Big Data Analytics in Healthcare:
Applications and Challenges

Jaimin Navinchandra Undavia, Smt. Chandaben Mohanbhai Patel Institute of Computer Applications, India

https://orcid.org/0000-0002-6854-3904

Atul Manubhai Patel, Smt. Chandaben Mohanbhai Patel Institute of Computer Applications, India

ABSTRACT

The technological advancement has also opened up various ways to collect data through automatic mechanisms. One such mechanism collects a huge amount of data without any further maintenance or human interventions. The health industry sector has been confronted by the need to manage the big data being produced by various sources, which are well known for producing high volumes of heterogeneous data. High level of sophistication has been incorporated in almost all the industry, and healthcare is one of them. The article shows that the existence of huge amount of data in healthcare industry and the data generated in healthcare industry is neither homogeneous nor a simple type of data. Then the various sources and objectives of data are also highlighted and discussed. As data come from various sources, they must be versatile in nature in all aspects. So, rightly and meaningfully, big data analytics has penetrated the healthcare industry and its impact is also highlighted.

KEYWORDS

Big Data Analytics, Big Data Applications in Healthcare, Healthcare Challenges, Hidden Patterns, Predictive Analysis

INTRODUCTION

Huge amount, high speed, versatility converts traditional database into Big Data. These characteristics make big data more difficult and challenging for compilation. The techniques used to compile traditional data; it cannot be used to compile such big data as it is not always collection of structured data (Undavia 2018). Such informative mass of data prove crucial for purpose of analytics.

This is because of availability of sophisticated data storage, which stores this huge amount of data. You cannot imagine a world without such data store where versatile information like details of person, organization, transaction performed, etc. are stored. This data can be used to extract valuable information and knowledge for the growth of the particular organization or improved activity. In the recent time, all aspects of information is available related to a customer. It ranges from customer name, its details about the purchase, his social connection, professional association, etc., This proves that, in current time data is the founding stone and vital element of any organization (Elragal 2014).

The way Big Data is defined is changed against its aspect of application. In the recent time, it is characterized as a collection of data elements whose size, speed, type and complexity is changed frequently (Belle 2015). These frequent changes makes one to seek, adopt and invent new hardware and software mechanisms to manage, store, visualize and analyse these form of data.

The three Vs are Velocity, variety and volume, playing very important role in all various applications of such data. Healthcare is one of the prime user of this Big Data technology. In health care, data is widely spread among multiple health care systems, health insurers, researchers, government entities etc (Manyika 2011). The data repositories are not always sophisticated enough which can provide platform of transparency for global data. As these kind of data have huge size, versatility in variety and rapid changing in nature, new type of data analytics, well-structured storage and accurate analysis methods are required. This huge amount of data then can be analysed properly and desired information can extracted from them.

Application of the Big Data technology in healthcare has large number of positive outcomes and such outcomes have life-saving phenomenon. Like other fields, in healthcare also, vast quantities of data and information is created/generated through digitization, automatic sensor data, etc. and then these mass have to consolidated and analysed by some specific technologies (Rehermann n.d.).

The technological advancements have also inferred the way healthcare systems work. Treatment models, data capturing of population or a person, deciding the treatment once a disease is diagnosed, the model of diagnosis, etc. have changed drastically with the advent use of the recent technologies. This all changes are taking place based on the data collected of the same domain so that proper corrective and improved steps can incorporated in recent and newer technologies. Most of these changes are data driven so the healthcare domain also become conscious about the data collection and storage. For the physician who are treating their patients, they always

eager to know as much as possible about his patient. This early stage of detection will help the doctors to decide mode of the treatment especially in case of some serious illness or disease. It is already proved that curing of such critical illnesses in early stage is more effective and less expensive as well.

When healthcare industry attempt to use the big data comprehensively, micro level analysis can be done for patient as huge amount of versatile data will be made available about the patient's history, way of treatment of other patients of same decease, other related parameters to cure this decease.

A tailored package of the data can be provided to doctors for better life-saving outcomes.

In past not only in healthcare, but in all industry gathering, storing and maintaining of data was very costly and time consuming too. Now in current time, there is availability of improved technology, it becomes very easy and providing very critical insight of data which can offer better understanding and usage of stored data.

The core purpose of healthcare data analytics is listed as under:

- Collect, maintain and store huge amount data;
- Collect critical insights from that stored data;
- Explore and evaluate ways of better care through historical data;
- Use data-driving findings to solve some critical problems;
- Assess methods and offer faster treatment;
- Maintain better track of inventory.

So, the technology Big Data Analytics in healthcare is the recent application of big data technology which is emerged as an enhancement in the field of healthcare. The data available of healthcare domain is in extraordinarily high volume, moves at high velocity and highly variable in nature. These 3 Vs are the founding stone of Big Data in Healthcare (Catalyst 2018).

BACKGROUND

The main objective of the article is to show case various applications of Big Data analytics especially in the field of healthcare. The article also highlights the insights of the challenges and applications of the Big Data analytics in the domain of Healthcare.

Big Data and Healthcare

As described in the introduction section, Big Data technology has impact in almost all types of industries. The healthcare is also one of them. When Big Data in healthcare is focused, basically it focuses the data from hospitals, pharmacist, druggist, pathologist, automatic data generated through sensors and web based applications which are responsible to manage healthcare data (Raghupathi 2014). Describing Big data technologies used in healthcare refers to healthcare data that is available online or electronically in large amounts such as terabytes and petabytes, which makes it

almost impossible to maintain the data using traditional hardware or software. The main challenge for big data in healthcare is the volume of data as well as the diversity of data types along with the accuracy and speed at which it needs to be managed (Anantharam n.d.).

Objective of penetration of the big data analytics in the area of healthcare varies across the aspect and nature of the application. Healthcare is substantially become costly over the year and so as the cost can be one of the prime objective of the research (Rehermann n.d.).

The healthcare is system, which is multidimensional in nature, and the primary objective of the healthcare system is to prevent, diagnosed and find best suitable treatment for health-related issues (Dash 2019). The major components are responsible to generate and provide data for this health related repository, which can be used for further use of data analytics. Mainly these components are health facilities like hospitals, medical stores, pharmacists, treatment technologies, doctors, etc. The health professionals belong to various health sectors like dentistry, medicine, midwifery, nursing, psychology, physiotherapy, and many others. Healthcare is required at several levels depending on the urgency of situation (Dash 2019).

Traditional health analysis was carried out before the advent of computers and other recent technologies. Before the big data technology, the traditional health analysis was carried out where healthcare industry had to be dependent on another industry for the big data analysis purpose. The stack holders of healthcare industries now become more inclined towards the results generated through such big data analysis. Like other industries, healthcare is also facing the challenge to handle huge amount of data, which is versatile and dynamic in nature, and it is increasing at rapid speed. In this situation, big data analytics is proved to be one of the most prominent solution to handle this dynamicity of data with power to find some interesting patterns within it. Data stored in hard copies of healthcare industries are required to be digitized for the defined purpose of analytics (Kumar 2019).

Figure 1 explains more insights in the field of big data analytics in the domain of healthcare.

LITERATURE REVIEW

Many researchers have indulge in the field of Big data analytics and later many of them have explore their research in the area of healthcare:

● Ashwin Belle and many have published their article titled "Big Data Analytics in Healthcare" and they have concluded that the data in healthcare is structured, semi-structured and unstructured. The data source which provides such collection of variety data is going to play vital role in future for healthcare practices in future. They have focused especially medical image analysis, physiological signal processing and integration of physiological and genomic data (Belle 2015);

**Figure 1. Insight of Big Data in Healthcare**
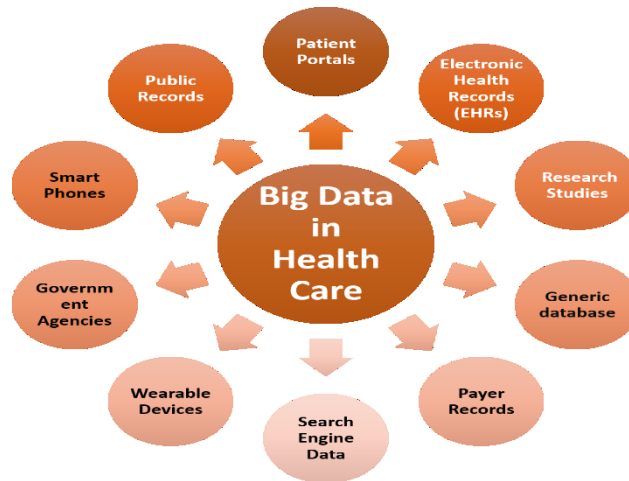
- Nada Elgendy and Ahmed Elragal have claimed in their article named "Big Data Analytics: A Literature Review Paper" that effective data source can help for the betterment in predictive analysis and other analytical purposes (Elragal 2014);
- In McKinsey Global Institute, James Manyika and others have highlighted importance of big data in their article titled "Big Data: The Next Frontier for innovation, competition and productivity. The team of authors have narrated that how big data creates values in several ways. Authors have also shown the way to enable experimentation to discover needs, expose variability and improve performance. Segmenting populations to customize actions were also narrated in the article (Manyika 2011);
- A book chapter titled "Security Issues and Challenges Related to Big Data" in the book series titled " Big Data Management and the Internet of Things for Improved Health Systems" narrated security issues while implementing big data in healthcare systems Undavia 2018);
- AISHWARYA ANANTHARAM wrote in the article titled "EXPLORE THE APPLICATIONS OF BIG DATA TECHNOLOGIES IN HEALTHCARE" that Big data analytics in healthcare is evolving by improving the cost reductions involved in the projects. The author to represent the real-time models that are being utilized in industries to provide patients with better treatment and keep track of these humongous amounts of data in an orderly manner conducted a case study (Anantharam n.d.);
- Journal research paper titled "Improving Healthcare Using Big Data Analytics" authored by Revanth Sonnati contains details of possible improvements in the field healthcare. The paper claimed that geographic location is also important to analyse healthcare data. Moving towards the big data analytics makes the system more reciprocal, effective and efficient when it comes in the healthcare domain (Sonnati 2017);

- A recent research article by Abhilash Shukla and Atul Patel explained the image segementation technique to detect bone cancer from the images of MRI and X-ray. They have also narrated Image segmentation technique like sobel, prewitt, canny, K-means and Region Growing are described in this paper which can be stimulated for X-Ray and MRI image interpretation (Shukla 2020).

These are the brief of the survey papers referred by us to find use case of the big data analytics in healthcare industry. Figure 2 also depict the same to narrate the use cases.

Figure 2. Use cases of Big Data Analytics in Healthcare
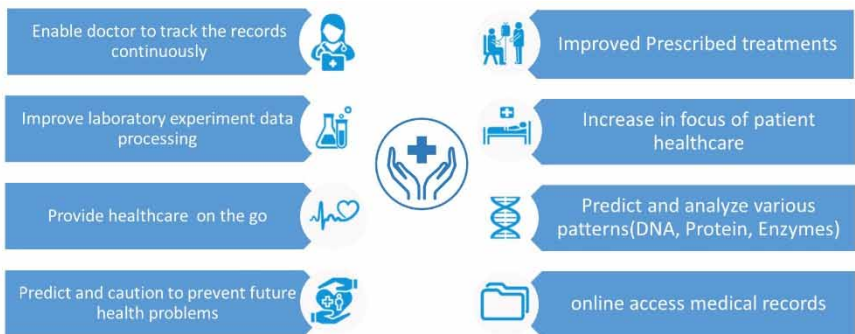
Use Case of Big Data Analytics in Healthcare

Applications of big data analytics can improve the patient-based service, to detect spreading diseases earlier, generate new insights into disease mechanisms, monitor the quality of the medical and healthcare institutions as well as provide better treatment methods.

As discussed in previous sections, big data has use cases in almost all the industry and it has very crucial use cases in healthcare domain too.

There are many computing solutions available, which are feasible through big data technology in healthcare domain. The main objective of this combination of healthcare and big data analytics is to find new insights of this big data analytics, expand research horizon in the field of healthcare, offer measurable benefits in the outcome of healthcare research, offer better and improved healthcare facilities to the society, etc. As a concluding point of this, the abstract aim of this combination is to benefit the society with the advent use of advanced technology to improve and enhance healthcare (Sonnati 2017).

There is a wide spectrum of Big Data applications in healthcare domain. Currently big data analytics is used to predict the possible effect of the given treatment against the historical data of the patient, to find best suitable medication, decide treatment plan of the patient, etc. However, to manage this complex data of healthcare domain is again a challenge itself (Kumar 2019; Shukla 2020). In this section, general methods

have been discussed which has impact of big data analytics in the healthcare domain (Anantharam n.d.; Kumar 2019):

- Medical Signal Analytics:
  - To keep track of patient's health over a period, reaction against the various medicine, etc., are belong to Telemetry and Physiological signal monitoring system;
  - Such automated data generation systems produce huge amount of data of wide variety. These heterogeneous data requires high analytical end to analyse it for its meaningful use (Dash 2019);
- Prediction in Healthcare:
  - Regardless to the nature of the business, analytics become an integral part of all type of business to predict the future trends of the business. This predictive analysis in the healthcare industry is very important use case of big data analytics;
  - As discussed in preceding sections, healthcare industry too generating wide range of data, which is maintained by computer systems. These big data can be subject to various analytics techniques like text mining, video mining, data mining, clustering of data, association rule mining to find out interesting insight of data, which is useful for future;
  - In recent times, machine learning and deep learning techniques can also be applied over these data. Current trend, future trend and past trend can be compared and analysed with the help of these techniques;
- Machine Learning/Deep Learning in Healthcare:
  - In recent times the techniques used in data mining is now inclined with machine learning techniques. These techniques are again responsible to discover hidden patterns from the huge amount of data;
  - In machine learning, the machine alters the execution of the program by learning from the hidden patterns within the data (Herland 2014);
- EHR (Electronic Health Records):
  - Electronic Health Records means the health related data, which is stored and maintained electronically for future purpose;
  - Such sophisticated record keeping helps the healthcare related resources to be more accurate and predictive. Such records can be shared in public and private sector for its effective use;
  - Once the data is available, there are versatile ways to use it against more sophisticated applications of Big Data analytics (Mike 2014);
- Image Processing of Healthcare Data:
  - Image processing is an integral technology used in medical science where way to treatment, nature of the treatment, intensity of the decease, area of infected area, etc are to be determined and decided and disease states is considered;

- ◦ Under the image processing, mostly used techniques is image segmentation. Such image segmentation techniques are used to detect bone cancer through the image of MRI and X- ray (Shukla 2020);
  - ◦ Image enhancement is also used to get details of a particular disease and decide the mode and plan of treatment. Such healthcare data are kept as electronic health record (EHR). Such data can be kept for future use in order to improve overall healthcare system;
- Big Data Genomics:
  - ◦ AS human genes are great in number, so it is not possible to treat or maintain them through traditional database system;
  - ◦ So big data is a very good option to handle such huge amount of data, which is ultimately from the healthcare industry. Therefore, GeneBank, Gene Structure prediction is part of Big Data Genomics, which has a good power and potential in terms of healthcare applications. Such gene analysis is required sophisticated big data analytics techniques.

CONCLUSION

In this article, we have highlighted application areas of big data analytics in healthcare industry. Like other industries, healthcare is also producing huge amount of data through either manual process or automatic data capturing machines. Such huge amount of data is considered as big data as it possesses all the characteristics of it. Over this data, big data analytics techniques are used to get very useful insight of the data. Such hidden patterns and other outcomes are useful for stockholders of this healthcare industry. It is also highlighted that for a specific use, customized architecture can be developed for finite set out outcomes. Further, healthcare is directly in the human life and its survival, such important data insights are used for overall social welfare.

REFERENCES

Anantharam. (n.d.). *Explore the applications of big data technologies in healthcare*. Academic Press.

Belle, . (2015). Big Data Analytics in Healthcare. *Journal of Biomedicine & Biotechnology*, 2–37.

Dash, . (2019). Big data in healthcare: Management, analysis and future prospects. *Journal of Big Data*, 1–25.

Elragal, N. E. A. (2014). Big Data Analytics: A Literature Review Paper. *Springer International Publishing Switzerland*, *2014*, 214–227.

Healthcare Big Data and the Promise of Value-Based Care. (2018). Available: https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0290

Herland, T. M. K. R. M., Khoshgoftaar, T. M., & Wald, R. (2014). A review of data mining using big data in health informatics. *Journal of Big Data*, *1*(1), 2. doi:10.1186/2196-1115-1-2

Jaimin & Undavia. (2018). Security Issues and Challenges Related to Big Data. In *Big Data Management and the Internet of Things for Improved Health Systems* (pp. 86–101). IGI Global. doi:10.4018/978-1-5225-5222-2.ch006

Kumar. (2019). Big Data Analytics for Healthcare Industry: Impact, Applications, and Tools. Big Data Mining and Analytics, 2(1).

Manyika. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. Academic Press.

Mehmood, G. G. R., & Graham, G. (2015). Big data logistics: A healthcare transport capacity sharing model. *Procedia Computer Science*, *64*, 1107–1114. doi:10.1016/j.procs.2015.08.566

Mike. (2014). Available: http://ihealthtran.com/iHT2 BigData 2013.pdf

Raghupathi, W. V., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(1), 3. doi:10.1186/2047-2501-2-3 PMID:25825667

Rehermann. (n.d.). *Datapine*. Available: https://www.datapine.com/blog/big-data-examples-in-healthcare/

Shukla, . (2020). Bone Cancer Detection from X-Ray and MRI Images through Image Segmentation Techniques. *International Journal of Recent Technology and Engineering*, *8*(6), 273–278.

Sonnati, R. (2017). Improving Healthcare Using Big Data Analytics. International Journal of Scientific & Technology Research, 6(3).