REPERCUSSIONS OF BIG DATA ANALYTICS IN HEALTH CARE – AN UPDATED REVIEW

Dr. Devesh U Kapoor

Dr. Dayaram Patel Pharmacy College, Bardoli

ABSTRACT

Health care is going to be complex due to the arrival of novel treatments, enhancing roles of providers, and changes in legislation, payment methods and healthcare information technology. The generation of big data due to fast adaptation of technology from a change of resources which are simple to complex in their content. The stakeholders of the healthcare system must recognize the need of big data and utilization of its elements for developing the models of pharmacy practice. To accurately analyze the enhancing volume of organized or unorganized biological and biomedical data from various sources such as laboratories, hospitals, pharmaceuticals companies, data mining and machine learning methods are frequently employed. The objective of this article is to offer a viewpoint regarding the application of big data towards healthcare systems. Big data can also increase the patient centered pharmaceutical service because information of health care is continuously expanding in terms of content and becoming more integrated. The current research including deriving application of a particular framework for healthcare proposes that data belonging to electronic health records to medical images can be handled by diversified data analytical techniques. The modern healthcare care companies can probably transform the personalized medicine and medical treatment with the help of robust amalgamation of healthcare and biomedical data.

Keywords: Big data, Predictive analytics, Internet on things, big data tools, Healthcare framework

1. INTRODUCTION:

There is a big challenge for data scientists for cautious assimilation and enactment of huge volumes of medical data gathered through diverse platforms. To encourage personalized and effectual treatment, it is imperative to assemble together analytics, health informatics and bioinformatics which will revolutionize the healthcare system **(Raghupathi & Raghupathi, 2014).** Moreover, to assess significant information from the structured, unstructured, semi structured and complex data innovative approaches and technologies should be cultivated (**Giudice, Musarella, Sofo, & Ursino, 2019)**. The biggest advantage of big data is due to its immeasurable possibilities. The computational experts put more efforts to fabricate novel approaches to investigate and deduce the huge amount of medical data in a given time limit. There is a noteworthy impact of big data analytics on performance and decision making of the healthcare sector in a short span of time. The market of healthcare relied on big data is growing with an exponential rate anticipated by numerous healthcare consulting companies **(Dash, Shakyawar, Sharma, & Kaushik, 2019).** For building an improved predictive healthcare framework, medical data such as electronic health records (EHR), electronic medical records (EMR) is constantly assisting. A prospective solution for big data analysis is Quantum computing **(McGeoch, 2014)**. Quantum annealing is used for optimization of beamlet intensity through intensity modulated radiotherapy (IMRT). To enhance the ability of signal separation, a recurrent quantum neural network (RQNN) is implemented. To attain diverse goals during various stages of the drug discovery and development process, data mining methods can be

employed. The algorithms based on medical data mining are used to narrow down the search area and offer suggestions to domain professionals for hypothesis formation, analysis and experiments. Predictive analytics can also be employed in recognizing patients who can get maximum advantage of pharmacist intervention and also in assessing medical outcomes. Pharmacists can have an enhanced understanding of the risks associated with medication problems of patients by predictive analytics. The role of predictive analysis will be better, as additional patient data will become available. Healthcare analytics are classified into four types, descriptive analytics deals with the data gathered from past offers understandings about trends and benchmarks **(Raghupathi & Raghupathi, 2013)**. The predictive analytics employs forecasting and modeling to decide what is going to happen in future. The prescriptive analytics use machine learning to recommend a strategy by considering several inputs. In discovery analytics, machine learning used to investigate the raw data to establish connections, patterns and outliers **(Muneeswaran et al., 2021).**

2. CONCEPTS OF BIG DATA:

It is defined as an enormously huge and complex database that accumulates dissimilar types and scale of information, collected from multiple sources. Big data related to healthcare include clinical trial data, electronic health report, administrative claims etc. Big data processing includes data feeding to the system, data preserving in storage, data computing and analyzing, outcome visualization **(Senthilkumar, Rai, Meshram, Gunasekaran, & Chandrakumarmangalam, 2018).** The clustering software of big data cartel the smaller machines resources, offering numerous advantages such as pooling of resources, High availability, and easy scalability. Due to the healthcare information digitalization, there is a surge of healthcare big data & value based care, stimulating the healthcare industry to employ data analytics to make professional decisions strategically **(Pramanik, Lau, Demirkan, & Azad, 2017**). Healthcare industry is dealing with numerous challenges regarding the volume, variety, velocity & veracity of healthcare data. Big data plays a vital role in stimulating health care innovation which is affected by numerous financial models such as the need of patient, motivation of provider & technology advancement **(Kumari, Tanwar, Tyagi, & Kumar, 2018)**. Big data is efficiently providing the best treatment option to patients and providers which are based on population statistics.

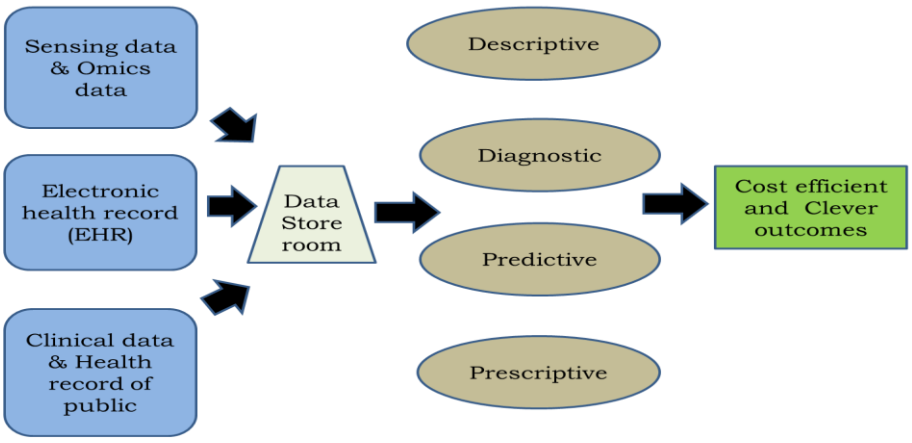

Figure 1: The working concept of big data analytics

2.1 Types of big data:

It is found in three forms:

- Structured data: The data which is accessed, stored & processed in fixed form format called as 'structured data'.

- Semi-structured data: This types of data consists both the forms such as fixed and unknown format.
- Unstructured data: The data which is available in unknown format or scattered structure known as 'unstructured data' **(Kitchin & McArdle, 2016**).

2.2 Tools of big data for healthcare system:

2.2.1 Data integration:

It provides elite prescriptive analytics for huge amounts of data, facilitating organizations for choosing cleverer decisions. Data lakes and cloud provide a platform for data integration which performs automated pipelines of data at a quicker rate (Palanisamy & Thirunavukarasu, 2019). It offers facilitation of data integration by employing on premise applications and saas based cloud services in a smart manner **(Dong & Srivastava, 2013).**

2.2.2 Visual data analytics:

It offers interactive analytics which are visual in nature at large scale with the help of information extraction by one or more sources. It gives quicker, superior interactive dashboards for the projection of extracted patterns. It provides operational and clinical data with the help of visual analytics for insights discovery **(Keim, Mansmann, Schneidewind, Thomas, & Ziegler, 2008)**.

2.2.3 Stream data processing:

There is a tool available for in memory computing of data regarding real time analytics. (SAP Hana 2017). It offers a tool which can process the unbounded stream data in scalable, dispersed and fault lenient manner. (Neumeyer et al 2010). There is a tool which gives help for the generation of distributed streaming applications that supply real time data integration, ingestion and analytics **(Cortés, Bonnaire, Marin, & Sens, 2015)**

2.2.4 Machine learning:

To process scalable mining algorithms, it provides a distributed machine learning library. There is a tool which performs advanced analytics and performs machine learning of gigantic data at extraordinary speed **(Char, Abràmoff, & Feudtner, 2020)**. It is offering solutions by using predictive analytics for big data cases. There are various machine learning techniques for big data of healthcare such as Supervised learning, Unsupervised learning, Reinforcement learning, Semi supervised learning, Self-supervised learning, Multi instance learning, Inductive learning, Transductive learning, Online learning, Transfer learning, Ensemble learning and Deep learning. Amongst these all learning, Deep learning is widely used for drug discovery and development, medical insurance fraud investigation, analyzing disease with the help of MRI, ECG and CT Scan, Genome understanding and early detection of Alzheimer's disease. Along with that there is some other learning techniques include kernel based learning, transfer learning; active learning, representation learning distributed and parallel learning are also play vital role in big data analysis **(Dhillon & Singh, 2019**).

2.2.5 Searching and processing:

It offers worthy programming constructs along with a superior performance distributed execution engine. It provides complements which are based on map computations backed by Hadoop for nested data processing with extraordinary scalability. There is a search engine available for accomplishing full text search on various platforms having superior performance **(Höög, Lysholm, Garvare, Weinehall, & Nyström, 2016**).

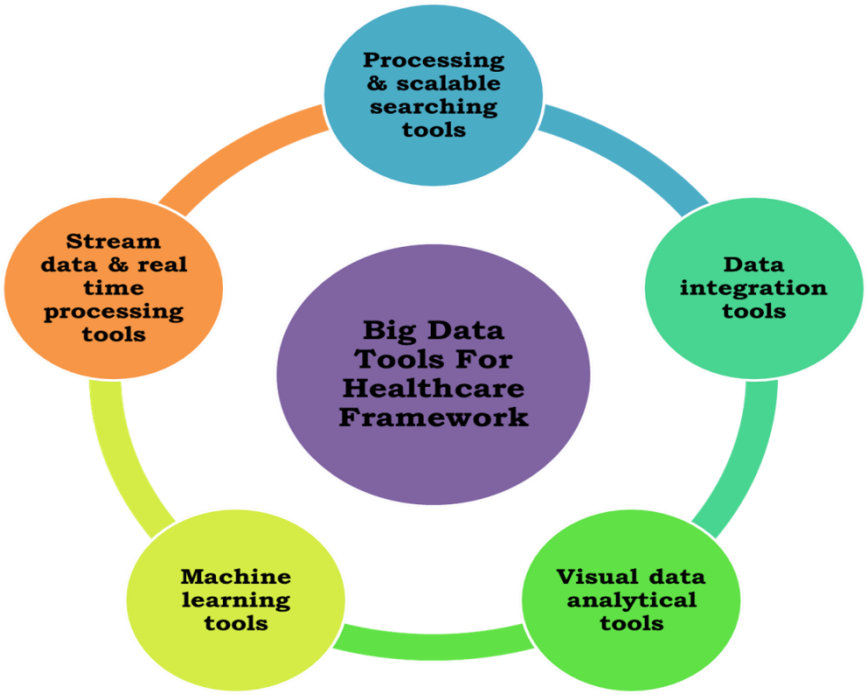

Figure 2: Different tools of big data used for development of healthcare system

3. **Electronic health records (EHRs):**
The national institute of health is collecting more than one million electronic health records of patients which include environmental data, medical imaging data and socio behavioural data. EHRs are offering numerous benefits in the management of data related to the recent healthcare system **(Yadav, Steinbach, Kumar, & Simon, 2018**). The medical practitioner has an enhanced access regarding the patient's medical history including prescriptions, medical diagnosis, allergies data, demographics, clinical descriptions and data gathered from numerous laboratory tests. Due to decline in such errors, there is diminishing in the cases of allergies by dropping errors in dose and frequency of medication. EHRs can diminish or abolish the confusion in the management of billing and claims. EHRs also give significant data of health insurance programs for quality care of the patients and also provide assistance to minimize the cost of health insurance. This electronic platform facilitates medical practitioners to develop their consultation by employing automatic reminders, vaccinations, abnormal laboratory outcomes, screening of cancer. These periodic involvements in between patients and providers of healthcare will enhance the superior care of patients **(Haas, Wohlgemuth, Echizen, Sonehara, & Müller, 2011**).

4. **Healthcare and Internet on things (IoT):**
IoT is an emerging player of the healthcare industry and has become an integral part of industry **(Yeole & Kalbande, 2016)**. The information & communication system such as near field communication (NFC) and radio frequency identification tags (RFID) are widely employed in the healthcare industry **(Khosravi, Karbasi, Shah, Brohi, & Ali, 2016)**. One can get critical information after analyzing the data from these devices which can be beneficial in the field of healthcare. IoT devices are chief contributors of big data in terms of observing the people's health. The data obtained from these reliable sources can be utilized to offer trustworthy and effectual treatment to patients with a chronic illness. There is a noteworthy decline in healthcare costs by intervention of IoT. Various IoT devices such as health tracking devices, biosensors, clinical services for evaluating vital statistics are used frequently by healthcare professionals. We can assimilate this data with already existing data such as electronic medical records (EMRs), resulting in the better forecasting of health status of patients **(Yuehong, Zeng, Chen, & Fan, 2016).**

Big data also helps in finding ways for specific outbreak of disease. An advanced hardware and updated software needed for the analysis of data through IoT. The stakeholders related to the healthcare system are trying to reduce the cost and enhance the quality of care by using big data analytics. Various organizations are using artificial intelligence to analyze the published reports to make big data analytics easier. Watson health of IBM is serving a platform for data sharing and analyzing to hospitals and researchers. In case of cancer research, Flatiron health is offering diverse technology related services towards healthcare big data analytics.

Table 1: Various software available for healthcare analytics (Xu et al., 2018)

Name of company	Solution provided	Url
Enlitic	Offer deep learning healthcare diagnosis by employing huge scale data sets from clinical tests	www.enlitic.com
Linguamatics	Offer platform for useful information through unorganised healthcare data by mining of text	www.linguamatics.com
Roam analytics	Gives infrastructure for natural language processing for contemporary medical systems	www.roamanalytics.com
Ayasdi	Offer analytics platform based on artificial intelligence (AI) for risk management, clinical changes	www.ayasdi.com
Lumiata	It drives data transformation so one can better predict costs and manage risks.	www.lumiata.com
Optum Health	It is committed to provide best health services, leading way to better experience at lower cost for you.	www.optum.com
Apixio	Our AI solutions for risk, quality, and clinical insights unlock actionable information from administrative data and unstructured clinical information.	www.apixio.com

5. Analysing healthcare big data by various approaches of data mining:

In health care and pharmaceuticals, a massive amount of data is generated every day. As the enhancing awareness regarding the data as an imperative asset, novel data mining approaches are coming to market to get the benefit of large volumes of data. There are some approaches which are utilizing data analysis such as classification, clustering, regression analysis and rules of data association **(Mdaghri, El Yadari, Benyoussef, & El Kenz, 2016)**.

5.1 Classification of data:

It is the most efficient and effective process for data organization. It is the most widely used method for data mining in the healthcare sector. Classification consists of two steps, training and testing. In 2015 Macrae *et al* employed a pattern recognition method to identify respiratory conditions of infants by forming an algorithm for unstructured data classification written by medical practitioners at primary care consultations **(MacRae et al., 2015)**. In 2015 Azar *et al* provided a new neuro fuzzy classifier having particular features for selection, reduction in dimensionality and classification **(Azar & Hassanien, 2015)**. In 2012 Estella *et al* fabricated an innovative system for classification of medical resonance imaging (MRI) images of patients suffering from neurodegenerative disease. The chief reason is to fabricate this system to get assistance in decision making for classification tasks.

5.2 Clustering of data:

It is a different method from classification. There is no predefined classification in this type of data mining. Clusters are, when a huge database is distributed into small no of groups. The data

is divided on the basis of similarities. The algorithms of clusters that have the same cluster consist of data, are more similar in nature to each other in comparison to other groups **(Kerr, Ruskin, Crane, & Doolan, 2008)**. In 2014 Jaskoviak *et al* investigated the clustering methods employing gene annotation to evaluate the quality of gene expression **(Jaskowiak, Campello, & Costa, 2014**). In 2015 Kar *et al* also studied the data of gene expression by employing categorized clustering methods by employing genetic algorithms **(Kar, Sharma, & Maitra, 2015).** In 2016, Liao *et al* applied various clustering approaches to ascertain the expenditure patterns of patients, started hemodialysis for end stage renal disease (ESRD) **(Liao, Li, Kianifard, Obi, & Arcona, 2016)**.

5.3 Data association:

It plays a vital role to find out the relationship between disease, human health condition and disease symptoms in the healthcare industry. The competences of data mining techniques are enriched by employing assimilated approach of classification and association techniques. Rashid *et al* (2014), have employed association rules of data mining to find out the disease pattern carried by a patient. They have fabricated a prototype system for clinical state correlation prediction (CSCP), which mines the data from healthcare database of patients, then converts the online transactional processing (OLTP) data into data warehouse by creating rules of association **(Rashid, Hoque, & Sattar, 2014**).

Lakshi *el al* (2017) employed multiple-criteria decision analysis to choose the precise rule of association mining algorithm for association rules extraction from medical records **(Lakshmi & Vadivu, 2017).**

This unified methodology is employed to determine the database rules and an efficient classifier is developed on the basis of these rules.

5.4 Regression analysis:

It is commonly employed method in investigating big data of healthcare for assessing the relationship between properties or variables. It consist various researches concern such as data sequence forecasting, associations among data and data sequence trend features.

In 2017, Wan *et al* mentioned factors contributing to patient satisfaction and also offer an algorithmic method to evaluate it from analysis of healthcare big data **(Wan & Alagar, 2017).**

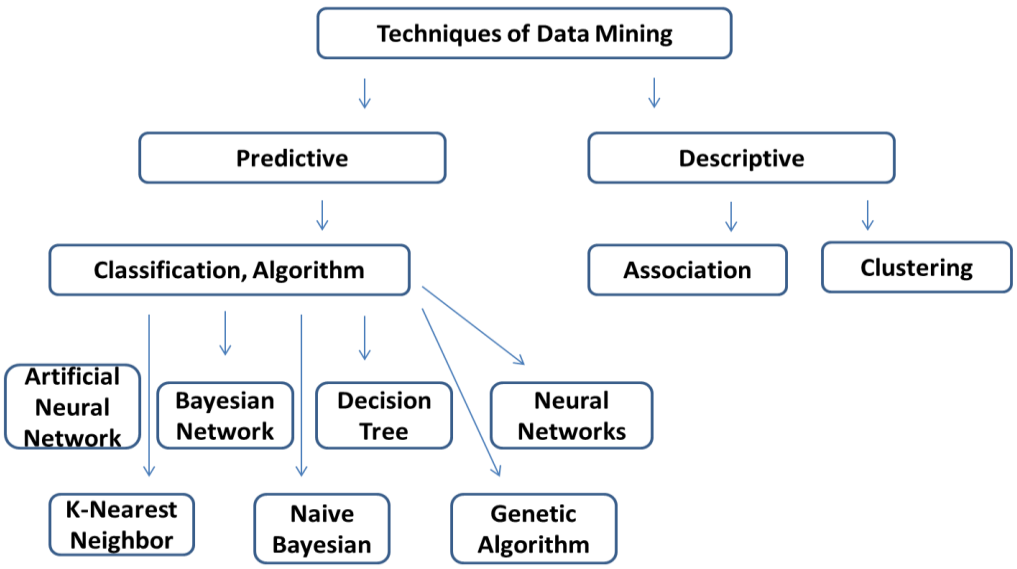

Figure 3: Classification of data mining approaches

6. Big data and patient care optimization:

The role of big data analytics is vital for implementation of EBM in true manner. For proper implementation of EBM, diverse data analytical methods such as Spark, Hadoop, Mapreduce, Mahout is used **(Roy, Rautaray, & Pandey, 2018)**. Amongst all Hadoop is widely employed open source distributed platform of data analytics for data distribution. There are two types of data sources internal and external. Usually clinics provide the data known as internal data sources. These data proofs are extracted from clinical trials and decision support system. The government sources, insurance company, laboratories are chief source of external data source. These data derived in different formats such as .csv, flat files, relational tables and ASCII/text. The big data is also play crucial role for pharmacovigilance involving new electronic approaches that are employed to investigate big and surging volume of information regarding adverse drug reaction IN database of spontaneous reporting system (SRS). Data mining is widely employed to identify the fresh drug-adverse drug reaction relations regarding drug safety investigation process **(Ventola, 2018)**.

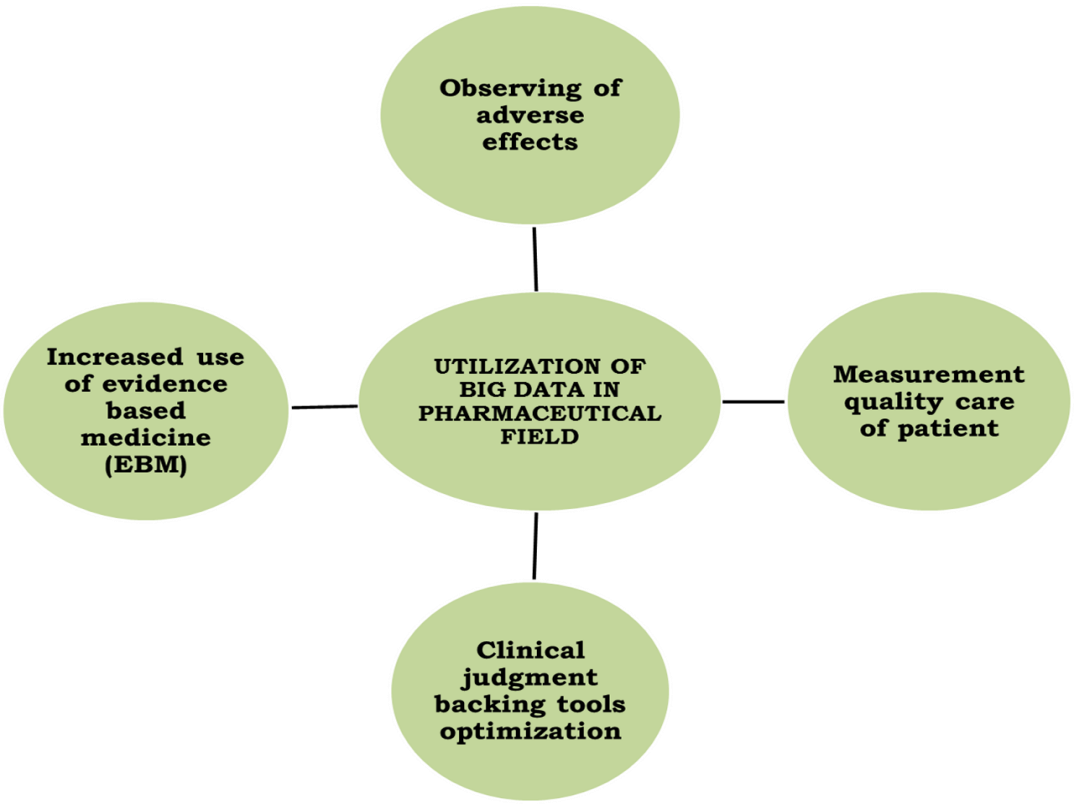

Figure 3: Management and operations of pharmaceutical system using big data

7. **Challenges associated with big data analytics:**

There is a big challenge related to collection, protection and sharing of healthcare data. There is the possibility of data transformation by using sophisticated technologies of big data analytics. There are some concerns related to privacy, security, governance and standards should be addressed properly, to mitigate the wrong use of this data **(Acharjya & Ahmed, 2016).**

7.1 Data analyst/Data scientist:

It is still a strenuous task to find people having expertise in computer science or information technology along with statistics. Additionally encouraging the growth of healthcare big data

analytics requires analysts having sound knowledge of clinical and analytics. We need personnel who are masters in data processing technology and medical data management. They can apply the specific management model to build the infrastructure of data, an unremitting platform of application & research, confirm steadiness and attain cross cutting cooperation **(Asamoah, Sharda, Hassan Zadeh, & Kalgotra, 2017).**

7.2 Retention and management of big data:

The data related to healthcare should be available for at least 5 years. There is a need for long term data tracking accessed for which purpose and by whom it is accessed. The collected healthcare data require proper formatting, checking for accuracy and made available for diverse purposes such as billing, clinical & administrative. The management of huge volume and velocity of this data is very challenging **(Ahmed et al., 2017).**

7.3 Challenge related to privacy & security of big data:

The data issues related to privacy & security are very imperative for healthcare business. Successful theft of healthcare data is very beneficial for criminals, at the same time tremendously detrimental for healthcare institutions. The health records consist of personal data ranging from lab tests, diagnosis and details of credit card too. So data breaching will be highly damaging to the individual. There are some other challenges such as data encryption, masking of data and data protection stringent methods are facing by healthcare providers **(Peek, Holmes, & Sun, 2014).**

7.4 Mismanagement between clinical and administrative systems:

The medical records and billing details retained by healthcare personnel of hospitals should be accurate when dealing with insurance claims. But it is quite often that there is a gap found in treatment codes of patients between clinical and administrative systems regarding management of such data **(Lee & Yoon, 2017).**

7.5 *Patient data sharing to other stakeholders:*

For easy and rapid access, maximum records of patients are kept in a centralized database. The actual challenge encountered when this data is to be shared to other stakeholders other than medical practitioners. It is a cumbersome task for pharmacies & external healthcare providers to get such kind of data **(Mittelstadt & Floridi, 2016).**

7.6 Data simulation and modelling:

Big data is ideal for modeling and data simulation, but the prerequisite is to recognize, arrange and pool the appropriate data so it can be employed to diverse models. It is very challenging to envisage and investigate the results and precise data extraction without proper data structuring **(Giabbanelli, 2019).**

8. Applications to analyze healthcare big data

Big data offer extensive sustenance to analyze the healthcare data. It has done the progress towards diverse areas such as public health, analysis of disease pattern, personalized medicine and medical information services. There are some applications which provide excellent results in clinical decision support system **(Kumar & Singh, 2018).**

8.1 The Help system:

It is the first ever clinical decision making system which offers the health characterization by employing logical processing system. This system takes decisions by utilizing clinical data of diverse sources which is stored in its clinical database. This system consist accounting system, a knowledge base, processor for decision making, time and data driver, repository of patient data and alerts of data reviews. This system is also capable in making decisions which are time driven.

8.2 The MYCIN system:

This system offers exceptional results for the diagnosis and treatment of infection related to central nervous system. There are three subsystems exists in this namely consultation, interpretation and rules. It make clinical recommendation and assisting physicians to determine the species of bacteria by imitating knowledgeable reasoning process which is based on clinical indication and laboratory outcome of patients.

8.3 The QMR system:

This system is one kind of clinical decision support system (CDSS) which provides help to medical practitioners by employing Internist-1 or Caduces knowledge base. This knowledge base is quite similar like a medical book which contains 6000 clinical symptoms, 40,000 disease relationships and 1000 disease. It is one of the earliest clinical support system employ probability raking system and artificial intelligence. One limitation of this system is that the knowledge base of QMR has to be updated regularly.

8.4 The ILIAD system:

This system is developed by School of medicine, University of Utah known as medical expert consultation system. This system comprises 4 main components such as the data driver, the information algorithm, user interface and engine of inference. It is widely employed consulting tool for teaching and simulation training of CDSS. The Bayestan and Boolean frames is utilized to represent the knowledge of this system. The mentioned frames allow the usage of specificities and sensitivities to explain the disease relationship to its signs which deliver a rational to explain its outcomes. The consultant working with Iliad uses various inferencing mechanisms to match the medical practitioner strategy on his patient.

9. Conclusion:

There is superb potential in big data which can alter the healthcare scenario related to patient safety management, drug discovery, perfection in clinical evaluation, effectiveness in treatment, personalization care of patients. As there are huge data records in the medical industry so it is necessary to employ data mining techniques to get support in decision making and prediction to recognize the type of disease. One can get the specific information from these mining techniques which will be beneficial for diagnosis of disease.

10. References:

Acharjya, D. P., & Ahmed, K. (2016). A survey on big data analytics: challenges, open research issues and tools. *International Journal of Advanced Computer Science and Applications*, *7*(2), 511-518.

Ahmed, E., Yaqoob, I., Hashem, I. A. T., Khan, I., Ahmed, A. I. A., Imran, M., & Vasilakos, A. V. (2017). The role of big data analytics in Internet of Things. *Computer Networks*, *129*, 459-471.

Asamoah, D. A., Sharda, R., Hassan Zadeh, A., & Kalgotra, P. (2017). Preparing a data scientist: A pedagogic experience in designing a big data analytics course. *Decision Sciences Journal of Innovative Education*, *15*(2), 161-190.

Azar, A. T., & Hassanien, A. E. (2015). Dimensionality reduction of medical big data using neural-fuzzy classifier. *Soft computing*, *19*(4), 1115-1127.

Char, D. S., Abràmoff, M. D., & Feudtner, C. (2020). Identifying ethical considerations for machine learning healthcare applications. *The American Journal of Bioethics*, *20*(11), 7-17.

Cortés, R., Bonnaire, X., Marin, O., & Sens, P. (2015). Stream processing of healthcare sensor data: studying user traces to identify challenges from a big data perspective. *Procedia Computer Science*, *52*, 1004-1009.

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, *6*(1), 1-25.

Dhillon, A., & Singh, A. (2019). Machine learning in healthcare data analysis: a survey. *Journal of Biology and Today's World*, *8*(6), 1-10.

Dong, X. L., & Srivastava, D. (2013). *Big data integration.* Paper presented at the 2013 IEEE 29th international conference on data engineering (ICDE).

Giabbanelli, P. J. (2019). *Solving challenges at the interface of simulation and big data using machine learning.* Paper presented at the 2019 Winter Simulation Conference (WSC).

Giudice, P. L., Musarella, L., Sofo, G., & Ursino, D. (2019). An approach to extracting complex knowledge patterns among concepts belonging to structured, semi-structured and unstructured sources in a data lake. *Information Sciences, 478*, 606-626.

Haas, S., Wohlgemuth, S., Echizen, I., Sonehara, N., & Müller, G. (2011). Aspects of privacy for electronic health records. *International journal of medical informatics, 80*(2), e26-e31.

Höög, E., Lysholm, J., Garvare, R., Weinehall, L., & Nyström, M. E. (2016). Quality Improvement in large healthcare organizations: Searching for system-wide and coherent monitoring and follow-up strategies. *Journal of health organization and management*.

Jaskowiak, P. A., Campello, R. J., & Costa, I. G. (2014). *On the selection of appropriate distances for gene expression data clustering.* Paper presented at the BMC bioinformatics.

Kar, S., Sharma, K. D., & Maitra, M. (2015). Gene selection from microarray gene expression data for classification of cancer subgroups employing PSO and adaptive K-nearest neighborhood technique. *Expert Systems with Applications, 42*(1), 612-627.

Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual analytics: Scope and challenges *Visual data mining* (pp. 76-90): Springer.

Kerr, G., Ruskin, H. J., Crane, M., & Doolan, P. (2008). Techniques for clustering gene expression data. *Computers in biology and medicine, 38*(3), 283-293.

Khosravi, M., Karbasi, M., Shah, A., Brohi, I. A., & Ali, N. I. (2016). *An adoption of halal food recognition system using mobile Radio Frequency Identification (RFID) and Near Field Communication (NFC).* Paper presented at the 2016 6th International Conference on Information and Communication Technology for The Muslim World (ICT4M).

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society, 3*(1), 2053951716631130.

Kumar, S., & Singh, M. (2018). Big data analytics for healthcare industry: impact, applications, and tools. *Big data mining and analytics, 2*(1), 48-57.

Kumari, A., Tanwar, S., Tyagi, S., & Kumar, N. (2018). Fog computing for Healthcare 4.0 environment: Opportunities and challenges. *Computers & Electrical Engineering, 72*, 1-13.

Lakshmi, K., & Vadivu, G. (2017). Extracting association rules from medical health records using multi-criteria decision analysis. *Procedia Computer Science, 115*, 290-295.

Lee, C. H., & Yoon, H.-J. (2017). Medical big data: promise and challenges. *Kidney research and clinical practice, 36*(1), 3.

Liao, M., Li, Y., Kianifard, F., Obi, E., & Arcona, S. (2016). Cluster analysis and its application to healthcare claims data: a study of end-stage renal disease patients who initiated hemodialysis. *BMC nephrology, 17*(1), 1-14.

MacRae, J., Love, T., Baker, M. G., Dowell, A., Carnachan, M., Stubbe, M., & McBain, L. (2015). Identifying influenza-like illness presentation from unstructured general practice clinical narrative using a text classifier rule-based expert system versus a clinical expert. *BMC medical informatics and decision making, 15*(1), 1-11.

McGeoch, C. C. (2014). Adiabatic quantum computation and quantum annealing: Theory and practice. *Synthesis Lectures on Quantum Computing, 5*(2), 1-93.

Mdaghri, Z. A., El Yadari, M., Benyoussef, A., & El Kenz, A. (2016). *Study and analysis of data mining for healthcare.* Paper presented at the 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt).

Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: current and foreseeable issues in biomedical contexts. *The ethics of biomedical big data*, 445-480.

Muneeswaran, V., Nagaraj, P., Dhannushree, U., Lakshmi, S. I., Aishwarya, R., & Sunethra, B. (2021). A Framework for Data Analytics-Based Healthcare Systems *Innovative Data Communication Technologies and Application* (pp. 83-96): Springer.

Palanisamy, V., & Thirunavukarasu, R. (2019). Implications of big data analytics in developing healthcare frameworks–A review. *Journal of King Saud University-Computer and Information Sciences, 31*(4), 415-425.

Peek, N., Holmes, J. H., & Sun, J. (2014). Technical challenges for big data in biomedicine and health: data sources, infrastructure, and analytics. *Yearbook of medical informatics*, *23*(01), 42-47.

Pramanik, M. I., Lau, R. Y., Demirkan, H., & Azad, M. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems with Applications*, *87*, 370-383.

Raghupathi, W., & Raghupathi, V. (2013). An overview of health analytics. *J Health Med Informat*, *4*(132), 2.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, *2*(1), 1-10.

Rashid, M. A., Hoque, M. T., & Sattar, A. (2014). Association rules mining based clinical observations. *arXiv preprint arXiv:1401.2571*.

Roy, C., Rautaray, S. S., & Pandey, M. (2018). Big Data Optimization Techniques: A Survey. *International Journal of Information Engineering & Electronic Business*, *10*(4).

Senthilkumar, S., Rai, B. K., Meshram, A. A., Gunasekaran, A., & Chandrakumarmangalam, S. (2018). Big data in healthcare management: a review of literature. *American Journal of Theoretical and Applied Business*, *4*(2), 57-69.

Ventola, C. L. (2018). Big data and pharmacovigilance: data mining for adverse drug events and interactions. *Pharmacy and therapeutics*, *43*(6), 340.

Wan, K., & Alagar, V. (2017). *Analyzing healthcare big data for patient satisfaction.* Paper presented at the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD).

Xu, S., Rogers, T., Fairweather, E., Glenn, A., Curran, J., & Curcin, V. (2018). Application of data provenance in healthcare analytics software: information visualisation of user activities. *AMIA Summits on Translational Science Proceedings*, *2018*, 263.

Yadav, P., Steinbach, M., Kumar, V., & Simon, G. (2018). Mining electronic health records (EHRs) A survey. *ACM Computing Surveys (CSUR)*, *50*(6), 1-40.

Yeole, A. S., & Kalbande, D. R. (2016). *Use of Internet of Things (IoT) in healthcare: A survey.* Paper presented at the Proceedings of the ACM Symposium on Women in Research 2016.

Yuehong, Y., Zeng, Y., Chen, X., & Fan, Y. (2016). The internet of things in healthcare: An overview. *Journal of Industrial Information Integration*, *1*, 3-13.