International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT20665

Importance of Big Data In Healthcare System-A Survey Approach

M. Pavithra[1], Dr. E. S. Shamila[2], G. Krishna Priya[3], G. VijiPriya[4], R. Ashwini[5]

Assistant Professor, Department of C.S.E, Jansons Institute of Technology, Coimbatore, India[1]

Professor, Department of C.S.E, Jansons Institute of Technology, Coimbatore, India[2]

Assistant Professor, Department of C.S.E, Jansons Institute of Technology, Coimbatore, India[3]

Assistant Professor, Department of Artificial Intelligence & Data Science, Jansons Institute of Technology, Coimbatore, India[4]

Assistant Professor, Department of C.S.E, Jansons Institute of Technology, Coimbatore, India[5]

Article Info
Volume 6, Issue 6
Page Number: 47-56
Publication Issue :
November-December-2020

Article History
Accepted : 01 Nov 2020
Published : 06 Nov 2020

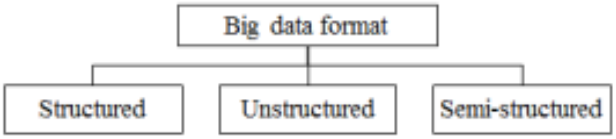
ABSTRACT

'Big data' is massive amounts of information that can work wonders. It has become a topic of special interest for the past two decades because of a great potential that is hidden in it. Various public and private sector industries generate, store, and analyze big data with an aim to improve the services they provide. In the healthcare industry, various sources for big data include hospital records, medical records of patients, and results of medical examinations, and devices that are a part of internet of things. Biomedical research also generates a significant portion of big data relevant to public healthcare. This data requires proper management and analysis in order to derive meaningful information. Otherwise, seeking solution by analyzing big data quickly becomes comparable to finding a needle in the haystack. There are various challenges associated with each step of handling big data which can only be surpassed by using high-end computing solutions for big data analysis. That is why, to provide relevant solutions for improving public health, healthcare providers are required to be fully equipped with appropriate infrastructure to systematically generate and analyze big data. An efficient management, analysis, and interpretation of big data can change the game by opening new avenues for modern healthcare. That is exactly why various industries, including the healthcare industry, are taking vigorous steps to convert this potential into better services and financial advantages. With a strong integration of biomedical and healthcare data, modern healthcare organizations can possibly revolutionize the medical therapies and personalized medicine.

Keywords : Big Data, Medical Examinations, Biomedical Research, Haystack, Healthcare Industry, Medical Therapies

I. INTRODUCTION

Information has been the key to a better organization and new developments. The more information we have, the more optimally we can organize ourselves to deliver the best outcomes. That is why data collection is an important part for every organization. We can also use this data for the prediction of current trends of certain parameters and future events. As we are becoming more and more aware of this, we have started producing and collecting more data about almost everything by introducing technological developments in this direction. Today, we are facing a situation wherein we are flooded with tons of data from every aspect of our life such as social activities, science, work, health, etc. In a way, we can compare the present situation to a data deluge. The technological advances have helped us in generating more and more data, even to a level where it has become unmanageable with currently available technologies. This has led to the creation of the term 'big data' to describe data that is large and unmanageable. In order to meet our present and future social needs, we need to develop new strategies to organize this data and derive meaningful information. One such special social need is healthcare. Like every other industry, healthcare organizations are producing data at a tremendous rate that presents many advantages and challenges at the same time. In this review, we discuss about the basics of big data including its management, analysis and future prospects especially in healthcare sector. The healthcare industry has generated large amount of data generated from record keeping, compliance and patient related data. In today's digital world, it is mandatory that these data should be digitized. To improve the quality of healthcare by minimizing the costs, it's necessary that large volume of data generated should be analyzed effectively to answer new challenges. Similarly government also generates petabytes of data every day. It requires a technology that helps to perform a real time analysis on the enormous data set. This will help the government to provide value added services to the citizens. Big data analytics helps in discovering valuable decisions by understanding the data patterns and the relationship between them with the help of machine learning algorithms (1). This paper provides an overview of big data analytics in healthcare and government systems. It describes about big data generated by these systems, data characteristics, security issues in handling big data and how big data analytics helps to gain a meaningful insight on these data set.



1. NEED FOR BIG DATA ANALYTICS IN HEALTHCARE

To improve the quality of healthcare by considering the following:

Providing patient centric services:

To provide faster relief to the patients by providing evidence based medicine-- detecting diseases at the earlier stages based on the clinical data available, minimizing drug doses to avoid side effect and providing efficient medicine based on genetic makeups(1). This helps in reducing readmission rates thereby reducing cost for the patients.

Detecting spreading diseases earlier:

Predicting the viral diseases earlier before spreading based on the live analysis. This can be identified by analyzing the social logs of the patients suffering

from a disease in a particular geo-location(1). This helps the healthcare professionals to advise the victims by taking necessary preventive measures.

Monitoring the hospital's quality:

Monitoring whether the hospitals are setup according to the norms setup by Indian medical council. This periodical check-up helps government in taking necessary measures against disqualifying hospitals.

Improving the treatment methods:

Customized patient treatment---monitoring the effect of medication continuously and based on the analysis dosages of medications can be changed for faster relief. Monitoring patient vital signs to provide proactive care to patients. Making an analysis on the data generated by the patients who already suffered from the same symptoms, helps doctor to provide effective medicines to new patients.

2. BIG DATA PLATFORMS

As in [9], big data uses distributed storage technology based on cloud computing rather than local storage. Some big data cloud platforms are Google cloud services, Amazon S3 and Microsoft Azure. Google's distributed file system GFS (Google File System) [21] and its programming model Mapreduce are the lead in the field. The performance of map reduce has received a valid amount of attention in large scale data processing. So many organizations use big data processing framework with map reduce. Hadoop, an influential aspect in big data was developed by Yahoo and it is an open-source version of GFS [29]. Hadoop enables storing and processing big data in distributed environment on large clusters of hardware. Enormous data storage and faster processing are supported by hadoop. Hadoop Distributed File System (HDFS) provides reliable and scalable data storage. HDFS makes multiple copies of each data block and distributes them on systems on a cluster to enable reliable access. HDFS supports cloud computing through the use hadoop, a distributed data processing platform. Another one, 'Big Table' was developed by Google in 2006 that is used to process huge amount of structured data. It also supports map reduce [6]. Amazon developed Dynamo [7], a key-value pair storage system. It is a scalable distributed data store built for Amazon's platform. It gives high reliability, cost effectiveness, availability and performance. Tom white [25] elaborates various tools for big data analytics. Hive, a framework for data warehousing on top of hadoop. It was built at Facebook. Hive with hadoop for storage and processing meets the scalability needs and is cost effective. Hive uses a query language called HiveQL which is alike on SQL. A scripting language for exploring large datasets is called 'Pig'. An opinion of map reduce is that writing of mappers and reducers, compiling the package and code are tough and so the development cycle is long. Hence working with mapreduce needs experience. Pig overcomes this criticism by its simplicity. It allows the developers to write simple Pig Latin queries to process the big data and thereby save the time. A distributed column oriented database Hbase [22] built on top of Hadoop Distributed File System. It can be used when we need random access of very large datasets. It speeds up the performance of operations. Hbase can be accessed through application programming interfaces (APIs) like REST (Representational State Transfer) and java. Hbase does not have its own queries, so it depends on Zookeeper. Zookeeper manages huge amount of data. This allows distributed process to manage through a namespace of data registers. This distributed service also has master and slave nodes like in hadoop. Another important tool is Mahout. It is a data mining and machine learning library. It can be categorized as collective filtering, categorization, clustering and mining. It can be executed by Mapreduce in a distributed mode. Big data analytics is not only based on platforms but also analytics algorithms plays a significant role.

3. BIG DATA ANALYTICS APPLICATIONS IN HEALTHCARE

With effective use of digitization, huge amount of information is available in healthcare system starting from the prescription of individual physician , group of doctors, to large hospitals and other organizations who come forward for take care of and provide facilities to patients. Such big data in healthcare organizations provide significant benefits to community. One of the promising benefit is that the disease can be detected at an early stage through the analysis of such huge information and proper care and treatment can be provided immediately in an effective way to an individual. Big data analytics in healthcare gives answers to abundant queries and investigations through the analysis of huge historical information about the patient family history, chronic diseases, nature of surgery or medication for the patient, benefits and side effects of the surgeries and medication, progress in health [3] , time to recover the healthy state after illness. Big data analytics [2] can provide various measures to be taken to save expenditure in healthcare by the people and to lead healthy life by taking initial care through predictable information. Big data analytics can show improvement and efficiency in the following areas: i) Analysis of disease patterns, tracking the disease and dealing with disease eruption. Such analysis can enhance the level of public health [9] and awareness and rapid actions to control the diseases. ii) Development of required vaccines by the medical researchers. iii) Conversion of large amounts of healthcare information into predictive models to recognize the requirements, services to be supplied, predict and impede health disaster for the benefit of people. . Other areas in which big data analytics [6] give enhanced profit are identifying the patient who uses maximum health resources and are at the greatest risk for adverse outcomes. Data analytics provides individuals with the information they need to make decisions and can take the measures to deal with their own health and can adopt and follow healthier behaviors like various programs that are not costlier but bring provable benefits. One of the advantages of big data analytics in healthcare is to supervise the public health [9] by detecting the vulnerabilities during the health disasters and bring together collectively various filed like medical, economic and effective data to act to effectively and productively in real life.

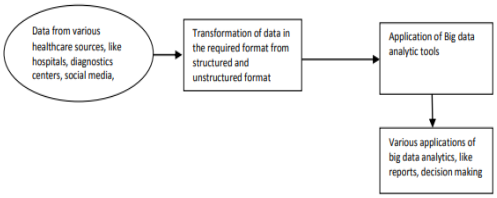
4. SIGNIFICANT FEATURES OF BIG DATA IN HEALTH CARE ENVIRONMENT

With the rapid development in internet technology, mobile devices, storage media, and fast processing technique facilitated the collection of huge amount of health information. Such information is growing day by day. Not only the healthcare system are providing the information relevant to patient, but other organizations which also support the financial benefit to the patient like insurance agencies and pharmacy industries are also contributing to this information growth. This accumulation of information will continue in coming years. The main aim of such bulk information is to analyze it, use it for the benefit and improve the health conditions in the society. Monetary benefit is not the main motive behind the collection and analysis of such information, but to utilize the effective tools, techniques and infrastructure of big data and consider the output results for decision making and providing better facility to the population and increase the conditions of health level. Healthy people will create healthy society and nation.

5. BIG DATA CHARACTERISTICS IN HEALTHCARE DATA

From the recent past years there is exponential growth in the data produced, collected, shared, by different organizations. Such Huge data cannot be

managed and processed by the conventional methods is called Big data. The features of big data are volume, variety, velocity, and veracity [1]. Also termed as 4 V's.



6. SUPPORT OF BIG DATA ANALYTICS IN HEALTH CARE DATABASES

There are number of reasons to implement big data analytics in healthcare databases. Data related to healthcare is generally available in different places, in difference formats, generally the data in the organization is assumed to be in the structured form, but in healthcare system data from various sources consist of diverse kind of data like images, X-rays, graphs, text, handwritten information which are obtained from diagnosis center, medicine details, doctor prescription etc. Such data are in the structured and unstructured form. Analysis on such diverse and complex kind of data is difficult for analysis with usual database management tools and techniques. With the fast development of computing technology, big data is the solution to effectively utilize the important value of such accumulated information. Big data is the best solution in the field of healthcare and improve the quality of human life. The main goal of big data in health care is not only to increase the profits and reduce time and other wastages, but to forecast the outbreak of diseases and its cure. This helps the society to lead a quality life. As the population is increasing day by day , the need of timely medical treatment and suggestion are very much essential and main source to fulfill these requirements is data .The data obtained from various sources like hospitals, diagnosis center, insurance companies etc. are very useful to assist the patient by giving an indication as warning message if the disease in in early stage and inform the person about the seriousness of disease and take preventive measures which will be less expensive compared the disease which is diagnosed at last stage and need costly treatment.

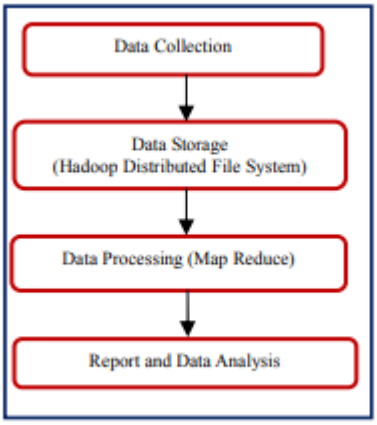
II. ALGORITHMIC TECHNIQUES

Big data mining is the method of winnowing hidden, unknown but useful information from massive amount of data. This information can be used to predict future situations as a help to decision making process. Helpful knowledge can be found by the usage of data mining techniques in healthcare applications like decision support system. The big data produced by healthcare organizations are very complicated and vast to be handled and analyzed by usual methods. Data mining grants the procedure to transform those bundles of data into useful information for decision support. Big data mining in healthcare is about learning models to predict patients' disease. For example, data mining can help healthcare insurance organizations to detect hypocrites and misuse, healthcare institutions make decisions of customer relationship management, doctors identify effective treatments and best practices, and patients get improved and more economical healthcare services. This predictive analysis is widely used in healthcare.

III. BIG DATA ARCHITECTURE

Big data project management is one of the challenging tasks as we need to collect the data from different source which have different varieties. Data nowadays are no better structured so that the old database management system can be utilized for knowledge extraction from the data. In big data we need to deal with structured, semi structured and unstructured data [6]. So the first step should be to collect all the data from the relevant sources and then

aggregate and store it into one common platform. Generally we use open source distribution of Apache Hadoop which provides us with Hadoop Distributed File System storage which takes care of distributed storage and fault tolerance. Once we have the data we need to process it with lowest computation time in distributed environment so that we can adopt Map Reduce [7] processing which can take the benefit of HDFS and inbuilt distributed processing to process data as quick as possible. In the processing layer we generally implement some of the machine learning algorithms which perform some intelligent analysis on the data and supply valuable knowledge which can be used to generate reports.



IV. CHALLENGES AND FUTURE DIRECTIONS

Big data analytics not only provides charming opportunities but also faces lot of challenges. The challenge starts from choosing the big data analytics platform. While choosing the platform, some criteria like availability, ease of use, scalability, level of security and continuity should be considered [27]. The other challenges of big data analytics are data incompleteness, scalability and security [1], [19]. Since cloud computing plays a major role in big data analytics, cloud security should be considered. Studies show that 90% of big data are unstructured data. But the representation, analytics and access of numerous unstructured data are still a challenge. Data timeliness is also critical in various healthcare areas like clinical decision support for making decisions or providing information that guides to take decisions. Big data can make decision support simpler, faster and more accurate because decisions are based on higher volumes of data that are more current and relevant. This needs scalable analytics algorithms to produce timely results [9]. However, most of the current algorithms are inefficient in terms of big data analytics. So the availability of effective analytics algorithms is also necessary. Concerns about privacy and security are superior, although these are increasingly being attempted by new authentication approaches and policies that better protect patient identifiable data.

V. BIG DATA STORAGE AND MANAGEMENT

One of the most important elements in dealing with and managing data is to know where and how this data will be stored once when it is collected. The traditional methods of storing and retrieving such data are not efficient anymore, since it was structured and stored in data warehouses and relational databases, after extracting and loading it from different outside sources. However, this data is transformed and classified before being ready to use and function (Bakshi 2012). Furthermore, Herodotou et al (2011) agreed with Bakshi (2012) when he said that there are many numbers of data sources now and that a huge amount of data has become available, so this growth of data will absolutely require an agile database which can deal with the data logically and through data synchronization in order to adapt to the rapid data evolution. On the other hand, Plattner and Zeier (2011) stated that databases only manage server memory data, therefore eliminating the option of managing other storage devices such as: disk and compact drivers. Accordingly, this will reduce the efficiency of database performance and real time response during the time.

BIG DATA IN HEALTH CARE

The types of data anticipated to be of use in BDA include:

1. **Clinical data** – up to 80 per cent of health data is unstructured as documents, images, clinical or prescribed notes;
2. **Publications** – clinical research and medical reference material;
3. **Clinical references** – text-based practice guidelines and health product (e.g., drug information) data; 4. Genomic data – represents significant amounts of new gene sequencing data;
5. **Streamed data** – home monitoring, telehealth, handheld and sensor-based wireless or smart devices are new data sources and types;
6. **Web and social networking data** – consumer use of Internet – data from search engines and social networking sites; and
7. **Business, organizational and external data** – administrative data such as billing and scheduling and other non-health data.

VI. OPPORTUNITIES OF BDA IN HEALTH CARE

Big data analytics represents a new approach to analytics. It does not yet have a large or significant footprint India or internationally. However, the continuing digitization of health records together with the interoperable electronic health record (EHR), presents new opportunities to investigate a myriad of clinical and administrative questions. There is potential to layer BDA-type applications, in a privacy-protective manner, on top of the foundational health IT infrastructure to derive value that might not otherwise be found. What follows are some innovative ideas and solutions.

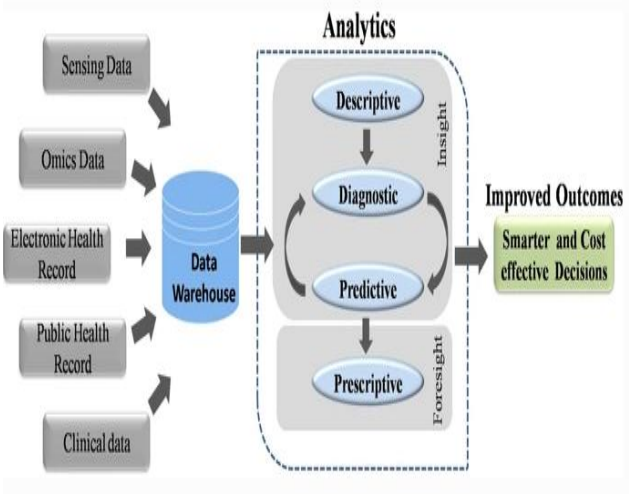
• **Clinical decision support** – BDA technologies that sift through large amounts of data, understand, categorize and learn from it, and then predict outcomes or recommend alternative treatments to clinicians and patients at the point of care.

• **Personalized care** – Predictive data mining or analytic solutions that can leverage personalized care (e.g., genomic DNA sequence for cancer care) in real time to highlight best practice treatments to patients. These solutions may offer early detection and diagnosis before a patient develops disease symptoms.

• **Public and population health** – BDA solutions that can mine web-based and social media data to predict flu outbreaks based on consumers' search, social content and query activity. BDA solutions can also support clinicians and epidemiologists performing analyses across patient populations and care venues to help identify disease trends.

• **Clinical operations** – BDA can support initiatives such as wait-time management, where it can mine large amounts of historical and unstructured data, look for patterns and model various scenarios to predict events that may affect wait times before they actually happen.

• **Policy, financial and administrative** – BDA can support decision makers by integrating and analyzing data related to key performance indicators.

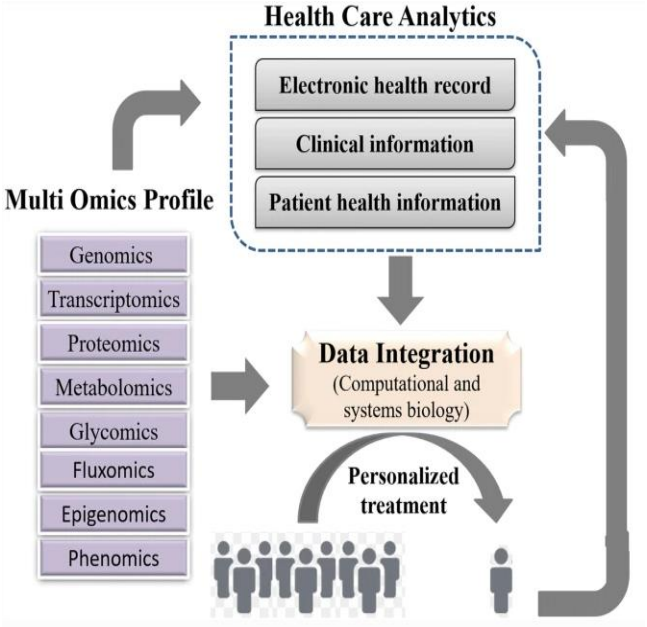


VII. BIG DATA CHALLENGES IN HEALTH CARE

• Leveraging the patient/data correlations in longitudinal records.

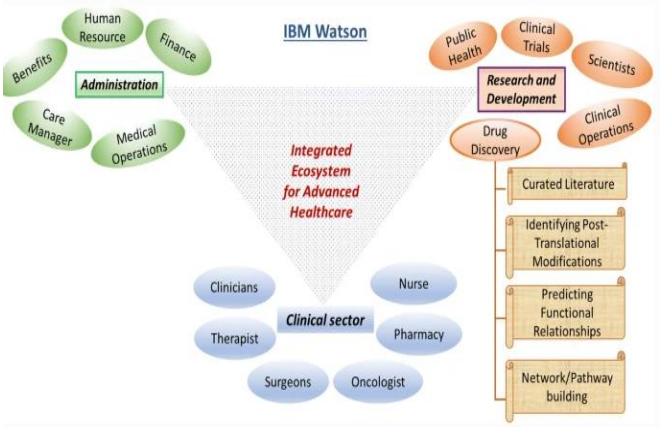
• Understanding unstructured clinical notes in the right context.

• Efficiently handling large volumes of medical imaging data and extracting potentially useful information and biomarkers.
• Analyzing genomic data is a computationally intensive task and combining with standard clinical data adds additional layers of complexity.
• Capturing the patient's behavioral data through several sensors; their various social interactions and communications.



VIII. NATURE OF THE BIG DATA IN HEALTHCARE

EHRs can enable advanced analytics and help clinical decision-making by providing enormous data. However, a large proportion of this data is currently unstructured in nature. An unstructured data is the information that does not adhere to a pre-defined model or organizational framework. The reason for this choice may simply be that we can record it in a myriad of formats. Another reason for opting unstructured format is that often the structured input options (drop-down menus, radio buttons, and check boxes) can fall short for capturing data of complex nature. For example, we cannot record the non-standard data regarding a patient's clinical suspicions, socioeconomic data, patient preferences, key lifestyle factors, and other related information in any other way but an unstructured format. It is difficult to group such varied, yet critical, sources of information into an intuitive or unified data format for further analysis using algorithms to understand and leverage the patients care. Nonetheless, the healthcare industry is required to utilize the full potential of these rich streams of information to enhance the patient experience. In the healthcare sector, it could materialize in terms of better management, care and low-cost treatments. We are miles away from realizing the benefits of big data in a meaningful way and harnessing the insights that come from it. In order to achieve these goals, we need to manage and analyze the big data in a systematic manner.



IX. MANAGEMENT AND ANALYSIS OF BIG DATA

Big data is the huge amounts of a variety of data generated at a rapid rate. The data gathered from various sources is mostly required for optimizing consumer services rather than consumer consumption. This is also true for big data from the biomedical research and healthcare. The major challenge with big data is how to handle this large volume of information. To make it available for scientific community, the data is required to be stored in a file format that is easily accessible and readable for an efficient analysis. In the context of healthcare data, another major challenge is the implementation of high-end computing tools, protocols and high-end hardware in the clinical

setting. Experts from diverse backgrounds including biology, information technology, statistics, and mathematics are required to work together to achieve this goal. The data collected using the sensors can be made available on a storage cloud with pre-installed software tools developed by analytic tool developers. These tools would have data mining and ML functions developed by AI experts to convert the information stored as data into knowledge. Upon implementation, it would enhance the efficiency of acquiring, storing, analyzing, and visualization of big data from healthcare. The main task is to annotate, integrate, and present this complex data in an appropriate manner for a better understanding. In absence of such relevant information, the (healthcare) data remains quite cloudy and may not lead the biomedical researchers any further. Finally, visualization tools developed by computer graphics designers can efficiently display this newly gained knowledge.

X. APPLICATIONS IN BIG DATA ANALYSIS

Quantum computing is picking up and seems to be a potential solution for big data analysis. For example, identification of rare events, such as the production of Higgs bosons at the Large Hadron Collider (LHC) can now be performed using quantum approaches [43]. At LHC, huge an amount of collision data (1PB/s) is generated that needs to be filtered and analyzed. One such approach, the quantum annealing for ML (QAML) that implements a combination of ML and quantum computing with a programmable quantum annealed, helps reduce human intervention and increase the accuracy of assessing particle-collision data. In another example, the quantum support vector machine was implemented for both training and classification stages to classify new data [44]. Such quantum approaches could find applications in many areas of science [43]. Indeed, recurrent quantum neural network (RQNN) was implemented to increase signal reparability in electroencephalogram (EEG) signals [45]. Similarly, quantum annealing was applied to intensity modulated radiotherapy (IMRT) beam let intensity optimization [46]. Similarly, there exist more applications of quantum approaches regarding healthcare e.g. quantum sensors and quantum microscopes [47].

XI. CONCLUSION

Big data analytics in healthcare is evolving into a promising field for providing insight from very large data sets and improving outcomes while reducing costs. Its potential is great; however there remain challenges to overcome. Big data analytics has the potential to transform the way healthcare providers use sophisticated technologies to gain insight from their clinical and other data repositories and make informed decisions. In the future we'll see the rapid, widespread implementation and use of big data analytics across the healthcare organization and the healthcare industry. To that end, the several challenges must be addressed. As big data analytics becomes more mainstream, issues such as guaranteeing privacy, safeguarding security, establishing standards and governance, and continually improving the tools and technologies will garner attention. Big data analytics and applications in healthcare are at a nascent stage of development, but rapid advances in platforms and tools can accelerate their maturing process. The importance towards the healthcare is increasing with the increase in population. Today the people are more conscious about their personal healthcare and try to access the information from various sources. Lot of information is easily accessible the importance from various sources like social media, internet, hospitals, diagnosis center, insurance companies. The cost reduction in computing process, storage and communication technology and mobile devices enabled the users to generate and access and store huge data. Such information can create value to it

and provide knowledge, comfort, productivity, better, healthy society and in turn economic benefit. Big data in healthcare provides very sophisticated and fast analytic tools which works on the massive and diverse kind of information. The output of big data analytics in healthcare [10] databases enables the user to obtain the true information, and also increase the profits in many business sectors like insurance companies, banking, and hospitals. Solutions offered by big data in healthcare system, helps the users by providing the useful information about the disease predictions, hospital information, symptoms of health, and details of insurance companies who come forward to provide the financial benefits to the patients. Such useful information to the public with less expenditure improves the quality of life. Today to utilize the efficiency of big data in healthcare system many organizations, like medicine industry, insurance companies, diagnostic centers, banks are working together and coming forward with various new ideas and attempts to achieve the objective to provide better healthy society with less cost to the population.

XII.FUTURE WORK

Large amounts of heterogeneous medical data have become available in various healthcare organizations. The rate of electronic health record (EHR) adoption continues to climb in both inpatient and outpatient aspects. Those data could be an enabling resource for deriving insights for improving patient care and reducing waste. Analyzing the massive amount of healthcare information that is newly available in digital format should enable advanced detection of powerful treatment, better clinical decision support and accurate predictions of who is likely to get sick

Data analysis and prediction area is entering into a new era with the tremendous growth of big data. The problem with the traditional machine learning algorithm is that they are not scalable and in some cases they take too much time for large data set. So we need to fine tune or modify the existing algorithm to make it suitable for the big data problem. In healthcare domain, there is a large amount of data collected and with prediction capability researcher has always made the problem simpler for patient. With big data coming into picture there is a tremendous growth in this area and a new dimension as now we can extract a more valuable information and hidden knowledge, which can provide a breakthrough in the medical cases.

XIII. REFERENCES

[1]. Raghupathi, Wullianallur, and Viju Raghupathi. "Big data analytics in healthcare: promise and potential." Health Information Science and Systems 2.1 (2014):

[2]. Hermon, Rebecca, and Patricia AH Williams. "Big data in healthcare: What is it used for?" (2014).

[3]. Groves, Peter, et al. "The 'big data 'revolution in healthcare." McKinsey Quarterly (2013).

[4]. Rama Satish, K. V., and N. P. Kavya. "Big data processing with harnessing hadoop-MapReduce for optimizing analytical workloads." Contemporary Computing and Informatics (IC3I), 2014 International Conference on. IEEE, 2014.

[5]. Kesavaraj, G., and S. Sukumaran. "A study on classification techniques in data mining." Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on. IEEE, 2013.

[6]. Tomar, Divya, and Sonali Agarwal. "A survey on Data Mining approaches for Healthcare." International Journal of Bio-Science and Bio-Technology 5.5 (2013): 241-266.

[7]. Cavallaro, Gabriele, et al. "On Understanding Big Data Impacts in Remotely Sensed Image Classification Using Support Vector Machine Methods." (2015).

[8]. Chandra, B., and Ratnesh K. Sharma. "Fast learning for big data applications using parameterized multilayer perceptron." Big Data (Big Data), 2014 IEEE International Conference on. IEEE, 2014.

[9]. Torres, Guadalupe J., et al. "A similarity measure for clustering and its applications." Int J Electr Comput Syst Eng 3.3 (2009): 164-170.

[10]. Tapia, José Juan, Enrique Morett, and Edgar E. Vallejo. "A Clustering Genetic Algorithm for Genomic Data Mining." (2009): 249-275.

[11]. Fahad, Adil, et al. "A survey of clustering algorithms for big data: Taxonomy and empirical analysis." Emerging Topics in Computing, IEEE Transactions on 2.3 (2014): 267-279.

[12]. Han, Jiawei, Micheline Kamber, and Jian Pei. Data mining: concepts and techniques: concepts and techniques. Elsevier, 2011.

[13]. Sasirekha, K., and P. Baby. "Agglomerative hierarchical clustering algorithm–A Review." Proceedings of the eleventh international conference on Information and knowledge management. 2002.

[14]. Alexandros Labrinidis and H. V. Jagadish, "Challenges and opportunities with big data," Proc. VLDB Endow. 5, pp. 2032-2033, August 2012.

[15]. Amir Gandomi, Murtaza Haider, "Beyond the hype: Big data concepts, methods, and analytics," International Journal of Information Management 35, pp. 137-144, 2015.

[16]. I.A.T. Hashem, et al, "The rise of "big data" on cloud computing: Review and open research issues," Information Systems, 2014.

[17]. Jawei Han and Micheline Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann

[18]. Publishers Inc., San Francisco, CA, USA, 2000.

[19]. Mark A. Beyer, Douglas Laney, "The importance of 'Big Data': A Definition," Gartner, retrieved on 21 June 2012.

[20]. S. Sagiroglu and D. Sinanc, "Big data: a review," in Proceedings of the International Conference on Collaboration Technologies and Systems (CTS '13), pp. 42-47, IEEE, San Diego, Calif, USA, May 2013.

Cite this article as :

M. Pavithra, Dr. E. S. Shamila, G. Krishna Priya, G. VijiPriya, R. Ashwini, "Importance of Big Data In Healthcare System A Survey Approach", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 6, pp. 47-56, November-December 2020. Available at doi : https://doi.org/10.32628/CSEIT20665
Journal URL : http://ijsrcseit.com/CSEIT20665