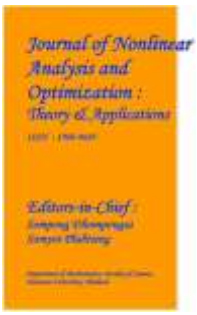
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 2 : 2023
ISSN : **1906-9685**



A STUDY OF CLOUD-BASED SOLUTIONS FOR DATA ANALYTICS IN THE HEALTHCARE BUSINESS

[#1]Ms. G. LAKSHMI, ***Assistant Professor,***

[#2]PALLA SWATHI, ***Assistant Professor,***

[#3]JITTAVENI SRINIVAS, ***Assistant Professor,***

Department of Computer Science & Engineering,

TRINITY COLLEGE OF ENGINEERING AND TECHNOLOGY, PEDDAPALLY

ABSTRACT: In the healthcare industry, a large amount of data is generated, which is critical for improving patient outcomes and scientific understanding. Traditional on-premises data processing and storage choices may be overwhelmed by the volume, variety, and speed of healthcare data. This study investigates the benefits and drawbacks of adopting cloud computing to store medical data. This article discusses the most recent advancements in AI and cloud computing. They are intended to inspire innovative thinking, improve patient care, and provide more accurate medical support, among other things. In this article, we will look at how Hadoop and Cloud Computing are employed in the healthcare profession for big data analytics in locations like the intensive care unit. There are numerous uses, techniques, and predictions for big data analytics that are discussed. Several cloud-based technologies, such as MMAP, can handle the aforementioned public health challenges rapidly, reliably, inexpensively, effectively, efficiently, and patient-centeredly. These tools can also be used to forecast how certain diseases would impact health on a national or regional scale. It may be simpler to transfer data with cloud computing, Hadoop, and artificial intelligence. As a result, healthcare management may perform the necessary calculations to identify the likely, observable, and verifiable patterns that influence catastrophe preparedness and program improvements.

Keywords: *Big Data; Healthcare; Cloud Computing; Data Analytics; Hadoop.*

1. INTRODUCTION

"Big data" technology and tools enable users to handle and examine massive amounts of data that are too huge for typical processing and analytical methods. There are numerous companies that provide a diverse range of healthcare datasets (HCD). Figure 1 depicts how learning about patient care from many sources improves our knowledge.

The majority of the government's health statistics come from the general public. Insect-borne diseases such as malaria and dengue fever are closely monitored by both the federal and state governments. They also monitor immunization rates and the findings of the COVID-19 trial. Furthermore, they collect health information for legal and regulatory purposes, such as creating prescription drug databases. Academics can make use of government-maintained health data sets. There is a wealth of information in the field of public health care that has resulted from scientific and biological breakthroughs.

Both governmental and private research organizations collect massive volumes of data that can be utilized to understand more about how diseases spread.

Clinical data repositories, often known as general health databases, collect information from a variety of healthcare-related sources. These archives are mostly concerned with the diagnostic and treatment procedures employed by various medical specialists. These platforms are simple to use and contain a wealth of information that may be utilized to examine patient medical records using various statistical approaches and visualizations. EHRs, or electronic health records, are the new way for healthcare facilities such as hospitals and clinics to maintain a patient's medical information. RTMRs use identity access management (IAM) to ensure that healthcare personnel always have quick and secure access to patient information.

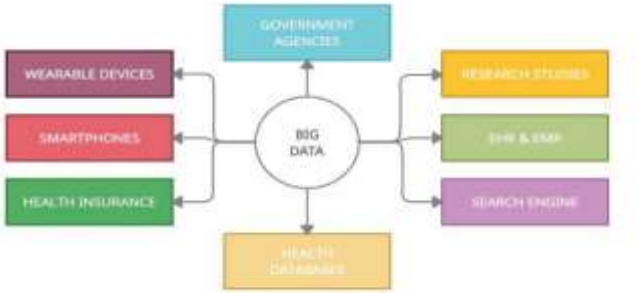


Fig. 1. There are numerous sources of health information.

Regular medical appointments generate computerized health records. This is where x-ray images, prescription pill notes, diagnoses, recommended treatments, and immunization and vaccination regimens are recorded. The internet has evolved into the best resource for learning about a population's health and well-being.

In the healthcare industry, the healthcare data lake can be utilized to store data from social networking sites. This data can be used for analytics and prediction. This data can be used to determine how popular a particular prescription or treatment is in a given area, as most individuals see the doctor because they are experiencing comparable symptoms. As if it were a worldwide epidemic. Wearable electronics, which collect data through sensors, are becoming increasingly common. These devices include monitors that measure temperature, heart rate, blood pressure, and the quantity of oxygen in saline solution (SPO2). This information can be provided to an IHI system to provide a more complete picture of a patient's health.

Health-related apps (HAD) on smartphones can save information from doctor visits and easy diagnosis. These shows can teach people a lot about how to eat, sleep, and exercise, and they may also have a significant impact on a patient's journey to recovery. We could predict potentially deadly health conditions by combining this knowledge with what we know from real-time diagnosis technologies. Similarly, health insurance information from companies that provide in-patient health insurance benefits to patients can be a gold mine of information regarding the procedures performed on the patient, the drugs administered, and the precautions and prescriptions to be followed following surgery. This information can be used in a variety of ways, including determining a patient's complete medical history, assisting with current care, and creating a credible database for research studies and clinical trials.

The three characteristics that distinguish this healthcare source system input data as "big data" are their quantity, variety, and speed. The amount of data generated by hospital network tools is enormous. This massive volume of data comes from a variety of sources and people. Smartwatches, fitness trackers, cellphones, AI assistants, and home sensors are all getting more popular. All of these factors contribute to the expanding amount of healthcare data. As a result, improved methods of evaluating and storing data are required.

The input source systems generate a wide range of data kinds. Many different forms of technology have resulted in various sorts of data and their qualities. Databases and spreadsheets are no longer required to keep medical information online. It can be kept in a variety of digital formats, including photos, sensor data, and online conversations. Because there is so much disorganized data and it is not all in one format, it is difficult to collect, organize, and evaluate it. It has been demonstrated that machine learning is an effective solution to this problem. Healthcare data that is available to the public is classified into three types: semi-structured data (such as X-rays), unstructured data (such as interaction and chat

data), and organized data (such as pathology reports and medical records). The usage of smart technology and the automation of formerly analog services are both on the rise. This generates new data streams that are potentially very interesting and can be gathered and investigated.

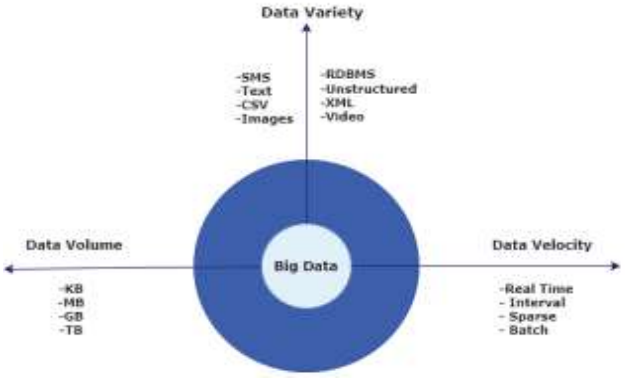


Fig. 2. examined data from the Big Three

Big data is rapidly expanding, bringing with it its own set of difficulties and requirements. The most significant factor in massive volumes of data is velocity, which is the rate at which data is created, processed, and made available for analysis and comprehension. Because sending and receiving real-time data happens so quickly, specific technologies are required. The rise of "big data" tools has made it much easier to cope with the massive volumes of data that are being generated at an increasing rate. Researchers discovered that more data has been created in the last few decades than in the entire history of humanity. In the new digital era, more gadgets and technologies will be employed. This will accelerate data entry into a data analytics platform. Figure 2 depicts the three categories of massive data discussed above.

2. LITERATURE REVIEW

Even though there are still some issues, big data analytics, which allows healthcare organizations to collect massive volumes of data, is the most important aspect of cutting costs and increasing results. Cloud computing is being used by an increasing number of medical facilities and practices to help them better identify and treat their patients. Many others, however, believe that the restrictions and issues associated with cloud computing will make the transition more difficult.

The Hadoop and MapReduce frameworks are being used to investigate the next wave of gene sequencing. They are getting more well-known in this area. We require innovative approaches to dealing with massive amounts of constantly changing data. Hadoop, NoSQL, and Map Reduce are three of the most significant tools for dealing with massive amounts of data. These techniques allow for the incredible organization of massive volumes of data. Experts have researched the fields that generate the most data, as well as how they store and arrange it. Cloud-based apps are becoming increasingly popular in the healthcare industry. Customers can take advantage of numerous advantages, including the option to personalize, the pay-as-you-go model, and widespread availability.

Because of the massive increase in medical data from many sources, such as physiological signal processing, genetic data processing, and medical picture analysis, the healthcare industry can now perform better big data analytics. This paper is divided into two sections. The first section discusses a method for data reduction that employs various technologies, including Sqoop and Hive. The second section provides a Hadoop map-reduce approach for handling large amounts of data. The purpose of this research is to look into new breakthroughs in computational image analysis approaches for leveraging digital pathology images to create prognostic models, with a focus on detecting, extracting, segmenting, and classifying tissues.

You must be able to analyze data if you want to work in healthcare analytics. The speed at which Apache Hadoop can process data is simply astounding. "Big data" is information that has been "map-reduced" and preserved in a distributed database like Hadoop. Large datasets and Hadoop are employed in medical investigations in this work. The author recommends a Hadoop-based design for healthcare big data analytics. This study investigates if big data analytics can be applied in intensive care units. The author of the paper discussed many ways that big data analytics might be utilized to help care for people in intensive care units, as well as potential future prospects for this discipline. The author discusses how big data can improve healthcare by making it more efficient, easier to manage, safer, and more

accurate.

This article discusses the Hadoop ecosystem and provides a list of the tools and components required. This article also discusses the issues that big data analytics has caused. AI and machine learning are widely used in the medical field. There had been applications for cardiac magnetic resonance (CVR). This article discusses the application of deep learning algorithms in cardiovascular magnetic resonance imaging. This article focuses on how machine learning can be utilized in healthcare settings such as translational, clinical, and public health. The value of genetic information, data exchange, and privacy is also discussed. Machine learning techniques are being used to assist doctors in the detection of contagious diseases. Now that machine learning has advanced, we can understand more about infectious disorders such as sepsis, HIV, and bacterial tolerance.

Machine learning may be valuable to researchers in better anticipating healthcare anomalies. This is due to the large number of datasets available for selecting and training a preliminary algorithm on. When there are many elements to consider and sophisticated, nonlinear interactions that affect the outcome, machine learning can create data-driven estimators that can be employed in a variety of circumstances. Cloud computing architecture facilitates data transmission and allows healthcare managers to perform analytical calculations that aid in the discovery of legitimate, relevant, and logical patterns required to avoid and prepare for calamities. The findings of the framework can be utilized by decision-makers in any country to examine and forecast how disasters would influence people's health, taking into account the country's geography and previous disasters.

When AI/ML and big data analytics are employed in healthcare, it is a significant step toward making healthcare truly intelligent. The purpose of this research is to conduct a complete examination of the m-healthcare system using big data analytics and artificial intelligence. Data mining is critical in this age of massive volumes of data. IT and data mining can make it easier to keep and access medical notes and information regarding patient follow-up. Looking for patterns and connections in medical data might help you learn more about identifying and treating patients. It can also help detect diseases earlier, provide a better prognosis, and speed up the healing process. Big data will have enough of an influence on the oral health care sector to fulfill the demand for precision in determining the optimal care at every follow-up by bringing in massive amounts of data and converting them into a series of events that improve treatment. There are various things that healthcare workers could do to improve the quality of their work. The novel ideas in this study, together with improved patient care and targeted medical support, will help the area of oral health advance significantly. This is where the cloud comes in: it saves and makes medical information readily available. This article describes a novel method for gathering large amounts of medical data. It employs a modified immunological technique. You can't create an information model in cloud computing until you understand what a massive data structure model is and how it operates. Using an immunological optimization strategy, we hope to get the optimal grouping of very big data sets.

Table 1 Healthcare statistics are relevant in a variety of contexts.

Sr. No.	Title	Year	Related Field	Application
1	W. Raghupathi et al. [9]	2014	Big Data Analytics	Healthcare
2	G. N Reddy et al. [10]	2014	Cloud Computing	Healthcare
3	E. A. Mohammed et al. [11]	2014	Hadoop	Clinical Big Data & Medical Health Informatics
4	J. Singh et al. [12]	2015	Hadoop	Big Data Management
5	L. Grubel et al. [13]	2015	Cloud Computing	Healthcare
6	Belle et al. [14]	2015	Big Data Analytics	Healthcare
7	Baste et al. [15]	2017	Big Data Analytics	Hadoop Mapreduce and Mitigation process
8	Madhubuchi et al. [16]	2017	Image Analysis & Machine Learning	Digital Pathology
9	Merla et al. [17]	2017	Hadoop	Cloud Platform AWS
10	S. Kumar et al. [18]	2018	Big Data and Hadoop	Big Data Analytics
11	L. NelsonSanchez-Pinto et al. [19]	2018	Big Data	Intensive Care Unit
12	Bahri, Safa, et al. [20]	2018	Big Data	Healthcare applications
13	Sarosha et al. [21]	2019	Big Data	Hadoop Ecosystem
14	Lerner, Tim, et al. [22]	2019	Machine Learning	Cardiovascular Magnetic Resonance
15	Norman, et al. [23]	2019	Machine Learning	Medical Informatics
16	Peiffer-Smadja et al.[24]	2019	Machine Learning	Clinical Decision support for Infectious Disease
17	Dinge et al. [25]	2019	Machine Learning	Healthcare Informatics
18	Modenian et al. [26]	2019	Cloud Computing & Big Data	Healthcare in Disaster
19	Alotaibi, Sultan Rafa et al. [27]	2020	Big data Analytics & Artificial Intelligence	Mobile Health

Various data processing technologies are utilized to obtain text data for healthcare applications. Medical text data can be handled using KNN, RNN, SVM, and other methods. Medical text data can be evaluated in a variety of ways, including through the use of connection rules, grouping, and classifying. To function, intelligent health care systems rely on insights from large amounts of data. Future healthcare systems will require machine learning and the Internet of Things to function. In this study, we examine more extensively at how machine learning approaches can be used to analyze very large amounts of healthcare data. Along with the benefits and drawbacks of current methodologies, potential concerns with future studies are discussed. This research will assist the government and healthcare organizations in keeping up with the newest developments in AI and machine learning-based healthcare analytics.

Processing massive amounts of data is difficult in and of itself. It takes a lot of time and effort to perform math on very big collections of data. The purpose of this article is to educate healthcare professionals and those new to the industry about bio-inspired technologies. Bio-inspired computer solutions are based on behaviors that have been utilized to tackle complex modeling, simulation, and optimization problems for thousands of years. There is compelling evidence that bio-inspired solutions should address the vast spectrum of difficulties that current technology can address. Cloud systems are becoming increasingly popular for developing specialized health apps such as MMAP. In MMAP, genetics can be used to produce phenotype predictions. This application is both highly accurate and simple to use. Table 1 provides an overview of the literature review conducted on a number of papers discussing the various applications of healthcare data analytics.

3. METHODOLOGY

Analytics in healthcare, when utilized effectively, can improve patient-centered care, cut costs, promote teamwork, and achieve better results with fewer resources. Companies that invest in healthcare analytics development can enhance patient outcomes, predict how diseases will be

treated in the future, and reduce the time it takes to see results.

Analytical tools are used to examine and comprehend data. This is what analytics is all about. It can be used to identify, analyze, and link relevant data examples. Information systems should also be utilized to analyze numerical and statistical data in a dynamic, organized, and enjoyable manner. A large amount of data from the recorder indicates that it performs admirably in certain conditions. Analytics success is measured using a combination of operations research, computer science, and statistics. Healthcare businesses are increasingly using analytics to uncover patterns in their data, explain what those patterns mean, and act on the results. Big data and sophisticated analytics can be used to improve clinical and study findings. These findings can then be used to address issues in the healthcare system and improve business procedures.

Fig. 3. Analytics models for application in the medical industry

Predictive analytics can reduce societal expenditures by identifying "higher risk patients" and initiating preventative measures before issues arise. Health care data can be utilized for descriptive analytics to uncover historical trends. Prescriptive analytics, on the other hand, employs specialized algorithms to determine what has to occur in order for a treatment plan to be effective for a certain patient. This is a compelling argument for merging data from many sources, such as the most recent information on sicknesses, health, wealth, and population. Figure 3 depicts multiple approaches of analyzing healthcare data.

There are numerous applications for data analytics in healthcare. Figure 4 depicts the various types of data, the analytical models that may be applied to them, data mining methods, and their applications in the healthcare industry. Web and social media data, device and monitoring equipment data, transaction data, and data from how individuals regularly act are the most common forms of data sources in healthcare analytics. Data mining techniques such as clustering, regression, classification, and detecting outliers are used by healthcare analytics professionals. Statistics are used in healthcare for a variety of purposes, including system management, medical decision-making, monitoring patient and general health, and ensuring drug safety.

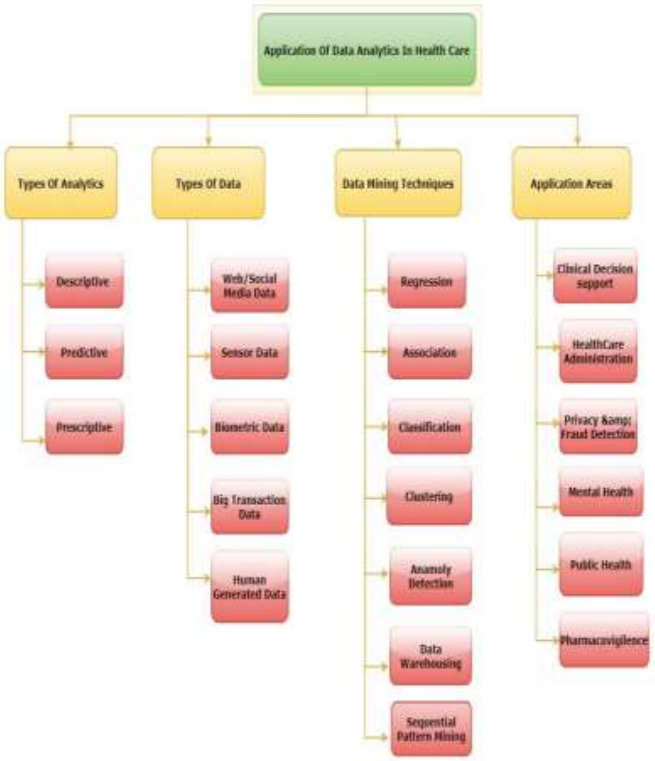


Fig. 4. How Can Analytics Be Used in Healthcare

4. CHALLENGES

The fact that cloud-based big data apps for healthcare analytics create a vast volume of data is a serious challenge. The data may be inaccurate if it comes from untrustworthy or biased sources. Before you begin using any healthcare analytics

tool that gets patients' or test subjects' health data, make sure they agree to it. It is also critical to include data from both internal and external systems in relevant analytics. It's difficult to process all of the many types of suspect information. When combining evidence from many sources, it is critical to reach clear and consistent conclusions. Keeping infrastructure costs down while managing massive amounts of data requires a lot of effort and planning. Problems arise as a result of the massive amount of information. Sending real-time information about global health necessitates extremely fast input/output rates and sophisticated computers, both of which are easily overlooked.

It is typical to require more than one chart to display different data sets on a graph and extract valuable information from the various data sources. Because clinical data is confidential, data security is critical when developing a cloud data tool for healthcare analytics. Policies, identification processes, and data encryption mechanisms must be implemented to ensure the security of data kept in the cloud.

5. CONCLUSION

Artificial intelligence (AI) solutions that employ HADOOP's Big Data Analytics features, as well as the expanding use of cloud computing, are causing healthcare centers and services to focus more on the patient. They also provide low-cost options for residents in the neighborhood to obtain urgent medical treatment. To handle highly large and intricate datasets (HCD) for next-generation sequencing analysis, the healthcare industry use the Hadoop with Map Reduce system. Tools like Sqoop, Hive, and others leverage readily available analytics models to assist with finding data, extracting features, categorizing it, and processing it during the data reduction process. When you employ the appropriate combination of analytics models, data mining techniques, and safeguarding heterogeneous data sets (HDS), you have the potential to dramatically transform the health care business and unlock a plethora of new benefits and applications. Every country's health policy can benefit from these models in order to become more open, collaborative, goal-oriented, effective, and flexible. For example, by merging all accessible data sets into the appropriate Prescriptive analytics system, new strategies to treat mental health problems could be discovered. Having a full model is also very useful for ensuring that pharmaceutical monitoring quality criteria are met globally. If difficulties with data storage, security, privacy, latency in data transmission, and network disruptions for critical apps can be resolved, the market for applications might increase extremely quickly.

REFERENCES

[1] Shilo, S., Rossman, H. & Segal, E. Axes of a revolution: challenges and promises of big data in healthcare. Nat Med 26, 29–38 (2020).https://doi.org/10.1038/s41591-019-0727-5.

[2] The Old Bailey and OCR: Benchmarking AWS, Azure, and GCP with 180,000 Page Images DocEng '20: Proceedings of the ACM Symposium on Document Engineering 2020 September 2020 Article No.: 19Pages 1–4https://doi.org/10.1145/3395027.3419595

[3] Banerjee, A., Chakraborty, C., Kumar, A., & Biswas, D. (2020). Emerging trends in IoT and big data analytics for biomedical and health care technologies. In Handbook of data science approaches for biomedical engineering (pp. 121-152). Academic Press.

[4] Bani-Salameh, H.; Al-Qawaqneh, M.; Taamneh, S. Investigating the Adoption of Big Data Management in Healthcare in Jordan. Data 2021, 6, 16. https:// doi.org/10.3390/data6020016

[5] D. E. O'Leary, "Artificial Intelligence and Big Data," in IEEE Intelligent Systems, vol. 28, no. 2, pp. 96-99, March-April 2013, doi: 10.1109/MIS.2013.39.

[6] Z. He, Z. Cai, Y. Sun, et al., Customized privacy preserving for inherent data and latent data, ACM Pers. Ubiquitous Compute. 21 (1) (2017) 43–54.

[7] Cloud-Based Big Data Analytics—A Survey of Current Research and Future

Directions Samiya Khan, Kashish Ara Shakil and Mansaf Alam© Springer Nature Singapore Pte Ltd. 2018 V.B. Aggarwal et al. (eds.), Big Data Analytics, Advances in Intelligent Systems and Computing 654, https://doi.org/10.1007/978-981-10-6620-7_57

[8] M.M.E. Mahmoud, J.J.P.C. Rodrigues, S.H. Ahmed, et al., Enabling technologies on cloud of things for smart healthcare, IEEE Access 6 (2018) 31950–31967

[9] M. Daniels, J. Rose, C. Farkas, Protecting patients' data: An efficient method for health data privacy, in: Proceedings of the 13th International Conference on Availability, Reliability and Security, ACM, 2018, p. 9.

[10] Raghupathi, Wullianallur, and Viju Raghupathi. "Big data analytics in healthcare: promise and potential." Health information science and systems 2.1 (2014): 1-10.