International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

A Review on Big Data Analytics and its Application in Medical Field

Shahana P N

Email id: *pnshahana[at]yahoo.com*

Abstract: ***To describe the promise and potential of big data analytics in healthcare and demand forcasting. The paper describes the nascent field of big data analytics in healthcare, discusses the benefits, outlines an architectural framework and methodology, describes examples reported in the literature, briefly discusses the challenges, and offers conclusions. Data from the health system, everyday data and genomic data have the potential to be used in predictive analytics models that help to improve forecasting of demand for health care services. A supportive infrastructure should be developed that allows interoperable multi - source data sets to be used in predictive analysis in a way that ensures individuals' data privacy requirements are met. Once this has been achieved, there may be the opportunity to use the insights from predictive models to inform network - level health system strategies. Big Data refers to the data or sets of records that are too large in volume to be operated using the existing database management tools and techniques. They have many important applications, such as search engines, business informatics, social networks, social media, genomics, medical field, meteorology, and weather forecast. The growing need to customer behavior analysis and demand forecasting is deriven by globalization and increasing market competitions as well as the surge in supply chain digitization practices. In this study, we performed a thorough review for applications of predictive big data analytics (BDA) in SC demand forecasting. Big data analytics in healthcare is evolving into a promising field for providing insight from very large data sets and improving outcomes while reducing costs. Its potential is great; however there remain challenges to overcome. The main objective of this paper is to give a brief introduction of Big Data, its application and analysis in future and challenges.***

Keywords: Big data, Analytics, Acquisition, Modeling, Globalization, Healthcare, Framework, Methodology

1. Introduction

Bigdata means a collection of data that is huge in size and yet growing exponentially with time. Big Data analytics has varied examples like stock exchanges, medical field, social media sites, jet engines, etc. Big Data could be 1) Structured, 2) Unstructured, 3) Semi - structured. In computing analysis, data is a fact that has been interpreted into a mode that is useful for progressing. Today's computer device and communication media, data is fact converted into digital binary form. Collection of standardized data it required a database [1].

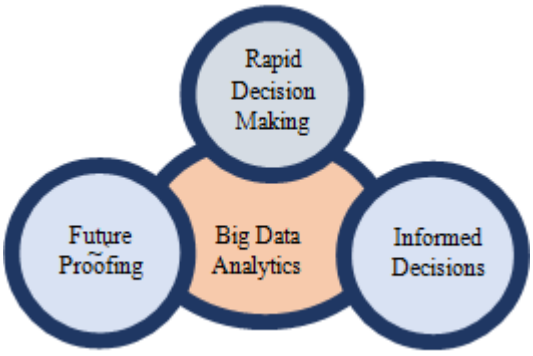

Figure 1: Big Data Analytics

In a relational database it gathers schemas, tables, queries, reports, views and other elements. Structured data is stored in database in sequential format [2]. Data mining is the computing method of detecting designs in huge information sets associating processes at the crossing of machine learning, statistics and database system. Actually, prediction means forecast about an uncertain event and it based upon a particular fact. For big data in industry of healthcare several sources incorporate medical examinations results, hospital records, devices and patient's medical records that are a piece of internet of things.

Relevant to public healthcare the biomedical research additionally produces a significant segment of big data. In order to determine significant data this data needs proper analysis and management [6]. The term often refers simply to the use of predictive analytics or other certain advanced methods to extract value from data, and seldom to a particular size of data set. For example, big data is usually unstructured and requires more time for analysis and processing. This development calls for new system architectures for data acquisition, transmission, storage, and large - scale data processing mechanisms. Big Data is data that are enormous in size and exceeds the processing capacity of regular or traditional database systems. Challenges include analysis, capture, data curation, search, sharing, storage, transfer, visualization, and information privacy. The data is so enormous and are generated so fast that it doesn't fit the structures of normal or regular database architecture. Electronic equipment and sensors spontaneously create diagnostic data that needs to be stocked and processed in real time. It is not only difficult to keep up with huge amount of data but also reasonably more challenging to analyze it, particularly when it is not matching with notions of conventional data structure, to recognize useful patterns and draw useful information. With this big data, there is every chance to transform business, science, government and day to day life. [3].

2. Characteristics of Big Data

Big Data Analytics hit an important role. Big Data having some characteristics like volume, velocity, variety, veracity and value. Here Big Data introduces a concept of 5V's as shown in Figure 2. Since the information is spreading immensely now - adays. Big Data defines both size and vision

from unstructured, composite, noisy, mixed, representation and volume of data [4].

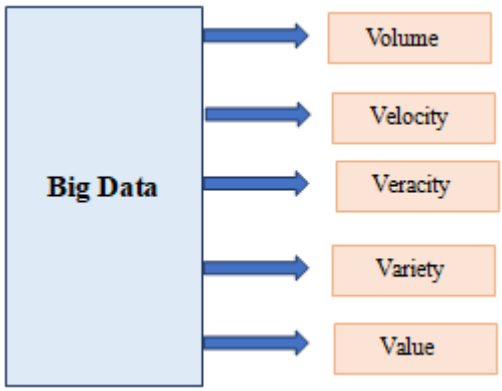

Figure 2: Components of Big Data Analytics

2.1 Volume

Big Data suggests a large weight of data. It used to be an individual's created data. Now - a - days data is generated by digitally on systems such as social media the volume of data to be analyzed is humongous. As amount of data volume increases, the value of different data records will decrease in proportion to age, type, richness, and quantity among all other factors [3].

2.2 Velocity

Speed at which data is being created is called Data Velocity which flows of data in the form of origin like professional systems, machines, organizations and communication of human along with stuff like social media, movable devices, etc [5]. The data is very large and constant in nature. Managing the Data velocity is much more and bigger than a band width issue; it is also an ingest issue (extract transform - load).

2.3 Veracity

The veracity concept in big data deals with bias, noise and unstructured. Big Data feels veracity in data analysis is the major issue when it compares to volume and velocity. This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the dat effectively [8].

2.4 Variety

Different types of data being created are called Data Variety. This concept is to direct the attention to a lot of origins and different categories of data which are structured and unstructured. We accustomed to supply data from sources like databases, file system and spreadsheets etc. Variety refers to heterogeneous source and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications [6]. Now - a - days data comes in the form of emails, photos, videos, pdf, audio etc. Incompatible data formats, incomplete data, non - aligned data strtures, and inconsistent data semantics represents significant challenges that can lead to analytic spread out over a large area in an untidy or irregular way [3].

2.5 Value

Importance of data or the value of information which includes data is called Data Value. The word value in Big Data plays an important role. It includes a massive volume and different varieties of data which are easy to access and delivers quality analytics that helps for king decision. It provides the actual technology. It has been noted that "the purpose of computing is insight, not numbers". [9] Data science is exploratory and useful in getting to know the data, but "analytic science" encompasses the predictive power of big data.

3. Types of Big Data

All data cannot be stored in the same way. The methods for data storage can be accurately evaluated after the type of data has been identified. A Cloud Service, like Microsoft Azure, is a one - stop destination for storing all kinds of data; blobs, queues, files, tables, disks, and applications data. However, even within the Cloud, there are special services to deal with specific sub - categories of data. Following are the types of Big Data: [10]

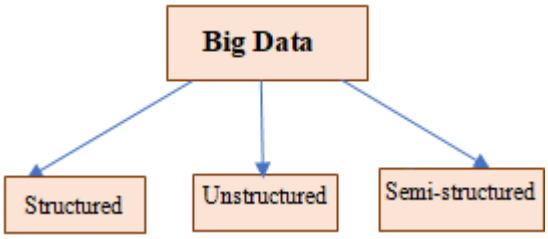

Figure 3: Types of Big Data

3.1 Structured

Structured data can be crudely defined as the data that resides in a fixed field within a record. It is bound by a certain schema, so all the data has the same set of properties. Structured data is also called relational data. Any data that can be stored, accessed and processed in the form of fixed format is termed as a 'structured' data. Over the period of time, talent in computer science has achieved greater success in developing techniques for working with such kind of data [3] (where the format is well known in advance) and also deriving value out of it. However, nowadays, we are foreseeing issues when a size of such data grows to a huge extent, typical sizes are being in the rage of multiple zettabytes.

3.2 Unstructured

Unstructured data is the kind of data that doesn't adhere to any definite schema or set of rules. Its arrangement is unplanned and haphazard. Photos, videos, text documents, and log files can be generally considered unstructured data. Even though the metadata accompanying an image or a video may be semi - structured, the actual data being dealt with is unstructured [5]. Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un - structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos

etc. Now day organizations have wealth of data available with them but unfortunately, they don't know how to derive value out of it since this data is in its raw form or unstructured format [10].

3.3Semi - structured

The data is not in the relational format and is not neatly organized into rows and columns like that in a spreadsheet. However, there are some features like key - value pairs that help in discerning the different entities from each other. Semi - structured data can contain both the forms of data. We can see semi - structured data as a structured in form but it is actually not defined with e. g., a table definition in relational DBMS. Example of semi - structured data is a data represented in an XML file [9].

4. Big Data Analytics in Medical Applications

Big data analytics for healthcare uses health - related information of an individual or community to understand a patient, organization or community. In the past, managing and analyzing healthcare data was tedious and expensive. More recently, technology has helped the healthcare sector make leaps and bounds to keep up with the flow of big data in healthcare. The big data is continually helps in the background of medical section. Actually, the new introduced technology increases the cost of medical care. It gathers some data, takes an action and then makes the appropriate decision. In health care business making a perfect decision that causes of life and death. Health care is a one of those very slow - moving industries with changing significantly. It is definitely behind area like finance, advertising completely in past time use [3].

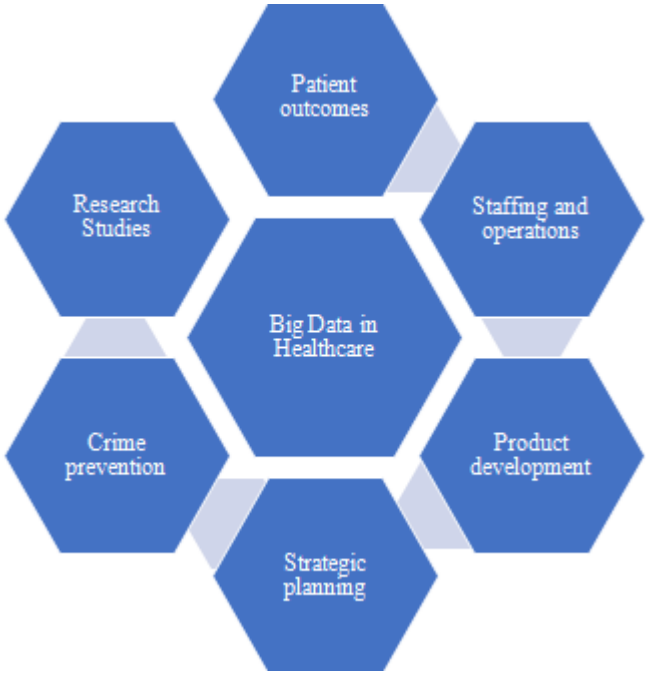

Figure 4: Examples of Big Data in Healthcare

The health care field generates an enormous amount of data every day. There is a need, and opportunity, to mine this data and provide it to the medical researchers and practitioners who can put it to work in real life, to benefit real people. The solutions we develop will be focused on preventing the onset of disease, improving diagnosis and enhancing quality of care. Further, there is the potential to lower health care costs, one of the greatest challenges facing our nation. Big data analytics helped to healthcare sector upgrade by implementing epitomize medicine and prescriptive analysis, hospital liability interference and predictive analysis, dissipation and responsibility, changeability reduction, automatic extraneous and constitutional exposure of patient record, regulated health conditionsand patient registries and disintegrated end solution [1].

Some area of enhancement is more endeavor than really carry out. The original step of information generation within medical application system is not superficial. Along the supplementary endorsement of mobile - Health, e - Health and usable techniques, the quantity of data will advance to gain. This involves electronic health record data (EHR), visualizing record, patient arrangement data, data related to sensor, and the additional patterns of crucial towards progress data. There is currently aconstant superior need for such environment to pay better attention to the data and worth of information [5]. "Big dataAnalytics is very usually meaning greasy data and the portion of data exaggeration boost along with the growthof data". For better treat disease and diagnose in medical the role of big data is one where it can construct better predictive models and health profiles around individual patients. In the pharmaceutical industry and medicine, the main limitations are the comprehension of the science of disease. Big data becomes an integral factoraround totaling increasingly more data around numerous scales for what establishes an ailment—from proteins, DNA and organisms, tissues, ecosystems, organs and metabolites to cells. By integrating big data these are the parameters of science that required for modeling. [7].

On medical database the most significant impact are as follows:

- In potential health problems detection even when there are no apparent symptoms it can do a wonderful job by wearable devices.
- It takes a 360 - degree view of health problem of patient in the institute of healthcare that helped by big data.
- How it delivered the healthcare refined by big data.
- When on the existing data the medical institutions apply big data then they can find useful, potential and new life saving Knowledge that data are in their database otherwise remained inert.
- To find more about patients the big data helps. It can find like organisms, lifestyle, family DNA, ecosystems, tissues, metabolites to cells, genetic structure, dietary habits, proteins and organs.
- The additional patterns, visualizing record, data related to sensor, EHR (electronic health record) data and patient arrangement data included in big data towards data progress.
- To reduce costs of predict outbreaks of epidemics, treatment, to improve the quality of life in general and avoid preventable diseases it has the potential by medicalanalytics.

4.1 Challenges of big data in healthcare

The main challenges [5] are interoperability, manageability, security, development, reusability andmaturity. For interoperability, integration of big data technologies with existing enterprise solutions is extremely important. Data ingestion, data modeling, data visualization using existing tools must be supported. For manageability, big data cluster management and monitoring is an important issue. It must be integrated with other management tools. As per as security is concerned all the security tools must be seamlessly integrated with the enterprise security tool. For development, there must be unified advancement device covering distinct big data plan.

- Through several sensors capture the behavioral data of patient, their several communications and social interactions.
- Many healthcare organizations lack adequate systems and databases.
- Huge amount of medical imaging data handles efficiently and extract potentially useful biomarkers and information.
- From complex heterogeneous sources of patient inferring knowledge [10].

5. Conclusions

Big data is very import in industry of healthcare, because several sources incorporate medical examinations results, hospital records, devices and patient's medical records that are a piece of internet of things. For treat disease and diagnose in medical the role of big data is one where it can construct better predictive models and health profiles around individual patients. The future trends of big data analytics in the healthcare system have the potential of enhancing and accelerating interactions among clinicians, administrator, lab director, logistic manger, and researcher by reducing costs, building a better efficiency based on result comparison, reducing risks, and improving personalized care. The followings are the future trends connected with medical big data analytics. Big data in health care has many challenges including but notlimited to: deducing knowledge from complex heterogeneous patientsources; leveraging the patient/data correlations in longitudinalrecords; understanding unstructured clinical notes in the appropriatecontext; competently managing huge volumes of medical imaging dataand mining potentially beneficial information and biomarkers; analyzing genomic data, a computationally rigorous task and mergingit with standard clinical data to increase layers of complexity; capturingthe behavioral data through multiple sensors with their various socialinteractions and communications [10].

References

[1] Shankar Krishnan, "Application of Analytics to Big Data in Healthcare", IEEE 32nd Southern Biomedical Engineering Conference, 2016.
[2] Yuxuan Jiang, Zhe Huang and Danny H. K. Tsang, "Towards Max - Min Fair Resource Allocation for Stream Big Data Analytics in Shared Clouds", IEEE Transactions on Big Data, Vol.4, no.1, pp.130 - 137, 2018.
[3] Hiba Asri, Hajar Mousannif, Hassan Al Moatassime, and Thomas Noel, "Big Data in healthcare: Challenges and Opportnities, International Conference on Cloud Technologies and Applications (CloudTech), pp.1 - 7, 2015.
[4] Murdoch, T. B., and Detsky, A. S. The inevitable application of big data to health care. Jama 309, 13 (2013), 1351 - 1352.
[5] Archenaa, J., and Anita, E. M. A survey of big data analytics in healthcare and government. Procedia Computer Science 50 (2015), 408 - 413.
[6] Naoual El aboudi and Laila Benhlima, "Big Data Management for Healthcare Systems: Architecture, Requirements, and Implementation", Advances in Bioinformatics, pp.1 - 10, 2018.
[7] St. KlimentOhridski, "Big Data Analytics in Medicine and Healthcare", Journal of Integrative Bioinformatics, pp.1 - 5, 2018.
[8] Nivedita Das, Leena Das, Siddharth Swarup Rautaray and Manjusha Pandey, " Big Data Analytics for Medical Applications", International Journal of Modern Education and Computer Science, Vol.2, pp.35 - 42, 2018.
[9] Nishita Mehta and Anil Pandit, "Concurrence of Big data Analytics and Healthcare: A Systematic Review", International Journal of Medical Informatics, Vol.114, pp.57–65, 2018.
[10] P. Ram Mohan Rao, S. Murali Krishna, and A. P. Siva Kumar, "Privacy preservation techniques in big Data Analytics: a survey. Journal of Big Data, pp.1 - 12, 2018.