Pakistan Journal of Engineering and Technology, PakJET
ISSN (p): 2664-2042, ISSN (e): 2664-2050
Volume: 5, Number: 3, Pages: 22- 27, Year: 2022

Applications of Big Data in Healthcare and COVID-19: A Survey

Zamia Ilyas*, Irshad Ahmed Sumra, and Tariq Mehboob

Information Technology Department, Lahore Garrison University, Lahore, 54000, Pakistan
Corresponding author: Zaima Ilyas (zaimailyas@gmail.com)
Received: 09/06/2022, Revised: 08/09/2022, Accepted: 22/09/2022

***Abstract*- Big data proves helpful in describing data sets in the medical field so that modern software can accurately store and analyze the data. Big data plays a vital role in storing data, storing it, and improving patient outcomes in healthcare. Big Data in healthcare mainly focuses on efficient and effective data source techniques, improving patient outcomes, and availing attention services at a low price. This paper describes the characteristicsof Big Data that have different Varieties, Velocity, Veracity, Validity, and Volatility. It will discuss the other application, techniques, and technology used to handle a large amount of data, explore security issues, and help generate accurate patient outcomes. The spread of any pandemic worldwide as COVID-19 has created an Immense Amount of data, which is expanding quickly. In this paper, big data can initialize the big problem of COVID-19 and its challenges. This paper comprehends the characteristics that prompted MYCIN's prosperity and sees how precisely the framework was organized and how the system was structured. This paper will focus on the application and tools used to analyze a large amount of data and clinical records, and this survey will examine the COVID-19 issues. Moreover, big data applications are used in healthcare, and detailed study about big challenges and methodologies is used to find different techniques to gather accurate data at a low price.**

***Index Terms*-- Big data, COVID-19 and MYCIN.**

I. INTRODUCTION

Big data analytics are generated data from different sources and helps businesses, healthcare, and organizations in the modern world. Healthcare is one of the business fields within which huge information can play an important role. Big data refers to large quantities of data- created by digitizing all types of info, together with health records that are large or complicated for ancient technology to be kept, processed, and remodelled into the price. To reduce the cost of healthcare, a large amount of data must be appropriately analyzed. Government can generate a large amount of data daily; it requires the technology and techniques that help perfume real-time analysis on the different data sets. It willdiscuss other principles of big data. [1].

1. This is to identify the risks of big data handling a large amount of data.
2. Highlight the priority-based Achievements to make.
3. Creating new clinical industry collaboration data characteristics will present the latest techniques.
4. Exploring the security issues in handling big data, the role of big data, and how data can be secure in the future.

The ubiquity of web crawlers and cell phones made more data in the organization. How we can be managed whoever could take the monstrous measures of crude, unstructured data would open a money box of bits of knowledge about purchaser conduct, business tasks, characteristic marvels, and populace changes never seen. Customary data distribution centers and social databases could not deal with the required undertaking development. In 2006, Hadoop and Apache were open-source projects processing massive amounts of data. It isthe principal contrast between customary versus big data investigation.

However, there is only a detailed study about big data in healthcare, techniques, and technology of big data, which is used to enhance the big data techniques in healthcare. Section II will present big data characteristics in healthcare. Section III defines the types of big data in healthcare. Section IV is its impactchallenges and health record-related work. Section V discusses collecting the data of COVID-19 patients and its impact in real time. Section VI discusses applications and software used in healthcare and COVID19.

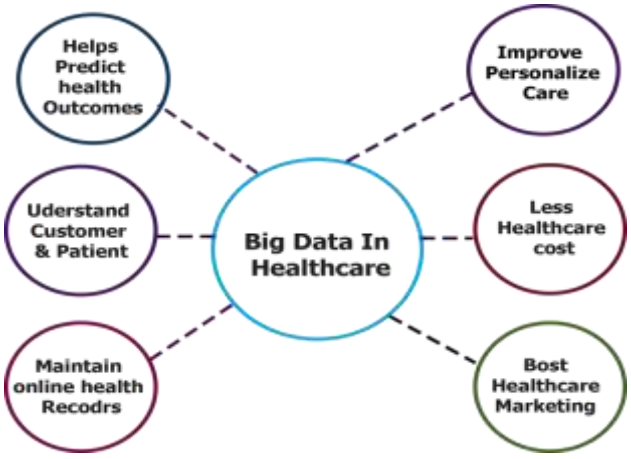

FIGURE1. Uses and benefits of big data in healthcare

Figure 1 discuss the benefits and uses of big data in healthcare. Define the major points which is highlight the opportunities of big data in healthcare.

II. BIG DATA CHARACTERISTICS IN THE HEALTHCARE SYSTEM

The characteristics of big data are a set of parameters Described in extensive data analysis. It can discuss the main feature of big data, referred to as 4vs. Big data analytics are generated data gathered from different sources and helps businesses, healthcare, and organizations in the modern world. Data is measured within the "3Vs and 5V. 5Vs refers to variety, velocity, veracity, validity, and volatility.

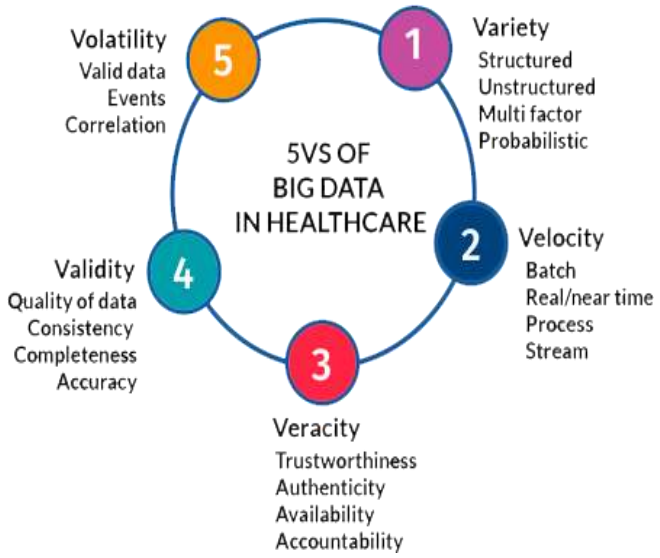

FIGURE 2. Characteristics of big data and 5vs

In Fig. 2, we define the characteristics of big data helpful in healthcare. Define the variety, velocity, veracity, validity, and volatility, which measures the data, the structure of data, and the best way of data collection. In 1st and 2nd vs. discuss the structure, unstructured, multi-factor, batch, and real-time data characteristics. In 3rd, 4th, and 5th vs. define the data quality, data consistency, complete data record and the correlated data.

A. VARIETY

In various collect, the data in different sources and ways use different techniques such as structured and unstructured data.

Structured Data: Data obeys outlined knowledge type data that follows summarized knowledge. For example, structure data include healthcare detail like the Patient record, patient information, lab values, patient demographic data, terminologies of various diseases, their symptoms, clinical service for the patient, billing information and fiancé detail.

Unstructured Data: Data that is not in the form of inherent structure. It cannot be easily organized in healthcare unstructured data examples such as radiology images oftext files, medical prescriptions of the patient written in human language, and clinical researcher's notes in the form of dynamic health records. In other words, unstructured data includes radiology images, clinical letters, etc. [2]. Furthermore, data exploration for healthcare to attain worthy insights for a heterogeneous partner (clinicians, hospitals, patients, pharmacies, etc.).

B. VELOCITY

Data must be created in real-time. In healthcare, create account information for an organization and data gathered worldwide. Velocity will be figuring out the internet of things, testing, medic actions, and machine learning.

C. VERACITY

The integrity of healthcare in big data is that accurate data should be gathered and emphasize the quality of generated data. Trust is also even more necessary than access once it involves the patient's care. The truthfulness of a data set is brutal to integrity; however, suppliers cannot utilize the insights data information that is incomplete, biased, or filled with noise.

D. VALIDITY

Validity is similar to integrity, and invalidity of data is a concern for clinicians ad researchers. A data set can be complete, but the important thing is that the gathered data is valid for which data you can use and the purpose of getting the data. Healthcare datasets must include correct information describing once, how, and by whom the information was created. The report helps to ensure that analysts perceive that their analytics are repeatable and that future knowledge scientists will question the data they're searching.

E. VOLATILITY

In healthcare, data can be changed quickly and daily patient. Data is saved in the record, so how to manage this data and which information is valuable, which metrics to include in analytics, how long data can be stored, and whichmetrics can be stored data can be deleted.

As the volume of data grows on a routine, these selections can become more and more necessary. The price of information storage could be an essential concern formany tending information technology institutions, sophisticated because HIPAA needs suppliers to retain certain patient information for a minimum of six years.

III. IMPORTANT TYPES, USES, AND ORIGINS IN HEALTHCARE

Big data is often mentioned in four primary sorts supported by information resources, i.e., medication in big data conjointly named in medical/clinical [3]. Data and medical experiments collect ample information in the medical field. Gather the various source of data like hospitals and clinics. It can summarize the major types of big data.

A. MEDICATION AND CLINICS IN BIG

Significant data sort medication and clinics records plus information collected that is achieved and summarized through all medical-related fields. In medication, big data is gathered from some background work and achievements made in medical history. As an example, it will assist in coming up with treatment methods.

The number of records and ID information of patients, diagnosis, prescription topic, doctor notes, and identifier data. Massive contribution from clinical exercises squares measures Electronic wellbeing record (EHR)/electronic anamnesis (EMR), individual wellbeing record (PHR), and clinical pictures. EMR incorporates organized and unstructured data containing all the patients' clinical action information and is usually utilized for treatment and treatment determinations.

B. BIG DATA AND MEDICAL RESEARCH

In the focal related deoxyribonucleic corrosive–sequencing, grouping sequencing might be a clinical investigation action getting the exact request of nucleotides inside deoxyribonucleic corrosive. This strategy prompts an unnecessary amount of data for recording corrosive deoxyribonucleic groupings. The clinical examination is generally performed by scientists in colleges, investigation foundations, and organizations. Their work plans to make a forward leap in people's cell, atomic, and physiological systems for medical care; rudimentary parts of it furthermore embrace natural science, clinical hereditary science, immunology, neurobiology, and cognitive science [4]. Omics information is the science data inside the sub-atomic level list that accepts hereditary science, proteomics, metabolomics, transcriptomics, epigenomes, lipidomics, immunomics, glycomics, and RNomics [5].In medical research, we can gather patient information, so using the different techniques and methods with the exported physicians significant the accurate result.

C. HOSPITAL OPERATORS SYSTEM (HOS)

Innovation for Big Data stockpiling and cycle very muchlike the Cassandra data set, has been applied; the most quality of this apparatus is that it will oblige concerning 2 million segments in a single line, making it a great deal of advantageous to address monstrous volumes of information [6]. In colossal information and medical services, one among the principal in style measure instruments, Hadoop, made by Apache, utilizes appropriation to deal with giganticvolumes of data [7][8]. In terms of information the executives, information stockrooms region units utilized for supporting dynamic, web-based managing measure (OLTP),and online investigation measure (OLAP) [9]. Moreover, AI in information handling appears to be the premier mechanical methodology in enormous informationinvestigation.

D. RESEARCHERS IN PHARMACY AND LABORATORY

In big data, researchers can get information from the laboratory and pharmacies. Practical health data analyses from heterogeneous significant data sources help medicine companies measure drug companies to live the result of designed medication with smaller and shorter trials [10]. By coupling in-memory computing technologies withautomatic systems in drug-producing units, pharmaceutical corporations will effectively integrate and analyze varied knowledge styles to create end-to-end product solutions.

In the laboratory, data can be generated, and test records. To use this data, disease and patient laboratory report analytics can be generated for future circumstances and the patient's treatment. Research of pharmacies and laboratories can develop an effective result for a better healthcare environment.

IV. CHALLENGES IN BIG DATA ANALYTICS IN HEALTHCARE

While this year, it can collaborate data in several domains like healthcare, business, scientific fields, biochemistry, public administration, and another field of sciences. Today the big challenge is collecting a large number of data, analyzing data, and finding accurate results using different techniques and models.

The complexity arises to gather different kinds of data from different sources. The big challenge is collecting the patient's data, medical records, clinical reports, telemedicine, and medical apps. Research is not only access for outcomes of a hospital.

E. BIG CHALLENGE OF DATA STORAGE AND MANAGEMENT

The big challenge in healthcare is to gather accurate data and how the data can be managed. The number of bytes of data produced daily from all over the world. The number of devices increases yearly, such as android devices, modern technologies, artificial intelligence, radiofrequency and sensory devices, etc. These devices connected systems and devices come to market. Front-line Clinics think about data storage because it's a critical costly, security, and record-saving issue for the information technology department. To implement the big data strategy, a companymust amend its approach to deciding and doing business. The {requirement} for information scientists and IT employees positively arises, which can require the corporate to change its IT infrastructure.

A. REVELATION KNOWLEDGE AND COMPUTATIONAL PROBLEMS IN HEALTHCARE

The big problem is that revelation discovers in several ways methods. Different organizations include medicine, catalogue, management, maintenance, information recovery, and representation. Other methods are being developed for big data investigation tasks, such as clustering, classification, a statistical and visual analysis.Recently, vast amounts of data are constantly generated and stored in the vaults in various research fields. The analyzer andresearcher face the big data problem because they study big data with different tools and methods. Many methods & tools used for routine data analysis are inappropriate for extensive data analysis [11]. Some techniques are not good characteristics, but at the same time, some methods have good scalability features over a parallel computer. The goal is to explore considerable data research in healthcare methods and tools to identify their strengths and weaknesses.

There are extensive data analyses and more computational complexities. This computational complexity has two types: Sample and computational complexity [12], [13], [14]. Analysis of algorithm is an integral part of a more extensive computational complexity theory which estimates the resources we need for an algorithm to solve a given computational problem.

B. CHALLENGE OF DATA GATHERING:

The main challenge of big data in the healthcare industry is that information is collected in different resources such as hospitals, clinics, government organizations, medical laboratories, and other pharmacies. The main task is to gather this data in the proper method and format, which is helpful for the Government and medical industry. So a lot of attention must collect to data used in better techniques. An extensive range of healthcare corporation units is unexpected to adopt big data strategy because it includes an immense potential to improve medicines records, technology, and finance within the healthcare sector. Gathered new data in

healthcare services in Pakistan. The challenge is that gather the big data simultaneously and find appropriate methods to find a large amount of data.

V. THE IMPORTANT BIT PART OF BIG DATA FIGHT AGAINST COVID-19 IN THE HEALTHCARE SYSTEM

COVID-19 infection increases day by day all over theworld. The world can face difficulties and not know how to manage the data and prevent these diseases. All the world is facing a crisis of health like World War Two. The world health organization collected data all around the world. They were able to get period alerts relating to the world health organization (WHO) may well be contaminated supported symptoms.

Additionally, in the present situation, QR Scanning and tools to collect a large amount of COVID19 patient data help analyze the accurate result and are less costly. All the world is facing this critical situation. In Pakistan, coronavirus confirmed cases 1,015 827. The fourth layer comes all over the world of the coronavirus—the death rate increases day by day of people due to COVID19. The scientist is successful in preparing the vaccine, which is helpful in to fight against COVID19. But the major issue is how to provide the vaccine and gather the data about the people. After COVID-19, they focused on analyzing the data and the infected death cases of COVID19.

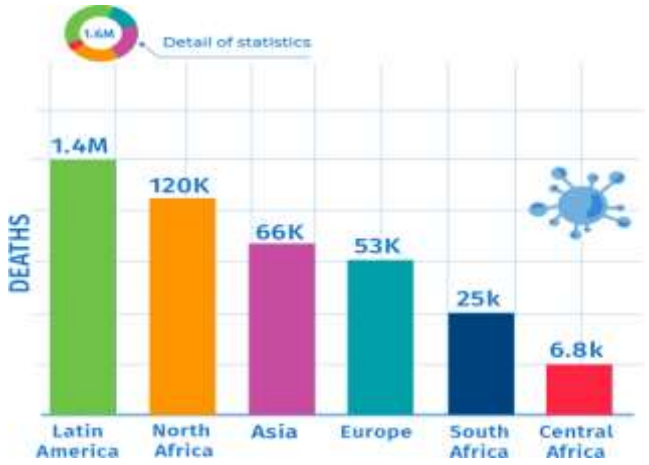

FIGURE 3: Death rate in the all-region of the world

Figure 3 all-region, which are affected due to COVID-19, have to define the total number of deaths. The coronavirus death rate increase daily.

Different resources of big data in healthcare include patient data, medical records, and laboratory examinations. In COVID-19, different apps are introduced, which help gather big data information and accurate results. In recent years, biomedical researchers have researched COVID-19 and generated significant results about appropriate public healthcare. Different resources are encountered in designing solutions against the epidemic of COVID-19.

A. PATIENT COORDINATE LEVEL

The patient data increases daily, and the number ofpatient data handling is a highly complex task. Also, sharingmovement and physiological data accumulated from wearables can add to building a predictive system.

Real-time analysis of hospital data identified with patient records and emergency clinic limits assists people with finding crowded hospitals. Moreover, AI models and Machine learning can assist with proposed function diseases by applying them to handle the pandemics and viruses of COVID-19. Incorporating such methodologies helps people pre-analyze respiratory Symptoms and the requirement for analysis by clinicians. For instance, during a review led in January 2020 [17], [18], just 37% of 4600 people without severe infection shared their data with medical services research associations. Mobile apps use to find the people who may have to contact each other about COVID19. Mobile apps can be beneficial for tracking symptoms and gaining health knowledge. Some apps are used to check the distance between different people.

B. GOVERNMENT LEVEL

Many applications can be used at the government level, which analyses the symptoms and public response to COVID19. Also, significant information science includes progressed AI procedures like profound learning and numerical and factual models. For example, autoregressive coordinated moving routine (ARIMA), improvement procedures, and molecule swarm streamlining (PSO). Furthermore, reproduction models, for instance, SEIR (Susceptible, Exposed, Infected, and Recovered states), can be utilized to precisely foresee the improvement of the episodes like COVID-19. Such a method helps gauge, control pandemics, estimate intercessions' effect, and control the virus.Social media stages fill in as a simple device for the person to share their perspectives and insights. Besides, big datacan likewise use it to get state-of-the-art data about the pandemic. The public authority can use these epic measures of information to follow individuals' perspectives about the arrangements and mindfulness of COVID-19. Some careful steps like lockdown, social separation, far-off work, and online training have secluded individuals. What's more, sometimes, it may bring about some mental medical problems. A few conclusion examination and assessment mining methods can empower to distinguish and analyze disease levels of the person preemptively.

C. HOSPITAL LEVEL

The analysis understanding checking data information can help assess the number of patients in a particular region to anticipate any normal expansion in the number of patientspast the emergency clinic limit. The work of computerized reasoning close by information investigation apparatuses has a job intending to this test. It can help in the extraction and portrayal of information progressively—the Savana system is a model. Many hospitals can use applications and methodologies to examine patients, screen remotely, and collect data about virus patients. The tool COVID-19 screening CDC informs you about your health and the discovery of the coronavirus.

VI. ANALYZING BIG DATA APPLICATIONS AND SOFTWARE USED IN HEALTHCARE

In healthcare, big data can provide many helpful aspects. Extensive data analysis discusses many applications to increase the efficiency of the work, such as electronic health records, CDS, big data in public health, medical information, and monitoring. These are the applications used in these steps.

A. EHR SOFTWARE

EHR system refers to gathering patient data and patient info regarding anything medical with the help of the electronic health record. This is always used and allowed for licensed users among a secure digital platform. So what is the purpose of the EHR code? It is to produce the associated economic plus those systems that are secure. This has become in trend and worthy as it helps a lot for integrating across utterly and differently for everydepartment regarding healthcare.

B. CDSS

A clinical choice emotionally supportive network clinic decision support system (CDSS) is a programming framework that upholds a clinician or medical care proficient dynamic. These frameworks are generally characterized as applications that present insightful information to help specialists or other clinical experts decide. A clinical support network is called a clinical decision help program (CDS program). The automated DSSs of cardiovascular is accessible in primary medical services units and clinics to satisfy the steadily expanding clinical necessities of visualization in coronary and cardiovascular sicknesses. Currently, PC-based choice help systems have been carried out in different fields of cardiovascular consideration [19]. The CDSS data-driven applications are QRM, Help System, MYCIN, Iliad, and DX Plain [20].

a) QRM/INTERNIST-1 SYSTEM

QRM is typically used in CDSS, which helps the physician using the internist1. QRM is used large community of people like practitioners and students. Internist 1 is almost contained in medical knowledge. The most effective CDSS system is used in artificial intelligence and ranking [21].

There will be three modes regarding HP that are called Health professionals. In its primary way, QMR is a system that gives recommendations like Internist-1 did (usinga similar content-associated evaluation scheme). There isalso a need for quick references regarding medical news that600 diseases may be related to sharing characteristics. Third, as a medical programmer, it will mix several features or conditions and confirm the implications.

b) HELP SYSTEM IN BIG DATA

This framework utilizes information based on choices through the information put away in its coordinated clinical data set. Time-driven dynamic abilities are additionally accessible inside the HELP framework.

c) MYCIN SYSTEM IN CDSS

MYCIN was first introduced into the field during the 1970s and was a framework intended to determine and give treatment ideas for patients who seemed to have a Blood/bacterial disease [22]. To comprehend the characteristics that prompted MYCIN's prosperity, see how precisely the framework was organized. By using creation rules and essentially a regressive approach, MYCIN had the option to get to data contained in its deduction motor and information base to produce likely patient results to be considered by medical care suppliers [23].

MYCIN is an intelligent master framework for determining and treating focal sensory system disease. It is made out of interviews and understanding. Furthermore, there is are also rules included as per lab. Consequences mimic master thinking measures, help clinicians decide proposals. The framework receives the strategy off chance that induction manages and delivers more than 400 epitomized information master judgment rules.

d) LEAD SYSTEM

Iliad System works as a specialist advisor to teach differential Diagnosis and an information-based patient case test system to educate and solve clinical test issues. Iliad is aclinical master counselling System created. It is utilized as a meeting device or a recreation preparing device instructing.

e) DX PLAIN

DX Plain is a very supportive system that covers everything about applied science. DX justifies why every one of those diseases can be thought of and suggests what more clinical data would be helpful to gather for every disease.

C. APPLICATION FOR BIG DATA:

BDA addresses a way out for the electronic and web-based media information to predict health outbreaks dependent on customers' research, social substance, and inquiry physicians.

Something else is known as the BDA's function in observing disease networking. Exploration shows that 33% of consumers use personal communication for medical services.

GBD provides a worldwide examination that evaluates mortality and incapacity from central infections, disease,and Risk factors. GBD is a coordinated effort of more than 1,800 specialists utilizing big data. All around the world, loose bowels were a primary source of death, just as a leading source to change.

It generally goes through direct medication security, especially ADRs, and distinguishes the powerless populace. ADR is characterized as an unsafe or terrible response coming about because of an intercession identified with the utilization of a therapeutic item [24]. Big data can utilize the ADR in the clinical organization and warrants anticipation, definitive treatment, adjustment of the measurements routine, and withdrawal of the item.

With its rise, there is also a rise in viewing HSMS as a quick information asset of the researcher for having direct ADR data, Inferable from the benefits discussed [25].

Alongside that displayed for the safety of drugs, BD can accomplish tremendous impacts in identifying defenceless populaces. An extensive collection of EHRs gathered by different clinical medicines gives a chance.

Here we have another form of HER that is called EPR.

D. APP OF CORONAVIRUS

The Government, hospitals, and clinics use the app to collect patient data and knowledge about the patient's health. The spread of the worldwide pandemic, COVID-19, has created an Immense Amount of data, which is expanding quickly. This data can be utilized by applying extensive information analytics techniques in different areas, including diagnosis, medical services dynamics, and the medical industry. Large data applications for COVID-19 are diagnosis, which screens the patient's ailment, including temperature, circulatory

strain, pulse, respiratory checking, glucose identification, and healthcare decision-making. Venture frameworks with capacities and usefulness for enormous information applications are known as large information data analytics stages. It helps organizations uncover recently disregarded Relationships, market patterns, and essential data from many massive data. Big information applications for COVID-19 are diagnosis, which screens the patient's medical issues, including temperature, pulse, respiratory monitoring, healthcare decision making, and estimate or predicts risk score.

VII. CONCLUSION

This research summarizes how much data canbe managed in the healthcare system. This paper introduces the basic concepts of big data in healthcare, the characteristics of big data, and the 5vs introduction, which is used in big data challenges, applications, and modern technology related to big data in healthcare. Additionally, we summarize the critical role of big data in the fight against COVID-19 in healthcare. In general, we study the big problem of COVID-19 and the challenges of big data in healthcare. This paper reviews various studies on the use of big data analytics in health care, along with the application and techniques in the medical sector. The focus is also placed on finding proposed methods to handle a large amount of data. The Current Research of Medicine isnot yet mature, and there are many problems we need to resolve. Gathering accurate data mining, analysis, and related ability is essential to exploit the significant examples in big data storage. In the end, we recommend focusing on the description of the recently proposed techniques, applications, and challenges that will help use big data in healthcare.

FUNDING STATEMENT

The authors received no specific funding for this study.

CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] M. K.Kakhani, S. Kakhani and S. R.Biradar, Research issues in big data analytics, International Journal of Application or Innovation in Engineering & Management,), vol. 2, no. 8, pp.228-232, 2015.

[2] A. Gandomi and M. Haider, Beyond the hype: Big data concepts, methods, and analytics, International Journal of Information Management), vol. 35, no. 2, pp.137-144, 2015.

[3] Yuen-Reed, G., & Mojsilović, A. The role of big data and analytics in health payer transformation to consumer-centricity. In C. Weaver, M. Ball, G. Kim & J. Kiel (Eds.), Healthcare information management systems, 2016, (pp. 399–420).

[4] X. Jin, B. W.Wah, X. Cheng and Y. Wang, Significance and challenges of big data research,Big Data Research, vol. 2, no. 2, pp.59-64, 2015.

[5] Abenstein, J. P., & Tompkins, W. J. A new data-reduction algorithm for real-time ECG analysis. IEEE Transactions on Biomedical Engineering, vol. 29, no. 1, 43–48, 1982.

[6] Abernethy, A. P., Wheeler, J. L., & Bull, J. Development of a health information technology-based data system in community-based hospice and palliative care. American 2011.

[7] Alyass, A., Turcotte, M., & Meyre, D. From big data analysis to personalized medicine for all: Challenges and opportunities. BMC Medical Genomics, vol. 8, no. 1, 33, 2015.

[8] Wang, Y., Kung, L., Ting, C., & Byrd, T. A.. Beyond a technical perspective: Understanding big data capabilities in health care. Proceedings of 48th Annual Hawaii International Conference on System Sciences vol. 48, pp.3044-3053, Hawaii, USA, 2015.

[9] Antonie, M. L., Zaïane, O. R., & Coman, A. Application of data mining techniques for medical image classification. Proceedings of the Second International Conference on Multimedia Data Mining, 94-101. doi:10.1.1.23.9742, 2001.

[10] Belle, A., Thiagarajan, R., Soroushmehr, S. M., Navidi, F., Beard, D. A., & Najarian, K. Big data analytics in healthcare. Biomed Research Internatioan, vol. 2015, pp. 370194, 2015.

[11] Chen, J., Qian, F., Yan, W., & Shen, B. Translational biomedical informatics in the cloud: Present and future. BioMed Research International, vol. 2013, pp. 658925. PMID:23586054, 2013.

[12] Belle, A., Thiagarajan, R., Soroushmehr, S. M., Navidi, F., Beard, D. A., & Najarian, K. Big data analytics in healthcare, 2015.

[13] Jee, K., & Kim, G. H.. Potentiality of big data in the medical sector: Focus on how to reshape the healthcare system, 2013.

[14] Christopher C. Yang, H. Y., Jiang, L., & Zhang, M. Social media mining for drug safety signal, 2009.

Detection. Proceedings of the international workshop on Smart health and wellbeing, 2012.

[15] Sepulveda, J. L., & Young, D. S.. The ideal laboratory information system. Archives of Pathology & Laboratory Medicine, vol. 137, no. 8, pp. 1129–1140, 2013.

[16] Poulymenopoulou, M., Malamateniou, F., Prentza, A., &Vassilacopous, G. Challenges of evolving PINCLOUD PHR into a PHR-based health analytics system. Paper presented at the Proceedings of the European, Mdediterranean & Middle Eastern Conference on Information Systems EMCIS, 2015.

[17] Kaisler, S., Armour, F., Espinosa, J.A., Money, W.: Big data: Issues and challenges moving forward, In 2013 46th Hawaii international conference on system sciences, IEEE, 2013, pp. 995-1004.

[18] Kim, G. H., Trimi, S., & Chung, J. H. Big-data applications in the government sector. Communications of the ACM, vol. 57, no. 3, pp. 78-85, 2014.

[19] Preen, D. B., Holman, C. D., Spilsbury, K., Semmens, J. B., & Brameld, K. J. Length of comorbidity lookback period affected regression model performance of administrative health data. Journal of Clinical Epidemiology, vol. 59, no. 9, pp. 940–946, 2006.

[20] Roberts, E. B., Health information systems. Clinics in Laboratory Medicine, vol. 23, no. 5, pp. 672–676, 1985.

[21] Braunstein, M. L, Health big data and analytics. Practitioner's Guide to Health Informatics Berlin, Germany: Springer International Publishing, 2015, (pp. 133–149).

[22] Schadt, E. E. The changing privacy landscape in the era of big data. Molecular Systems Biology, vol. 8, no. 1, pp. 612, 2012.

[23] Ward, J. C. Oncology reimbursement in the era of personalized medicine and big data. Journal of Oncology Practice vol. 10, no. 2, pp. 83–86, 2014.

[24] Wilson, A. M., Thabane, L., & Holbrook, A. Application of data mining techniques in pharmacovigilance. British Journal of Clinical Pharmacology, vol. 57, no. 2, pp.127–134, 2004.

[25] Fenderson, & Bruce.,A.. Molecular Biology of the Cell,5th Edition. Medicine & Science in Sports & Exercise, vol. 40, no. 9, pp. 1709, 2008.