iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** I **Month of publication:** January 2018

DOI: http://doi.org/10.22214/ijraset.2018.1481

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

A Review of Healthcare Disease Diagnosis using Big Data Analytics

Tawseef Naqishbandi[1], Zahid Gulzar Khaki[2]
[1]*Department of Computer Science and Engineering, IUST, Awantipora, J&K*
[2]*Department of Electronics and Communication, Engineering IUST, Awantipora J&K*

Abstract: Healthcare industry has been a significant area for innovative application of various technologies over decades. Being an area of social relevance governmental spending on healthcare have always been on the rise over the years. Big Data has been in use for many years for situational awareness and response generation. Computing technologies have played an important role in improvising several aspects of healthcare. Recently emergent technology paradigms of Big Data, Internet of Things (IoT) and Complex Event Processing (CEP) have the potential not only to deal with pain areas of healthcare domain but also to redefine healthcare offerings. This project aims to lay the groundwork for a healthcare system which builds upon integration of Big Data with further relevance to IOT and CEP.

In this proposed research paper, the patient health conditions are analyzed and most probable disease is diagnosed using big data analytics. This system will take a set of symptoms and present back possible diseases that could be the cause of those symptoms.

Keywords: Big Data, Healthcare, Disease, IoT, Analytics, Predictive Analytics

I. INTRODUCTION

The health care industry is extremely big incorporating several sectors that are dedicated to providing health care services and products. Some of these sectors depend significantly on data storage, computing and communication technologies. Healthcare Information Technology (HIT) [16] is ―the application of information processing involving both computer hardware and software that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making‖. In recent years a newer sector of Healthcare known as _Healthcare Informatics' has emerged and started to gain significant popularity. It is a discipline at the intersection of information science, computer science, and healthcare [17]. Healthcare finance and insurance is also another area which relies significantly on computing and communication technologies. Healthcare is an area of concern and importance for both developed and developing nations owing to its social relevance. Governments across the world are focusing on all aspects of healthcare like policy changes, legal statures, insurance, funding, and technology overhaul. Patient Protection and Affordable Care Act also popularly known as Obama care is ushering significant changes in US healthcare industry. Hearth insurers are expected to see substantial increase in their cost due to increased risk of covering more people and they cannot legally deny insurance to individuals based on prior health conditions. Affordable care ensures that maximum number of people is covered. This poses as a challenge to hospitals as they will witness increase in patients, which means increasing amount of data to be analyzed. The new era of health care will be driven by an explosion in health-related data from a growing range of public and private sources, and can be analysed by increasingly powerful number-crunching computers. This new era is moving from a world where illnesses can be treated to one where it can be predicted and can be prevented. Big Data technologies will play a key role in turning this idea into reality. Big Data is a term encompassing the use of techniques to capture, process, analyze and visualize potentially large datasets in a reasonable time frame not accessible to standard IT technologies [1]. It refers to the ability to crunch vast collections of information, analyze it instantly, and draw from it sometimes profoundly surprising conclusions. According to Gartner analyst Doug Laney [2] data growth challenges and opportunities as being three-dimensional, i.e. increasing volume (amount of data), velocity (speed of data in and out), and variety (range of data types and sources). Big data technologies deal with petabytes of records, files, transactional data either arriving as streams or in batches. Rise of disruptive technologies like social and mobile are contributing to increase in unstructured and semi structured data.

©IJRASET (UGC Approved Journal): All Rights are Reserved

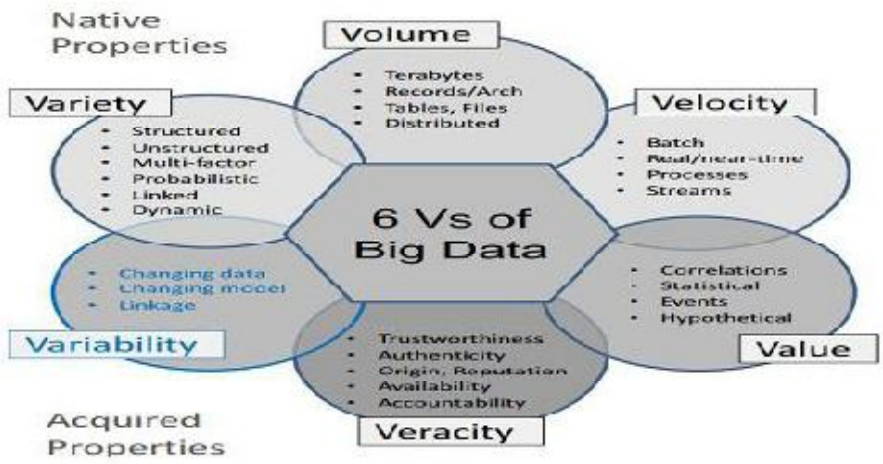


Figure 1: Big Data Properties[29]

II. BACK GROUND AND LITERATURE SURVEY

Big Data analytics are an exciting research direction analytics, predictive analytics and future selection. Prior analysis work has focused on privacy, security, architectural and predictive analytics. Healthcare informatics has emerged and started to gain significant popularity. It is a discipline at the intersection of information science, computer science and healthcare [1]. Providing patients with accurate and up-to-date information and guidance rather than just data will help them make better decisions and better adherence to treatment programs [2]. Report on Big Data Analytics [3] indicates that bis data solutions can help stakeholders personalize care, engage patients, reduce variability and costs, and improve quality of health delivery.

Sagiroglu et al [1],"Big Data: A Review" describes the big data content, its scope, methods, samples, advantages and challenges of Data. The critical issue about the Big data is the privacy and security. Big data samples describe the review about the atmosphere, biological science and research. Life sciences etc.

By this paper, we can conclude that any organization in any industry having big data can take the benefit from its careful analysis for the problem solving purpose. Using Knowledge Discovery from the Big data easy to get the information from the complicated data sets [6]. The overall Evaluation describe that the data is increasing and becoming complex. The challenge is not only to collect and manage the data also how to extract the useful information from that collected data. According to the Intel IT Center, there are many challenges related to Big Data which are data growth, data infrastructure, data variety, data visualization, data velocity.

Garlasu, D et al [b] "A Big Data implementation based on Grid Computing", Grid Computing offered the advantage about the storage capabilities and the processing power and the Hadoop technology is used for the implementation purpose. Grid Computing provides the concept of distributed computing. The benefit of Grid computing center is the high storage capability and the high processing power. Grid Computing makes the big contributions among the scientific research, help the scientists to analyze and store the large and complex data [5].

Mukherjee et al [19]"Shared disk big data analytics with Apache Hadoop" Big data analytics define the analysis of large amount of data to get the useful information and uncover the hidden patterns. Big data analytics refers to the Mapreduce Framework which is developed by the Google. Apache Hadoop is the open source platform which is used for the purpose of implementation of Google's Map reduce Model [2]. In this the performance of SF-CFS is compared with the HDFS using the SWIM by the facebook job traces .SWIM contains the workloads of thousands of jobs with complex data arrival and computation patterns.

Aditya B. et al [20]"Addressing Big Data Problem Using Hadoop and Map Reduce" reports the experimental work on the Big data problems. It describe the optimal solutions using Hadoop cluster, Hadoop Distributed File System (HDFS) for storage and Map Reduce programming framework for parallel processing to process large data sets[3]. Real Time Literature Review about the Big data According to 2013, facebook has 1.11 billion people active accounts from which 751 million using facebook from a mobile. Another example is flicker having feature of Unlimited photo uploads (50MB per photo), Unlimited video uploads (90 seconds max, 500MB per video), the ability to show HD Video, Unlimited storage, Unlimited bandwidth. Flickr had a total of 87 million registered members and more than 3.5 million new images uploaded daily.

©IJRASET (UGC Approved Journal): All Rights are Reserved

Wager et al. suggests that The healthcare sector is widely considered as one of the most important industries in information technology. More and more, information technology has been considered as a practice that facilitates healthcare performance through using data and information efficiently within the healthcare sectors. Therefore, Wager et al (2005) said that in order to understand the relation between information technologies and healthcare, we first need to understand what are the technologies used in healthcare. Information technology functions have developed over the last few years not only as a technology services provider, but also as a strategic provider that develops and integrates industries' infrastructures to facilitate and ensure quality of service (LeRouge et al 2007). In the mid-80,'s information technology changed the healthcare industry and brought many benefits when they used microcomputers, which were a small in shape, fast and very powerful for that time. Moreover, this allowed hospitals to develop clinical applications for various medical care settings. As a result, hospitals started to purchase and adopt information systems in the healthcare industries, and after that, challenges began to emerge when professionals tried to integrate data among these systems (Wager et al 2005).

However, Bhattacherjee et al and Castro et al, granted that information technology has improved healthcare industries, but they also highlighted some of the difficulties related to the use of information technology in healthcare sectors, as they noticed that it is hard to implement information technology in small clinics and organizations, with high costs due to reduced efficiencies of scale. Therefore, IT implementation requires long term training and retention of skilled professionals. On the other side of the debate …..

Abbott et al [21] believe that information systems and information technology occupy a high position in improving healthcare industries in general, and in electronic healthcare record (HER) in specific; for the reason that implementing such technologies can save costs and times associated with daily hospital data records, such as patients schedules and billing. This is in addition to improving healthcare performance and efficiency by eliminating manual data records and paper work, and alongside smooth and flexible tracking of patient details.

Healthcare Analytics & Data Mining Data Mining is described as a process by which data is gathered, analysed and stored in order to produce useful and high quality information and knowledge. This term also includes the way of how this data is gathered, filtering and preparation of the data for use and finally the processing of data to support data analytics and predictive modelling (Russom 2011).

Data collection The first stage of data mining is the process of gathering and collecting data. However, even before gathering the data, ideas and plans should be assumed to decide which data should be gathered in order to collect specific data as desired and use it efficiently (Lamont, 2010). Furthermore, Chordas (2001) added that a lot of projects fail and exceed estimated costs because of poor quality of gathered data which can result from poor data cleaning.

Healthcare Sectors & Big Data Analytics Big data storage and management One of the most important elements in dealing with and managing data is to know where and how this data will be stored once when it is collected. The traditional methods of storing and retrieving such data are not efficient anymore, since it was structured and stored in data warehouses and relational databases, after extracting and loading it from different outside sources. However, this data is transformed and classified before being ready to use and function (Bakshi 2012).

Furthermore, Herodotou et al (2011) agreed with Bakshi (2012) when he said that there are many numbers of data sources now and that a huge amount of data has become available, so this growth of data will absolutely require an agile database which can deal with the data logically and through data synchronization in order to adapt to the rapid data evolution.

On the other hand, Plattner and Zeier (2011) stated that databases only manage server memory data, therefore eliminating the option of managing other storage devices such as: disk and compact drivers. Accordingly, this will reduce the efficiency of database performance and real time response during the time.

Patients Role in Healthcare Analytics This section is concerned about how individuals (and patients in specific) can improve healthcare analytics through understanding the small and personal data, as well as educate themselves in how to collaborate with the healthcare data analytics to reach a high level of efficiency and accuracy (Luciano 2013).

Swan (2012) was discussing the same point when he identified the term "citizen science", where nonprofessional and educated individuals are skilled enough to conduct and support healthcare analytics system. Accordingly, this will require organizations to train individuals how to follow up and track their health information, as well as self-monitoring.

Principally, to perform good data analytics, first of all we should teach individuals how to understand and realize the importance of dealing with such data, for instance how to deal with breast cancer (Hanoch 2012).

However, Miron et al (2011) believed that whatever and how much our patients are educated and skilled to provide us with the data we expect, medical professionals still highly need to test and clarify this data to consider it and keep it on record. Also, he added that once when the data has been tested and clarified, we then need to find out how to change an individual's behavior, starting with

©IJRASET (UGC Approved Journal): All Rights are Reserved

parents and guardians who are responsible for raising their children. However, Kim (2013) says that sometimes being motivated for change and in understanding of information are not enough. Furthermore, patients should identify the risks and detect where to change, for instance; some patients know that they have a high level of blood pressure but they don't know how to deal with it and control it: should they change drugs? Change eating habits and life styles? Do more exercises? Turner (2011) agreed that social media and internet applications have a big influence on collecting patient information through filling and completing some online forms in order to keep track of their state of health, as well as to provide the suitable advice and treatment when needed. Moreover, patients can share some information with other patients, so they increase their knowledge, background and awareness in the healthcare analytics sectors regarding their conditions. Finally, patients who share their symptoms, diagnoses and results with others can gain benefit from the ability to understand their health conditions by comparing them with other patients (Brownstein & Wicks 2010)

Connectivity between Healthcare Analytics System and Individuals (Medical Staff and Patients) Connectivity approaches generate thoughts and ideas from connected networks of minds and leverages prior experiences with the utilization of technologies in our everyday life. (Siemens 2004) Moreover, Mc Horney (2009) has added that healthcare analytics is not solely regarding technology and the knowledge however; it is also regards how much individuals are attached to and familiar with medical care systems and their personal skills such as ability to learn and adopt such systems in their life, as different people have different attitudes and reasons for not accepting such technologies, especially older people. Healthcare Predictions and Decision Support System (DSS) Healthcare prediction is another data analytics method focusing on reducing future medical costs. Predictive technique uses patient medical history to evaluate all the potential health risks and predict a future medical treatment in advance (LexisNexis 2015).

Loginov et al 2012 stated that by retrieving and reviewing past patient details, information and diagnoses from the databases, predictive methods can take a place through forecasting, reducing time and costs. Parkland hospital in Dallas, Texas has launched a predictive system which scans all patients' details and information to identify potentials risks and outcomes. As a result, the hospital has saved more than half a million dollars, especially in heart failure and disease predictions in terms of performing patients' monitoring and avoiding future complications (Jacob 2012).

The most significant and obvious result of using such technology within the healthcare sectors is its results on costs. Because of cost, information is one of the main aspects that have a big effect on the cost of healthcare predictive analytics. Medical care systems have focused on increasing healthcare analytics performance as well as minimizing the cost by simplify unstructured clinical record and reducing irregular information. Consequently, large quantities of information then will be managed and controlled smoothly and efficiently (Bertsimas et al 2008).

Predictive analytics can assist to avoid and reduce inaccurate prediction costs plus time for the reason that it makes the data sourcing cost lower by specifying the desired and necessary data only, since the data is simplified, standardized and exists in historical clinical databases (Bradley & Kaplan 2010) Healthcare Analytics & Real Time Murphy (2013) believed that real time analytics produces more accurate results and information, since it evaluates current patients' history and conditions, therefore investigating patients' diagnosis correctly and offering the best treatment. Real time monitoring techniques guarantee to keep data up to date and increase the quality of information, as assumed so by Taylor (2010). She believed that real time matters in healthcare analytics are very significant for the reason that it generates accurate results, such as where diabetes patients can recover if their ailment is discovered and treated correctly in the earlier stages. Walker et al (2012) agreed with that however they also highlighted some of its disadvantages, such as high cost, its high required level of training and long time to complete.

III. TECHNOLOGY OVERVIEW

According to Gartner [1], Big data is high-volume, high-velocity and high-variety information that demand cost effective, innovative forms of information processing for enhanced insight and decision making. Gartner‟s definition of Big Data was refined by Yuri et. al. [2] as data intensive technologies that are targeting to process high-volume, high-velocity, high-variety data to extract intended data value and ensure high-veracity of original data and obtained information that demand cost-effective, innovative forms of data and information processing and analytics for enhanced insight, decision making, and processes control.

Definition by Yuri et. al. [2] captures the essential aspect in terms of the native and acquired properties of Big data. The native properties of Big Data (also known as three big Vs) are "Volume, Variety and Velocity" . Acquired properties (also known as three small Vs) are "Variability, Value and Veracity". Practitioners and researchers have further added upon other properties [11][12] of big data like validity, visibility and viability. Table 1 highlights the properties of big data.

©IJRASET (UGC Approved Journal): All Rights are Reserved

Property	What it means
Volume	Huge size of data
Variety	Various types of data
Velocity	Speed at which data keeps coming
Veracity	Accuracy of data; Conformity to facts
Value	Worth, usefulness, or importance of the data to its consumer
Variability	Indication of changing nature of data
Validity	Interpreted data having a sound basis in logic or fact
Visibility	State of being able to see or be seen
Viability	Practicality in applying for specific problems

Table 1. Properties of Big Data

Notions of volume, variety and velocity are easily understood unlike the other properties. Big data without veracity will lead to incorrect inferences being drawn. Data in and of itself has no value. Unless the data is of some worth to someone there is no meaning on focusing on any other aspect of data. Failure to understand the variable nature of data will lead inferences which will be short lived. Big data without validity is worthless. Data from different sources need to be brought together such that they are visible to the technology stack which forms Big Data. Viability is significance in determining the relevance and practicality of data before progressing into further modeling and analysis. Figure 1.2 illustrates the conceptual architecture of big data analytics.



Figure 1.2 Conceptual Architecture of Big Data Analytics

Big data analytics is made possible with an amalgamation of various platforms and tools like Hadoop, Hive, HBase, MapReduce, CouchDB, MySQL, NoSQL, UnSQL, Jaql, Cassandra, MongoDB, Pig, SOLR, flume, sqoop, Ambari, Oozie, ZooKeeper, R, Mahout and such others.

IV. FEASIBILITY, ECONOMICAL, TECHNICAL AND BEHAVIORAL STUDY

A. Feasibility Study

Feasibility study is a system proposal according to the work ability, impact on the organization, ability to meet users need and efficient use of resources. Three key considerations are involved in feasibility analysis, economical, technical & behavioural.

B. Economical Feasibility

The economical analysis of the present proposed system is necessary to evaluate if high investment of the system is beneficial. Actually the implementing changes for “Disease Diagnosis” are very minimal. It only needs window platform to run the program. The software is also not at all resource hungry.

C. Technical Feasibility

It is must that the proposed Strategy is technically feasible in the organization. The existing strategy is manual and hence it is necessary to bring around awareness of the computer but this may not require an in-depth technical knowledge as the system developed is simple and easy to understand. The result obtain should be true in real time conditions. Then only the software can be used in Hospitals for checking purpose.

©IJRASET (UGC Approved Journal): All Rights are Reserved

D. Behavioral Feasibility

Behavioural feasibility deals with the runtime performance of the software.

V. PROBLEM DEFINTION

The formulation of a research problem is the first and most important step of the research paper. This is more like identifying a destination prior to beginning a journey. A research problem is like the foundation of a building. The type and architecture of the building depends on the foundations. If the foundation is well designed and strong, you can expect the building to be strong as well. You must have a clear idea with regard to what it is that you want to find out but not what you think you must find.



Figure 3.1: Relationship between research problem (input) and quality of content(output)[30]

In this research paper Data analytics is an important building block for predictive analytics and future selection for disease diagnosis. The unstructured nature of data is highly challenging and all the insights depend upon the nature and availability of data. The analysis of vast amounts of patient data. For example, before healthcare organizations can implement pre-emptive care programs, they must first identify the relative risk of their patient. This is based on a variety of clinical, financial, and lifestyle factors, including:

A. Problem list of patients, especially chronic conditions
B. Procedures, medications and other hospital data
C. Claims information
D. Risk factors such as tobacco, alcohol and drug use
E. Availability and accessibility of health services and social support.

F. This research paper focuses on three things.

1) *Patient Data collection Gathering (structured as well as unstructured in nature):* Collecting data from and on behalf of medical patients is a critical component of healthcare, to provide the best insights and most proper diagnosis. The Internet has become a vast resource for healthcare organizations through the creation of digitized disease registries, which can process more data quickly and provide timely results to expedite treatment. This includes identifying each step in the process, from determining what data is collected to who is responsible for doing that, as well as when and where data will be gathered.
2) *Challenges Ahead In Data Gathering:* One challenge that our team faced, especially with a patient who depends on multiple medical organizations for care, is the collection and sharing of information. As, Some patients may be less comfortable sharing information electronically, or not have access to a secure method of transmission. Language barriers also can frustrate fluid communication, and even the race and ethnicity of patients can play a factor. Information such as race, ethnicity, language, sexual orientation, gender identity, as well as social and behavioral norms, can influence health. Collecting this information about patients allows providers to more easily identify disparities and achieve health equity, according to the California Pan-Ethnic Health Network.

©IJRASET (UGC Approved Journal): All Rights are Reserved

3) *Overcoming Challenges While Gathering Patient Data:* Those above challenges can be removed by following these three steps: to eliminate the logistical problems of the paper records by making clinical data immediately available to authorized users wherever they are – no more unavailable or undecipherable clinical records. to reduce the work of clinical bookkeeping required to manage patients – no more missed diagnoses when laboratory evidence shouts its existence, no more forgetting about required preventive car) to make the informational 'gold' in the medical record accessible to clinical, epidemiological, outcomes and management research."

IV. BIG DATA ANALYTICS

Big data analytics comes into play in order to unlock the value of data, organizations need Inferring knowledge from complex heterogeneous patient sources Leveraging the patient/data correlations in longitudinal records. Understanding unstructured clinical notes in the right context. Efficiently handling large volumes of medical imaging data and extracting potentially useful information and biomarkers. Analyzing genomic data is a computationally intensive task and combining with standard clinical data adds additional layers of complexity. Gathering the patient's behavioural data; their various social interactions and communications. Take advantage of the massive amounts of data and provide right intervention to the right patient at the right time. Personalized care to the patient Potentially will benefit all the components of a healthcare system i.e., provider, payer, patient, and management.

Big data technologies have to deal with all these varieties of data. Whereas volume, variety, velocity are the native properties of big data system, it has also three acquired properties of variability (indication of changing nature of data), value(significance based on statistics, hypothesis etc,), and veracity (trustworthiness of data, provenance etc.). `Big data is an idle fit for dealing with the technology challenges faced by the Health care industry. Increasing public health records with all the new sources of health data generated by wearable sensor devices, Wi-Fi enabled scales, smart phones and low-cost diagnostic kit, could provide a far more accurate picture of individual's health and the treatments they receive. In terms of big data for health care ―Volume‖ refers to the rapidly expanding size of the sets of data being generated in every area of activity in an healthcare enterprise, from revenue, to patient data, to supply and operations. ―Variety‖ includes the diversity of data collected. In a hospital, for instance, data includes patient records containing a variety of information like lab reports, scans, x-rays, prescription details and so on. Beyond diagnosis and treatment related data, all aspects of finance, patient scheduling and workflow, insurance data and medical outcomes etc are also available. Synthesizing and analyzing such disparate elements is challenging on its own. With newer sources of healthcare monitoring devices like personal health monitors, Body Area Networks (BAN) there is significant Velocity of incoming healthcare data. There used to be a time in healthcare industry where more in-patient days translated to more revenue. But now patients are increasingly demanding information about their healthcare options so that they understand their choices and can participate in decisions about their care. Patients are also an important element in keeping healthcare costs down and improving outcomes. Providing patients with accurate and up-to-date information and guidance rather than just data will help them make better decisions and better adherence to treatment programs [4]. As a result of this focus on meaningful information, all healthcare constituents are impacted by big data, which supports analytics that predict how these patients are likely to behave, encourage desirable behaviour and minimize less desirable behaviour. Report on Big Data Analytics [5] indicates that Big data solutions can help stakeholders personalize care, engage patients, reduce variability and costs, and improve quality of health delivery. Big data analytics can also contribute to providing a rich context to shape many areas of health care like analysis of effects, side-effects of drugs, genome analysis etc. Access to huge amounts of healthcare data coupled with insights provided by big data analytics, it will be possible to develop predictive algorithms that can forecast which demographics are likely to cost the most to treat in future, for conditions like diabetes and asthma. This will enable healthcare providers intervene earlier and redesign their services to cope with the expected massive increase in healthcare demand. There are also certain challenges particular to using big data for healthcare:

A. Accuracy

Human tendency to understate negative factors, such as smoking and failure to comply with treatment is of concern. People also tend to overstate positive factors, such as exercise. These biases need to be identified and corrected, or passive techniques need to be used in order to acquire data that does not have self-reported bias.

B. Privacy

People are reluctant to divulge personal information because of concerns about privacy. Providing guarantees for security and privacy, incentives will help to address these concerns.

©IJRASET (UGC Approved Journal): All Rights are Reserved

C. Consistency

Standards need to be developed and implemented to promote consistency, increase usefulness and facilitate data usage amongst all the stakeholders involved in various healthcare sectors.

D. Facility

Mechanisms need to be developed to make it easy for patients to accurately self-report data. This includes evolving passive mobile computing, wearable devices that require no effort. Techniques are needed to get data from healthy people to make the populations truly representative and not biased by the ill in any kind of statistical analysis.

E. Fragmentation

Healthcare data is significantly fragmented. There is also unwillingness for healthcare participants to share data

V. CONCLUSION, SUMMARY AND FUTURE SCOPE

The normal approach is where a patient tells doctor about different symptoms he/she is having or even if the doctor as him/her regarding certain related symptoms to diagnose the exact disease he/she is suffering from. This process has the tendency of possible human error to mistakenly forgetting asking or telling about some related symptom(s) which may lead to incorrect diagnosis. The approach discussed in this research paper works towards making the disease diagnosis much more effective and accurate. The idea is to link multiple symptoms to different diseases in order to help diagnose particular disease patient is suffering from based on exact symptoms he/she is having. The approach asks patient to choose multiple set of symptoms one at a time and then based on that symptom the software automatically shows possible diseases and asks user about other symptom based on the possible diseases. This process continues until user keeps on choosing another symptom and the process stops where a single disease is left possessing the selected set of symptoms. This process of suggesting related symptoms for disease based on previously selected symptom eradicates the process of skipping of some symptom that a patient or doctor can forget thus making disease detection much more effective and accurate. There are much more perspectives other than this one which can be taken into consideration to make Frame work and future system even better. The concept adapted makes disease detection of a patient more accurate and effective thus helps doctors to diagnose the patient properly. This way possibility of patient's cure increases to great extent & that too in quicker span of time.

The product as a deliverable will help its users to diagnose a disease a patient is suffering from based on multiple symptoms a patient is possessing. Further it will also keep a history of disease or diseases patient is possessing based on previous disease detection., it's not the end but surely a lot more can be done further ahead.

REFERENCES

[1] NESSI White Paper, December2012
[2] Big Data. [Online] http://en.wikipedia.org/wiki/Big_data (Accessed December 2017)
[3] Aleksandar Milenkovic, Chris Otto, and Emil Jovanov,"Wireless sensor networks for personal health monitoring: Issues and an implementation." Computer Communications (Special issue), Wireless Sensor Networks: Performance, Reliability, Security, and Beyond, 2006, vol. 29 pp.2521-2533
[4] Big Data is the Future of Healthcare, White Paper Cognizant 20-20Insights
[5] TDWI_Best Practices Report on Big Data Analytics, 4th Quarter,2011.
[6] Patient Protection and Affordable Care Act. [Online] http://www.gpo.gov/fdsys/ pkg/BILLS-111hr3590enr/pdf/BILLS-111hr3590enr.pdf (Accessed JULY2017).
[7] Bill Peer, Prakashraj,Naryan Chathanur, ―Complex event processing, unberdeningBigdatacomplexities‖,InfosysLabBriefingsVol1,2013
[8] Vanitha B P and Punam Borah. ―Achieving energy efficiency for random routing in wireless sensor network using sleep/wakeup protocol‖. World Journal of Science and Technology 2012, 2(10):44-47ISSN: 2231 –2587.
[9] Dave EvansThe Internet of Things How the Next Evolution of the InternetIs Changing Everything white paper Cisco Internet Business Solutions Group (IBSG) April2011
[10] Technology Trends.[online] http://www.gartner.com/technology/research/top-10 technology-trends (Accessed March2014)
[11] DavidLuckham,―AShortHistoryofComplexEventProcessing,‖Articleonthedevelopment of Complex Event Processing, 2007.
[12] David Luckham, ―The Power of Events: An Introduction to ComplexEvent Processing in Distributed Enterprise SystemsAddison-Wesley,2002
[13] MichaelEckertandFrancoisBry,―ComplexEventProcessing(CEP)‖ArticleonComplex Event Processing, Informatik-Spektrum, Springer2009.
[14] Tommy G. Thompson and David J. Brailer, ―The Decade of HealthInformation Technology: Delivering Consumer-centric and Information-rich Health Care‖, July,2014.
[15] Health Informatics. [online]http://en.wikipedia.org/wiki/Health_informatics (Accessed August 2017) .
[16] Imthyaz Sheriff and Angelina Geetha., ―A Comprehensive Framework for Complex Event Processing‖, International Journal of Recent Advances in Engineering & Technology (IJRAET), 2013,Vol 1, No. 3. pp.11-22.

©IJRASET (UGC Approved Journal): All Rights are Reserved

[17] Chris Otto, Aleksandar Milenković, Corey Sanders and Emil Jovanov, "System Architecture of a Wireless Body Area Sensor Network For Ubiquitous Health Monitoring", Journal of Mobile Multimedia, 2006, Vol. 1, No.4. pp.307-326.

[18] Sagiroglu, S.; Sinanc, D. ,(20-24 May 2013),"Big Data: AReview"

[19] Mukherjee, A.; Datta, J.; Jorapur, R.; Singhvi, R.; Haloi, S.; Akram, W., (18- 22 Dec.,2015) , "Shared disk big data analytics with ApacheHadoop"

[20] Aditya B. Patel, Manashvi Birla, Ushma Nair ,(6-8 Dec. 2012),"AddressingBig Data Problem Using Hadoop and MapReduce"

[21] Abbott, PA & Coenen, A 2008, 'Globalization and advances in information and communication technologies: The impact on nursing and health', NursingOutlook

[22] Bakshi, K. 2014, 'Considerations for big data: architecture and approaches',Aerospace conference, IEEE, pp.1-7.

[23] Bhattacherjee, A & Hikmet, N 2007, 'Physicians' resistance toward healthcare informationtechnology: a theoretical model and empirical test', European Journal of Information Systems: Including a Special Section on HealthcareInformation.

[24] Bradley, P., & Kaplan, J. (2010), 'Turning hospital data into dollars',HFM (Healthcare Financial Management), vol. 64, no. 2, pp. 64-68. Brownstein, CA. & Wicks, P. 2010, 'The potential research impact of patient reported outcomes on osteogenesis imperfecta', Clin orthop relat Res, vol. 468, no.10.

[25] Dean, J. and Ghemawat, S., "MapReduce: simplified data processing on large clusters", Communication of The ACM - 50th Anniversary Issue: 1958 – 2008, 51, 1, January, 2008, pp 107- 113.

[26] Dieste, O., Grimán, A. and Juristo, N., "Developing search strategies for detecting relevant experiments", Empirical Software Engineering, 14, 5, October, 2011, pp 513-539.

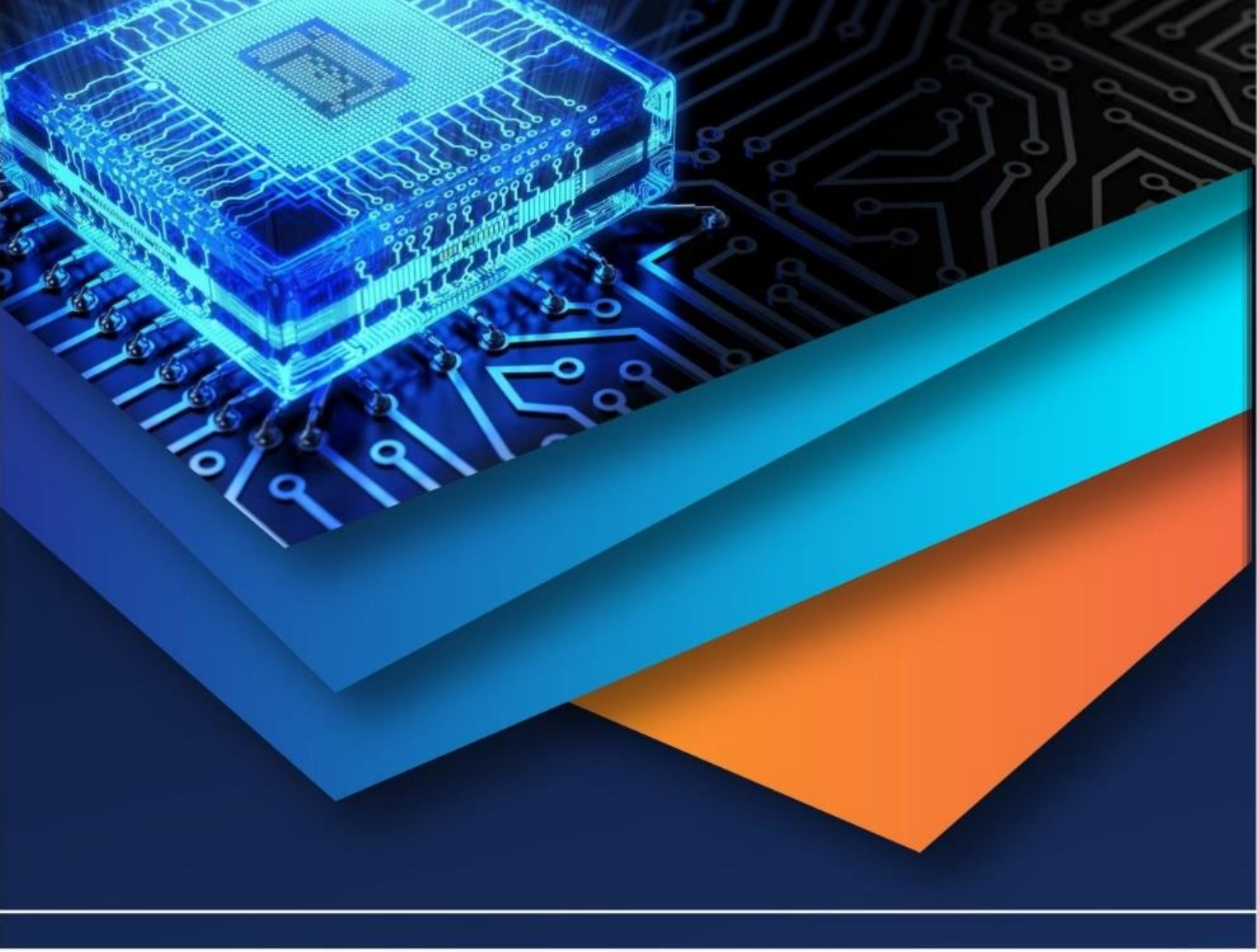
[27] Gabe,T.2015. "Poverty in the United States: 2013." https://www.fas.org/sgp/crs/misc/RL33069.pdf Retrieved 09 August,2017.

[28] Ganeshkumar, P., Arun Kumar, S. and Rajoura, OP., "Evaluation of computer usage in healthcare among private practitioners of NCT Delhi", Stud Health Technol Inform, 169, 4, 2006, pp960-964.

[29] Tawseef Ahmad Naqishbandi and Imthyaz Sheriff, C.: Big Data, CEP and IoT Redefining Healthcare Information Systems and Analytics. International Conference on Advances Research in Engineering and Technology (2014).

[30] www..goggle .com (accessed December 2017)

©IJRASET (UGC Approved Journal): All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)