International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

Healthcare Industry Data Enhancement using Bigdata Analytics

Anupama M[1], K. Raghuveer[2]

[1]*Student, MTech-Computer Network and Engineering, Department of IS&E, The National Institute of Engineering, Mysuru, Karnataka, India*
[2] *Professor, Department of IS&E, The National Institute of Engineering, Mysuru, Karnataka, India*

Abstract - *In the healthcare industry different sources of raw data includes hospital records, medical records of patients, results of medical examinations and devices that are a part of internet of things. The Biomedical researches are also generating a signifcant portion of huge data that is applicable to public healthcare. This data requires proper management and analysis in order to derive meaningful information whenever it's required. Seeking solution by analyzing the huge amount of data quickly becomes an easiest and cost effective task. An efficient management, analysis, and interpretation of this huge data can change the avenues for modern healthcare. This research proposing a framework as Big data analytics in health care. It provides the ability to store, explore, aggregate, model and analyze the contents from different sources of the data repositories.*

Key Words: Bigdata, Metadata, Bigdata Analytics, Data visualisation, Data storage

1. INTRODUCTION

Healthcare industry like hospitals and clinics will ensure the health condition of patients based on the continuous and consistent monitoring of the Records of the patient generated Previously. So taking care of the Patient Historical health record is as important as maintaining, the Current Data generated by them. Analyze services provides a framework for ingesting and managing information, executing ETL's and analytics applications and rapidly envisioning retrospective, prospective, predictive, and prescriptive data. Hosts HealthSuite Insights that offers analytics asset creation, deployment, and support, By this Consultant will come to know the problems that can occur in advance based on the insight of data so that they can take relevant actions to prevent the problem even before it occurs and can provide a quick solution to the problem if it occurs so that it would affect to the patients treatement. This Big data Platform Enables the development of big data analytics solutions to derive predictions and meaningful insights. Decreases the complexities of building and maintaining distributed big data computing environments by giving a common set of tools to ingest, transform, and extract data for advanced data visualization. This will noticeably reduces the service cost, Storage and provides Security services.

2. Problem Statement

In health care system, Generation of data increases rationally day by day leads massive amount of data collection but provide storage to this data and maintaining it's confidentiality with a secure mechanism ,processing large volume of data is a critical problem. To overcome from this complex scenario we come up with a framework that is Big data Analysis

Solution:

The first step towards resolving the problem is to find the root cause of the it, as the name suggests, 'big data' represents large amounts of data that is unmanageable using traditional software or internet-based platforms. It outperforms the traditionally used amount of storage, processing and analytical power. The Drawbacks of Existing System are:

1. Huge amounts of the data decreases the capacity and ability of traditional data storage, data management and data retrieval systems such as data warehouses.
2. Health care data today come in many formats, such as the structured and free-text data ,diagonize images and data streaming from social media and mobile applications
3. Most traditional health infrastructures are not able to process and analyze large amounts of freshly generated different formats of data in real time.

3. RELATED WORK

The Proposed system consist of an frame work called Big data analytics ,which helps to collect the huge data from different sources in different form locations in a secure way and converts it into a standard format that can be then used by the respective Analysis and processing to handle the further process.

Big data can assist patients with settling on the correct choice in a convenient way. From patient data, analytics can be applied to distinguish people that need "proactive care" or need change in their way of life to avoid health condition degradation. From quiet information, examination can be applied to distinguish people that need "proactive consideration" or need change in their way of life to evade wellbeing condition debasement. For example, patients in early stages of some diseases (e.g. heart failure often caused by some hazard factors such as hypertension or diabetes) should be able to get benefit from preventive care. The proposed system fulfils the drawbacks in the existing

system to provide an efficient method to get a better data utilization with the help of analysis.

Why Big Data Analytics in Healthcare?

The principal advantages of applying Big Data analytics in healthcare are:

A) Early discovery and check of epidemics
B) Accurate recognision and cure of diseases which have low treatment success
c) Research and discovery of new types of treatments using genomics and patients data.
D) Prevention of insurance and medi-claim fraud
Increase in profitability of healthcare institutions.
E) Based on the incoming medical data, the data scientist can tackle the big problems.

The enormous information investigation permits the patients to see a boundless number of sagacious data, which is close, transitional, and out of reach to the patients or doctors so as to make the correct recommendations.

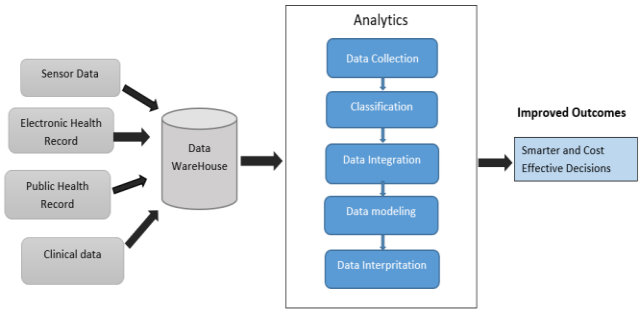

Fig -1: Data processing diagram

3.1 Phases of Big Data Analytics

This Section Discusses the distinct stages of Big data analytics in the advancement of social insurance applications and research. It is specially used to disease investigation with the support of analytical tools and techniques in order to gather, process, break down, examine and deal with the huge volumes of organized, unstructured and semi structured data created by current healthcare systems or organizations.

a) First Phase: Data collection, information Gathering, and storage

The first phase is principally used to acquire the heterogeneous volume of healthcare data from the billions of different sources (i.e. internal and external, individually). The aggregated data might be in different formats and types. At that point, it will be transferred to the system for analysis or either it is stored in databases or data distribution center. Actually, there are also a lot of challenges for taking care of healthcare data due to the absence of information conventions, information guidelines, adaptability, and information protection issues, etc. In addition, it is exceptionally hard to discover the correct metadata. To describe what sort of health data is stored and how it is estimated. However, this metadata acquisition system that depicts and gives data about different kinds of medical data with the end goal of knowledge discovery and identification but also can increase the burden of examining the metadata. Another significant issue is data cleaning. If the stored information isn't valuable, at that point it tends to be passed through the entire data analysis phase and increased the processing error.

b) Second Phase: Data cleaning, extraction, and classification

This subsequent stage is especially used to extract the medical information and stored on a solitary database. The data cleaning is the process of recognizing and evacuating incorrect wellbeing related records. Regularly, the gathered data from sensors, doctor's professionally prescribed, medical image data, and person to person communication data will not be incorrect format since the information should be in an organized form for performing the appropriate/ suitable analysis. For example, the received data may also include medical images (i.e X-ray, CT, PET/CT, and Ultrasound, respectively) and in such a case the data recovery is often highly application dependent and extremely hard to filter dependent on their structure. These information should be classified in the form of structured, semi-structured and unstructured data to play out the important examination.

c) Third phase: Data integration, aggregation, and portrayal

Fundamentally, the health-related records are extremely delicate in nature. Since the coordination of dynamic clinical data with previously existing static data isn't a simple assignment in the real-time environment. Simultaneously, healthcare professionals need the patient's data to be accurate and up-to-date in order to treat the diseases. Moreover, the representation of significant information should be manageable in size, even after the information decrease, noise removal as well as the significant attributes of the original health data should not be changed. Along these, the possibility of choosing a an information portrayal model may result in more meaningful information

d) Fourth Phase: Data modeling, analysis and query processing

The data modeling phase is utilized to process the unpredictable medical data into an effectively understandable form using diagrams, content, and symbols. It is mostly used to view the indistinguishable health-related

information and ensure that all procedures, elements, relationships, and data flows have been distinguished. There are a few unique information demonstrating approaches used including conceptual, physical and logical levels. Since, it guarantees data consistency in unique qualities, semantics, and security while ensuring the quality of the data. The analysis of the information is to discover valuable data from the healthcare data sets using various analytical methods and technologies for example, using data mining algorithms. One of the vital issues with current big data analysis is the lack of ability to utilize diverse database frameworks together effectively. However, once the medical information is coordinated and analyzed, the subsequent stage is to query the significant data. Query processing is an approach to react to the client level inquiries. It might be simple or the range from significant level queries to low-level queries from the physicians, families or even from people with respect to their health status. Base on the intricacy of the query, the big data analyst must choose the suitable platform and analysis tools.

e) Fifth Phase: Data interpretation, delivery, and feedback

The last phase of big data analytics is Data interpretation, data delivery, and feedback. Data interpretation is significant after finished all the above processing steps in the data sets. The understanding of the healthcare data results should be very clear. If it is not clear, then the patients or other healthcare professional can't comprehend the deciphered outcomes provided by the data analyst or decision maker or even computer systems. Besides, this interpretation includes investigating all the assumptions made and following the data analysis. However, a decision-maker has to inspect difficulties on the many assumptions at different phases of analysis. The data delivery phase assists with creating health report based on the previous information model. This model will help the caregivers or medical doctors to take the essential treatment to avoid any further inconveniences. At the last phase, the feedback will be acquired from the patients and decision makers with regards to improving the quality of patient's care.

4. CONCLUSION

This Research mainly focuses on the importance of data analysis and processing various formats of data in the healthcare industry for the better utilization. The healthcare system's aim is to provide the best service for the well being of the people. The drawback of the existing data management model is it's difficult to manage and provide security to critical and healthcare data. This drawback can be overcome using Big data analytics methods and makes the data insight in time consuming and cost effective way.

REFERENCES

[1] Big data in healthcare: management, analysis and future prospects.
[2] Big Data Technologies in Healthcare Needs, opportunities and challenges.
[3] Application of Analytics to Big Data in Healthcare.
[4] Big Data Analytics in Healthcare.
[5] Big data analytics: a survey.