Review Article

Big Data Insights and Comprehensions in Industrial Healthcare: An Overview

Weiwei He,1 Shah Nazir,2 and Zahid Hussain3

1Sichuan Academy of Medical Sciences & Sichuan Provincial People’s Hospital, Chengdu, China
2Department of Computer Science, University of Swabi, Swabi, Pakistan
3Department of Information Technology, Quaid-e-Awam University Nawabshah, Nawabshah, Pakistan

Correspondence should be addressed to Weiwei He; weiweihehot@sina.com and Shah Nazir; snshahnzr@gmail.com

Received 22 December 2020; Revised 17 January 2021; Accepted 20 January 2021; Published 31 January 2021

1. Introduction

With the rising technology and digital innovation, the healthcare technology is refined at a high rate. Healthcare data are continuously growing, and their volume is drastically increasing. Such data can make it harder to identify useful form of data producing meaningful information. Big data analytics are a useful delivery of data, producing valuable and useful insights for huge volume of data. Several advantages are there with the use of big data such as identification of patient disease in the early stage, delivering enhanced services of healthcare, reducing cost, and producing quality care. Various sources of big data are available in healthcare. These sources include electronic healthcare data, biometric data, registration, medical imaging, biomarker data, patient-reported data, clinical data, and administration data. Such data can be visualized in order to extract meaningful and valuable information. Analytic tools of big data play a significant role in analysing and integrating huge volume of unstructured, semistructured, and structured data providing diverse hospital, clinical, and many other decision services. Diverse applications of big data in healthcare are presented as follows [1–4].

Panda et al. [5] concentrated on mobile-based development of application through which the patient can send query of symptoms to the doctors. The application of mobile will be fitted out with instruction of first aid based on the severity and nature of the symptoms. Huge bulk of data was gathered from doctors and users, and the data were used for training the machine for automating the task to some level. Beneficial insights for improving efficiency and quality of care were gathered from the collected data. The patient is able to reach out for the solution of health with easiness and cheap. The approach is used as an analytic tool for gathering data for the delivery purpose and extracting useful information. Mian et al. [6] described the real-time exploration of data for datasets of healthcare through in-memory databases. Two databases of in-memory were considered for
comparing and studying the ability and responsiveness for handling complexity of classic data of health for exploring tasks. The study outlined the issues which need addressing for future developments of big data. Gonzalez-Alonso et al. [7] presented an approach for addressing the technologies of big data analytics incorporated into the care provider. Katsis [8] described the applications of big data in efficient use of public health discoveries. The study showed that the analytic approaches classically applied in public health are not up to the data task currently existing. The results elaborated the big data case study which was carried out in San Diego, California. The study analyzed thousands of variables associated to health for gaining useful insights on the determination of various health outcomes. The contribution of the proposed study is given as follows:

(i) To deliver a review report of the existing approaches for big data insights in the industry of healthcare.
(ii) To provide an in-depth review of the available sources from different point of views by publishing the papers in different areas of big data in healthcare.
(iii) To provide a gateway on the basis of which researchers can devise new approaches, techniques, and tools to facilitate healthcare.

The organization of the paper is as follows. Section 2 provides the literature study to the big data insights in the industry of healthcare. Section 3 shows the library-based study process of the searching materials and their analysis. The paper concludes in Section 4.

2. Literature Review

Numerous approaches are in practice to deal with diverse perspectives of big data. The data can be viewed to extract meaningful insights for delivering quality services. Cha et al. [9] proposed the technical architecture of the data analytic platform for health offering a methodological solution for analysis of big health data which are produced from various sources. Data standardization was used as the main feature of the architecture, and SNOMED-CT was considered as a standard for standardizing data of health from various sources. With the help of the case study involved, huge volumes of data based on the laboratory are incorporated and analyzed through this platform.

Li et al. [10] introduced the big data analytics proof-of-concept framework for development of the risk-tuning patient expenditure model. The approach is based on the strategy of “divide and conquer” for exploiting the big-yet-rich data for enhancing the accuracy of the model. The MapReduce tool was used for implementation of the machine learning algorithms for the dataset. The framework determined the predictive analytics' effectiveness through the random forest algorithm and distributed-computing platform efficiency. Straton et al. [11] presented an approach for evaluation of predictions and post engagement on Facebook. The study focused on insights into associated indicators leading to the engagement with posts of healthcare on Facebook. The approaches of unsupervised and supervised are used. The study elaborated the strategy of healthcare for engaging users and building the prevention mechanism by contents posted on Facebook about informative healthcare. Sterling [12] reviewed the associated concepts from the literature of organizational communications and elaborated technology acquisition theories. The study described the possible benefits and the issues in the area. The study has further emphasized on the issues of resources like cost of human, patient privacy, and thoughtful of technology interventions.

Pandey and Subbiah [13] proposed architecture of big data storage containing the storage cluster and application cluster for facilitating read, write, or update data optimization as well as speedup. The application cluster facilitates effective storage and retrieval function from the user, while the services of storage are provided by the storage cluster. Ambigavathi and Sridharan [14] analyzed the properties of big data, analytics tools of big data, and various phases through healthcare economy from the collection of data for the stage of data delivery. The study reported the challenges of research with possible findings. Zhang et al. [15] examined the three knowledge modes achieving beneficial use of big data analytics with electronic health records. The authors surveyed 580 nurses of a hospital in China in 2019. For examining the association of electronic health records and knowledge, structured equation modelling was used. The study contributed to the available literature of big data and digital health through exploring the analytical tools to electronic health records from various modes of knowledge to extract useful big data analytics. Mande et al.’s [16] goal was establishing of a collaborating processing environment with patient data through technologies of open source software. The structures of distributed data were produced from hospital explicit metadata for extracting side effects in the patient through the query of whole clinical trial data to develop, repurpose, and reposition drugs. Harerimana et al. [17] identified the data sources, technologies, techniques, main issues, and future directions of healthcare big data analytics. The study elaborated a do-it-yourself review delivering a simplified, easily understandable, and holistic view of different technologies’ applications for developing incorporated analytic applications of health. Nazir et al. [18] presented a comprehensive and detailed report on the big data visualization in cardiology. The study is based on the last 10 years from 2009 to 2018. 53 studies were filtered based on the inclusion and exclusion criteria, quality assessment, and some derivations derived. They elaborated that the study is associated to cardiology and will help researchers and practitioners in the area for devising new solutions based on the existing evidences. Several other studies are of the same authors related to big data are available [19–22].

Singh et al. [23] explained the usefulness of real-time data for analysis and prediction of severe emergency cases. They explained that without approaches of Hadoop cluster and analytics of big data, the data become useless. Xu and Kumar [24] presented a framework based on machine learning for analytics of big data and tested for improvement of performance and quality of auxiliary power-unit services of health monitoring. They developed and employed useful
technologies and practical analytics of big data for application of industries in aviation and aerospace. Ojha and Mathur [25] gave an overview of how the patient record can be saved as digital data in the form of electronic health record and how useful information can be produced from the saved records through approaches of analytics and tool supporting in saving money and time of both the doctors and patients. The study was done in the hospital of Maharaja Yeshwantrao, Indore, Madhya Pradesh, India. From this hospital, huge amount of heterogeneous data is generated from various sources such as laboratory tests, patient health records, electronic medical requirements, social media, health insurance data, genome research, drug research, clinical outcome, and transaction from Mahatma Gandhi Memorial Medical College, working under Maharaja Yeshwantrao. Analytics of big data can be used for making useful information for retrieval. Sabharwal et al. [26] focused on the applications and analytics of big data and issues in the IoH care. Rao et al. [27] presented different possible solutions of security for harnessing the probable of big data affecting healthcare in vastly regulated surroundings. Ajayi et al. [28] proposed a particle swarm optimization approach for selection of feature and tuning parameters of the gradient boosting machine approach on 1,349,239 data points of an accident dataset. The ability of prediction of the proposed approach was compared to the traditional tree-based approach discovered about accurate prediction of the model on test data. Khenou et al. [29] explored efficiently the appropriate adaptation of analytical tools for electronic health records for upgrading the usage through practitioners of health. A case study of the implementation process of electronic health record relies on open electronic health record and identifies analytics of health adoption in every phase of the approach.

Gaitanou et al. [30] performed a systematic literature review of the applications of big data in healthcare systems for improving clinician behaviour, experiences of the patient, and quality of care facilitated to the patient. The PubMed database was searched for the search process and identifying associated research papers. 12 research papers were selected based on the inclusion and exclusion criteria from the 108 research papers searched in the library. The derivations of the study show that 9 papers were showing the positive effects of big data, while some papers resulted negative. The key advantages of the applications of big data include enhanced usability, positive behaviour change, and effective decision support, while issues were traced for acceptance of technology. The key issue happened in the systems’ process heterogeneous datasets, outcomes reported by patients, and in motion data, as different to electronic health records. The study highlighted round about areas of research where exploration is needed for understanding the applications of big data in healthcare and improving its success. Talukder [31] pointed out that a huge amount of patient data is available in hospitals. Apart from this, large medical knowledge of body exist in the digital form in various repositories. The focus of the study has addressed the issues arising from such data with the analytic solutions combining the knowledge bodies and data by technology of big data integrated with mathematical models, artificial intelligence, and translation medicine "Evidence Based Precision Medicine—the perfect decision outcome with perfect knowledge backing." Several benefits exist for various stakeholders such as cost of payer are reduced considerably. The medical decision accuracy in improved and enhanced productivity of hospital, reduced burden of disease, decreased medical errors, and reduced wastage and fraud.

3. Big Data Insights and Comprehensions in Industrial Healthcare Based on Famous Libraries

Insights of big data is the essential process of mining and analysing big data for producing operational and business information at exceptional scale and specificity. Han et al. [32] presented a nonsystematic review for identifying associated gray and scientific literature related to the current state and restrictions of assessment of health involvement and infrastructure of health data in the US. The study comprised the literature on nations with unified data systems. The aim of the study was producing in-depth analysis of the available endeavors for use and standardized data collection through encounters of patient with doctors for the purpose of public health. Ravikumar et al. [33] elaborated the fields of big data analytics developing in healthcare, giving details of the advantages, and presented architectural framework and approach. The sources of data were identified and the tools and platforms for technologies, approaches, and analytics were recommended, and the issues and derivations were discussed. The study elaborated that analytics of big data in healthcare is an encouraging area for facilitating in-depth insights from huge volume of datasets. Aceto et al. [34] systematically surveyed the adoption of technologies of Industry 4.0 applied for the domain of health. The study provided details of the key paradigms and technologies’ applications to Healthcare 4.0, discussed the key application scenarios, and examined benefits, new cross-disciplinary issues, and lessons learned. Istepanian and Al-Anzi [35] provided the associated issues of big data from the perceptions of mobile health. The study also presented the associated issues of rapprochement of analytics of big m-health data. Galetsi et al. [36] surveyed a systematic review associated to big data analytics in healthcare. The study focused on different resources utilization for creating capabilities/values for organization; based on the publications selected, the study discussed the big data types’ classifications associated to healthcare, the analysis approaches, stakeholders value, the tools and platforms to handle data of big health, and forthcoming aspects in the area. Pragmatic examples were given for showing the advancements of healthcare.

Wang et al. [37] developed the transformation model for analytics of big data accordingly to the practice-based view revealing the casual association among analytics of big data abilities, benefits’ dimensions, IT-enabled practices of transformation, and business values. The system was validated in the environment of healthcare. Wami et al. [38] evaluated the possible utility of Mosaic groups for the purpose of research in health through identifying...
the associations with indices of multiple, driven to the population of British. Various popular libraries were searched for identifying associated studies. Figure 1 depicts year-wise distribution of papers in the library of Sciencedirect. Figure 2 depicts article types with the ratio of papers based on its types. Figure 3 shows the publications’ titles with the total of papers in the library. Figure 4 presents the subject areas with the total of papers. The library of IEEE was searched for finding the related materials. Figure 5 shows the publications’ topics in the library. Figure 6 presents the total papers in the conferences held. Figure 7 depicts the title of publications with the total number of papers. Figure 8 shows the publication types in the given library.

The library of Springer and PubMed were searched for related research studies. Bates et al. [39] elaborated the
insights’ types such as emerging from clinical analytics, types of analytics for obtaining insights, algorithms, monitoring devices, evaluation scores, registries, and analytics infrastructure needed for organization for performing the essential analysis and implementing changes for enhancing care at low costs. Nanayakkara et al. [40] surveyed the biomedical big data in healthcare advancements, issues, and limitations with the focus on oral health and elaborated the possible forthcoming applications of big data for oral health for improving the efficiency and quality of personalized healthcare. Madanian et al. [41] identified and examined the issues associated to the existing status of India’s healthcare system with the emphasis on mHealth and technologies of big data analytics. For addressing these issues, the study proposed a framework for incorporating the produced mHealth big data and employed the results in healthcare of India. The study was conducted with the use of electronic sources between December 2018 and February 2019 with the publications’ language as English. Brennan and Bakken [42] presented a study consisting of investigation of emergent initiative of federal big data and explored exemplars from informatics of nursing research for benchmark where nursing are previously poised for
participating in the revolution of big data. Reflections and observations were provided on experiences in the initiatives of emerging big data. Luo et al. [43] reviewed and elaborated the applications of big data in four key sub-disciplines of biomedical including the clinical informatics, bioinformatics, imaging informatics, and public health informatics. The study reviewed the advances and breakthroughs of the applications of big data in the domains of healthcare and surveyed the issues, opportunities, and gaps for improving and advancements of healthcare big data applications. Sukumar et al. [44] presented a study with the aim of highlighting the quality of data issues in context of healthcare big data analytics. The insights described in the study were results of analytics research performed in various organizations on different datasets of health. The datasets included Medicaid claims and Medicare, provider datasets from electronic health records and private and public sources from health centers available through partnership with claims of healthcare under privacy-protected policies of health. Slime et al. [45] presented a study focusing on application enhancements of the technology in the field of health.
Figure 7: Title of publications with total papers.

Figure 8: Types of papers.
Figure 9: Publication topics and total of papers.

Figure 10: Publications types with ratio of papers.
aspects of big data are discussed, giving some implementation and elaborating various architecture of big data in healthcare. Figure 9 depicts the topics of publications with total of publications. Figure 10 depicts the publications’ types with the ratio of papers. Figure 11 represents the languages in which the papers are published. Figure 12 shows the publications in the given years.

4. Conclusion

With the growing technology and digital revolution, the technology of healthcare is getting advanced at a high rate. Healthcare data are always rising and drastically increasing the volume of healthcare data. Such data can make it tougher to recognize useful form of data generating meaningful information. The data are always playing a significant role in organization and industry for their daily activities to function smoothly. This rise in big data volume can be a challenging task concerning analysing the data for normal industry and IoT. Numerous tools, approaches, and techniques are existing in literature for supporting, to handle, and manage the process of big data. A review of the literature is needed in order to gather current evidences from the literature to show which method or tool works for which particular situation. The proposed study presents a review of the existing techniques of big data insights and scientific programming in the industry of healthcare. The report presents the summary of the literature. The study collects evidences from the existing literature and organizes it through the process of literature review with some derivations. This review will benefit practitioners to identify the right techniques for their specific purpose of research.

Data Availability

The data are not available.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References


[41] S. Madanian, D. T. Parry, D. Airehrour, and M. Cherrington, “mHealth and big-data integration: promises for healthcare...
system in India,” *BMJ Health & Care Informatics*, vol. 26, no. 1, 2019.


