6 AI-Assisted Big Data Analytics for Smart Healthcare Systems

Akruti Sinha, Devika Sapra, Gaurav Srivastava, Mahin Anup, and Deepak Sinwar
Manipal University Jaipur

CONTENTS

6.1 INTRODUCTION

The recent COVID-19 pandemic severely strained the entire world's healthcare system, exposing flaws in our traditional healthcare system that were previously overlooked. It exemplified many of the issues plaguing the healthcare system. The most serious issues were the high cost of healthcare and the lack of access to high-quality care. One of the major reasons for this is the complexity of the healthcare system. It is a highly interconnected system, which is something that can be exploited by hackers. The interconnectedness of the system also increases the data volume exponentially

DOI: 10.1201/9781003326182-6

with each patient, and the number of patients is growing at the moment. There are solutions to these issues, and the ongoing wave of innovation has demonstrated that artificial intelligence (AI) and big data analytics (BDA) are the two fundamental components with the potential to lead the smart healthcare transformation.

While AI can assist with predictive analysis, intelligent judgements, and creative solutions to age-old problems [1], BDA can assist with the first part of the problem, that is, mining information and extracting meaningful data from the massive amount of available data [1] that is growing exponentially by the minute.

Despite its recent popularity, there have not been many foolproof frameworks, and platforms that use the power of both BDA and AI to provide a better smart healthcare system (SHCS) in terms of reliability, security, access, and other important features. There are several challenges to properly deploying such a framework in today's ever-changing, interconnected healthcare systems. First and foremost is the problem of unstructured data. Healthcare is an industry that generates a lot of data, but a lot of that data is not easily consumable. This is primarily because of the variety of different formats in that data is stored. For example, the medical notes of patients, doctors' reports, images, and various other types of documents. As the authors of [2] have stated, the problems of unstructured data, which brings its own set of problems such as correlated and redundant data, result in the consumption of more energy and bandwidth by networks, which are already quite constrained in terms of the resources they can access. This eventually leads to network congestion.

The rest of the chapter is organised into different subsections. Section 6.2 deals with an overview of a few related works. The background of AI and big data is presented in Section 6.3. Section 6.4 deals with the importance of AI and BDA for SHCSs. Common methods and their implications are discussed in Section 6.5. Section 6.6 discusses the most used frameworks in the field of SHCSs that integrate both AI and BDA. Emerging trends in big data and AI with reference to the SHCS are presented in Section 6.7. Section 6.8 discusses various challenges and issues followed by a conclusion in Section 6.9.

6.2 RELATED WORKS

Scholars in the field of healthcare have published numerous papers that integrate both AI and BDA to propose architectures and frameworks that can be implemented at a higher level to create a smart and efficient healthcare system. The authors of [3] have meticulously compiled a comprehensive survey that not only reviews but also enumerates the strengths and weaknesses of all such proposed frameworks and architectures. They have also identified a few challenges and unresolved issues as we have done later in the chapter as well. Reference [3] has provided excellent insight for all scholars and aids in keeping up with current trends in the field.

The authors of [3] hold [4] in high regard, and it is entirely feasible that it deserves such praise. The authors of [4] proposed a one-of-a-kind system that derives its power and functionality from that of the Internet of things (IoT) to not only recognise an emergency but also aid in the quick communication of healthcare personnel amongst themselves. As the authors correctly point out, one of the benefits of such a system is that it significantly reduces the possibility of human errors caused by individuals' delayed reaction time. It also provides doctors and physicians with more time to make accurate decisions.

While the idea is admirable in every way, and the architecture is nearly flawless, it should be noted that the results generated by [4] may vary greatly depending on person to person and situation to situation. In this case, testing with a large dataset is recommended.

In [5], one finds a very lucid explanation of the crucial areas of BDA needed for successful implementation in the healthcare sector. Reference [5] also discusses the challenges that usually encompass the applications of BDA in the field of healthcare organisations. Due to issues like confidentiality, privacy, and unstructured data, the authors of [5] suggest three major components of BDA that have a significant impact on it:

a. **Data validation**: This ensures the validity of data and checks to see if it is corrupted or not to ensure its accuracy. The Hadoop Distributed File System (HDFS) is used for the validation process.
b. **Process validation**: This component ensures that the business logic is correct and that the key-value pair that is created is accurate.
c. **Output validation**: This final one determines whether the data has been distorted after it has been processed in the repository. The data is validated by comparing it to an HDFS file.

As [6] correctly points out, BDA can be applied in varied ways. According to the authors of [6], the various sources of BDA include patient demographics, treatment history, and diagnostic reports. This information can give healthcare providers a more complete picture of their patients and allow them to customise treatment. It can also be used to enhance the quality of care patients receive. For example, a patient with a history of cancer may require a different type of treatment than a patient without such a history.

According to the authors of [7], advanced analytics must be used to gather important information from electronic health records (EHRs) because BDA can reveal important patterns that can improve the overall value of the healthcare system. Reference [7] also suggests that in the future, artificial neural networks may be used in conjunction with BDA. These algorithms can be trained to recognise patterns that indicate fraud or other problematic behaviours. This can help to reduce fraud and improve healthcare security and reduce the time spent on detecting fraud and other problematic behaviours.

GEMINI [8] is one of the most extraordinary proposals for a real-time service that provides health services. GEMINI begins by extracting data from different types of sources, which includes both structured and unstructured data, for each patient and storing it for future use. The stored information forms a patient's profile graph, allowing it to provide a detailed view of the health information of the patient. In addition, the profile graph enables GEMINI to infer and extract information that may not be easily understood to perform predictive analysis. Another notable part of the proposed model is that it gives a "self-learning knowledge base" [8], which also keeps all medical professionals in the loop to determine findings and conclusions.

While innovations in the field of BDA and AI are constantly being developed, it is equally important to ensure their dependability and trustworthy use. The authors of [9] propose a few possible solutions. As a first step, it has been suggested that a clear and effective data governance framework be developed. To ensure that the design and frameworks are human-centred, all steps must be taken. The following step is to define a clear process that explains how, why, and which information can be used. The third step should focus on empowering those whose data is being used by understanding and enforcing strict rules for how this data can be managed and governed more effectively. As a fourth and final step, all government agencies must collaborate to ensure that the process is not only transparent but also accurate.

The steps outlined above can be used effectively to ensure that trustworthiness and equality continue to exist in the field of healthcare [9]. Reliability is especially important in the field of healthcare due to the very sensitive nature of the data. The misuse of medical data can wreak havoc on the healthcare system because it can lead to discrimination and has the potential to cause serious harm to individuals.

6.3 AI AND BIG DATA BACKGROUND

6.3.1 Big Data in Healthcare

According to [10], big data encompasses several statistical techniques for data evaluation and assessment [11,12]. "Big Data" has been a phrase often coined to define huge data obtained through data interchange across diverse systems [13,14]. Even though it is indicated as a "huge" database, its relevance has been driven by the ability to gather little information regarding the given issue under investigation [15,16]. Reference [17] outlined the features of big data with "seven Vs", that is, "Volume, Variety, Velocity, Veracity, Value, Variability, and Valence". With the progressively increasing quantity of big information inside the given environment, data analysing applications within medical services are steadily developing [18,19]. Diverse inferences, such as patient characteristics [20], clinical history [21], and examination results, are among the potential big data origins in the health sector [22]. Such information, according to [23], might be organised (for example, genome, phenotype, or genomics data) or unstructured. Patient information frequently necessitates the development and collection of high-quality actual statistics [24,25]. Experienced professionals in medical institutions may use big data to get useful insights and take measures [25,26]. Technology is being used by medical institutions to cope with the different intricacies of data processing [27,28]. Furthermore, big data in medical services may be used to integrate diverse sectors to further investigate a condition holistically [29]. In conclusion, all the aforementioned features of big data may be observed in the case of health care.

When leveraging big data, observational, prospective, and adaptive statistical tools may be used to upgrade the effectiveness of many areas of medical care [30]. Clinical diagnostics, universal healthcare, clinical observation, and clinical management are some of the opportunities presented by big data in the health industry.

6.3.2 AI in Healthcare

According to [31], one of the most attractive prospect sectors for the implementation of automated systems is healthcare [32]. Doctors can now diagnose cancers more quickly and sooner than ever before, detect illnesses before they emerge, and discover genetic problems that may harm us in the future. Similarly, analytics can improve backroom procedures, enhancing the patient hospital experience while also reducing expenses by decreasing inefficiencies and money that can then be spent on improving patient outcomes. Nevertheless, in the medical industry, technologies must adhere to the legislation, norms, and security standards in ensuring that technology serves the public good. Technology can help not only in the discovery of valuable medications but also in enhancing the efficiency of current products. Even though AI systems in medical services are now limited, the health and financial benefits are too tremendous to ignore. High-cost industries will become opportunities for forward-thinking healthcare organisations to adopt cutting-edge technology and maintain their competitive advantage over their competitors. These are only a few instances of what is possible when AI is implemented in health systems to its utmost advantage. The opportunities are limitless, and coordination between government and private-sector entities is vital to accomplishing goals. As worldwide societies remain healthy and the prevalence of acute illness grows, the costs of medical care will continue to be a focal area amongst healthcare entities.

The use of AI for medical services is both exciting and desirable. Collaboration between physicians and robots might mark a watershed moment in our abilities to combat ailments and enhance human health. The advantages will undoubtedly be essential, ranging from accuracy and focused treatment to backend administration and simpler procedures, from assistance for transitional housing for the aged to advanced testing abilities.

It might be time to seek the assistance of computers. In spite of the fantastic innovations, they should be examined in the context of our current medical transition. The arrangement of modern healthcare has a huge influence on the medical industry, and the actions we take now will impact fairly long ramifications according to how we treat patients.

6.4 IMPORTANCE OF AI AND BDA IN SHCSs

As previously stated and by researchers in [33], the current wave of innovation will be one of the most crucial for the global healthcare system, owing to the increased potential of tools that will enable us to address the most pressing concerns in the healthcare sector. These tools will include novel, affordable, and easy-to-use devices that will deliver diagnostics and therapeutics to the patient faster than ever before.

AI and BDA have long been considered potential solutions to the problems that currently plague the healthcare system. However, as the authors of [34] correctly point out, at this point, we have the required technology including a large AI talent pool, making this future not only plausible but also likely. The impact on healthcare sustainability is difficult to predict, but one thing is certain: change is coming, and it is coming soon.

The improved tools at our disposal only increase BDA's ability to provide better results and not only assist healthcare personnel in their daily tasks of documentation and others but also provide better, clear, accurate image analysis, patient monitoring, and perhaps even fully automated life-supporting devices used extensively in SHCSs. However, it is not all that easy. The most difficult task, and one that has long been a source of concern, is an accurate interpretation of data in the form of an image or video [33]. The challenge, of course, is much greater for images taken in poor lighting conditions or a noisy environment. However, there have been many innovations in the current era, and with an increased number of professionals in the industry, BDA used cautiously and diligently can solve this issue in the coming days as well.

When we shift our focus to BDA, we find that it is just as critical as, if not more important than, AI. First and foremost, BDA can be defined as "the process of analysing a huge amount of data from various sources of data and different formats to convey the perception of enabling a decision-making process in real-time" [35].

Many people were unsure what role BDA could play in the healthcare industry other than "analysing" patient data. BDA can do much more than that, thanks to the massive amount of data being appended every passing second. BDA applications in the smart healthcare industry today include managing, storing, and examining clinical data from doctors and prescriptions. BDA can also help organise and interpret computed tomography (CT) scans, magnetic resonance imaging (MRI) images, X-rays, laboratory data, Electronic Patient Records (EPR) data, and other types of data. When implemented intelligently and utilised, BDA can also aid in the analysis of accurate and meaningful data from drugstore documents, insurance-related files, and other sources. This data can then be used to enhance healthcare services, ensure lower expenses, and increase efficiency. Even though scientists have access to a large amount of data, they have not been able to develop a suitable framework to aid in the achievement of all of the aforementioned applications [35].

As previously stated, this decade will be very eventful for the healthcare field, especially because we now have improved tools to help us solve problems that were once thought unsolvable. With increased efficiency comes new challenges such as security, interoperability, and other issues, but with BDA on the scene to assist, the SHCS is poised for a bright future.

6.5 COMMON METHODOLOGIES AND THEIR IMPLICATIONS

There have been instances where certain common techniques in the domains of IoT and information analysis were recommended for the creation of advanced medical systems. According to [36], a predictive model known as ViSiBiD may correctly detect harmful clinical events using patient monitoring based on data gathered from various patients with similar symptoms. (i) efficient and scalable data acquisition in all parameters combining "Pearson's correlation coefficients" and "wavelet transform", as well as quick metrics; (ii) using numerous data mining tools, such as "random forest (RF)", "J48 Decision Tree (J48)", as well as "sequential minimal optimisation (SMO)" [37], can improve initial clinical irregularity prediction.

Intel has launched the "A-wear" initiative, which is centred on a portable gadget that connects predictive analysis and portable tech. "Apple" has been striving for analysing blood sugar levels through tears, as well as creating motion medical equipment for blood tracking through the skin. Google is focusing on cancer detection, imminent strokes, and cardiovascular problems, among other things. Samsung is collaborating with medical doctors from the "University of California, San Francisco" to accelerate the development and commercialisation of sensor technologies, computations, as well as medical technology. The given measures sparked rivalry in the business for wearable technology, as well as technological advancements that have energised the sensors and wearable business [38].

A medical scheme based on power harvesting calculation for monitoring systems and the deployment for information analysis throughout medicare was proposed. Resource harvesting and generation of data, pre-processing, data analysis, as well as data application, are the three levels of the strategy. It highlighted the efficacy of the medical IoT with power generation, as well as reliable samples that have been utilised through the "Hadoop server" for method standardisation using threshold limit values (TLVs) [39]. Wearable technology systems to constantly monitor individuals' conditions have been built using a semantic centralised database to incorporate diverse wearable medical datasets [40]. Wearable biosensors enable personalised mHealth and eHealth, making it easier for healthcare professionals to provide personalised and premium services [41].

The following are some of the issues of massive data and data analytics in healthcare [42–46]: (i) data with a massive quantity, several dimensions, or a wide range of formats; (ii) unbalanced data, inconsistent data, and data of low performance and durability; (iii) Healthcare professionals (particularly hospitals) who are responsible for aggregating and analysing fragmented or siloed data; (iv) Unstructured or semi-structured data aggregation and analysis; (v) filtering and analysing continuous data streams, particularly rapidly moving data streams; (vi) algorithmically difficult jobs in genetic data analysis; (vii) massive data visualisation; and (viii) data ownership, information management, confidentiality, and cybersecurity are all important considerations. Even if privacy protection is guaranteed, many healthcare professionals are unwilling to disclose their information owing to market rivalry. Preserving an optimal balance between safeguarding patient data and maintaining information availability or integrity of data is often challenging. There are problems with public access, formalisation, and the incorporation of usable and legible data [44].

6.6 TYPES OF AI-BDA FOR SHCSs

6.6.1 Recommendation Systems

Numerous recommendation systems use the power of AI, BDA, and IoT to recommend either an appropriate diagnosis or steps for disease monitoring. Other types of recommendation systems exist as well.

The authors of [47] proposed a rather viable recommendation engine for personalised IoT solutions provided to any individual. The engine's process, as proposed by the authors of [47], is as follows: first, the diseases to which a person is predisposed