Big Data Analytics in Healthcare Industries

M. Anita Rajkumar
Assistant ProfessorAnna Adharsh College for Women, Anna Nagar, Chennai

Abstract— The healthcare industry has large amounts of data, driven by record keeping, medicine details, new era of medical needs, compliance & regulatory requirements, other stakeholders in the healthcare delivery and patient care. While most data is stored in paper records, the digitization of these large amounts of data would be the first and foremost activity.

Keywords— *Big Data analytics; Electronics Record; Genomic Analytics; Veracity* .

1. Introduction

By definition, large amount of data in healthcare refers to the above electronic health data sets which are so large and complex. They are difficult to manage with traditional software and hardware. Big data in healthcare is blooming because of its volume and diversity of data types and the velocity at which it must be managed.

The enormous amount of data related to patient healthcare and well-being make up "big data" in the healthcare industry. It includes,

- clinical information and clinical decision support systems (Doctor's written notes and prescriptions, medical imaging, laboratory test results, medicine records from pharmacy, patient's insurance coverage, and other administrative data)
- Patient's information in electronic patient records
- Machine generated/sensed data, such as from monitoring vital signs.
- Doctor's conference details, Social media posts, status updates on Face book and other platforms, and web pages.
- Less patient-specific information, including emergency care data, news feeds, and articles in medical journals.

By discovering the associated details, understanding trends and recent activities within the data, Big data analytics in health care can improve care, save lives and lower costs. Thus, big data analytics applications in healthcare can make better informed decisions, and can be considered as a research category.

This article will provide an overview of big data analytics in healthcare as a discipline.

- Definition of advantages and characteristics.
- Volume, Velocity, Variety and Veracity of Data in Healthcare industry.
- Big Data - Complex Architectural Framework Overview.
- Examples of big data analytics in healthcare.
- Challenges identified
- Conclusions and future directions

Big data would be the next era under IT service management.

2. Advantages to Healthcare

By moving to globalization or by making the information worldwide by combining and effectively using big data, healthcare organizations ranging from small team of physician dispensaries and multi-provider groups to large hospital networks and accountable care organizations stand to realize significant benefits .

Potential benefits with moving to Big data include,

- Detecting diseases at earlier stages when they can be treated more easily and effectively;
- Managing specific individual and population health and detecting health care fraud more quickly and efficiently.
- Numerous questions can be addressed with big data analytics.
- Certain developments or outcomes may be predicted and/or estimated based on vast amounts of historical data such as
  - length of stay (LOS)
  - patients who will choose elective surgery
  - patients who likely will not benefit from surgery
  - complications
  - patients at risk for medical complications
  - Illness/disease progression patients at risk for advancement in disease states
  - Causal factors of illness/disease progression

As healthcare reimbursement models are changing, meaningful use and pay for performance are emerging as critical new factors in today's healthcare environment.

3. Advantage in Clinical Operations

By digitizing the data into Big data, The Comparative effectiveness research to determine more clinically relevant and cost-effective ways to diagnose and treat patients with various expert advices and with past experiences.

4. Research & Development

- Predictive modeling to produce an R & D pipeline in drugs and devices
- Statistical tools and algorithms to improve better

match treatments to individual patients and speeding new treatments to market
- Analyzing clinical trials and patient records to identify follow-on indications and discover adverse effects before products reach the market.

In addition, Big data analytics in healthcare can contribute to improve in the following daily routines of the physician.

- Evidence-based medicine
- Genomic analytics
- Pre-adjudication fraud analysis
- Device/remote monitoring
- Patient profile analytics

Data Analytics would be the vast area for research and development. Next section will cover the 4Vs of Big Data Analytics.

5. Volume, Velocity, Variety and Veracity of Data in Healthcare industry

5.1 Volume of Data

Over time, health-related data will be created and accumulated continuously, resulting in an incredible *volume* of data. The already daunting volume of existing healthcare data includes personal medical records, radiology images, clinical trial data FDA submissions, human genetics and population data genomic sequences, etc.

5.2 Velocity of Data

Due to the various advances in data management, particularly virtualization and cloud computing, data is accumulated in real-time and at a rapid pace, or *velocity*.

5.3 Variety of Data

Most healthcare data are paper files, x-ray films, and scripts. Velocity of mounting data increases with data that represents regular monitoring, such as multiple daily diabetic glucose measurements, blood pressure readings, and EKGs.

5.4 Veracity of Data

Data Assurance will be the next goal/challenge for the medical industry as the Data quality issues are of acute concern in healthcare. Main causes for the above un-achieved goal is as below.

Life or death decisions depend on having the accurate information Quality of healthcare data as they cannot be structured is highly variable and all too often incorrect. (Inaccurate "translations" of poor handwriting on prescriptions are perhaps the most infamous example). The constant flow of new data accumulating at unpredicted velocity with unidentified types presents new challenges. In the next section, Big data framework has been discussed.

6. Big Data - Complex Framework in Big Data Analytics

In a normal health care analytics project, the analysis can be performed with a business Intelligence (BI) tool installed on stand-alone systems. But, Big data processing is broken down and executed across multiple nodes. By moving to Big data, health care providers can gain insights from their large data repositories of various health related decisions. As the technological and conceptual architecture of Big data, such as Hadoop / Map Reduce are available on the cloud, industries have encouraged the application of big data analytics in healthcare. Big data analytics tools are extremely complex, programming intensive, and require the application of a variety of skills. As Figure 1indicates, the complexity begins with the data itself.

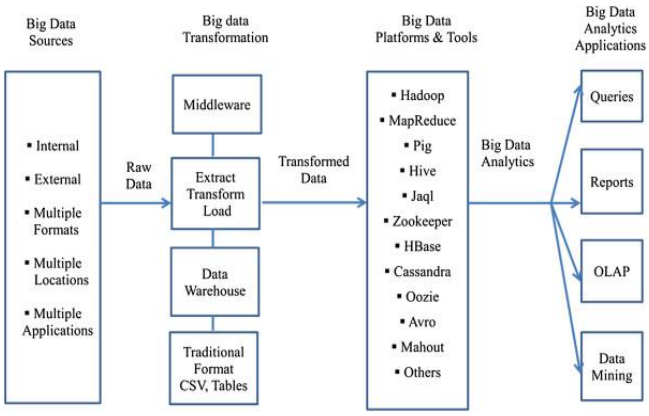

Fig.1: Big Data

7. Examples of big data analytics in healthcare

U.S. healthcare alliance network, has more than 2,700 members, hospitals and health systems, 90,000 non-acute facilities and 400,000 physicians and is reported to have data on approximately one in four patients discharged from hospitals. As the data suggested, the network has assembled a large database of clinical, financial, patient, and supply chain data, with which the network has generated comprehensive and comparable clinical outcome measures, resource utilization reports and transaction level cost data.

These outputs have informed decision-making and improved the healthcare processes at approximately 330 hospitals, saving an estimated 29,000 lives and reducing healthcare spending by nearly $7 billion.

North York General Hospital, a 450-bed community teaching hospital in Toronto, Canada, reports using real-time analytics to improve patient outcomes and gain greater insight into the operations of healthcare delivery.

North York is reported to have implemented a scalable real-time analytics application to provide multiple perspectives, including clinical, administrative, and financial. The next section briefly identifies some of the key challenges in big data analytics in healthcare.

8. Challenges Identified

A Big data analytics platform in healthcare must support the key functions necessary for processing the data. The criteria for platform evaluation may include availability, continuity, ease of use, scalability, ability to manipulate at different levels of granularity, privacy and security enablement, and quality assurance.

In addition, while most platforms currently available are open source, the typical advantages and limitations of open source platforms apply.

To succeed, big data analytics in healthcare needs to be packaged so it is menu-driven, user-friendly and transparent. Real-time big data analytics is a key requirement in healthcare. The lag between data collection and processing has to be addressed. The dynamic availability of numerous analytics algorithms, models and methods in a pull-down type of menu is also necessary for large-scale adoption. The important managerial issues of ownership, governance and standards have to be considered. And woven through these issues are those of continuous data acquisition and data cleansing. Health care data is rarely standardized, often fragmented, or generated in legacy IT systems with incompatible formats. This great challenge needs to be addressed as well.

9. Conclusions and Future Directions

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 highlighted above, 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.

Reference

[1] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4341817/
[2] https://www.sas.com/content/dam/SAS/en_us/doc/research2/big-data-analytics-105425.pdf
[3] Vinesh Kumar and Prof.Jyant Shekhar, Compendious Sensible Data Structures for Top-p Completion in Big Data Store, Special Issue of International Journal of Linguistics and Computational Applications, Proceedings of the UGC Sponsored - National Conference on "Recent Statistical Computing Techniques and their Applications", March 11-12, 2016 pp. 9-20
[4] Aneeshkumar A.S and Jothi Venkateswaran C, "Estimating the Surveillance of Liver Disorder using Classification Algorithms", International Journal of Computer Application, Volume 57, Issue 6, pp.39-42, November 2012.
[5] M.Gomathi and, V.Sujatha, Information Security Issues in Big Data and Data Mining, Special Issue of Engineering and Scientific International Journal pp.48-5, May 2016.
[6] Aneeshkumar A.S and Jothi Venkateswaran C, "A novel approach for Liver disorder Classification using Data Mining Techniques", Engineering and Scientific International Journal, Volume 2, Issue 1, pp.15-18, March 2015.
[7] Dr. Arpita Aggarwal, Purnima Khurana, Ishan Rathi and Kashish Singh, Big Data : Real Challenge, Special Issue of Engineering and Scientific International Journal, March 11-12, 2016, pp.44-49.
[8] Jimeng Sun and Chandan K. Reddy, Big Data Analytics for Healthcare, Tutorial presentation at the SIAM International Conference on Data Mining, Austin, TX, 2013, pp.1-112.
[9] Samiya Khan, Kashish Ara Shakil and Mansaf Alam, Cloud-Based Big Data Analytics – A Survey of Current Research and Future Directions, https://arxiv.org/ftp/arxiv/papers/1508/1508.04733.pdf.
[10] https://www.ericsson.com/res/docs/whitepapers/wp-big-data.pdf