



A Survey on Security and Privacy Issues of BigData in Healthcare Industry and Implication of Predictive analytics

Parimala.S

Assistant Professor, Dept. of Computer Science, SRM University, Chennai, India

ABSTRACT: A colossal amount of data is being generated every day from Big data. This paper explores the impact of Big data in Healthcare information. Big data analytics uses the Hadoop Tools which plays a major role in the real-time analysis to organize the unstructured and heterogeneous data which comes in immense volume. This paper portrays the big data analytics in Healthcare, its benefits and about the implication of predictive analytics in healthcare and also enlightens the data security and privacy issues in health care industry.

KEYWORDS: Big Data, Analytics, Healthcare, Hadoop, Predictive Analytics, privacy, security

I.INTRODUCTION

The healthcare industry historically has generated large amounts of data, driven by record keeping, compliance & regulatory requirements, and patient care [1]. Healthcare is generating huge amount of data which are both structured and also unstructured to keep the number of patient records. Driven by mandatory requirements and the potential to improve the quality of healthcare delivery meanwhile reducing the costs, these massive quantities of data (known as 'big data') hold the promise of supporting a wide range of medical and healthcare functions, including among others clinical decision support, disease surveillance, and population health management [2, 3, 4, 5] . This paper enlightens the impact of Healthcare along with its security and privacy issues in Big data analytics.

II.RELATED WORK

In [19] the authors have mentioned about the various types of challenges on Big data in healthcare industry. They have also enlightened about the uses of Bigdata in Healthcare. In [20] benefits of using the Big data in healthcare organization for the individuals' and for the patients, for the Government etc. Even some huge promises of big data in healthcare like the quality of data, missing data are some challenges that need to be taken up. In [21] the authors mentioned that by the massive set of data in the health care industry provides the possibilities to do the predictive analysis and find the solutions to many problems. It is also mentioned that Predictive Analysis where we use various statistical method, machine learning techniques and, data mining approaches to process, analyse data and predict the outcome for the unknown bag of data. Healthcare domain is still in early stage to take up the new possibilities, which can be offered by big data solution and use it to do effective decision-making [22].In [23] the author has designed a prediction algorithm which collects the data, aggregates, apply the case attribute capture the performance which can be implemented in healthcare organizations. In [24] the authors have mentioned how the predictive analysis is being achieved in Healthcare and what are the tools and techniques that are being used.

III. IMPACT OF BIG DATA IN HEALTHCARE INFORMATION

What exactly is big data? A report delivered to the U.S. Congress in August 2012 defines big data as "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information". [7]



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 4, April 2017

By digitizing, combining and effectively using big data, healthcare organizations ranging from single-physician offices and multi-provider groups to large hospital networks and accountable care organizations stand to realize significant benefits [6].

There are some areas in which the big data and the Analytics can enhance the health care organization:

1.1 Targeting the right people:

Identifying the people who are at risk and who could benefit from additional treatments and screenings for contrary consequences. Data analytics can also optimize the health status.

When dealing with large populations, it becomes even more important to know who can potentially benefit from interference as a way to improve health and lower costs.

During the management of the people, we can understand who can get the benefit from the interference and also it is one of the way to enhance the patient's health and cost can be reduced.

1.2 Right intervention:

The ability to deliver the right intervention at the right time will improve as people begin to understand their own risks, monitor their health and share pertinent information with their care providers.

The big data analytics has the capability to produce the right intrusion makes the people to understand the risks and monitor the health.

1.3 Perceiving the spreading diseases in advance: 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 [8]. This helps the experts who are working in health care organizations to take some precautionary actions.

1.4 Identifying the cost for treatments:

Pinpointing the treatments for the diseases which is very expensive and by using some analytics and with some effective analytics we can replace with lower costs.

IV. ARCHITECTURAL FRAMEWORK

We can easily understand the concept of frame work for Analytics in Big data project which is similar to conventional method of health informatics. The main difference between the health informatics and Big data Analytics is the way in which the processing take place.

The analysis are performed with the help of the business Intelligence tools are used in the health information processing. Hadoop/Map Reduce are the two open source platforms available on the cloud which stimulates the big data Analytics in the department of Healthcare.

The interfaces of analytical tools which are used in Conventional Method and Big data are totally different from one another but the algorithms and other models are same. The tools in the conventional method are apparent and user friendly.

The tools which are used in the Bigdata Analytics are very complex to understand and requires exhaustive programming knowledge and skills to apply those tools.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 4, April 2017

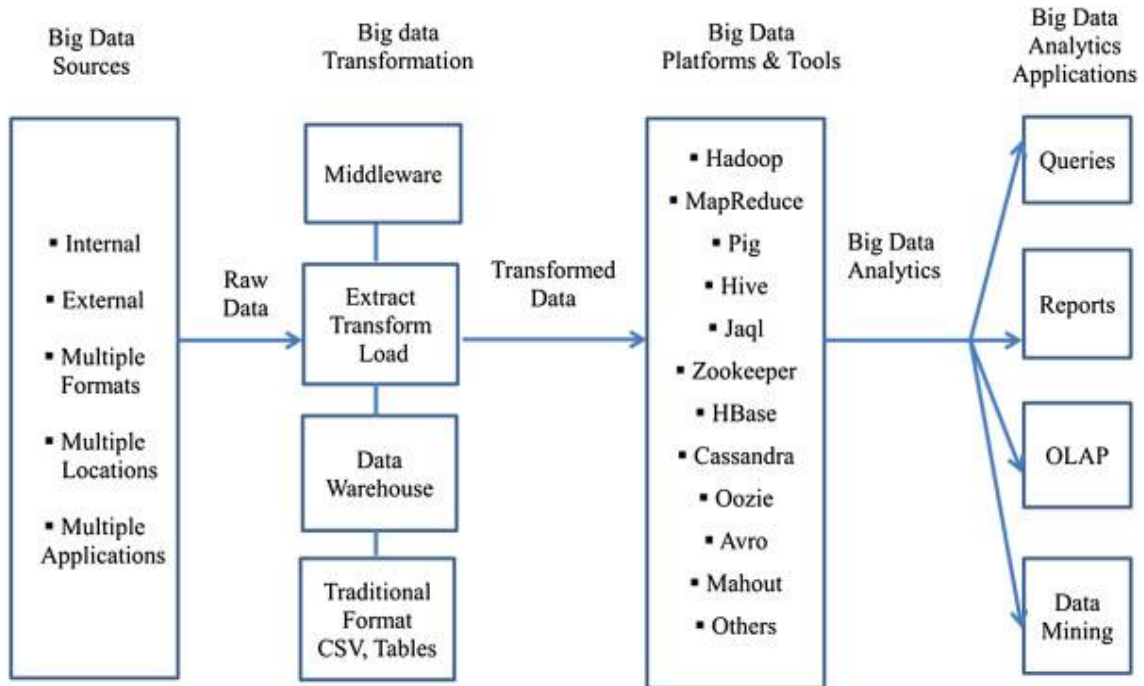


Fig: 1 An applied conceptual architecture of big data analytics

V. PREDICTIVE ANALYTICS IN HEALTH CARE

Predictive analytics (PA) uses technology and statistical methods to search through massive amounts of information, analyzing it to predict outcomes for individual patients. That information can include data from past treatment outcomes as well as the latest medical research published in peer-reviewed journals and databases. [18]

Predictive algorithms are used to diagnose the individual patient's accurately by the physicians. When the patient's come with the complaint of chest pain to the healthcare organization it is difficult to predict whether that patient should be hospitalized or not. If the details about the patient's and their health status are entered properly in a system then the Predictive algorithm will evaluate the patient's condition that he should be hospitalized or could be sent home.

Over the last five years, electronic health records (EHRs) have been widely implemented in the United States, and health care systems now have access to vast amounts of data. While they are beginning to apply "big data" techniques to predict individual outcomes like post-operative complications and diabetes risk, big data remains largely a buzzword, not a reality, in the routine delivery of health care [13]

Predictive analytics encompasses a variety of statistical techniques from predictive modeling, machine learning, and data mining that analyze current and historical facts to make predictions about future or otherwise unknown events.[10][11]

Predictive analytics and machine learning in healthcare are rapidly becoming some of the most-discussed, perhaps most-hyped topics in healthcare analytics. [9]

Some measurements to be taken for patient's care:

- ✓ To increase the way of diagnosing the diseases:

By predictive algorithms the physicians can make accurate diagnose of the patient's diseases. Health care and insurance costs. It is predicted that in 10 years' people in every world region will suffer more death and disability from non-communicable diseases than infectious diseases [12].

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 4, April 2017

- ✓ Increasing the care for patients:

Health care industry is taking much effort for anticipatory methods for treating the patients. For Instance, the patients can be classified as the one who will develop prolonged conditions and the other who will immediately respond to some sorts of treatments.

- ✓ Resource Optimization :

Based on the estimation of the admissions of the patients, the utilization of bed and the duration of the stay can also be analyzed for future database.

VI. BIG DATA SECURITY AND PRIVACY ISSUES.

Information about Health and Exchanges and Electronic Health Records:

- ✓ In 2013, Kaiser Permanente (one of the healthcare providers in US) notified its 49,000 patients that their health information had been compromised due to theft of an unencrypted USB flash drive containing patient records [14]. This kind of healthcare offers Health Information exchanges (HIEs) and Electronic Health Records (EHRs) to share with the patients. This lead to the thieves to steal the data when amassing quantities of data are being shared between multiple providers within the network.
- ✓ A study on patient privacy and data security showed that 94% of hospitals had at least one security breach in the past two years. In most cases, the attacks were from an insider rather than external [15]
- ✓ Data Standardization and Structure issues: The data which are available in healthcare industry are in an unstructured format which will be in the form of graphs, charts, tables, images etc. Structured form of data will be heterogeneous. Understanding those kinds of data is a challenging issue.



Fig2: Architecture of Electronic Health Record.

Imminent of Bigdata in Health care:

The usage of Bigdata and its acceptance throughout the world has provided different scopes to different observation in real-time activities.

The following are some of the aspects that should be deliberated in future for the of Healthcare industry.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 4, April 2017

- Big Data has become an emerging force for the growth of IOT. Gartner estimates 26 billion IOT devices will be functional by 2020 and the amount of traffic generated by such devices will be large enough to place it in the category of big data [17].
- If the strategies in Big data are leveraged to its fullest potential, the Global Institute McKinsey estimation seems to be \$100Billion increase of Profit annually

VII. CONCLUSION

This paper enlightens the overview of big data in healthcare, the significant impact of big data in the healthcare organization. There are many opportunities and as well as many challenges are there in the healthcare industry. Predictive analytics and algorithms are used in healthcare for saving the lives of the patients. It is also believed the predictive analytics plays a major role in big data especially in healthcare organization. Even though many technologies are available some uncertainty of data or information from the electronic health record could lead to mortality or the patient's diseases cannot be diagnosed properly. Healthcare predictions are mostly benefited by the emergency care to build tools with more potential variables as input which are available in Electronic Health Record.

REFERENCES

1. .Raghupathi W: Data Mining in Health Care. Healthcare Informatics: Improving Efficiency and Productivity. Edited by: Kudyba S. , Taylor & Francis, 211-223,2010
2. Burghard C: Big Data and Analytics Key to Accountable Care Success. , IDC Health Insights ,2012
3. Dembosky A: "Data Prescription for Better Healthcare." Financial Times, December 12, 2012, p. 19., Available from: <http://www.ft.com/intl/cms/s/2/55cbca5a-4333-11e2-aa8f-00144feabdc0.html#axzz2W9cuwajK>,2012
4. Feldman B, Martin EM, Skotnes T: "Big Data in Healthcare Hype and Hope." October 2012. Dr. Bonnie 360. 2012,<http://www.west-info.eu/files/big-data-in-healthcare.pdf>, [Google Scholar](#)
5. Fernandes L, O'Connor M, Weaver V: Big data, bigger outcomes. J AHIMA. , 38-42,2012
6. Burghard C: Big Data and Analytics Key to Accountable Care Success. , IDC Health Insights,2012
7. IHTT: Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry. 2013, IHTT: Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry. 2013,<http://ihealthtran.com/wordpress/2013/03/iht%C2%B2-releases-big-data-research-report-download-today>
8. Yanglin Ren, Monitoring patients via a secure and mobile healthcare system, IEEE Symposium on wireless communication,2011
9. <https://www.healthcatalyst.com/predictive-analytics>
10. Nyce, Charles , Predictive Analytics White Paper(PDF), American Institute for Chartered Property Casualty Underwriters/Insurance Institute of America, p. 1,2007
11. Eckerson, Wayne , Extending the Value of Your Data Warehousing Investment, The Data Warehouse Institute.,2007
12. WHO, (2011). Global Health and Ageing, National Institute on Ageing and National Institute of Health, U.S. Department of Health and Human Services.
13. Ravi B. Parikh, Ziad Obermeyer, David Westfall Bates, <http://hbr.org/2016/04/making-predictive-analytics-a-routine-part-of-patient-care>
14. E. McCann, "Kaiser reports second fall data breach," Healthcare IT News, 2013.
15. P. Institute, "Third Annual Benchmark Study on Patient Privacy and Data Security," Pokémon Institute LLC, 2012.
16. P. Groves, B. Kayyali, D. Knott and S. V. Kuiken, "The 'big data' revolution in healthcare," McKinsey & Company, 2013.
17. P. Middleton, P. Kjeldsen and J. Tully, "Forecast: The Internet of Things, Worldwide," Gartner, 2013
18. Linda A. Winters-Miner, PhD Posted on 6 October 2014<https://www.elsevier.com/connect/seven-ways-predictive-analytics-can-improve-healthcare>
19. Priyanka Prof Nagarathna Kulennavar,. "A Survey On Big Data Analytics In Health Care International Journal of Computer Science and Information Technologies, Vol. 5 (4) , pp.5865-5896,2014
20. Jaslene 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, pp, 430-435 April 2016
21. Raol Priyanka Ajaysinh, Hinal Somani, "A Survey on Machine learning assisted Big Data Analysis for Health Care Domain" IJEDR |Volume 4, Issue 4, pp-550-554,2016
22. Dharavath Ramesh, Pranshu Suraj, Lokendra Saini, "Big data Analytics in healthcare: A Survey Approach", IEEE, pp. 1 – 6, DOI: 10.1109/MicroCom.2016.7522520,2016
23. R. Sathiyavathi," A Survey: Big Data Analytics on Healthcare System", Contemporary Engineering Sciences, Vol. 8, no. 3,pp- 121 – 125,2015