WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Healthcare Applications and Challenges in Big Data Analytics

Dr. T. Nikil Prakash

Assistant Professor, Department of Information Technology, St. Joseph's College, Thiruchirappalli, affiliated to Bharathidasan University, Thiruchirappalli.
Email: nikilprakash1992@gmail.com

ABSTRACT

Big data is a massive amount of data that can produce user activities on the internet via websites, social media, blogs, and applications, etc. Big data can manage various public and private sectors to store and provide information. In the healthcare system, big data is used in many ways such as maintaining patient's information, medicine data, treatments, doctor's information, and disease, etc. it also maintains billing information, medical records, and hospital records, etc. Big data can change the adventure of modern healthcare systems and handle easy customer supports. Bio-medical research also generates a significant portion relevant to public healthcare. This proposed work analyses challenges and applications in the big data healthcare system. Additionally, this paper focuses on the tool of the big data healthcare system and the performance of big data analytics in the healthcare system.

Keywords: Healthcare, Data Analytics, Big Data, Sentiment Analysis, Machine Learning, and Lexicon Based Approaches.

1. Introduction

Big Data is essential to every major healthcare organization. Learn about the problems, applications, and the brilliant potential for big data in healthcare. Big data tools refer to a lot of quality data created by the digitalization system. This system provides specific technologies for developing and analyzing Healthcare systems [1]. Big data tools are applied to Healthcare systems to use specific health data that particularly or individually and potentially help to prevent epidemics, cure disease, and reduce cost, etc. [2].

Nowadays treatment techniques are changed and most of the treatment changes are driven by data. Doctors want to analyze patient information and identify any treatment or disease at an earlier stage, this is very simple and less expensive [3]. Healthcare data analytics system is used for managing every patient's treatment information, disease, stage of disease cured. Additionally, it is used to details of doctors, hospital (facilities, ward, etc.), clients, and surgeries, etc. with the unfeasibility to connect properly. The healthcare data analytic tools easily gather and manage a large amount of data for medical usage that help to reduce time-consuming and cost [4].

The new Healthcare systems collect and manage such relevant data's but besides critical insights and used to provide better health care [5]. Big data tools involve better treatment for patients to access quicker and involve the patient's health and authorization [8].

2. Literature Review

Aswin belle et al [8] proposed big data analytics in healthcare challenges in the medical research area. They identify the challenges in medical research images, signals, and genomics. They combined multi-modal medical data set from various sources.

Wullianallur Raghupathi et al [9] proposed the promise and potential of big data analytics in healthcare. They described the benefits of a big data healthcare group using data analytics and capabilities. These capabilities are predictive, traceability, decision support, and unstructured analytical. Additionally, they focused on the benefits of big data analytics in IT infrastructure, operational, organizational, managerial, and strategic fields.

Yichuan wang et al [12] proposed capabilities and resource-based theory model for healthcare systems using big data analytics. They identified the relationship between big data capabilities and business value. They found business transformation in big data analytics and described much information for big data analytics in healthcare systems.

Jianbi Yuan et al [13] described the prediction of textual contents and images in sentiment analysis. They predicted mid-level attributes in sentiments and classified low-level attributes/ additionally, they described edge-face-basic facial expressions in sentiment analysis. They compared mid-level attributes and image sentiments.

Marie Katsurai et al [14] proposed the opportunities of latent correlations among textual views and sentiment views. The views are contributed in SentiWordNet. The results found latent embedding space and considered complementary information from different views. The use of these views described multiple modalities and achieves sentiment classification methods. They found a single modality and state of art method.

Donglin Cao et al. [15] proposed microblog topic image model analysis in sentiment analysis. They obtained visual sentiment features and build visual sentiment topic models using the microblog topic model. They found the performance for better results and given the approach of the visual-based model.

3. Applications of Healthcare Systems in Big Data Analytics

3.1 Patient Predictions

It is a common source to predict the patients' health conditions daily and hourly. Many hospitals are expected to predict the patient's information. It can reduce the waiting time and predict all the works for hospital systems. It can provide a high number of visitors and provide better hospitality. It's solved only a few hospital systems. It helps doctors make a decision within seconds and improve the treatments for patients. It is particularly useful for complex and critical patient's medical histories suffering multiple health conditions [16].

3.2 Electronic Health Records (EHR)

The medical records of every patient are digitally recorded which are medical history, laboratory test, allergies, and surgeries, etc. The records are showed both public and private sectors with secured information. EHRs are real-time, patient-centered records that make information available instantly and securely to authorized users [18].

3.3 Real-Time Monitoring

It is one of the critical functions of big data analytics. The patient's data analyzed in real-time or on the spot and provides health descriptions with prescriptive decisions. For example blood pressure.

It gives patients access to personal care. The mobile device identifies the medical condition using aggregated data and can recommend patients seek medical care based on input to the application. The device can identify the patient's cause and notify the medical attention. The hospital tracks the patient's location and sends the report to the patient's mobile phone. It can easily understand their symptoms and take measures to prevent attacks.

3.4 Medical Research

Medical research can use a large amount of data treatment plans and recovery tasks of cancer patients, to find the treatment and provide better success rates in the real world.

3.5 Fraud and Security

Patient data is extremely valuable and portable on the block markets. Many organizations started to use analytics to support prevent security threats by identifying network traffic or any other actions that reflect cyber-attacks.

3.6 Keeping patients healthy

It helps avoid the illness and disease start at the front of any priority list. Patients keep track of their fitness activities and physical level activities and report specific health-related activities to doctors.

3.7 Cost Reduction

Medical clinics, facilities, and treatment focus squander a lot of cash in money-related administration, for the most part from understaffing or overstaffing. With prescient examination, this issue can be unraveled by foreseeing the confirmation rate and guaranteeing that the fitting staff is accessible to address persistent issues.

4. Challenges of the healthcare system in Big data analytics

4.1 Data Aggregation

First, patient and financial data are often spread across many payers, hospitals, administrative offices, government agencies, servers, and file cabinets. Pulling it together and arranging for all data producers to collaborate in the future as new data is produced requires a lot of planning.

4.2 Policy and Process

Once data is validated and aggregated, various process- and policy-related issues need to be addressed. The HIPAA regulations demand that policies and procedures protect health information. Access control, authentication, security during transmission, and other rules complicate the task. This multifaceted issue has been solved to some extent by cloud service providers, perhaps most notably Amazon AWS, which offers cloud services that comply with HIPAA and Protected Health Information (PHI).

4.3 Management

It is realizing the promises of big data analytics in healthcare that requires organizations to adjust their ways of doing business. Data scientists will likely be needed along with IT staff that have the required skills to run the analytics.

4.4 Capture

All data come from somewhere, but unfortunately for many healthcare providers, it doesn't always come from somewhere with impeccable data governance habits. Capturing data that is clean, complete, accurate, and formatted correctly for use in multiple systems is an ongoing battle for organizations, many of which aren't on the winning side of the conflict.

4.5 Cleaning

Healthcare suppliers are personally acquainted with the significance of tidiness in the facility and the working room, however may not be very as mindful of the fact that it is so imperative to purge their information, as well.

Dirty data can quickly disrupt a big data analytics project, especially when bringing together disparate data sources that may record clinical or operational elements in slightly different formats. Data cleaning – also known as cleansing or scrubbing – ensures that datasets are accurate, correct, consistent, relevant, and not corrupted in any way.

4.6 Storage

Front-line clinicians rarely think about where their data is being stored, but it's a critical cost, security, and performance issue for the IT department. As the volume of healthcare data grows exponentially, some providers are no longer able to manage the costs and impacts of on-premise data centers.

4.7 Security

Data security is the number one priority for healthcare organizations, especially in the wake of a rapid-fire series of high-profile breaches, hackings, and ransomware episodes. From phishing attacks to malware to laptops accidentally left in a cab, healthcare data is subject to a nearly infinite array of vulnerabilities.

The HIPAA Security Rule includes a long list of technical safeguards for organizations storing protected health information (PHI), including transmission security, authentication protocols, and controls over access, integrity, and auditing.

4.8 Querying

Robust metadata and strong stewardship protocols also make it easier for organizations to query their data and get the answers that they are expecting. The ability to query data is foundational for reporting and analytics, but healthcare organizations must typically overcome several challenges before they can engage in meaningful analysis of their big data assets.

4.9 Reporting

After providers have nailed down the query process, they must generate a report that is clear, concise, and accessible to the target audience.

Once again, the accuracy and integrity of the data have a critical downstream impact on the accuracy and reliability of the report. Poor data at the outset will produce suspect reports at the end of the process, which can be detrimental for clinicians who are trying to use the information to treat patients.

4.10 Visualization

At the point of care, a clean and engaging data visualization can make it much easier for a clinician to absorb information and use it appropriately.

Color-coding is a popular data visualization technique that typically produces an immediate response – for example, red, yellow, and green are universally understood to mean stop, caution, and go.

4.11 Updating

Healthcare data is not static, and most elements will require relatively frequent updates to remain current and relevant. For some datasets, like patient vital signs, these updates may occur every few seconds. Other information, such as a home address or marital status, might only change a few times during an individual's entire lifetime.

Understanding the volatility of big data, or how often and to what degree it changes, can be a challenge for organizations that do not consistently monitor their data assets.

4.12 Sharing

Few providers operate in a vacuum, and fewer patients receive all of their care at a single location. This means that sharing data with external partners is essential, especially as the industry moves towards population health management and value-based care.

Data interoperability is a perennial concern for organizations of all types, sizes, and positions along the data maturity spectrum. The industry is currently working hard to improve the sharing of data across technical and organizational barriers.

5. Future for Big Data in Healthcare

Commerce and industrial sectors declare their big data initiatives have been successful and transformational, the outlook for healthcare is even more exciting. Below are a few areas where big data is destined to transform healthcare.

5.1 Precision Medicine

It is envisioned by the National Institutes of Health, which seeks to enroll one million people to volunteer their health information in the All of Us research program. That program is part of the NIH Precision Medicine Initiative. According to the NIH, the initiative intends to "understand how a person's genetics, environment, and lifestyle can help determine the best approach to prevent or treat disease. The long-term goals of the Precision Medicine Initiative focus on bringing precision medicine to all areas of health and healthcare on a large scale."

5.2 Wearable's and Internet of Things (IoT) Sensors

The potential to revolutionize healthcare for many patient populations and to help people remain healthy. A wearable device or sensor may one day provide a direct, real-time feed to a patient's electronic health records, which allows medical staff to monitor and then consult with the patient, either face-to-face or remotely.

5.3 Machine learning

It is a component of artificial intelligence, and one that depends on big data is already helping physicians improve patient care. IBM with its Watson Health computer system has already partnered with Mayo Clinic, CVS Health, Memorial Sloan Kettering Cancer Center, and others. Machine learning, together with healthcare big data analytics, multiply caregivers' ability to enhance patient care.

6. Tools for Big data Analytics

Tool	Description
The Hadoop Distributed File System (HDFS)	HDFS enables the underlying storage for the Hadoop cluster. It splits the data into smaller parts and allocates it across the Different servers/nodes.
MapReduce	MapReduce conveys the edge for the transmission of sub-undertakings and the social event of yields. At the point when tasks are performed, MapReduce tracks the preparation of each server/hub
PIG and PIG Latin	Pig programming language is designed to captivate a wide range of information (organized/ unstructured, and so on.). It is included two key modules: the language itself, called PigLatin, and the runtime adaptation in which the PigLatin code is executed.
Hive	Hive is a runtime Hadoop support design that uses Structure Query Language (SQL) with the Hadoop stage. It grants SQL software engineers to create Hive Query Language (HQL) articulations similar to ordinary SQL proclamations.
HBase	HBase is a section arranged database the board framework that sits on the head of HDFS. It utilizes a non-SQL approach.

Cassandra	Cassandra is likewise a conveyed database framework. It is assigned as a high-level task displayed to deal with large information disseminated across numerous utility servers. It similarly gives solid assistance with no specific purpose of dissatisfaction (http://en.wikipedia.org/wiki/Apache_Cassandra) and it is a NoSQL framework.

Conclusion

Big data has become more significant in healthcare due to three major shifts in the healthcare industry that is, the vast amount of data available, growing healthcare costs, and a focus on consumerism. This paper focused on big data health systems to turn the challenges into opportunities to provide personalized patient journeys and quality care and also described applications and tools that are used in big data health care systems. In the future, healthcare organizations will adopt big data in greater numbers as it becomes more crucial for success. Healthcare big data will also continue to help make marketing touchpoints smarter and more integrated. Additionally, the amount of data available will grow as wearable technology and the Internet of Things (IoT) gains popularity. Constant patient monitoring via wearable technology and the IoT will become standard and will add enormous amounts of information to big data stores.

References

[1]. B Ristevski, M Chen, (2018) "Big data analytics in medicine and healthcare", Journal of integrative bioinformatics, DOI: https://doi.org/10.1515/jib-2017-0030, Vol.15, issue. 3.

[2]. Kankanhalli A, Hahn J, Tan S, Gao G. (2016) "Big data and analytics in healthcare: introduction to the special section" Inform Syst Front, Vol.18, pp. 233–235.

[3]. Wu PY, Cheng CW, Kaddi CD, Venugopalan J, Hoffman R, Wang MD. (2017) "Omic and Electronic Health Record Big Data Analytics for Precision Medicine". IEEE Trans Biomed Eng, Vol. 64, pp. 263–73.

[4]. Viceconti M, Hunter P, Hose R. (2015) "Big data, big knowledge: big data for personalized healthcare" IEEE J Biomed Health Inform, Vol. 19, pp. 1209–15.

[5]. Wang Y, Kung LA, Wang WY, Cegielski CG. (2017) "An integrated big data analytics-enabled transformation model" application to health care. Inf Manag, Vol. 55, pp. 64–79.

[6]. Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma & Sandeep Kaushik. (2019) "Big data in healthcare: management, analysis and future prospects" Journal of Big Data, Vol. 6, pp.1-25.

[7]. Suraj Patil, (2019) "An Overview of Big Data Analytics in Healthcare" Towards Data Science, https://towardsdatascience.com/an-overview-of-big-data-analytics-in-healthcare-5a599f1bb48c.

[8]. Ashwin Belle, Raghuram Thiagarajan, S. M. Reza Soroushmehr, Fatemeh Navidi, Daniel A. Beard, and Kayvan Najarian. (2015) "Big Data Analytics in Healthcare" BioMed Research International, Vol. 2015, pp.1-15.

[9]. Wullianallur Raghupathi and Viju Raghupathi. (2014) "Big data analytics in healthcare: promise and potential" I Health Information Science and Systems, DOI: 10.1186/2047-2501-2-3, Vol. 2, pp. 1-10.

[10]. Yichuan Wang, LeeAnn Kung, Terry Anthony Byrd. (2018) "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations" Technological Forecasting & Social Change, Doi: https://doi.org/10.1016/j.techfore.2015.12.019, Vol. 126, pp. 3-13.

[11]. Yichuan Wang, LeeAnn Kung, William Yu Chuang Wang, Casey G. Cegielski. (2014) "Developing a Big Data-Enabled Transformation Model in Healthcare: A Practice-Based View" Thirty-Fifth International Conference on Information Systems, DOI: 10.13140/2.1.2843.3601.

[12]. Yichuan Wang, LeeAnn Kung, Terry Anthony Byrd. (2018) "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations", Vol. 126, pp. 3-13.

[13]. Jianbo Yuan, Sean Mcdonough, Quanzeng You, Jiebo Luo. (2013) "Sentribute: image sentiment analysis from a mid-level perspective, ACM, DOI: 10.1145/2502069.2502079, pp. 1-8.

[14]. Marie Katsurai and Shin'ichi Satoh. (2016) "Image sentiment analysis using latent correlations among visual, textual, and sentiment views" DOI: 10.1109/ICASSP.2016.7472195.

[15]. Donglin Cao, Rongrong Ji, Dazhen Lin and Shaozi Li. (2016) "Visual sentiment topic model based microblog image sentiment analysis" Multimedia Tools and Applications, Vol. 75, pp. 8955–8968.

[16]. Bernard Marr. (2016) "Big Data In Healthcare: Paris Hospitals Predict Admission Rates Using Machine Learning" Fobes article.

[17]. Health IT Feedback Portal, Official Website of the Office of the National Coordinator for Health Information Technology.

[18]. T. Nikil Prakash and A. Aloysius, "Textual Sentiment Analysis using Lexicon Based Approaches" Annals of the Romanian Society for Cell Biology, ISSN:1583-6258, Volume. 25, Issue. 4, pp. 9878 - 9885, 2021.

[19]. T. Nikil Prakash and A. Aloysius, "Hybrid Approaches Based Emotion Detection in Memes Sentiment Analysis" International Journal of Engineering Research and Technology, ISSN 0974-3154 Volume 14, Number 2, pp. 151-155, 2021.

[20]. T. Nikil Prakash and A. Aloysius, "Lexicon Based Sentiment Analysis (LBSA) to Improve the Accuracy of Acronyms, Emoticons, and Contextual Words", Statistics and Applications, Society of Statistics Computer and Applications, E-ISSN 0974-3154, 2021