

# Data Mining im Dienste der Gesundheitsförderung



**Dr. Farshideh Einsele**  
Dozentin Wirtschaftsinformatik, BFH



**Dr. Leila Sadeghi**  
Dozentin aF&E Ernährung  
und Diätetik, BFH



**Dr. Helena Jenzer**  
Dozentin aF&E Ernährung  
und Diätetik, BFH

Viele chronische Erkrankungen sind ernährungsbedingt. Wir verfolgen bei unserer multidisziplinären Forschung einen neuartigen Ansatz, um mittels Data-Mining-Methoden aussagekräftige Regeln über den Einfluss von Ernährung auf chronische Krankheiten zu entdecken.

Ernährungsmuster spielen eine wichtige Rolle für die Gesundheit der Menschen. Aus diesem Grund gewinnen gesundheitliche Belange bei Supermärkten, Gesundheitspraktikern, Sportorganisationen und Regierungen zunehmend an Bedeutung. Vermehrt ernähren sich Menschen gesundheitsbewusst, aber sie sind sich der Eigenschaften, Einschränkungen und vor allem der Bestandteile ihrer Nahrung kaum bewusst. Es gibt mittlerweile viele Ernährungsschemata und -programme, denen Einzelpersonen folgen und die ihnen helfen festzustellen, wie gesund sie sind. Die Programme beinhalten aber oft keine kritische Analyse darüber, wie die Ernährung die chronischen Krankheiten beeinflusst.

Unsere Forschungsidee besteht darin, den Einfluss der Lebensmittel auf chronische Krankheiten mittels Analyse grosser Datenmengen aufzuzeigen. Dazu wird eine umfangreiche Datenbank aus den erhobenen Ernährungsdaten oder Warenkorbdaten des Lebensmittelmarktes mit den demografischen und den Gesundheitsdaten aus derselben Region verknüpft. Die daraus resultierende Datenbank wird verwendet, um Ernährungsmuster zu erkennen, die zu modernen chronischen Zivilisationskrankheiten führen. Um solche versteckten Muster zu extrahieren und zu entdecken, verwenden wir Data-Mining-Techniken für unsere Forschung. Data Mining wird zunehmend im Bereich der Datenanalyse eingesetzt, die in jüngerer Zeit aus der Informatik hervorgegangen ist und sich von traditionellen statistischen Analysen unterscheidet. Klassische statistische Analysetechniken werden häufig entwickelt, um Beweise zur Unterstützung oder gegen eine

Hypothese aus einem begrenzteren Datensatz zu bilden. Somit wird üblicherweise bei der statistischen Analyse die Gültigkeit der Hypothese untersucht, indem statistische Tests anhand von Daten durchgeführt werden, die möglicherweise zu diesem Zweck erhoben wurden. Data-Mining-Techniken hingegen werden nicht hauptsächlich dazu verwendet, um Vertrauen in eine Hypothese zu bilden, sondern vielmehr, um unbekannte Beziehungen zu extrahieren, die im Datensatz vorhanden sind. Data Mining ist daher eine Hypothesen-freie Datenanalyse-Methode, die statistische Methoden als Werkzeuge verwenden kann, aber zunächst nicht davon ausgeht, dass eine Hypothese verifiziert oder zurückgewiesen wird.

Das «Proof of Concept» unserer Forschungsidee wurde bereits aufgrund einer Vorstudie untersucht. In dieser Vorstudie wurde vorerst eine US-amerikanische Konsumdatenbank von der Westküste mit den demografischen und den Gesundheitsdaten derselben Bevölkerung integriert. Anschliessend wurden unter Anwendung unserer Data-Mining-Analysenmethoden interessante Regeln, die die Verbindung zwischen Ernährungsgewohnheiten und chronischen Krankheiten demonstrierten, entdeckt.<sup>1</sup>

## Was sind chronische oder Lebensstilkrankheiten?

Die wachsende Epidemie der Lebensstilkrankheiten, von denen sowohl Industrie- als auch Schwellenländer betroffen sind, hängt mit Veränderungen der Ernährung und des Lebensstils zusammen. Veränderungen in der Weltnahrungsmittelwirtschaft spiegeln

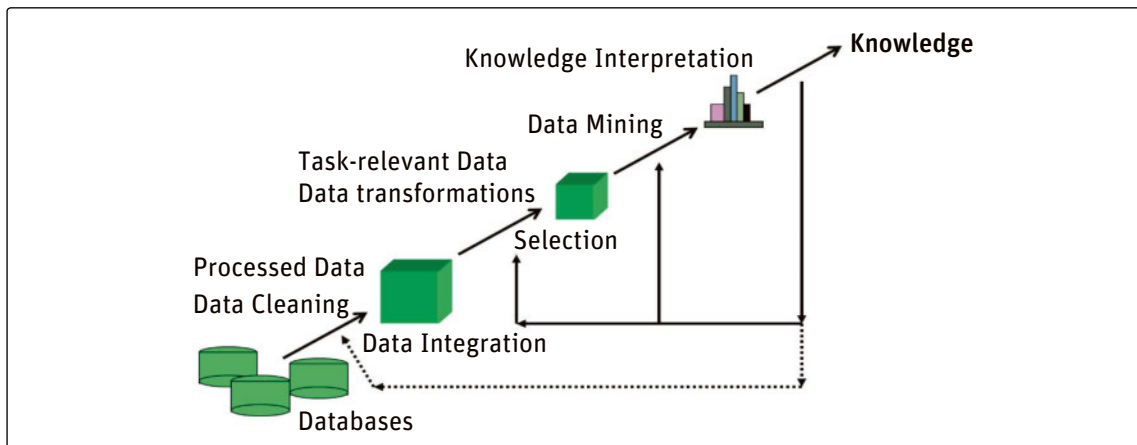


Abbildung 1: Data-Mining-Schritte

sich in sich verändernden Ernährungsgewohnheiten wieder, beispielsweise durch erhöhten Verzehr energiereicher Nahrung mit hohem Fettgehalt, insbesondere gesättigten Fettsäuren, und geringen Mengen nicht raffinierter Kohlenhydrate. In diesen Ländern zeigt sich ein Wechsel vom Muster «hohes Vorkommen von Infektionskrankheiten», die mit Mangelernährung zusammenhängen, zum Muster «hohes Vorkommen von chronischen oder Lebensstilkrankheiten», die mit der Ernährung in urbanen und industriellen Gebieten verbunden sind. Dieses Phänomen wurde bereits durch die WHO erkannt.<sup>2</sup>

Zu den Lebensstilkrankheiten gehören Fettleibigkeit (Adipositas), Bluthochdruck (Hypertonie), Herzerkrankungen, Typ-2-Diabetes, Krebs, psychische Störungen und viele andere.

### Data Mining in der Ernährungsforschung

Daten werden mit einer exponentiellen Rate produziert, zumal die Speicherfähigkeit heute praktisch unbegrenzt ist. Data Mining wird zunehmend in der Datenanalyse verwendet als ein aufstrebendes multidisziplinäres Feld von: Statistiken, maschinellem Lernen, Datenbanken, Informationsrückgewinnung, Visualisierung. Gemäss Fayyad et al.<sup>3</sup> kann Data Mining in vier Schritten aufgeteilt werden: Datenquellen sammeln und reinigen, bereinigte Datenbanken integrieren, Data-Mining-Methoden einsetzen, um neue Regeln zu entdecken, und Wissensinterpretation, um neues Wissen zu schaffen. Abbildung 1 zeigt diese Schritte auf.

### Gesundheitsförderung als Ziel

Unser Ziel ist die Prävention und Vorhersage von Lebensstilkrankheiten in Industrie- und Schwellenländern mithilfe der Analyse von Lebensmittelkonsum- und demografischen und Gesundheitsdaten in der Schweiz, aber auch im Ausland. Die Neuheit unseres Ansatzes besteht darin, verborgene Muster und die damit verbundenen Regeln auf der Grundlage einer Hypothesen-freien Forschungsmethode zu entdecken. Die Verknüpfung von Lebensmittelkonsumdaten – sei es aus dem Warenkorb oder aus der statistischen Haus-

haltungserhebung mit demografischen Daten und Gesundheitsstatistiken mithilfe von Data-Mining-Methoden – ist ein aussagekräftiger Weg, um zu bestimmen, ob eine bestimmte Bevölkerungs-Untergruppe aufgrund ihrer Lebensmittelkonsummuster besonders durch bestimmte chronische oder Lebensstilkrankheiten gefährdet ist.

In Zukunft wollen wir uns mit der Beschaffung von Lebensmitteldaten einer Schweizer Supermarktkette oder Nutzung der nationalen Erhebungsdaten und deren Verknüpfung mit den Gesundheitsdaten aus derselben Region befassen. Damit wollen wir ein besser fundiertes Verständnis für die Verbindung zwischen Lebensmittelverzehr, Gesundheitszustand und Volksbeweglichkeit erhalten. Dieses Verständnis wird helfen, Fortschritte im Bereich der Gesundheitsförderung von Konsumenten zu erzielen. Die entwickelte integrierte Datenbank ist anwendungs- und marktorientiert. Die Synergie zwischen Wirtschaftsinformatik und Public-Health-Ernährung zu nutzen, ist eine Erfolg versprechende und angemessene Vorgehensweise, um multidisziplinäre Ernährungsforschung in der Epidemiologie zu betreiben. Epidemiologie befasst sich bekanntlich mit der Verbreitung sowie den Ursachen und Folgen von gesundheitsbezogenen Zuständen und Ereignissen in Bevölkerungen. Wir sind sehr zuversichtlich, dass wir auf diesem Weg dazu beitragen, klinische Langzeitergebnisse, Volksgesundheit und die damit verbundenen finanziellen Ergebnisse zu verbessern.

### Kontakt

– leila.sadeghi@bfh.ch  
– farshideh.einsele@bfh.ch

### Literatur

- <sup>1</sup> F. Einsele, L. Sadeghi, R. Ingold, H. Jenzer, A Study about Discovery of Critical Food Consumption Patterns Linked with Lifestyle Diseases using Data Mining Methods, Proceedings of the International Conference on Health Informatics (HEALTHINF-2015);17: pp.239-245.
- <sup>2</sup> WHO, World Health Organization Geneva 2003, Diet, Nutrition and the Prevention of Chronic Diseases, Report of a Joint WHO/FAO Expert Consultation.
- <sup>3</sup> U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, From Data Mining to Knowledge Discovery in Databases, AI Magazine Volume 17 Number 3 (1996), <https://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1230/1131> (20.8.2018)