

Investigating the dimensions, components, and key indicators of the use of big data in the health industry

Hamed Nozari¹, Javid Ghahremani-Nahr²

¹ Faculty of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran;

² Faculty Member of Academic Center for Education, Culture and Research (ACECR), Tabriz, Iran

ARTICLE INFO

Received: 10 April 2022

Reviewed: 12 May 2022

Revised: 18 May 2022

Accepted: 23 May 2022

Keywords:

Big Data Analysis, Health Industry, Data Production, Intelligent Technologies in the Medical Industry.

ABSTRACT

Using big data analytics in healthcare has positive as well as life-saving results. Big data refers to the vast amounts of information generated by the digitization of everything that is synthesized and analyzed by specific technologies. Here Big Data uses health services to use specific health data of a population (or a specific individual) and potentially help prevent disease pandemics, treat diseases, reduce costs, and more. In the field of health, big data covers a wide range of information, including physiological, behavioral, molecular, clinical, medical imaging, disease management, medication history, nutrition, or exercise parameters. Big Data Analysis In the field of health, it is a complex process of examining big data to discover information. This information includes hidden patterns, market trends, unknown correlations, and customer preferences. Information analysis can help organizations make informed business and clinical decisions. The medical data-driven industry is the most complex among industries. Not only is this data available from a variety of sources, but it must also comply with government regulations. This process is difficult and delicate and requires some level of security and communication. Due to the importance of this issue, in this article, after introducing the types of data available in the health industry, the characteristics and sources of big data in health are defined and an analytical model for the use of large data in the health industry is presented. This model helps to understand the dimensions, components, and key elements of using big data in the health industry.

¹ Corresponding Author: javid.ghahremani@yahoo.com
<https://doi.org/10.52547/jda.1.1.1>

1. Introduction

Collecting large amounts of data for medical purposes has been costly and time-consuming for many years. With today's ever-improving technologies, it becomes easier not only to collect such data but also to turn it into relevant insights that can be used to provide better care. Using big data analytics in healthcare has positive as well as life-saving results. Big data refers to the vast amounts of information generated by the digitization of everything that is synthesized and analyzed by specific technologies (Aliahmadi et al., 2013). Here Big Data uses health services to use specific health data of a population (or a specific individual) and potentially help prevent disease pandemics, treat diseases, reduce costs, and more. Therapeutic models have now changed, and many of these changes are driven by data. Doctors want to know about a patient's life as soon as possible so that they can quickly detect the warning signs of a serious illness when it occurs. It is much easier and cheaper to treat any disease in the early stages (Au-Yong-Oliveira et al., 2021). By analyzing health care data, prevention is better than cure, and management allows insurance to provide a more appropriate cost package to paint a comprehensive picture of a patient (Dash et al., 2019).

Health data is divided into three categories: structured, semi-structured, and unstructured data. Structured data include clinical data; the focus of this data is on genotype, phenotype, genomic, and ICD information, and it is easily collected, stored, and processed by a specific device. These data make up only 5 to 10% of the health system data. Most data is unstructured or semi-structured. Emails, photos, videos, audio files, clinical notes, prescriptions, medical imaging, EHRs, lifestyle data, and environmental and health economics are all part of this type of information. The challenge of big data analysis is to address these heterogeneous data in order to gain insight into improving health care outcomes (Ghayvat et al., 2021).

Most data is unstructured or semi-structured. Emails, photos, videos, audio files, clinical notes, prescriptions, medical imaging, EHRs, lifestyle data, environmental and health economics are all part of this type of information. The challenge of big data analysis is to address these heterogeneous data in order to gain insight into improving health care outcomes. From predicting treatment outcomes to curing cancer and effective patient care, data science has proven to be a very important part of the industry's future. There are three main factors that drive health innovation (Lotfi et al., 2016):

- Advances in technology
- The growth of digital consumerism
- Need to fight rising costs



Fig. 1. Big data sources in healthcare

Data science provides tools and methods for extracting real value from unstructured patient information (Miah et al., 2022). This knowledge will ultimately contribute to greater efficiency, accessibility and personalization of health care. With the opportunities created by the digital and information revolution, the healthcare industry can take advantage of the potential of big data. Big data analysis provides increasing value for the future of health by improving the quality of care and outcomes, and the provision of cost-effective care services. The predictive nature and diagnostic aspect of the Big Data Analysis model in health allow for a shift from experience-based medicine to evidence-based medicine.

2. How does Big Data help the healthcare industry grow and develop?

The advanced analysis provided by Big Data provides tremendous opportunities for most stakeholders in the healthcare industry (patient, provider and payer). Some of these opportunities are as follows:

- **Improving the quality of care:** By changing the model of health services from cost-based care model to value-based models, data analysis becomes the main function of daily operations. Data analysis in this method can improve the performance and quality of care, reduce patient waiting time and increase their satisfaction, reduce readmission rate, reduce human error, provide personalized therapies and improve the overall patient experience (Mohammadi et al., 2015).
- **Population Health Management:** Population health management has shifted health care services by focusing more on anticipation and prevention in public health rather than response and treatment. With predictive analyzes, health systems can identify patients at high risk for chronic disease in the early stages of disease progression. This provides an opportunity to avoid

long-term health problems that lead to costly care and hospitalization. Analyzes of possible epidemics allow for prediction and intervention (Nahr et al., 2021).

- **Early diagnosis of diseases:** Tracing healthy behaviors and monitoring the disease can help in early diagnosis of diseases. Big data analysis can be effective in preventing a wide range of deadly diseases or epidemics and in managing and monitoring personalized treatment (Nozari et al., 2021).
- **Quality, structure and availability of data:** Big data provides the ability to quickly record data and convert raw and unstructured data into meaningful information. Unwanted data can be deleted and effective data can be reused. The quality of data in this way is optimally enhanced. Open source technology increases data accessibility and transparency (Pablo et al., 2022).
- **Improving decision making:** Using Big Data, it is possible to use evidence-based medicine and health care providers can make more informed decisions. Remote monitoring, patient profile analysis and genomic analysis also influence the decision-making process. This feature makes the decision-making process faster, simpler and more accurate (Parimanam et al., 2022).
- **Cost reduction:** Big Data helps reduce medical costs by anticipating and diagnosing the disease early, making smart decisions, implementing cost-effective treatments, monitoring adherence to treatment, and reducing transportation costs. Also, proper and intelligent management of the steps and process of service delivery and optimization of staff allocation and budget, will save costs.
- **Patient-centered care:** The increasing use of technology has changed the orientation of health care from patient-centered to patient-centered care. Big data allows information to be provided directly to the patient and the patient to play an active role in self-care. Increasing communication between patients and service providers helps patients make informed decisions.
- **Personalized medical upgrades:** Access and processing of large volumes of data enables the provision of personalized services to the patient.
- **Globalization:** With the globalization of data, access to more information and service providers can share their information and knowledge and be informed and benefit from new knowledge and methods (Park et al., 2022).
- **Creating Advantages for Health Payers and Detecting Fraud:** Health insurance companies are subject to changing laws. By collecting and interpreting data through analytical solutions, payers can identify and target individuals by analyzing profiles and demographics. Evaluate hospital claims and identify the possibility of fraud using predictive analysis. Use pricing data to identify service providers with the highest value and lowest cost.
- **Diagnose Health Threatening Problems:** This Big Data feature has the opportunity to enhance capabilities for rapid and accurate threat detection, and is especially useful for government use.

3. What challenges does the health system face in using Big Data?

Despite the many potential benefits of big data analytics, the health industry is in the early stages of using this technology. Despite the large amount of data available, lack of knowledge, lack of infrastructure, and the need for huge initial investment make using Big Data more difficult. Challenges in the use of big data in the health industry can be categorized into 9 areas (Senthilkumar et al., 2018):

- **Data structure:** Health data is much more heterogeneous than data from other domains. Problems in this area include scattered data, inconsistent structures, heterogeneous data, raw and unstructured databases, large volume, high variability and speed, and lack of transparency.
- **Data Security:** There are many privacy concerns about using big data analytics. Open source data is freely available and therefore very vulnerable. Health data also faces the issue of confidentiality. On the other hand, the centrality of this data makes them vulnerable to attacks.

- **Data standardization:** With the globalization of data, big data must cope with a variety of standards, language barriers and different terms. Enterprise platforms are very different, and data storage formats are not compatible with all applications and technologies (Shilo et al., 2020).
- **Storage and transmission:** Data production is inexpensive compared to storage and transmission. Cloud-based health information technology adds an extra layer of security for extracting, converting and uploading patient data.
- **Management problems:** Governance issues and ownership issues are other problems in the field of health big data. Data governance should be on the list of top priorities of organizations and should be treated as a core asset rather than a by-product. Data ownership and data monitoring are creating new roles in business (Shyamala Susan et al., 2022).
- **Lack of skills:** Due to lack of knowledge about the best algorithms and tools for analysis and the unavailability of trained clinical scientists and Big Data managers to interpret Big Data results, health care is still far from realizing the potential of Big Data analysis.
- **Inaccuracies:** In healthcare, self-report data is widely used, which raises issues such as incompatibility and inaccuracy.
- **Compliance with Legal Regulations and Concerns:** Organizations that use big data analytics as part of their information systems must meet significant standards and standards of compliance with specific health care regulations.
- **Real-time analysis:** One of the key needs in healthcare is the ability to use big data in real time.

Personal analytics devices, formerly known as business intelligence buzzwords, are constantly gathering the ability to become part of a patient health delivery strategy and sending this information to the cloud. In addition, this information is available to the Public Health Status Database, which allows physicians to compare this data in socioeconomic contexts and, accordingly, modify treatment guidelines, health care institutions, and care managers. Sophisticated tools to monitor this flow use huge amounts of data and react whenever the result is alarming. For example, if the patient's blood pressure is alarming, the system warns the patient at the right time, after which it reaches the patient and takes the necessary steps to reduce the pressure. Another example is Asthmapolis, which is starting to use radiators with active GPS trackers to detect asthma at the individual level, using this data to create better treatment plans for asthma (Sreedevi et al., 2022).

4. Big data values in healthcare

One of the most interesting achievements of big medical data processing is the right lifestyle. Imagine that our medical information is analyzed and made available to various software developers with privacy. In this case, these companies will be able to produce software that monitors the health status on a daily basis and send us the necessary suggestions or warnings during daily activities. Patients should receive the most timely and accurate treatment. With the correct processing of the data, no doctor is left unaware of the decisions of his colleagues, and according to the available records, which have been categorized in detail, the necessary treatment will be provided as soon as possible (Sujitha & Seenivasagam, 2021). The framework presented in **Figure (2)** shows the most important parameters influencing the analysis of big health data.



Fig. 2. Effective parameters in big health data analysis

Judgment time ends based on job name and title. At the time of illness, the records of all doctors and medical centers are available, and the patient can put his life in the hands of a specialist based on real and verified information and not on the basis of advertisements. Also, there may be medical activities that nurses can do without the need for medical attention at all. Under these conditions, medical service providers and insurance companies can operate entirely on the basis of real values. For example, an insurance company considers the condition of repaying the amount to the healthcare organization to be the complete treatment of the patient, which can only be done by examining the patient's condition moment by moment (Surantha et al., 2021).

Large stakeholders in the medical industry need a lot of creativity and innovation to survive. They will have to place their R&D department at the top of the organization and imagine modern medicine.

5. Conclusion

In an age where even our smallest behaviors - like social media or a mouse on a monitor screen - are considered valuable and analytical data, falling behind the global caravan is an irreparable mistake. In fact, collecting large amounts of data for medical purposes has been costly and consuming-time for many years. With today's ever-improving technologies, it becomes easier not only to collect such data, but also to turn it into relevant insights that can be used to provide better care. The purpose of health care data analysis is to use data-driven findings to predict and solve problems before it is too late. This tool allows the treatment methods and methods to be evaluated faster and better monitoring of individual health can be done. Due to the increasing growth of data production sources and the importance of using this data in the medical industry, in this study, the basic dimensions and components of a treatment system based on big data analysis were examined and analyzed. Also showed the main factors in this treatment system. Using this framework can help to better understand the effective factors and the high importance of this field.

Conflicts of Interest

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

- Aliahmadi, A., Jafari-Eskandari, M., Mozafari, M., & Nozari, H. (2013). Comparing artificial neural networks and regression methods for predicting crude oil exports. *International Journal of Information, Business and Management*, 5(2), 40-58.
- Au-Yong-Oliveira, M., Pesqueira, A., Sousa, M. J., Dal Mas, F., & Soliman, M. (2021). The potential of big data research in healthcare for medical doctors' learning. *Journal of Medical Systems*, 45(1), 1-14.
- Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, 6(1), 1-25.
- Ghayvat, H., Pandya, S. N., Bhattacharya, P., Zuhair, M., Rashid, M., Hakak, S., & Dev, K. (2021). CP-BDHCA: Blockchain-based Confidentiality-Privacy preserving Big Data scheme for healthcare clouds and applications. *IEEE Journal of Biomedical and Health Informatics*.
- Lotfi, F. H. Z., Najafi, S. E., & Nozari, H. (Eds.). (2016). *Data Envelopment Analysis and Effective Performance Assessment*. IGI Global.
- Miah, S. J., Camilleri, E., & Vu, H. Q. (2022). Big Data in healthcare research: a survey study. *Journal of Computer Information Systems*, 62(3), 480-492.
- Mohammadi, H., Ghazanfari, M., Nozari, H., & Shafiezzad, O. (2015). Combining the theory of constraints with system dynamics: A general model (case study of the subsidized milk industry). *International Journal of Management Science and Engineering Management*, 10(2), 102-108.
- Nahr, J. G., Nozari, H., & Sadeghi, M. E. (2021). Green supply chain based on artificial intelligence of things (AIoT). *International Journal of Innovation in Management, Economics and Social Sciences*, 1(2), 56-63.
- Nozari, H., Fallah, M., Kazemipoor, H., & Najafi, S. E. (2021). Big data analysis of IoT-based supply chain management considering FMCG industries. *Бизнес-информатика*, 15(1 (eng)).
- Pablo, R. G. J., Roberto, D. P., Victor, S. U., Isabel, G. R., Paul, C., & Elizabeth, O. R. (2022). Big data in the healthcare system: A synergy with artificial intelligence and blockchain technology. *Journal of Integrative Bioinformatics*, 19(1).
- Parimanam, K., Lakshmanan, L., & Palaniswamy, T. (2022). Hybrid optimization based learning technique for multi-disease analytics from healthcare big data using optimal pre-processing, clustering and classifier. *Concurrency and Computation: Practice and Experience*, e6986.
- Park, Y. T., Kim, D., Koh, S. J., Kim, Y. S., & Kim, S. M. (2022). Patient Factors Associated with Different Hospice Programs in Korea: Analyzing Healthcare Big Data. *International Journal of Environmental Research and Public Health*, 19(3), 1566.
- Senthilkumar, S. A., Rai, B. K., Meshram, A. A., Gunasekaran, A., & Chandrakumarmangalam, S. (2018). Big data in healthcare management: a review of literature. *American Journal of Theoretical and Applied Business*, 4(2), 57-69.
- Shilo, S., Rossman, H., & Segal, E. (2020). Axes of a revolution: challenges and promises of big data in healthcare. *Nature medicine*, 26(1), 29-38.
- Shyamala Susan, V., Juliana Gnana Selvi, K., & Bambang Sugiyono Agus Purwono, I. (2022). Big Data in Healthcare: Applications and Challenges. *Advanced Analytics and Deep Learning Models*, 351-363.
- Sreedevi, A. G., Harshitha, T. N., Sugumaran, V., & Shankar, P. (2022). Application of cognitive computing in healthcare, cybersecurity, big data and IoT: A literature review. *Information Processing & Management*, 59(2), 102888.
- Sujitha, R., & Seenivasagam, V. (2021). Classification of lung cancer stages with machine learning over big data healthcare framework. *Journal of Ambient Intelligence and Humanized Computing*, 12(5), 5639-5649.
- Surantha, N., Lesmana, T. F., & Isa, S. M. (2021). Sleep stage classification using extreme learning machine and particle swarm optimization for healthcare big data. *Journal of Big Data*, 8(1), 1-17.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).