



BIGDATA: DATENMANAGEMENT UND DATENANALYSE IN MEDIZIN UND PRODUKTION

DIE AUFGABE

Datenmanagement und der verantwortungsvolle Umgang mit Versorgungsdaten sind wichtige Themen unserer Zeit, nicht nur im Bereich Medizin und Life Sciences. Die Verwaltung medizinischer Daten, beispielsweise Befundergebnisse, kostet Forscher bis zu einem Viertel ihrer Arbeitszeit. Zudem erzeugen die verschiedenen Diagnostik- und Laborgeräte unterschiedliche Datentypen, dabei geht die Übersichtlichkeit schnell verloren.

In großen, komplexen medizinischen Datensätzen sind wichtige Informationen enthalten, die nicht mit herkömmlichen technischen und analytischen Verfahren extrahiert werden können. Der Begriff »BigData« beschreibt Datenstrukturen, die in ihrer Größe, Diversität und Komplexität neue Datenverarbeitungs- und Analysetechniken erfordern, um daraus verborgenes Wissen zu gewinnen. Solche BigData-Lösungen werden am Fraunhofer IWS Dresden im Laborbetrieb erprobt.

In der medizinischen Forschung und Versorgungsforschung stellt die Nutzung von medizinischen Routinedaten eine wertvolle Ressource dar, um Versorgungs- und Risikomuster besser zu verstehen und darauf aufbauend individualisierte Präventions- und Therapiestrategien zu entwickeln. Große Datenmengen bedürfen neuer technischer Möglichkeiten der Speicherung, Prozessierung und Analyse, um die medizinische Forschung voranzutreiben. BigData-Verfahren bieten speziell für die Versorgungsforschung und personalisierte Medizin neue Möglichkeiten und können wichtige, bisher unbekannt Informationen zu Erkrankungsrisiken, Verläufen, therapeutischem Ansprechen und zur Arzneimittelsicherheit liefern.

UNSERE LÖSUNG

Im Rahmen von verschiedenen öffentlich geförderten Projekten wendet das Fraunhofer IWS Dresden erstmals BigData-Analyseverfahren zur Verarbeitung und wissenschaftlichen Analyse von Routinedaten der Krankenhäuser und gesetzlichen Krankenversicherung an. Zur Identifikation von Krankheitsprädiktoren in medizinischen Routinedaten kommen multiparametrische Methoden des maschinellen Lernens bereits zur Anwendung. Nun werden erstmals bioinformatische BigData-Methoden zur Mustererkennung eingesetzt.

In fachlicher Zusammenarbeit mit dem Universitätsklinikum Carl Gustav Carus Dresden und dem Verband der Universitätsklinika Deutschlands e. V. werden dem Anwender BigData-Konzepte erläutert, implementiert sowie anhand vieler praktischer Beispiele gezeigt.

Das Kompetenzprofil der BigData-Gruppe am Fraunhofer IWS umfasst folgende Facetten:

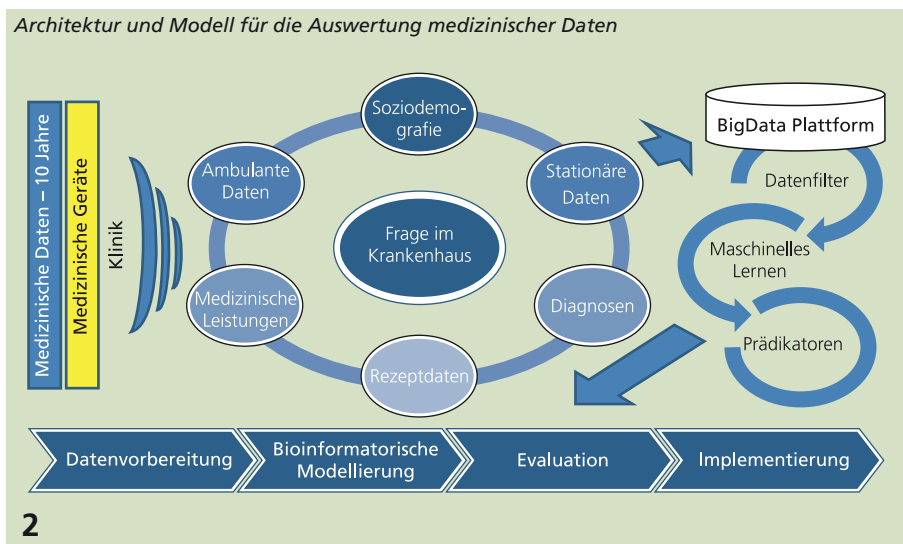
- Forschung an der BigData-Plattform und Anpassung an Anwendungen,
- Prozessoptimierung für Verwaltung und Speicherung in Datenbanken,
- Visual-Analytics
- multiparametrische und statistische Datenauswertung, Visualisierung, Management,
- automatische Echtzeit-Datenbearbeitung und -verwaltung
- Entwicklung von Bilddatenbanken und Technologien,
- Bildbearbeitung und 3D-Visualisierung,
- Entwicklung biomedizinischer Bildformate,
- Softwareentwicklung für Planung, Modellierung, Architektur, Analyse,

- Entwicklung von Technologie- und Datenstandards für die Verwaltung von digitalen Bildern und Metadaten,
- professionelle Softwareentwicklung und Beratung im Hinblick auf objektorientierte Programmiersprachen, professionelle Projektverwaltung sowie die Modellierung von IT- und Softwarearchitektur.

ERGEBNISSE

Der BigData-Einsatz in medizinischen Studien dient als Quelle für die Generierung neuer Hypothesen zu Erkrankungsrisiken und -verläufen bei Zivilisationskrankheit sowie für neue Hypothesen zu Ursachen, Zusammenhängen und Folgen der Krankheiten. Daraus resultiert die Entwicklung individualisierter Präventions- und Therapiemaßnahmen. Darüber hinaus kann BigData zur Ermittlung von Prädiktoren für die Wirksamkeit und Sicherheit von medikamentösen Therapien und Versorgungspfaden für Patienten eingesetzt und somit für eine bessere Versorgung genutzt werden.

Der neue Ansatz der Nutzung von bioinformatischen Methoden der Mustererkennung in medizinischen Routinedaten ist auf zahlreiche andere Krankheitsbilder und Fragestellungen der Versorgungsforschung übertragbar. Zur Erkennung neuer, bisher unbekannter Zusammenhänge zwischen den Zielvariablen der untersuchten Einflussgrößen (Expositions faktoren) werden anhand nichtparametrischer Verfahren und mittels



Methoden des maschinellen Lernens sehr viele unterschiedliche Ausgangsdaten zu Gruppen gebündelt und auf Ähnlichkeiten untersucht. Dieser Ansatz ist auf zahlreiche Beispiele in der industriellen Fertigung und Analytik übertragbar.

1 Cluster Server für BigData Plattform

KONTAKT

Prof. Karol Kozak
 +49 351 83391-3717
 karol.kozak@iws.fraunhofer.de

