

Big Data Analytics for Medical Applications

Nivedita Das

Kiit University, Bhubaneswar, India
Email: niveditads26@gmail.com

Leena Das, Siddharth Swarup Rautaray and Manjusha Pandey

Kiit University, Bhubaneswar, India
Email: {ldasfcs, siddharthfcs, manjushafcs} @kiit.ac.in

Received: 17 November 2017; Accepted: 10 January 2018; Published: 08 February 2018

Abstract—Big Data is an accumulation of data sets which are abundant and intricate in character. They comprise both structured and unstructured data that evolve abundant, so speedy they are not convenient by classical relational database systems or current analytical tools. Big Data Analytics is not linearly able to expand. It is a predefined schema. Now big data is very helpful for backup of data not for everything else. There is always a data introducing. It also helps to solve India's big problems. It also helps to fill the data gap. Health care is the conservation or advancement of health along the avoidance, interpretation and medical care of disorder, bad health, abuse, and other substantial and spiritual deterioration in mortal. Health care is expressed by health experts in united health experts, specialists, physician associates, mid-wife, nursing, antibiotic, pharmacy, psychology and other health. This paper focuses on providing information in the area of big data analytics and its application in medical domain. Further it includes introduction, Challenging aspects and concerns, Big Data Analytics in use, Technical Specification, Research application, Industry application and Future applications.

Index Terms—Big data Analytics, HIV/AIDS Prediction, Healthcare System, R Programming, Statistical Analysis, Bioinformatics Application.

I. INTRODUCTION

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 [4]. In a relational database it gathers schemas, tables, queries, reports, views and other elements. Structured data is stored in database in sequential format. 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. The framework is actual layered structure indicating what kind of work done in this paper.

Big Data is actually an enormous size of data that cannot be gathered and handled using classical access within a given time limit. There is a lot of misinterpretation; we are revering the big data analytics. We usually purpose the data to introduce to the data either in GB/TB/PB/EB/anything that greater than size. That does not characterize the term "Big Data" entirely. A like a small number of data that can be indicated to as a Big Data provided by the context it is being used [13]. Let assume an example and try to explain this Big Data Analytics to you. For instance, if you try to attach a record, i.e., of 100MB in size to an E-mail we would not be capable to do so. As the E-mail system would not support the attached document with respect to E-mail can be reverred to as Big Data Analytic.

Dissimilar types of data are exploded as well as a way of machinery into the space obtainable for functioning with data. Big Data Analytics is represented and appropriate objections linked to the operation along with data which are explained. In 2011, Peter Sondegaard, Senior-Vice President and Global Head of Research for World Wide Information Technology Research Company Gartner stated that "Information is the oil of the 21st Century in health care and logic is the combustion engine". The Institute of Medicine (IOM) in 2012 entitled "Excellent responsibility at low cost: The way to moderately learning health care in America" established that the American medical system has become far too complicated and too expensive to extend business as expected. The basic efficiency and ability to manage a rapidly depending upon a base and the world system focus on to keep patience in all the improvements, Security and Feature of concern and the nation economic balance and go comparative.

Today's hospital data uses to be accessible, logically to support improved health care transmission [3]. Big Data makes more conquer new structure of judgment by calculating and observing process diagonally. We can correlate data more efficiently like in size specialty towards workflow grows efficiency and improved in health care. Systematic analysis of extensive data can have to detect pattern so that the proper treatment for individuals and could track outcomes. A digital network gathers knowledge sharing, delivering contextual

information and able to more efficient decision making. Healthcare is only be fitted by Big Data when it is made structure, relevant, smart and accessible.

In economics sense, a patient goes to a hospital or clinic when they feel ill, then they have taken care by doctors or nurses and they pay for this service. And most of the prevention they didn't serve as an increment will be done that means a doctor would take care of a patient and after curing the patient will pay some amount [14]. There are a large number of diseases in the entire world. The main feature is Government. The Government of different countries has taken different approaches for their health care system and tries to provide the best results. The entire health care system is shown in Figure 1.

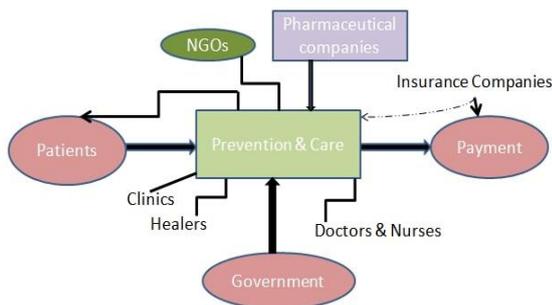


Fig.1. Health Care System

Big Data is an expecting thinking and generalizing value from very large data sets. They can be analyzed to traditional computing techniques. We can produce lots of data [19], e.g., Social Media, Public shipment and GPS. Particularly of data daily we can apply 55 new features, 340 thousand tools and 1 billion records and enter these we can produce 2.5 quintillion bytes per a day, that enough zeros specidulous we know this "Big Data", which is very important. We can carry out the methods of a data using huge numbers of computers. People in cloud and analyze number of more services and powerful algorithms. In this way we can interpret the overall million of species of data in minutes and gets accurate results. In the forthcoming, we can uniformly practice DNA of Big Data to resolve the splendid and best medical care [20]. Thus, this wise overcomes the disastrous disease like cancer would become more easily and that just started.

A. Why Big Data Analytics for Medical Applications

Big Data used in a various number of cases and it is very complex. It gathers some data, takes an action and then makes the appropriate decision [12]. 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. In health care just getting to the point, a beginning to happen to change upon the mantel looks to the doctor is doing the work which is meeting for patients, diagnosis, twitting patient beside they want to do prescribe and keeping up significant people doing well as per the mentally changing. Due to prior in the field of Big

Data Analytics, brand-new and contemporary applications are being apprehended in health care sector as well as sector of medical application. In the health care sector appliances pasturing from apparatus management to patient management are being refined.

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 conditions and 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 a constant superior need for such environment to pay better attention to the data and worth of information. "Big data Analytics is very usually means greasy data and the portion of data exaggeration boost along with the growth of data". Personal research about the big data analytics extent is absurd and there is a determined demand in medical service for creative tools for veracity and believability discipline and handling of info missed [21]. While expanded info in health care is now computerized, it is appropriate under the big data umbrella as most is disorganized and hard to use.

B. Applications of Big Data

Now-a-days big data play a vital role in different fields. Big data applications are creating a new generation in all different sectors. It also makes an easier life style for human being [11]. All applications of big data is given below Figure 2.

Banking: Big Data is helpful for detecting fraud cases in this sector. It identifies the misuse of credit cards and debit cards etc.

Medical Care: The big data is continually helps in the background of medical section. Actually the new introduced technology increases the cost of medical care. Big data is a best solution for solving this problem.

Communications, Media and Entertainment: Big Data collects analyses and utilizes customer vision. It influences mobile and social media willing. It also understands the patterns of real-time, media willing usages.

Education: Big data plays a vital role in higher edification. It is used to measure lectures potential to assure a better knowledge for both students and lectures.

Insurance: To provide consumer observations for clear and straightforward commodities by analyze and predict consumer action via data consequent from social media, GPS-enabled devices and CCTV footage. Big Data Analytics also allows for better consumer custody from insurance departments.

Transport: In this case big data helps in credit authority, technical improvement, coordination and for ambitious advantage (through packages and advance shipment act).

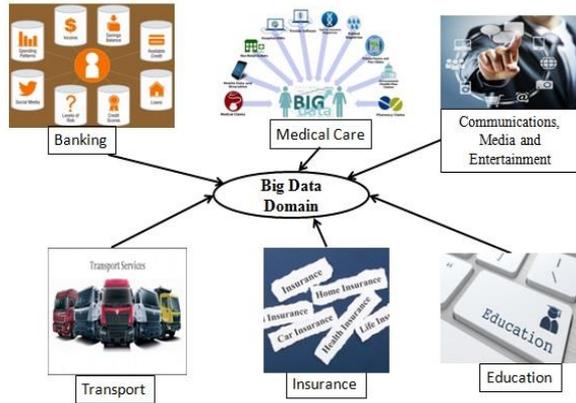


Fig.2. Applications of Big Data

C. Components of Big Data Analytics

In the field of medical sector 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 3. Since the information is spreading immensely now-a-days. Big Data defines both size and vision from unstructured, composite, noisy, mixed, representation and volume of data. These 5V's are discussed in details in below.

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.

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. The data is very large and constant in nature.

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.

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. Now-a-days data comes in the form of emails, photos, videos, pdf, audio etc.

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 making decision. It provides the actual technology.



Fig. 3. Components of Big Data Analytics

D. To whom is it good?

When an advanced technique or tool is imported to communal benefit, the primary query is how will it benefit me or what can it do for me? Here we explain who aids from utilizations of Big Data Analytics in the medical application. From society perspectives the biggest impact on healthcare [28]. So Big data analytics scratching the data in the context of smart cities all about how do you care, for aging, population, the kind of health care systems at another hand side of migration of people, health and care. I think smart cities, improving their health care by using Internet of Things (IOT), different types of analytics project opportunity. The paper is organized as follows: section 2 includes Challenging aspects and concerns, section 3 includes Big Data Analytics in use, section 4 includes Technical Specification, section 5 includes Research application, section 6 includes Industry application and section 7 includes Future application.

II. CHALLENGING ASPECTS AND CONCERNS

In this section, we will be discussing about the many objections that are experienced when arranging advanced Big Data Analytics based clarifications in the healthcare sector. A few threats experienced when arranging Big Data Analytics based explanations consist of issues related to Science, entertainment, retail, government, social, health and medicine, finance, transportation and issues related to learning.

The main challenges [5] are interoperability, manageability, security, development, reusability and maturity. 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. The experimental apparatus aiding data feature issues must be addressed. As per as reusability is concerned, must have the ability to reuse applications/ Scripts/ metadata across multiple

environments [9]. For maturity, we all know the big data technology is still rapidly evolving technology and the skills are available.

The concerns in Big Data are some of the theoretical points that should be accepted by the management to start the technique adequately. Big Data concerns are need not be disorganized with complications, they are valuable to know and crucial to handle.

III. BIG DATA ANALYTICS IN USE

Big Data Analytics has altered the approach we handle, figure out and advantage data in any management. One of the best up-and-coming fields where big data can be applied to form an adjustment is health care. Healthcare analytics have the possible to cut down amounts of hospitalization, anticipate outbreaks of infectious, evade avoidable epidemics and boost the nature of activity in general. The regular human life period is developing along world community, which acts as new objections to present hospitalization transmission approach.

A. Use of Big Data Analytics - Current & Past

The change in the world today is immense data as many as every individual are digital footprint. In this section we will discuss about some present and past big data analytics technique based technologies used in the field of medical uses.

Heart Attack Prediction: It commits the common [3] information on big data analytics on medical application, detecting heart attacks and unpredictable medical study of the particular using IOT techniques and Hadoop. It is an end-user of supports online specialist project. Hereby goal of application that grants user to bring immediate guidance on their heart disease through an intelligent system online.

Brain Disease Prediction: Brain Disease is also called brain disorders [18]. Brain injuries caused by severe mental pain and that includes blood clots, contusions, concussions, stroke, cerebral edema. The symptoms are vomiting, revulsion, speech difficulty, insensibility, memory lapses no concentration and paralysis. Data mining algorithms are used for identify the condition that reveal to the Parkinson's disease voice data set is used for diagnosing the human disease.

Prediction of Disease Outcome: It has been observed [6] that by discarding the misclassified data after clustering, the performance of classifiers is significantly enhanced. The categorization efficiency, awareness and particularity acquired by the recommended framework are found to be superior to those acquired by other competing techniques on all eight medical datasets of UCI machine learning data repository.

Using Hive & R Analyzing Diabetics: A proper description of the diabetic data set was to achieve smoothly with the assistance of hive and R [17]. The clues which were disclosed during the method can be used for establishing some forecasting models. Here only the investigation is achieved but the data which was

disclosed, it can be further used to establish efficient predictive models.

Analysis of Coronary Artery Heart Disease: The intention of this analysis was the existing conclusions of special research and proposed conclusion of sufferer with coronary artery disease (CAD) and angina pectoris (AP) in useful cardiology. High-resolution HRV analysis is an explanatory and mindset approach in the analysis and authority of CAD and AP.

Infectious Disease Outbreak Prediction: It focuses on an efficient review of dengue fever and measles outbreak prediction specifically on the data types [10], sources and factors used for the prediction. The review found that limited studies have been done on measles outbreak prediction while the existing studies of measles outbreak do not use datasets from variety of sources.

Tuberculosis Prediction: Tuberculosis caused by Mycobacterium [23] Tuberculosis. It deals with big data resources and applications, big data storage and management, big data limitations and challenges, and big data and its prospective future works with respect to Tuberculosis prediction and its cost effective treatment.

Early Stage Heart Attack Detection: An early heart attack detection system base on medical [24] biosensor is proposed. This system can detect heart attack some hours before it actually happens.

Heart Disease Prediction: In present day, many nursing homes handle health care information using health care information system [25]; when the system contains enormous number of data, used to excerpt the secret clue for preparing an imaginative clinical analysis. The major objective of this analysis is to frame a creative Prediction System of Heart Disease that provides analysis of heart disease using factual heart related data. To establish this system, clinical conditions such as sexuality, blood pressure, and cholesterol like 13 aid aspects are used. To bring extra suitable consequences, two more aspects i.e. fatness and smoking is used, as these aspects are treated as valuable aspects for heart disease.

Intelligent Heart Disease Prediction System: The accessible dataset of Heart contamination from UCI Machine Learning Repository has been prepared, pre-processed and insolvent to develop it for arrangement mechanism [26]. Coactive Neuro-fuzzy designing was recommended as a reliable and hefty approach matured to establish a nonlinear connection and calibrating between the distinct conditions. It has been demonstrated that of GA is a very advantageous facility for auto tuning of the Coactive Neuro-Fuzzy Inference System (CANFIS) frameworks and pick of optimum characteristics set. The evidence is that data processing machine cannot change humans and by analyzing the computerized disclosure conclusions with the pathologic conclusions, doctors can determine more about the perfect method to calculate the fields that computerized disclosure features.

Diagnose of Chronic Kidney Disease: Kidney is one of the key organs of Human body. The essential authority of Kidney is to drain out the unnecessary products from metabolic activities [27]. Kidney plays an important role in the removal of foreign elements like urea and

creatinine from human body. This disease is also known as chronic kidney failure. This disease is due to decrease the nephron in kidney.

HIV/AIDS Disease Prediction: The HIV virus is a type of organism. Once someone contaminated with HIV it stays in the body for the whole life. This virus damages the Immune System [8]. The immune system defects the body against the diseases. HIV virus is the source of AIDS. The full form of AIDS is Acquired Immune Deficiency Syndrome, which means deficient immune system. The difference between big data and relational databases is that the traditional table-and-column figure doesn't have by big data but that relational databases have. In classical relational databases, a schema for the data is required. The previous work done in the area of HIV/AIDS as shown in Table 1.

Table 1. Selected Work Done in the Area of Hiv/Aids

Sl. No	Author	Paper Title	Objectives	Tools/ Techniques
1	Sean D. Young et. al	A big data approach to HIV epidemiology and prevention [22]	This paper addresses new tools and techniques to HIV infections.	Bio-Informatics
2	Jokonya Osden	Towards a Big Data Framework for the prevention and control of HIV/ AIDS, TB and Silicosis in the mining industry [7]	To predict and control the mining industry	Big Data
3	S. Packiyam et. al	Aids Detection system Using Big Data Analytics [34]	In this paper this tool is helped for predicting HIV/AIDS disease in very active manner than fresh access	Cluster analysis, Hadoop
4	S. Packiyam et. al	Big Data Analysis for Aids Disease Detection System using Clustering Technique [15]	The proposed approaches for collecting and storing Big Data for analytics presented in this paper show how important it is to select the technology migration.	Hadoop, Cluster analysis

IV. TECHNICAL SPECIFICATION

R is a programming language and software framework for statistical analysis and computer graphics. It's an efficient information administration and repository facility. It is a collection of operators for evaluations on arrays and matrices [29]. It is a broad, consistent, unified group of mechanisms for information scrutiny. It is a graphical facility for information scrutiny and show either on-screen or on hard copy. It is a well-developed, smooth

and efficient encoding words which includes limited, loops, user-defined recursive function and input and output facilities.

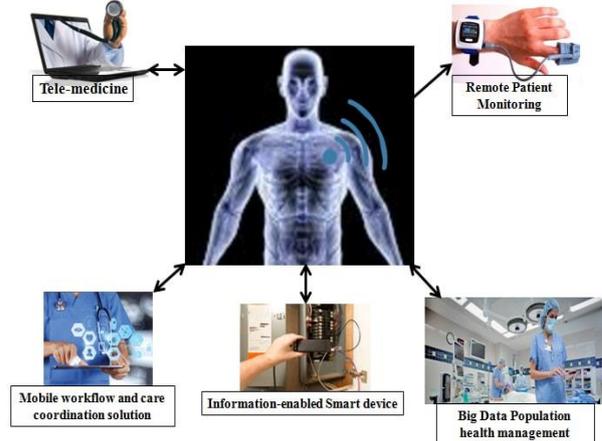


Fig.4. How Big Data revolutionize the healthcare system

In this Figure 4, the remote patient monitoring helps to organize patient basic signs smoothly in the medical and home. Using this big data population health management influence of data to direct our most challenging healthcare problems [30]. The information-enabled smart device helps to monitor and manage how we cling to our treatment. The mobile workflow and care coordination solution helps to drive best hospital outcomes and patient care. The Tele-medicine is replacing the way we interact with our care givers.

V. RESEARCH APPLICATION

A. Statistics Analysis

Statistics includes the concept of the collection, settlement, scrutiny, representation, manipulation, clarification and presentation of significant information. It compromised with all forms of data, including the preparation of data accumulation in terms of the construction of inspections and experiments [31]. There are two statistics techniques: descriptive and inferential statistics.

Descriptive statistics: A phenomenon of statistic is a total that is deriving from information, for example an intend or a standard error. It can be exactly advantageous when exploring data to get a correct set of applicable pictorial statistics. Specially, it can be very alluring to correlate statistics accessed from dissimilar (but similar) columns, or between levels of a factor. This gives an idea of the similarities or disparity between the data.

Inferential Statistics: After info outbreak, helped by judgment and depiction approaches, particular will need to recognize what legal numerical analysis approach is appropriate to inspect the data additional and to tie broad consequences. A broad number of numerical approaches have been advanced to manage many distinct types of data and construct relation between them. The origins are: Wikipedia, Training goal and eHow.

B. Applications of Bioinformatics

Actually Bioinformatics is the combination of computer science and biology [32]. In Bioinformatics and health care big data analytics brings the areas of biology, machinery and antibiotic in order to present an encyclopedia application on the developed data processing utilizations fundamental in the field of electronic medical record management. Investigation in bioinformatics evaluates biological system variations at the microscopic level. With modern trends in personalized medicine, there is an expansion needs to produce, collect and resolve these huge datasets in a convenient time frame.

VI. INDUSTRY APPLICATION

The immense application of the normal origin on this earth is asking for the immense size as well as the speed of big data. Similarly, a considerable load of information beginning the augmented industry is not being used. The untapped data averts state-of-art prominence of commodity, capacity, capability, trustworthiness, and enhanced earnings barriers [33]. In the normal wealth industry, big data analytics allows for diagnostic designing to assist intelligence establishment that is used to absorb and include enormous quantities of data from geological information, pictorial information, document and historical statistics.

Big data analytics is excessively used in the discovery of fraud in the sectors of bank. In banking sectors as the big data analytics is achieved, it discovers out all the evil assignments finished. It identifies the wrong use of credit cards and debit cards, actual of research tracks, endeavor credit threat analysis, business accuracy, and customer statistics amendment, public analysis for business, IT activity analysis, and IT planning achievement analysis. The Securities and Exchange Commission helps this big data analytics in order to manage a track of all the economical market changes.

VII. FUTURE APPLICATION

Big Data has a consequential hidden to unbolt the entire healthcare worth sequence. Big data analytics exchanged the common viewpoint of healthcare structures from discovery new drugs to patient-central healthcare for better medical outcome and increased effectiveness. 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 efficiencies based on result comparison, reducing risks, and improving personalized care. The followings are the future trends connected with medical big data analytics.

A. Text analytics applications used for E-Hospital, E-Antibiotic, and identical case recovery functions

A huge amount of health related information is unorganized as records, representations, medical or duplicate notes. Investigation articles, survey articles, medical references, and system guidance are wealthy sources for content analysis functions that intent to disclose ability by excavating these forms of content based information.

B. Genetic Applications

Genetic related information symbolizes important amounts of genetic material arrangement information and utilizations are needed to investigate and figure out the series in regards to better understanding of patient treatment.

C. Excavate and study of Biosensors function

The current information about controlling home, tele-health, handheld and sensor-established Wi-Fi is well settled information origin for medical information.

D. Applications Related to Social Media Investigation

The Social media will boost the connection between patients, specialist and society. Therefore, analysis is enforced to resolve this info to indicate emerging epidemic of disease, comfort of patient, and consent of patient to medical controls and analysis.

E. Business and Organizational Designing Applications

Regulatory information like billing, organizing, and other harmful information exist an aggressively expanding origin of data. Analysis and optimization of this type of information can accumulate enormous amounts of money and increase the continuity of a healthcare capability.

REFERENCES

- [1] Rabia Masood, S.A. Khan, M.N.A. Khan, "Plants Disease Segmentation using Image Processing", International Journal of Modern Education and Computer Science(IJMECS), Vol.8, No.1, pp.24-32, 2016.DOI: 10.5815/ijmeecs.2016.01.04
- [2] Shivam Goyal, Jaskirat Singh, "Two-Level Alloyed Branch Predictor based on Genetic Algorithm for Deep Pipelining Processors", International Journal of Modern Education and Computer Science(IJMECS), Vol.9, No.5, pp.27-33, 2017.DOI: 10.5815/ijmeecs.2017.05.04
- [3] Alexander, C., and Wang, L. Big data analytics in heart attack prediction. J Nurs Care 6, 393 (2017), 2167-1168.
- [4] Archena, J., and Anita, E. M. A survey of big data analytics in healthcare and government. Procedia Computer Science 50 (2015), 408-413.
- [5] Assuncao, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., and Buyya, R. Big data computing and clouds: Trends and future directions. Journal of Parallel and Distributed Computing 79 (2015), 3-15.

- [6] Chen, C. P., and Zhang, C.-Y. Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Information Sciences* 275 (2014), 314-347.
- [7] Drey, N., Roderick, P., Mullee, M., and Rogerson, M. A population-based study of the incidence and outcomes of diagnosed chronic kidney disease. *American Journal of Kidney Diseases* 42, 4 (2003), 677-684.
- [8] Jokonya, O. Towards a big data framework for the prevention and control of hiv/aids, tb and silicosis in the mining industry. *Procedia Technology* 16 (2014), 1533-1541.
- [9] Kelly, J. A., Murphy, D. A., Sikkema, K. J., McAuliffe, T. L., Roffman, R. A., Solomon, L. J., Winett, R. A., Kalichman, S. C., and Collaborative, T. C. H. P. R. Randomised, controlled, community-level hiv-prevention intervention for sexual-risk behaviour among homosexual men in us cities. *The Lancet* 350, 9090 (1997), 1500-1505.
- [10] Labrinidis, A., and Jagadish, H. V. Challenges and opportunities with big data. *Proceedings of the VLDB Endowment* 5, 12 (2012), 2032-2033.
- [11] Mohapatra, C., Das, L., Rautray, S. S., and Pandey, M. Map-reduce based modeling and dynamics of infectious disease. 895-898.
- [12] Murdoch, T. B., and Detsky, A. S. The inevitable application of big data to health care. *Jama* 309, 13 (2013), 1351-1352.
- [13] Patil, B. M., Joshi, R. C., and Toshniwal, D. Effective framework for prediction of disease outcome using medical datasets: clustering and classification. *International Journal of Computational Intelligence Studies* 1, 3 (2010), 273-290.
- [14] Prajapati, Vignesh. *Big data analytics with R and Hadoop*. Packt Publishing Ltd, 2013.
- [15] Raghupathi, W., and Raghupathi, V. Big data analytics in healthcare: promise and potential. *Health information science and systems* 2, 1 (2014), 3.
- [16] S. Packiyam, A. P. Big data analysis for aids disease detection system using clustering technique. *International Journal of Computer Trends and Technology* 48 (2017), 85-92.
- [17] Jeffrey, Annah M., Xiaohua Xia, and Ian K. Craig. "When to initiate HIV therapy: a control theoretic approach." *IEEE transactions on Biomedical Engineering* 50.11 (2003): 1213-1220.
- [18] Sadhana, S., and Shetty, S. Analysis of diabetic data set using hive and r. *International Journal of Emerging Technology and Advanced Engineering* 4, 7 (2014), 626-9.
- [19] Shamli, N., and Sathiyabhama, B. Parkinson's brain disease prediction using big data analytics. *International Journal of Information Technology and Computer Science (IJITCS)* 8, 6 (2016), 73.
- [20] Srivathsan, M., and Arjun, K. Y. Health monitoring system by prognostic computing using big data analytics. *Procedia Computer Science* 50 (2015), 602-609.
- [21] Zahid Ullah, Muhammad Fayaz, Asif Iqbal, "Critical Analysis of Data Mining Techniques on Medical Data", *International Journal of Modern Education and Computer Science (IJMECS)*, Vol.8, No.2, pp.42-48, 2016. DOI: 10.5815/ijmecs.2016.02.05
- [22] Tomar, D., and Agarwal, S. A survey on data mining approaches for healthcare. *International Journal of Bio-Science and Bio-Technology* 5, 5 (2013), 241-266.
- [23] Young, S. D. A big data approach to hiv epidemiology and prevention. *Preventive medicine* 70 (2015), 17-18.
- [24] Gardy, Jennifer L., et al. "Whole-genome sequencing and social-network analysis of a tuberculosis outbreak." *New England Journal of Medicine* 364.8 (2011): 730-739.
- [25] D'Agostino Sr, Ralph B., et al. "Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation." *Jama* 286.2 (2001): 180-187.
- [26] Palaniappan, Sellappan, and Rafiah Awang. "Intelligent heart disease prediction system using data mining techniques." *Computer Systems and Applications, 2008. AICCSA 2008. IEEE/ACS International Conference on. IEEE, 2008.*
- [27] Fouque, D., et al. "A proposed nomenclature and diagnostic criteria for protein-energy wasting in acute and chronic kidney disease." *Kidney international* 73.4 (2008): 391-398.
- [28] John Walker, Saint. "Big data: A revolution that will transform how we live, work, and think." (2014): 181-183.
- [29] Davenport, Thomas H., Paul Barth, and Randy Bean. "How big data is different." *MIT Sloan Management Review* 54.1 (2012): 43.
- [30] Lazer, David, et al. "The parable of Google Flu: traps in big data analysis." *Science* 343.6176 (2014): 1203-1205.
- [31] McAfee, Andrew, Erik Brynjolfsson, and Thomas H. Davenport. "Big data: the management revolution." *Harvard business review* 90.10 (2012): 60-68.
- [32] O'Driscoll, Aisling, Jurate Daugelaite, and Roy D. Sleator. "Big data", Hadoop and cloud computing in genomics." *Journal of biomedical informatics* 46.5 (2013): 774-781.
- [33] Patel, Aditya B., Manashvi Birla, and Ushma Nair. "Addressing big data problem using Hadoop and Map Reduce." *Engineering (NUiCONE), 2012 Nirma University International Conference on. IEEE, 2012.*
- [34] S. Packiyam, et al. "Aids Detection System Using Big Data Analysis." *IJANA (International Journal of Advanced Networking & Applications)*, (2017):105-109.

Authors' Profiles



Nivedita Das, is a CSI Accredited Student. She is continuing her studies in School of Computer Engineering as a M.Tech postgraduate at KIIT University, Bhubaneswar. Her research areas include Data Analytics, Big Data and Data Mining. Her paper "A Big Step for Prediction of HIV/AIDS with Big Data Tools" has been accepted in "IC4S 2017", SCOPUS Indexed. She can be reached at niveditads26@gmail.com.



Prof. Leena Das is an Assistant Professor, life-time member in Indian Science Congress and IET at the School of Computer Engineering, KIIT University, Bhubaneswar. She has more than a decade of teaching experience. Her area of interest is Real Time Systems and Software Engineering. She can be reached at lidasfcs@kiit.ac.in.



Siddharth Swarup Rautaray, PhD (Computer Science), Member IEEE is Professor at the School of Computer Engineering, KIIT University, Bhubaneswar. He has more than a decade of teaching and research experience. Dr Rautaray has published numbers of Research Papers in peer-reviewed International Journals and

Conferences. His areas of interest are Image Processing, Data analytics, Human Computer Interaction. He can be reached at siddharthfcs@kiit.c.in.



Manjusha Pandey, PhD (Computer Science), Member IEEE is Professor at the School of Computer Engineering, KIIT University, Bhubaneswar. She has more than a decade of teaching and research experience. Dr Pandey has published numbers of Research Papers in peer-reviewed International Journals and Conferences. Her areas of interest

WSN, Data analytics. She can be reached at manjushafcs@kiit.ac.in.

How to cite this paper: Nivedita Das, Leena Das, Siddharth Swarup Rautaray, Manjusha Pandey, " Big Data Analytics for Medical Applications", International Journal of Modern Education and Computer Science(IJMECS), Vol.10, No.2, pp. 35-42, 2018. DOI:10.5815/ijmeecs.2018.02.04