Research Journal of Computer and Information Technology Sciences ISSN 2320 – 6527
Vol. **5(7),** 5-7, September **(2017)** *Res. J. Computer and IT Sci.*

Short Communication

Research scopes in health care analytics

Kauleshwar Prasad[1*], Arpana Rawal[1] and Ani Thomas[2]
[1]Department of Computer Science and Engineering, Bhilai Institute of Technology, Durg, India
[2]Department of Information Technology, Bhilai Institute of Technology, Durg, India
kauleshwarprasad@gmail.com

Available online at: www.isca.in
Received 10th April 2017, revised 12th September 2017, accepted 19th September 2017

Abstract

In the field of health care practices and research, big data analytics are playing a vital role. It has provided tools to collect, grip, scrutinize, and integrate large volumes of dissimilar, structured, and unstructured data produced by current healthcare systems. Big data analytics has been recently applied towards aiding the process of care delivery and disease exploration. This paper focuses on introduction of big data and its characteristics like volume, velocity, variety, veracity and value. It also focuses on different types of analytics in health care and life science. Various applications of big data in health are also discussed.

Keywords: Volume, Velocity, Variety, CPOE, EPR, HER.

Introduction

Big Data Analytics (BDA) is the process of extracting knowledge from sets of Big Data. The healthcare industry is highly data intensive. Worldwide digital healthcare data was estimated to be equal to 500 petabytes (10^{15} bytes), and is expected to reach 25 exabytes (10^{18} bytes) in 2020. Global growth in data will become critical issues to clinical applications and data management in health care organization. Since the growth of data is so fast, large, complex and distributed that it is difficult to manage, maintain and analyze the data by using common database management system and traditional data analysis methods. For this Big data analytics will be helpful. The summation of patient's data and wellbeing shows the application of big data in healthcare and life science. This includes clinical data from Computerized Physician Order Entry and clinical decision support systems, data in electronic patient records, machine generated/sensor data, Social data and information related to patient. The power of Big Data analytics in life sciences and healthcare lead to improve care of patients at lower costs and save lives. As per the data taken from CRISIL report on Hospital industry in India, it has been found that Us is very rich in building secondary and tertiary care hospital and the cost is around 25 lakhs. For effective and proper caring it requires highly skilled doctors, nurses, lab technicians, pharmacists etc. India is lacking in both capital investment and manpower. We generally see that there is shortage of beds having only 0.7 per 1000 of population. As per the data given by WHO the average ratio is 1 bed per 1000[1].

Health care analytical realms

Clinical data analytics: Clinical data contains information about patients, such as diagnosis, treatments administered, medicines prescribed, procedures and lab tests conducted, and hospitalizations. Third parties can be provided data and services by synthesizing and aggregating patients clinical records. Mobile applications can be used for the management and care of patients and outcomes can be easily increased[2]. Monitoring medical devices, including wearables, to capture and analyze in real-time large volumes of fast-moving data, for safety monitoring and adverse event prediction, enabling payers to monitor adherence to drug and treatment regimens and detect trends that lead to individual and population wellness benefits.

Medical image analytics: Medical imaging gives large volume of information on structure and functioning of organ. It also gives information about the states of diseases. Moreover, it is used for explanation of organ, identifying tumors in lungs, diagnosis of deformity in spinal, detection of artery stenosis, detection of aneurysm, and so on. In these applications, image processing techniques such as enhancement, segmentation, and denoising in addition to machine learning methods are employed[1]. For the analysis of medical imaging following challenges arises: i. Preprocessing: It is a method of noise reduction, artifact removal; missing data handling, contrast adjusting. ii. Compression: It is a method of reducing the volume of data. iii. Parallelization / real time realization: Developing scalable/parallel methods and frameworks to speed up the analysis/processing. iv. Registration/mapping: Aligning consecutive slices/frames from one scan or corresponding images from different modalities. v. Sharing/security/ anonymization: Integrity, privacy, and confidentiality of data must be protected. vi. Data integration / mining: Finding dependencies/patterns among multimodal data and/or the data captured at different time points in order to increase the accuracy of diagnosis, prediction, and overall performance of the System. vii. Validation: Assessing the performance or accuracy of the system/method. Validation can be objective or subjective. For the former, annotated data is usually required.

Medical signal analytics: Analysis of continuous waveform (signal varying against time) and related medical record information developed through applied analytical disciplines are called Medical Signal Analytics.

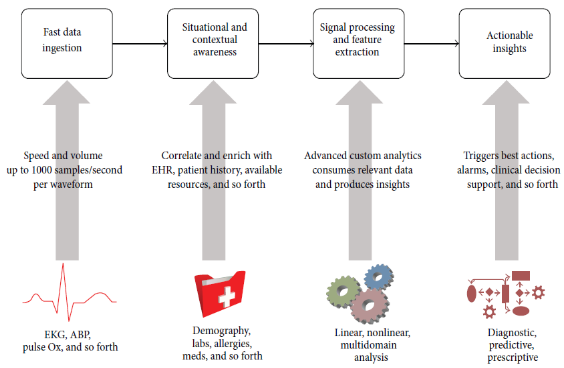

Figure-1: Medical Signal Analytics[1].

In Figure-1, it has been demonstrated by Ashwin Belle et al.[1] that analytics can be done on real time streaming waveforms in health care fields. Firstly data acquisition and ingestion is done on large and high speed waveform data. After this, dynamic waveform is integrated with static data of EHR. After the extraction of features the result is inputted to machine learning model to produce an actionable insight. These actionable insights could either be diagnostic, predictive, or prescriptive. These insights could further be designed to trigger other mechanisms such as alarms and notification to physicians.

Genomics: Genomics is an interdisciplinary branch which focuses on genomes. It can be said as branch of biology which discuss on function, structure and evolution of genomes. Genome analysis is the process of analyzing various complex diseases.

Major milestones

Various breakthroughs are already on the roads of success in performing big data analytics taking enormous streaming digital datasets from health care information systems. The put forth survey attempts to show case major highlighted initiatives taken up across the globe as described in the following survey reports.

Karamjit Kaur et al.[3] proposed extended version of A Smart Polyglot big data solution in Healthcare. Polyglot persistence is the amalgamation of different databases within an healthcare application. Polyglot programming, which refers to software that uses more than one programming language. An impedance mismatch exists between various data stores in terms of data structures, granularities, query languages, and backup and restore strategies, the challenge for polyglot-persistent solutions is to execute queries that require data from more than one data store to give a complete answer. Polyglot HIS, is a solution for managing healthcare data that can choose the appropriate data store based on the type of data it handles best, thus achieving overall performance improvements.

LO'AI A. Tawalbeh et al.[4] proposed a work on Mobile Cloud Computing Model and Big Data Analysis for Healthcare Applications. Variety of useful tasks such as scheduling meetings, ordering food, booking flights, buying cars online, real-time navigation, etc. are performed by people in daily life. It shows wide application of mobile devices in our daily life. Mobile cloud computing maximizes the utilization of mobile devices capabilities to run intensive-computing applications. These intensive jobs are executed in the mobile cloud computing infrastructure overcoming the mobile device limitations, saving energy, and providing better throughput. In the MCC model the cloudlets are placed nearby the hospital and cover an area that can be accessed by authorized people who can access the patients' information and follow their status remotely. Here big amount of patients data being generated and need to be analyzed.

Wells et al.[5] proposed the idea of leveraging big data in population health management sector. In this context big data sets were exploited under a population health program for individuals. This study oriented to prove upon the hypothesis that older adults who purchases medigap coverage live healthier lives. The policy plan enrollees had an access to a wellness program, holistic care coordination programs, two telephone-based advice lines, concierge support for insurance and medical care needs, and a program designed to help reduce unnecessary emergency room (ER) visits. Program management, reporting, and evaluation processes generated additional data which, when analyzed, continues to refine program implementation and quality. Future improvements to this program may include enhanced integration of social service programs that will generate their own data streams for analyses designed to further improve health and wellbeing.

Anwaar Ali et al.[6] further emphasized that Big Data For Developmental initiatives (BD4D) should be explored to resolve humanitarian emergencies in top priority. These emergencies can be perceived in health care sector at first glance. According to a healthcare update reported by Stefano Palomba et al.[7] upon Pregnancy complications in women with Poly Cystic Ovary Syndrome (PCOS) and subsequent health outcomes in the new born children. These complications are found to be governed by various clinical and biomedical parameters which if regularly monitored during gestation periods till the time of labor, may reduce the risk of obstetric and neonatal complications. Similar to the mentioned context, a case study was reported by Ogwang S. et al.[8] on the use of partogram, a universal tool for monitoring labour. The study was done after taking input data survey from eight health units of Rukungiri District, Uganda, South Africa.

However, the reports inferred only 30% use of the partograms and concluded that only 2% of partographs sufficed the complete monitoring of foetal heart rate due to the incomplete tracking mechanisms on gestation parameters for pregnant women with partograms.

Conclusion

Today big data health care monitoring projects have been successfully deployed as a part of either government or NGO big data analytics initiatives. These projects have leveraged legions of disparate, structured, and unstructured data sources that play a vital role in practicing intelligent healthcare information systems in near future. One can also already see a spectrum of application domains where the analytics is being utilized viz. natural disasters, hunger, poverty, agriculture, healthcare, migrant crisis, education and finance. The analytical tools aiding in the decision making and performance of healthcare personnel and patients can again be classified into different types. One possible scope of research can be to monitor the prognosis periods of pregnant patients by keeping track of their demographic, geographic, clinical, pathological, gestation and labor parameters so that they can be navigated to suitable health care unit in nearest vicinity at the time of delivery emergencies.

References

1. Belle Ashwin, Thiagarajan Raghuram, Soroushmehr S.M., Navidi Fatemeh, Beard Daniel A. and Najarian Kayvan (2015). Big Data Analytics in Healthcare. *BioMed Research International*, 1-17.

2. Reisner A.T., Khitrov M.Y., Chen L., Blood A., Wilkins K., Doyle W., Wilcox S., Denison T. and Reifman J. (2013). Development and Validation of a Portable Platform for Deploying Decision-Support Algorithms in Prehospital Settings. *Applied clinical informatics*, 4(3), 392-402.

3. Kaur Karamjit and Rani Rinkle (2015). A Smart Polyglot Solution for Big Data in Healthcare. *IT Pro,* IEEE, 17(6), 48-55.

4. Lo'ai A. Tawalbeh, Mehmood Rashid, Benkhlifa Elhadj and Song Houbing (2016). Mobile Cloud Computing Model and Big Data Analysis for Healthcare Applications. *IEEE*, 4, 6171-6180.

5. Wells Timothy S., Ozminkowski Ronald J., Hawkins Kevin, Bhattarai Gandhi R. and Armstrong Douglas G. (2016). Leveraging big data in population health management. *Big Data Analytics*, 1-14.

6. Ali Anwaar, Qadir Junaid, Rasool Raihan ur, Sathiaseelan Arjuna, Zwitter Andrej and Crowcroft Jon (2016). Big data for development: applications and techniques. *Big Data Analytics*, 1-24.

7. Palomba Stefano, Wilde Marlieke A De, Falbo Angela, Koster Maria P.H., Sala Giovanni Battista La and Fauser Bart C.J.M. (2015). Pregnancy complications in women with polycystic ovary syndrome. *Human Reproduction*, 21(5), 575-592.

8. Ogwang Simon, Karyabakabo Zepher and Rutebemberwa Elizeus (2009). Assessment of partogram use during labour in Rujumbura Health Sub District, Rukungiri District, Uganda. *African Health Sciences*, 9(1), 27-34.

9. Altena Allard J., Moerland Perry D., Zwinderman Aeilko H. and Olabarriaga Sílvia D. (2016). Understanding big data themes from scientific biomedical literature through topic modeling. *Journal of Big Data* , 3, 23.

10. Toga Arthur W. and Dinov Ivo D. (2015). Sharing big biomedical data. *Journal of Big Data* , 2, 7.

11. Fitzhenry Fern, Resnic F.S., Robbins S.L., Denton J., Nookala L., Meeker D., Ohno-Machado L. and Matheny M.E. (2015). Creating a Common Data Model for Comparative Effectiveness with the Observational Medical Outcomes Partnership. *Applied Clinical Informatics*, 6(3), 536-547.

12. He Hui, Du Zhonghui, Zhang Weizhe and Chen Allen (2016). Optimization strategy of Hadoop small file storage for big data in healthcare. *J Supercomput*, 72(10), 3696-3707.

13. Ma Linh Van, Kim Jisue, Park Sanghyun, Kim Jinsul and Jang Jonghyeon (2016). An efficient Session_Weight load balancing and scheduling methodology for high-quality telehealth care service based onWebRTC. *J Supercomput*, 72(10), 3909-3926.

14. Rodríguez-Mazahua Lisbeth, Rodríguez-Enríquez Cristian-Aarón, Sánchez-Cervantes José Luis, Cervantes Jair, Luis García-Alcaraz Jorge and Alor-Hernández Giner (2016). A general perspective of Big Data: applications, tools,challenges and trends. *J Supercomput,* 72(8), 3073-3113.

15. Kuo Mu-Hsing, Chrimes Dillon, Moa Belaid and Hu Wei (2015). Design and Construction of a Big Data Analytics Framework for Health Applications. *IEEE*, 631-636.