

Big Data Analytics for Health Care Applications Using Cloud Computing- A Study

^{1*}Subhadra K., ²N Kavitha

¹Research scholar Nehru arts and science Coimbatore, India

²Head of the department Computer science Nehru arts and science college Coimbatore, India

Available online at: www.ijcseonline.org

Abstract- This paper focuses the art of big data in medical field general background of big data is discussed first and then its related areas, such as cloud computing, data centers, internet of things and Hadoop. The value chain four phases are discussed here that is data generation, data acquisition, data storage and data analysis. For each phase here we are introducing general background technical challenges and review the latest advantages. Big data is a concept which defines the difference between itself and “massive data” or “very big data”. Three v s of big data are volume, velocity and variety which are defined by Doughty Laney in 2001.

Keywords- Big data analytics, information management, literature review, health care, data driven application

I. INTRODUCTION

In health care data coming by large volumes are collected in health care database and it is not high enough to require for the circumstances. Here is a collection of very similar data objects and which defines new use cases supporting genomics cloud computing is a major concept in big data which creates interesting possibilities for sharing information about patients and their treatment process. They are log on and view their details and can treat them self and even play more of their medical area. It will be very useful and it is a way to educate current or future patients and how to get better and to stay healthy. After this the cost of health care is big issue now a days. Cloud computing acts as a mass storage device for the patients and the cost of storing and processing information. Cloud services solutions can make a hospital work together efficiently and accurately.

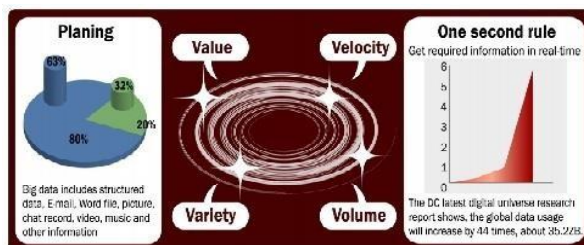


Fig. 1

II. RELATED WORK

Tablet computers on hand to access the cloud during operations has common. Surgeons have great deal of knowledge and having access to the cloud can allow surgical texts in an electronic format as we needed. Cloud computing in hospitals creating possibilities for share the information about patients and their condition at present and the treatment procedure they adopt. The name cloud was inspired by the symbol that is used to represent in flow charts and diagrams.

Hace theorem

Big data starts with large volume heterogeneous autonomous sources and it is distributed, decentralized control. This theorem helps in better understanding of the underlying arrangements as well as how well data is arranged within a cluster. Heterogeneous data is same data represented in same format. For example in case of medical field the information of patients can be stores differently in different departments. In neuroscience departments the information is stored as gene sequences, in the orthodontics department store the information as the difference in bone marrow, in eye specialist might store the data retina or lens power, thus details of the same person is represented in same format. In the case autonomous it is a standalone working environment and there are centralized servers.

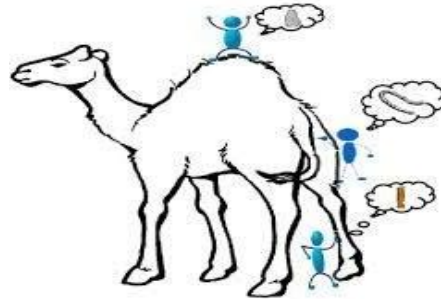


Fig.2

In the theorem, we can imagine that a number of blind men are trying to size up a giant camel. The main objective of each blind man is to extract the information of the camel according to the part of information during the procedure. Here each individual is expected that the blind men will conclude independently depending on the part each of them is limited to. Hace theorem suggests that the main characteristics of big data are.

- Large volume of miscellaneous dimensionalities.
- Being autonomous sources and with circulations and disperse control.
- Complex and evolving associations.

III. DATA ANALYTICS

The art of examining raw data with the purpose of drawing or inspecting, cleaning, modeling and transforming data with the purposes of highlighting useful information's is called data analytics .

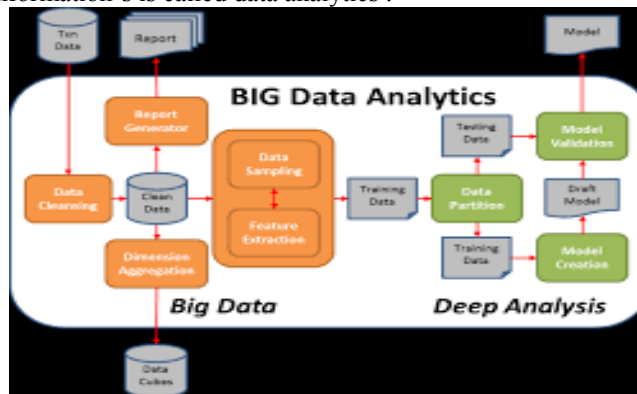


Fig.3

It is divide into exploratory data analytics(EDA) and confirmatory data analytics(CDA), qualitative data analytics(QDA). The term analytics is used by business intelligent software vendors to explain different functions. It is well defined that spread sheets are the established data collection and data analysis tools in the case of technical computing. Excel is an example and which not support computation of large scale data analytics. Researchers of developers are developed a loud data analytics which is based on Daytona. Daytona is a Map reduce runtime optimized technique for data analytics. Here it act as a bridge between the client and cloud, user can use this service to discover and import data from the cloud and uses some algorithms to extract information's from big data sets.

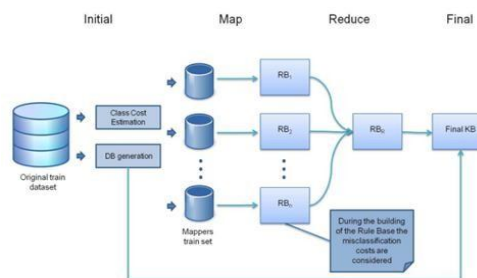


Fig.4

IV. TOOLS FOR DATA ANALYSIS

Different data tools are available for data analysis the following tools are available

- Data wrangler is a web based service and it is designed for cleaning and rearranging data.
- Google refine is a spread sheet based tool and it can import data in different formats.
- The R project for statistical computing is a general statistical analysis platform.
- Google fusion tables, is one of the simplest way to turn data into chart or map

Top benefits of big data in the healthcare industry

- Advanced patient care:** Electronic health records help in collecting demographic and medical data such as lab test, clinical data, diagnoses, and medical conditions, which helps healthcare practitioners to provide quality care.
- Improve operational efficiency:** Healthcare companies use big data as a part of their business intelligence strategy to examine historical patient admission rates and to analyze staff efficiency. Healthcare companies can cut down on healthcare cost and provide better care with the help of predictive analytics. Big data performance, and reduce readmissions.
- Finding a cure for diseases:** A particular medication seems to work for some people but not for others, and there are many things to be observed in a single genome. It is not possible to study all of them in detail. But big data can help in uncovering unknown correlations, hidden patterns, and insights by examining large sets of data. By applying machine learning, big data can study human genomes and find the correct treatment or drugs to treat cancer.

V. CONCLUSION

The healthcare industry faces multiple challenges, ranging from new disease outbreaks to maintaining an optimal operational efficiency. Big data analytics can help in solving these healthcare challenges. With the vast amount of data available in the healthcare sector like financial, clinical, R&D, administration and operational data, big data can derive meaningful insights to improve the operational efficiency of the industry.

REFERENCES

- [1]. W. Raghupathi, V. Raghupathi, "Big data analytics in healthcare: promise and potential", Health Information Science and Systems, vol. 2, pp. 1-10, 2014. WHO. Mobile phones help people with diabetes to manage fasting and feasting during Ramadan. Features. 2014.
- [2]. David Houlding, MSc, CISSP. Health Information at Risk: Successful Strategies for Healthcare Security and Privacy. Healthcare IT Program Office Intel Corporation, white paper. 2011.
- [3]. Giangregorio LM, Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, et al. FRAX underestimates fracture risk in patients with diabetes. J Bone Miner Res. 2012;27(2):301-8.
- [4]. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform. 2007;2:59-77.
- [5]. Witten IH (Ian H., Frank E, Hall MA (Mark A, Pal CJ. Data mining: practical machine learning tools and techniques. 621 p.
- [6]. Maglogiannis IG. Emerging artificial intelligence applications in computer engineering: real world AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies. IOS Press; 2007. 407 p.
- [7]. Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. Am J Epidemiol. 2005;163(3):262-70.
- [8]. Nemes S, Jonasson JM, Genell A, Steineck G. Bias in odds ratios by logistic regression modelling and sample size. BMC Med Res Methodol. 2009;9(1):56.
- [9]. Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. Am J Epidemiol 2005;161(1):81-88.
- [10]. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710-8.