Chapter 16
Overview of Big Data in Healthcare

Mohammad Hossein Fazel Zarandi
Amirkabir University of Technology (Tehran Polytechnic), Iran

Reyhaneh Gamasaee
Amirkabir University of Technology (Tehran Polytechnic), Iran

ABSTRACT

Big data is a new ubiquitous term for massive data sets having large, more varied and complex structure with the complexities and difficulties of storing, analyzing and visualizing for further processes or results. The use of Big Data in health is a new and exciting field. A wide range of use cases for Big Data and analytics in healthcare will benefit best practice development, outcomes analysis, prediction, and surveillance. Consequently, the aim of this chapter is to provide an overview of Big Data in Healthcare systems including two applications of Big Data analysis in healthcare. The first one is understanding disease outcomes through analyzing Big Data, and the second one is the application of Big Data in genetics, biological, and molecular fields. Moreover, characteristics and challenges of healthcare Big Data analysis as well as technologies and software used for Big Data analysis are reviewed.

INTRODUCTION

The last decade has experienced significant advances in the amount of data which is generally generated and stored in almost everyday activities, as well as the capability of utilizing technology to analyze and comprehend that data. The massive amount of data generated in healthcare systems is identified as Big Data and the ability to analyze that data is named Big Data analytics. Big Data supports businesses in various industries to become more productive and efficient. One of those industries in which Big Data is very useful is healthcare. Although health data is not always Big Data, there are some types of health data sets which are categorized as Big Data. Those data sets include massive data obtained from high-volume laboratory information system, electronic medical record (EMR), biomedical and biometrics data, test usage data, and gene expression data. In addition to profit boosting and decreasing squandered overhead, healthcare Big Data is being used to predict epidemics, cure disease, improve quality of life

and avoid preventable deaths. Both of the existing health data and the behavioral data could significantly increase opportunities to forecast long-term health conditions. Big Data also provides better diagnostics techniques, disease prevention, and enhance access and decrease healthcare costs.

The applications of Big Data are not limited to proposing more efficient healthcare systems. Big Data and analytics are applied in the healthcare industry showing very encouraging results. Integration and digitizing of Big Data as well as impressively using it, provide many advantages for physician offices, hospitals, and care organizations. Those advantages include disease diagnosis at its earlier stages providing the opportunity of easily treatment, controlling health conditions of individuals and groups, and detecting health care fraud promptly and effectively. Big Data and analytics support healthcare systems to achieve their goals. Two main categories for the "Big Data in Healthcare" have been reviewed in literature as follows.

The first category includes three important issues [IMIA]: (I) Big Data extracted from the health system such as health and medication history, lab reports, and pathology results, where these analyzes are aimed at improving physicians understanding of disease outcomes and their risk factors, decreasing health system costs, and enhancing its efficiency; (II) Massive data sets of biological and molecular fields are known as "Omics" data, genomics, proteomics, microbiomics, and metabolomics, where the goal of analyzing these data sets is to comprehend the mechanisms of diseases and expedite the medical treatments; (III) Data collected from social media along with the signs and behaviors of people who use Internet and software applications, for improving their health conditions (Hansen et al., 2014).

The second category includes five important issues according to [Big Data in healthcare]:

(I) Big Data which is used to manage healthcare delivery costs; (II) Big Data which is used for clinical decision making; (III) Big Data for extracting clinical information. (IV) Big Data for demographical analysis by investigating behavior and consumer category; (V) Big Data which is used as support information; this type of Big Data application is not categorized in one of the above categories (Groves et al., 2013; Hermon & Williams, 2014). Several researchers have studied those categories related to Big Data in healthcare.

As the first example, Archenaa and Anita (2015) reviewed the role of government in increasing quality of healthcare systems by reducing its costs which lies in the first category. This cost reduction is due to analyzing clinical Big Data of healthcare systems. For instance, diseases are detected at the earlier stages and the best medicine is prescribed by analyzing clinical Big Data and using genetic makeups. This leads to reduction of readmission rates, and as a result, it causes cost reduction for patients (Archenaa & Anita, 2015). Moreover, Hay et al. (2013) used statistical techniques such as Boosted Regression Tree method to predict the risk of disease outbreak in different geographical locations. That study is classified in the first category of Big Data analysis in healthcare.

In addition, Murphy et al. (2016) uses clinical Big Data to identify delays in follow-up of abnormal chest imaging results. That research is classified in the second category since its focus is on extracting clinical Big Data and clinical decision making. Another research which lies in the second category has been conducted by Bibault et al. (2016). In that study, machine learning techniques such as decision trees, naïve Bayes, support vector machines, neural networks, and deep learning which can be applied to aggregate clinical Big Data of radiation oncology were reviewed. Reviewing literature of both categories of Big Data analysis in healthcare shows that understanding disease outcomes, which lies in the first

Related Content

Multi-Criteria Decision Making: A Cast Light Upon the Usage in Military Decision Process

Tolga Temucin (2021). *Research Anthology on Decision Support Systems and Decision Management in Healthcare, Business, and Engineering (pp. 48-76).*

www.irma-international.org/chapter/multi-criteria-decision-making/282579

Developing a DSS for Allocating Gates to Flights at an International Airport

Vincent F. Yu, Katta G. Murty, Yat-wah Wan, Jerry Dannand Robin Lee (2009). *International Journal of Decision Support System Technology (pp. 46-68).*

www.irma-international.org/article/developing-dss-allocating-gates-flights/1744

Fuzzy Multi-Objective Linear Programming Problem Using DM's Perspective

Vishnu Pratap Singh, Madhura Deshmukhand Kirti Sharma (2022). *International Journal of Decision Support System Technology (pp. 1-17).*

www.irma-international.org/article/fuzzy-multi-objective-linear-programming-problem-using-dms-perspective/286695

Radio-Frequency Identification Adoption Trends in Retail: A Delphi Study

Mithu Bhattacharya, Gregory W. Ulfertsand Terry L. Howard (2020). *International Journal of Strategic Decision Sciences (pp. 49-64).*

www.irma-international.org/article/radio-frequency-identification-adoption-trends-in-retail/269689

Systematic Literature Review: XAI and Clinical Decision Support

Thomas M. Connolly, Mario Soflanoand Petros Papadopoulos (2023). *Diverse Perspectives and State-of-the-Art Approaches to the Utilization of Data-Driven Clinical Decision Support Systems (pp. 161-188).*

www.irma-international.org/chapter/systematic-literature-review/313785