IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Unconventional Tasks and Challenges Faced by Big Data in Medical Science

Sukhpreet Singh[1], Manjeet Kaur [2], Inderjeet kaur[3]

Research Scholar, M. Phil (Comp. Sc), Punjabi University, Patiala, Punjab, India[1,3]

Dept. of Comp. Science, Gurur nanak college of management & technology (GNCMT), Punjab, India[2]

Abstract: Big Data can be combined with new technology to bring about positive conversion in the health care segment. A technology aimed at making Big Data analytics a certainty will act as a key element in transforming the way the health care industry operates today. The study and analysis of Big Data can be used for tracking and managing population health care effectively and efficiently. This also plays a vital role in delivering preventive care. Health care will change a great deal as it becomes a data-driven industry. But the size of the data, the speed at which it's growing and the threat it could cause to individual privacy mean mastering it is one of biomedicine's most critical challenges. Hiding within this huge amount of data is knowledge that could change the life of a patient. Health information that can be extracted from patient records is another rich resource of data. In my research I would explore how big data can enhance healthcare segment in country like India which has a huge population and huge medical complexities. I addition to this I would also explore problems faced by big data, obstacles in using big data in the health industry, how big data analytics can take health care to a new level by enhancing the overall quality of patient care. Today, a significant proportion of the cost and time spent in the drug development process is attributable to unsuccessful formulations. By enabling researchers to identify compounds with a higher likelihood of success, Big Data can help reduce the cost and the time to market for new drugs. Also, by integrating learning from medical data in the early stages of development, researchers will now be able to customize drugs to suit aggregated patient profiles.

Keywords: Big Data, framework, medical science, EMR, HIS.

INTRODUCTION

Big data [1, 2] are rapidly all over the place. Everyone seems to be collecting, analyzing, and making money from it. No matter whether we are talking about analyzing zillions of Google search queries to predict flu outbreaks, or zillions of phone records to detect signs of terrorist activity, or zillions of airline stats to find the best time to buy plane tickets, big data are on the case.
By combining the power of modern computing with the enormous data of the digital era, it promises to solve virtually any problem like crime, public health, the evolution of grammar, etc.

The goal of big data management is to ensure a high level of data quality and accessibility for business intelligence and big data analytics applications. Corporations, government agencies and other organizations employ big data management strategies to help them contend with fast-growing pools of data, typically involving many terabytes or even petabytes of information saved in a variety of file formats. Effective big data management helps companies locate valuable information in large sets of unstructured data and semi-structured data from a variety of sources, including call detail records, system logs and social media sites.

Most big data environments go beyond relational databases and traditional data warehouse platforms to incorporate technologies that are suited to processing and storing non-transactional forms of data. The increasing focus on collecting and analyzing big data is shaping new platforms that combine the traditional data warehouse with big data systems in a logical data warehousing architecture. As part of the process, it must be decided what data must be kept for compliance reasons, what data can be disposed of and what data should be kept and analyzed in order to improve current business processes or provide a business with a competitive advantage. This process requires careful data classification so that ultimately, smaller sets of data can be analyzed quickly and productively.

ISSUES RELATED WITH BIG DATA CHARACTERISTICS

Data Volume – With the increase in volume, the worth of different data records will decrease in proportion to age, type, richness, and quantity among other factors.

Data Velocity – Data is being generated at tremendous speed with each minute passing. The velocity at which this data is being generated is beyond the handling power of traditional systems.

Data Variety - Mismatched data formats, non-aligned data structures, and inconsistent data semantics represents significant challenges that can lead to analytic collapse.

Data Value – Often it is witnessed that there is a huge gap in between the business leaders and the IT professionals. The main concern of business leaders is to just add value to

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

their business and to maximize their profit. On the other hand, IT leaders deal with technicalities of the storage and processing.

Data Complexity - Data scientists have to link, match, cleanse and transform data across systems coming from various sources. It is also necessary to connect and correlate relationships, hierarchies and multiple data linkages or data can quickly spiral out of control [3].

Data Veracity - Veracity refers to the messiness or trustworthiness of the data. With many forms of big data quality and accuracy are less controllable. [4 , 10].

BIG DATA IN MEDICAL SCIENCE

Big data in health care is different from big data in marketing and product development because of regulation medical ethics, privacy and the diversity of data source. Goals differ as well understanding these differences will be important to unlocking its hidden value. Health data volume is expected to grow dramatically in the years ahead. Although profit is not and should not be a primary motivator, it is vitally important for healthcare organization to acquire the available tools, infrastructure, and techniques to leverage big data effectively or else risk losing potentially millions of dollars in revenue and profit Health care is one of the top social and economic issues in many countries, such as the India, the UK, South Korea, the United States and even middle income countries.

In India, Health care Sector Suffers from Underfunding and bad governance. No doubt, India has made huge improvements since independence; majority (70%) of the effort has been led by the private sector still India accounts for 21% of the world's burden of disease. The term big data [1, 2] refers to the collection of data sets so large and complex that it becomes difficult to process using readily available database management tools. In actual, big data refers to the situation where more and more portions and objects of everyday life are available in digital form, like personal profiles or company profiles, social network and blog postings, health records etc. through which huge amount of data gets dynamically produced particularly on the Internet and on the Web.

Unstructured data forms close to 80% of information in the healthcare industry and is growing exponentially. Getting access to this unstructured data such as output from medical devices, doctor's notes, lab results, imaging reports, medical correspondence, clinical data, and financial data is an invaluable resource for improving patient care and increasing efficiency.

In the last few years there has been a move toward evidence-based medicine, which involves making use of all clinical data available and factoring that into clinical and advanced analytics. The outcomes of this movement include improved ability to detect and diagnose diseases in their early stages, assigning more effective therapies based on a patient's genetic makeup, and adjusting drug doses to minimize side effects and improve effectiveness.

ROLE PLAYED BY BIG DATA IN MEDICAL SCIENCE

These are some example use cases that illustrate how big data is being used in healthcare, helping to increase Efficiency and improve patient care.

Personalized Treatment Planning

Personalized treatment planning is a way to customize treatment for a patient to continuously monitor the effects of medication. The dose can be modified or the medication can be changed based on how the medication is working for that particular individual.

Assisted Diagnosis

Big data being able to access a broad combination of Knowledge across multiple data sources aids in the accuracy of diagnosing patient conditions. Assisted diagnosis is accomplished using expert systems that contain detailed knowledge of conditions, symptoms, medications and side effects.

Fraud Detection

Healthcare organizations need to be able to detect fraud based on analysis of anomalies in billing data, procedural benchmark data or patient records. For example, they can analyze patient records and billing to detect anomalies such as a hospital's over utilization of services in short time periods, patients receiving healthcare services from different hospitals in different locations simultaneously, or identical prescriptions for the same patient filled in multiple locations.

Monitor Patient Vital Signs

Healthcare facilities are looking to provide more proactive care to their patients by constantly monitoring patient vital signs. The data from these various monitors can be used in real time and send alerts to nurses or care providers so they know instantly about changes in a patient's condition.

Digitization of Data

Till date most of the data in the health care industry are stored in the form of hard copy, but the current trend is toward rapid digitization of these large amounts of data. The medical community has accepted big data as a research tool and beliefs that the society's enormous amount of diverse health information has the potential to help solve some of medicine's most troublesome problems. By discovering relations and understanding patterns within the data, big data analytics has the potential to improve care, save lives and lower costs [5, 11, 13].

TECHNOLOGY USED BY BIG DATA

Big Data needs a framework for running applications on large clusters of commodity hardware which produces huge data and to process it. One such framework is Hadoop. Hadoop includes two main components. First one is HDFS (Hadoop Distributed File System) and the second one is

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

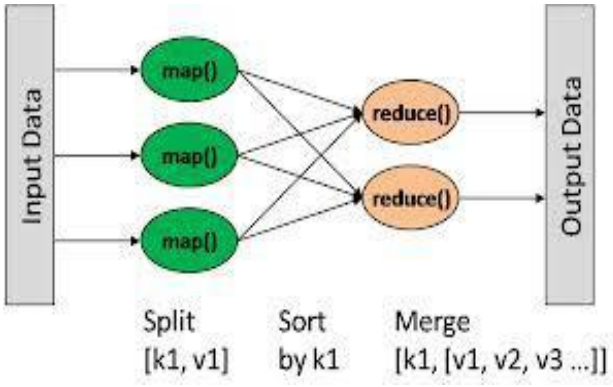
Map/Reduce technology. The process starts with a user request to run a MapReduce program and continues until the results are written back to the HDFS . As MapReduce is an algorithm, it can be written in any programming language.

Hadoop map reduce works in three stages:

First Stage: mapping: In this stage, a list of elements is provided to a 'mapper' function to get it transferred into pairs. The mapper function does not modify the input data, but simply returns a new output list.

Intermediate stages: Shuffling and Sorting: After the mapping stage, the program exchanges the intermediate outputs from the mapping stage to different 'reducers'. This process is called shuffling.

Final Stage: Reducing: In the final reducing stage, an instance of a user-provided code is called for each key in the partition assigned to a reducer. In particular, we have one output file per executed reduce task



IMPACT OF BIG DATA ON HEALTH CARE SYSTEM

RIGHT LIVING - The right-living pathway focuses on encouraging patients to make lifestyle choices that help them remain healthy.

CORRECT CARE - It involves ensuring that patients get the timely and appropriate treatment available.

ACCURATE PROVIDER - It proposes that patients should always be treated by high-performing professionals that are best coordinated to the task and will achieve the best outcome.

PRECISE VALUE - To fulfill the goals of precise value, providers will continuously enhance healthcare value while preserving or improving its quality [6]. *RIGHT*

INNOVATION - It involves the identification of new therapies and approaches to delivering care, across all aspects of the system, and improving the innovation engines themselves.

SECURITY ISSUES IN BIG DATA

A major challenge to healthcare cloud is the security threats including tampering or leakage of sensitive patient's data on the cloud, loss of privacy of patient's information, and the unauthorized use of this information.

The main security and privacy requirements for healthcare clouds are discussed below

Authentication: In a healthcare cloud, both health care information offered by CSPs (cloud service providers) and identities of users should be verified at the entry of every access using user names and passwords assigned to users by CSPs.

Authorization: It is a crucial security requirement that is used to control access priorities, permissions and resource ownerships of the users on the cloud.

Non-repudiation: It implies that one party of a transaction cannot deny having received a transaction nor can the other party deny having sent a transaction.

Integrity and Confidentiality: Integrity means preserving the precision and consistency of data. In the healthcare system, it refers to the fact that EHRs (electronic health records) have not been tampered by unauthorized use.

Availability: For any EHR system to serve its purpose, the information must be available when it is needed. High availability systems aim to remain available at all times, preventing service disruptions due to power outages, hardware failures, and system upgrades.

CONCLUSION

The real issue is not that we are acquiring large amounts of data. It's what you do with the data that counts. Today, a significant proportion of the cost and time spent in the drug development process is attributable to unsuccessful formulations.

By enabling researchers to identify compounds with a higher likelihood of success, Big Data can help reduce the cost and the time to market for new drugs. Also, by integrating learning from medical data in the early stages of development, researchers will now be able to customize drugs to suit aggregated patient profiles.

Second, for medical data that cuts across departmental boundaries, a top-down approach is needed to effectively manage and integrate big data. Governments and healthcare stakeholders should establish big data control towers to integrate accumulated datasets, whether structured or unstructured, from each silo.

Third, real-time analysis of in-motion big data should be carried out, while protecting privacy and security. Fourth, leading big data governments appear to have different goals and priorities; therefore, they use different sets of data management systems, technologies, and analytics.

Finally, this study is limited in that the practical applications of big data for investigating healthcare issues have not yet been fully demonstrated due to the dearth of practice.

DOI 10.17148/IJARCCE.2016.5321

REFERENCES

[1]. http://radar.oreilly.com/2012/01/what-is-big-data.html
[2]. www.forbes.com/sites/lisaarthur/2013/08/15/what-is-big-data/ http://www.business2community.com/digital-marketing/4-vs-big-data-digitalmarketing- 0914845#!bgCHyQ
[3]. http://dashburst.com/infographic/big-data-volume-variety-velocity/
[4]. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3848064/
[5]. http://www.informationweek.in/informationweek/cio-blog/175124/health carecompulsion- choice
[6]. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3717441/ http://www.datanami.com/2013/07/19/big_data_emerges_in_indian_health_care/
[7]. http://eandt.theiet.org/magazine/2013/03/journey-to-the-centre-of-big-data.cfm
[8]. http://itbussinessconsulting.jimdo.com/2014/08/11/how-big-data-and-analytics-canreshape- healthcare/
[9]. http://www.hissjournal.com/content/2/1/

DOI 10.17148/IJARCCE.2016.5321