Daten spezialisiert.

ANZEIGE

BESCHLEUNIGE DEINE KARRIERE
Studiere an der HFT Mittelland

hftm.ch

Dipl. Techniker/In HF
in den Fachrichtungen
Elektrotechnik, Informatik,
Maschinenbau, Systemtechnik,
Unternehmensprozesse

Nachdiplome NDS HF
und weitere attraktive
Weiterbildungsangebote

HF Technik Mittelland

facebook.com/hftmittelland