International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 11 | Nov 2018 www.irjet.net p-ISSN: 2395-0072

Integration of Big Data Analytics in Healthcare Systems

Ruchika Patil[1], Prof. Dilip Motwani[2]

[1]Ruchika Patil, Dept. of Computer Engineering, Vidyalankar Institute of Technology, Maharashtra, India
[2]Prof. Dilip Motwani, Dept. of Computer Engineering, Vidyalankar Institute of Technology, Maharashtra, India

---

Abstract - *Healthcare information is the most complicated and numerous data produced in the todays world. The healthcare area generates petabytes of data, information from Electronic Healthcare Records (EHR), research studies, experiments and medical records is often unstructured and due to the large volume of it, can't be processed manually by human being. The common database architectures are no longer up to the challenge with huge data and create a big havoc. Big Data Analytics plays a major role in solving issues and challenges arises in healthcare area. Big Data plays an important role in the predictive analysis in the healthcare sector. The use of Big Data in the field of healthcare has reduced waste and error, great potential for better quality of concern and reducing care costs. This paper explains Big Data's analytical techniques and technologies in healthcare and their benefits. Big Data Analytics can be most useful in the field of health care.*

***Key Words*: Big Data, Big Data Analytics, Healthcare**

1. INTRODUCTION

Big Data is a collection of large and complex data sets that traditional data processing applications cannot efficiently process. Big Data Analytics have a widen role in handling the data which is generated from various resources to provide quality information. Big information plays a major role in healthcare. Big data has features such as volume, speed, variety and veracity. Big Data helps in cost reduction, faster, better decision making in healthcare.

Big Data analytics has the ability to analyze a wide variety of complex data and generate valuable insights which otherwise would not have been possible. Big Data means there is big opportunities and challenges. Big data has many possible future effects on patients, providers, researchers, payers and other healthcare constituents. It specifies to the tools and techniques to collect, store, process, analyze and find patterns.

Nowadays, the lot of information generated in healthcare sector is fast growing. An aim to solve old problems in better way and new problems. Various developing countries are still uses paper based forms but there is required to be collected data moved to a computerized system. Most of universal health decisions are depend on historic data. Nowadays health organizations use more developed tools to deal with the large bulk of information created day by day, to track their information and use it for analysis. In real cases uses personal patient's information, molecular and genetic data for taking personal medical decisions.

There are various analytical approaches such as data mining, natural language processing, artificial intelligence and predictive analytics can be employed to analyze, contextualize and visualize the data.

2. LITERATURE REVIEW

Priyanka K, Prof Nagarathna Kulennavar [3] proposed brief introduction about how we can uncover additional value from health information used in health care centers using a new information management approach called as big data analytics. Including big data analytics in health sector provides stakeholders with new insights that have the potential to advance personalized care, improve patient outcomes and avoid unnecessary costs. They define big data analytics and its characteristics, comments on its advantages and challenges in health care.

A. Rishika Reddy, P. Suresh Kumar [1] give an overview of storing and retrieval methods, Big Data tools and techniques used in healthcare clouds, role of Big Data Analytics in healthcare and discusses the benefits, outlooks in nascent fields of predictive analytics, faces challenges and provides solutions.

Shankar Krishnan [7] objective is to review a few applications of analytics of Big Data in the healthcare field and the associated outcomes. Using Big Data Analytics properly in healthcare can lead to improved care and significant cost savings. Concurrent challenges to be addressed include accessibility, privacy, security, usability, implementation costs, transportability, interoperability, and standardization. Jasleen Kaur Bains [6] give a wide insight and know how about the various Big Data analytics (BDA) initiatives taken to improve healthcare worldwide. It also explains the various phases involved in BDA process and depicts its benefits and challenges with focus on healthcare industry.

Sushmitha.S.P., Prof Renuka Devi D. [8] objective is to provide deep study on the field of Big Data in healthcare sector. BDA is to analyze the immense quantity of dataset. They propose the conceptual architecture of BDA in healthcare sector and highlight the applications of BDA, its potential challenges and future directions in healthcare domain.

Nishita Mehta, Anil Pandit [2] aims to determine the scope of Big Data analytics in healthcare including its applications and challenges in its adoption in healthcare. It also aims to identify strategies for addressing the challenges.

3. BIG DATA ANALYTICS IN HEALTHCARE

The use of Big Data Analysis in health care has many positive and life - saving results. Big data refers to the vast amount of information generated by digitizing everything, which is consolidated and analyzed through specific technologies. Applied to health care, it uses population-specific health data and can help prevent epidemics, cure diseases, reduce costs. Doctors want to understand a patient as much as possible and to take warning signs of serious illness as soon as possible–it is much simpler and cheaper to treat any disease at an early stage. With healthcare data analytics, prevention is better than cure and managing to draw a comprehensive picture of a patient will let insurances provide a tailored package.

The purpose of using healthcare data analytics is to use data-driven findings to predict and solve a problem before it is too late, but also assess methods and treatments faster, keep better track of inventory, involve patients more in their own health [4]. Health information as well as clinical decisions plays an important role in healthcare data. This healthcare data can be analyzed with Big Data Analytics to predict patient diseases and to suggest suitable drugs needed. Predictions will provide new ways of discoveries and new set of patterns that can help in solving problems of different fields such as hospitals, banks, social networking sites and manufacturing [1].

3.1 Sources of Healthcare Big Data

Health care data is disorganized and distributed from different sources with different structures and forms. Healthcare Big Data includes data on physiological, behavioural, molecular, clinical, environmental, medical imaging, disease management, history of prescription drugs, nutrition or exercise parameters [2].

- Clinical Information Systems: These are traditional sources of clinical data that health care providers are accustomed to viewing.
- Claims Data from Payers: Paying public (e.g. Medicare) and private payers have large claims repositories on their recipients. Some health insurers are now offering incentives to share your health information.
- Research Studies: Research databases contain information about study participants, experimental treatments, and clinical outcomes.
- Public Records: The government keeps detailed records of events related to health, such as immigration, marriage, birth, and death.
- Web Knowledge: Less patient-specific information, such as emergency care data, news feeds, and articles in medical journals.
- Social Media: Social media posts, including Twitter feeds, blogs, Facebook status updates and other platforms, and web pages can reflect and demonstrate a person's health and mood.
- The Internet of Things (IoT): Massive troves of health-related information are also collected and stored on mobile and home devices [5].

3.2 Big Data Analytical Techniques and Technologies in Healthcare

Multidimensional health data – medical images (X - rays, MRI images), biomedical signals (EEG, ECG, EMG, etc.), audio transcripts, handwritten prescriptions and structured EMR data and their dynamic and complexity make it difficult to analyze them. There is paucity of analytical strategies that can handle such heterogeneous data and facilitate decision-making.

By incorporating descriptive and comparative analytics, healthcare organizations have seen improved quality of care. Some of the Big Data Analytical Techniques used in healthcare are shown in Table 1.

Table -1: Big Data Analytical Techniques in Healthcare

Big Data Analytical Technique	Healthcare Application
Cluster Analysis	Determination of obesity clusters for identifying high-risk groups;
	Determination of population clusters with specific health determinants for treatment of chronic diseases
Data Mining	Bio-signal monitoring for health-related abnormalities;
	Determination of epidemics;
	Inductive reasoning and exploratory data analysis in healthcare
Graph Analytics	Analysis of hospital performance across various quality measures
Machine Learning	Prediction of disease risk;
	Assessment of the hospital performance;
	Determination of epidemics
Natural Language Processing (NLP)	Improvement of efficiency of care and controlling costs;
	Providing training, consultation and treatments;
	Identification of high-risk factors;
	Extraction of information from clinical notes;

	Reducing likelihood of morbidity & mortality
Neural Networks	Diagnosis of chronic diseases;
Pattern Recognition	Prediction of patients' future disease; Improvement of public health surveillance
Spatial Analysis	Extracting meaningful population-level insights by using visual, spatial and advanced analytics

MapReduce has the ability to improve the performance of common signal detection algorithms for pharmacovigilance at approximately linear speedup rates. Algorithms based on the Hadoop distributed platform can refine protein structure alignments more accurately than existing algorithms. MapReduce based algorithms can improve the performance of neural signal processing. Image reconstruction algorithms accelerate the reconstruction process.

Some of the Hadoop- based Big Data processing tools like Oozie and Pig which can be used for batch processing; and non-Hadoop processing tools like Storm, Spark, Hive and GraphLab which can be used for streaming data analysis. Regardless of these potential applications, there is a need for analytical tools to offer parallelization, in order to enable the timely processing of data [2]. Table 2 shows comparison of big data analysis tools in healthcare.

Table 2: Comparison of Big Data Analysis Tools in Healthcare

Parameters	Hadoop MapReduce	Apache Spark
Speed	MapReduce reads and writes from disk, as a result, it slows down the processing speed.	Spark is lightning fast cluster computing tool.
Difficulty	In MapReduce, developers need to hand code each and every operation which makes it very difficult to work.	Spark is easy to program as it has tons of high-level operators.
Real-time analysis	MapReduce fails when it comes to real-time data processing as it was designed to perform batch processing on voluminous amounts of data.	It can process real time data i.e. data coming from the real-time event streams at the rate of millions of events per second, e.g. Twitter data for instance or Facebook sharing/posting.
Cost	MapReduce is a cheaper option available while comparing it in terms of cost.	As spark requires a lot of RAM to run in-memory. Thus, increases the cluster, and also its cost.
Security	Apache Hadoop MapReduce is more secure because of Kerberos and it also supports Access Control Lists (ACLs) which are a traditional file permission model.	Spark is little less secure in comparison to MapReduce because it supports the only authentication through shared secret password authentication.

4. CHALLENGES IN BIG DATA ANALYTICS IN HEALTHCARE AND STRATEGIES TO OVERCOME THEM

The healthcare industry is faces different set of challenges each day, ranging from new disease outbreak to maintaining optimal operational efficiency. While big data analytics is useful for patient care, challenges such as loss of privacy, data security, user friendliness, implementation costs, transportability and interoperability are encountered. In order to ensure the privacy and infallibility of big data and BDAH, adherence to standard guidelines and procedures is required [7].

There are some Challenges associated with it which are as follows: -

- Protecting the Patient's Privacy.
- Data Aggregation.
- Cost Incurred for Establishment of Big Data Architecture.
- Requirement of Expert Knowledge in Big Data Analytics.
- Quality of data.

In order to overcome these challenges, Various Strategies were found in the literature. The strategies for curbing the aforementioned issues include:

- Implementing (big) data governance: Due to poor governance, healthcare organizations incur huge financial costs in IT investment. With appropriate data governance, the enterprise-wide data resources can be leveraged effectively to create business value.
- Developing an information sharing culture: Information sharing and aggregation of data can address the issue of interoperability and enable effective utilization of the Big Data analytical and predictive capabilities.
- Employing security measures: Strong encryption of data, validation of source of data, access control and authentication and deidentification are some of the

measures for securing the data and maintaining confidentiality.

- Training key personnel to use Big Data analytics: In order to extract meaningful insights and valuable information from Big Data, healthcare professionals should be trained with Big Data analytics competencies. This is critical for healthcare, because incorrect interpretation of the reports generated could lead to unanticipated consequences.
- Incorporating cloud computing into the organization's Big Data analytics: The challenge of storage of voluminous data can be tackled by making use of cloud computing. This would enable small and medium sized hospitals and care organizations to eliminate cost and data storage issues [2].

5. BENEFITS OF HEALTHCARE BIG DATA ANALYTICS

- Advanced patient care: Electronic health records (EHR) collects all related demographic and medical data consisting of lab tests, clinical data, diagnoses, medical conditions and allergy information.
- Faster time to treatment: Big data analytics tools help expedite the process by factoring in unique circumstances, such as lifestyle choices and demographics, along with the patient's symptoms to help physicians make a more accurate diagnosis and formulate the best treatment regimen in real-time.
- Reducing energy costs: Increased monitoring capabilities as part of a Big Data solution make it so healthcare facility staff and managers can pinpoint areas of with high energy use, and determine where it is really required and where energy is being wasted, and then take steps to combat those inefficiencies.
- For personage/patients/Individuals: With the help of BDA, the specific treatment is given for a patient based on his genomic data, location, weather, lifestyle, medical history, response to certain medicines, allergies, family history etc. Then the specific line of treatment can be constructed for every individual. By using the BDA techniques, it is easier to predict the disease in earlier stage and treated effectively.
- Detecting fraud: Medical abuse and insurance fraud are two most significant problems facing the healthcare industry. The efficient use of big data in healthcare can solve these problems. By analyzing a larger dataset of claims history, fraud patterns can be identified and even predicted before they occur. Health analytics can process large amounts of data effectively in order to reduce fraud, waste and medical abuse.
- Supply Expenditure: With the use of Big Data Analytics, predictions based on past year's records can be made as to what the estimate for this year will be. Predictive analytics allow hospitals to save a large portion of their money by accurately predicting demand for medical supplies.
- Reduced hospitalizations and readmissions: Predict the patients staying and readmission information. New healthcare plans will be developed to prevent hospitalization. Various questions can be answered by analyzing the data using BDA tools and techniques regarding disease treatment. The hospital management can take and manage administrative decisions in the better way.
- Risk stratification: This data analytics tool helps hospitals track and identify the sickest and often the costliest patients in a proactive way.
- Finding cure for diseases: No two persons in the world would have the same genetic sequences, which is the reason why particular medication seems to work for some people but not for others. Since in one genome there are millions of things to observe, it is almost impossible to study them in detail. Big data in healthcare has, however, revolutionized the field of genomics. Big data analytics can uncover unknown correlations, hidden patterns, and insights by examining large data-sets [6].

6. CONCLUSION

Big Data analytics increasingly provide value to healthcare by improving healthcare quality and outcomes and providing cost-effective care. Big Data analytics integration in the healthcare system will improve the tools and techniques and also many challenges can be overcome by various strategies. Using Big Data analytics in healthcare there will be growth in the healthcare sector, advanced patient care and preventing human errors.

REFERENCES

[1] Rishika Reddy, P. Suresh Kumar, 'Predictive Big Data Analytics in Healthcare', 2016 Second International Conference on Computational Intelligence & Communication Technology.

[2] Nishita Mehta, Anil Pandit,'Concurrence of big data analytics and healthcare: A systematic review', International Journal of Medical Informatics 114 (2018).

[3] Priyanka K, Prof Nagarathna Kulennavar, 'A Survey On Big Data Analytics in Health Care', (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4), 2014, 5865-5868.

[4] 'Big data in healthcare' available at https://www.datapine.com/blog/big-data-examples-in-healthcare/.

[5] 'Sources of Big Data' available at https://www.verywellhealth.com/sources-of-big-data-in-health-care-1739184.

[6] Jasleen Kaur Bains 'Big Data Analytics in Healthcare- Its Benefits, Phases and Challenges' International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 4, April 2016.

[7] Shankar Krishnan, 'Application of Analytics to Big Data in Healthcare ',2016 32nd Southern Biomedical Engineering Conference, Ph.D. Biomedical Engineering,

Wentworth Institute of Technology, 550 Huntington Avenue, Boston MA, 02115 USA.

[8] Sushmitha.S.P., Prof.Renuka Devi D., 'A Review On Leveraging Big Data Analytics in Health Care Sector for Enhanced Diagnosis and Competent Patient Health Care Monitoring System', International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Special Issue 11, September 2017.