



**International Journal of Biology, Pharmacy
and Allied Sciences (IJBPAS)**

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

A HYBRID BIG DATA ANALYTICS AND IOT FRAMEWORK FOR HANDLING PATIENT HEALTHCARE DATA IN SECURE MANNER

SARANYA R^{1*}, POTHIRAJ S², JAIKUMAR K³ AND VIVEK R⁴

-
- 1: Assistant Professor in Computer Science at Dr SNS Rajalakshmi College of Arts and Science(Autonomous),Coimbatore, India
- 2: II M.Sc in Computer Science at Dr SNS Rajalakshmi College of Arts and Science, Coimbatore, India
- 3: II M.Sc in Computer Science at Dr SNS Rajalakshmi College of Arts and Science, Coimbatore, India
- 4: II M.Sc in Computer Science at Dr SNS Rajalakshmi College of Arts and Science , Coimbatore, India

*Corresponding Author: Saranya R; E Mail: saranvars26@gmail.com

Received 24th July 2021; Revised 26th Aug. 2021; Accepted 30th Sept. 2021; Available online 1st Nov. 2021

<https://doi.org/10.31032/IJBPAS/2021/10.11.1103>

ABSTRACT

Demographic changes, as well as physical difficulties all around the globe, have contributed to the growth in the popularity of digital care delivery. The growth of the Internet of Things (IoT), data, science, plus deep learning has personally helped those technologies. Furthermore, as a consequence of such advancements, huge volumes of data were generated, rendering information systems processing a big concern. The enormous complexity, roughness, plus sparseness among those information make more efficient analysis challenging to execute. Analysis can help with some of these issues. Researchers present a unique statistical methodology with massive healthcare data gathered from IoT portable tech or stored medical record photos throughout this research. Utilizing intermediary among multiple databases as well as MapReduce Hadoop clusters, the suggested solution could adequately solve the information heterogeneity challenges. Moreover, the suggested architecture allows therefore for the usage of

mist computers using internet systems to address issues with physical and digital information processing, storing, including categorization. This also ensures that patient medical data was known safely and reliably.

Keywords: Internet of Things; health care data; analytic framework; patient medical data; Big data analytics

INTRODUCTION

Whereas the definition of e-health was in its adolescence around the globe, these have grown in popularity in several countries. For instance. Several academics have subsequently become encouraged to tackle the digitalization of medical information, especially to the development of information technology. Electronic health records (EHRs) include substantial quantities of diverse medical information throughout many medical institutions [1-3]. Information from electronic health records should be gathered, merged, cleansed, maintained, analyzed, as well as presented promptly to assure quality and reduce network delay. Making those decisions in healthcare systems necessitates a thorough examination of vast volumes of true information tracking including fitness bands [4]. Those choices have such a big effect on the well-being of the patients. An intelligent town's healthcare organization, for instance, would automatically access individuals as well as assist if they participate in behaviors that could harm your wellbeing. Several scientists should use the Internet of Things

(IoT) wearable technology to strategically aim which could be used to diagnose various ailments [5]. Deep learning methods, and from the other side, were gaining popularity as a result of much higher quality. The medical study utilizes grouping algorithms, including judgment forests to figure out why certain illnesses resurface in individual members, and yet it investigates the events that take place as well as the elements that cause occurring [6]. Demand forecasting uses a variety of machine learning techniques to create a forecasting model that can forecast unexpected occurrences. The scenario plan aims to achieve the best selections possible by recommending successful changes that result in patient-specific care [7].

The growing volume of health information necessitates the development of advanced analytics platforms including appropriate analysis software. Networking, data processing, data, finding, data gathering, plus data protection both have been the subject of a considerable study looking toward the goal [8-10]. The study proposes a

novel method to e-health that analytical hierarchical data extraction including deep learning algorithms to augment IoT, data storage, plus cloud-enabling huge healthcare data gathering. Rather than centralized healthcare data transmission, which would be a barrier to big data processing, cloud technology allows everyone to access information through a dispersed network of cloudlets. In addition, that essay goes into the framework of technology's architectural principal components in painstaking depth.

Related works

The main purpose of the research would be to enhance the productivity and efficacy of hospital services [11-14]. Researchers want to achieve sustainability in IoT portable tech as well as reduce network cost here between networked cloud and edge planes throughout this article. The suggested e-health framework would profit from either the effective communication revolutions to integrate EHRs that allow specialized clinical information sharing by utilizing the greatest controlled processing infrastructure. Furthermore, the current proposal intends to improve communication skills by merging EHR plus visual information gathered during 24/7 surveillance, as well as to improve medical services by forecasting and maybe attempting to prevent disorders. Furthermore,

the suggested system will give clients rapid assistance suggestions. That would also relieve physician stress, assist population ages in most need of lengthy treatment, increase internet connectivity, plus lower medication mistakes. As a result, socialized health care resources would have been less stressed. Several scientists have also been compelled to create large datasets including topics data management in recent times. Simultaneously, several mathematical techniques capable of extracting as well as analyzing big quantities of information were also used to assist judgment. Systems have just been created to aid in the analysis of massive volumes of data sources. Furthermore, owing to the continuous properties and high complexity of health information, algorithmic difficulties had emerged in medicinal business intelligence. The approach was largely meta data-centric, with no autonomous search or splitting. That approach works best with unorganized datasets. A platform for genuine big information processing. That paradigm solely addressed storage capacity as well as speed, but it did all of this without taking into account a diversity of information sources. The Oriented Feature Selection Support Vector Machine (OFSSVM) seems to be a binary classifier for prognosis utilizing

biological information [15]. Researchers proved the efficacy of different classifiers' categorizations depending on genetic ordering. Furthermore, the recommended elements throughout this research failed general compatibility. C-means grouping method towards health science purposes. The model can benefit both professionals as well as consumers. Researchers analyzed illnesses, including associated characteristics using the University of California– Irvine (UCI) deep learning library using MapReduce. An uncontrolled fully convolutional training technique utilizes a cloud IoT network. The goal of these programs was to create a generic human depiction using EHRs. Researchers employed randomized woodland models that had been developed for a population of 200,000 patients. A large data, digital health architecture approach that highlights the influence of security aspects. In biomedical research methodology, pictures, sounds, especially genetic information synthesis all contribute. Intelligent healthcare would be an informatics platform aimed at minimizing IoT communication delays by cloud loading as well as addressing the difficulties of massive healthcare data using Information And Communications Technology (ICT) [16]. In the latter days, online systems have

evolved into a data-sharing system, allowing individuals with similar wants to access their knowledge.

MATERIALS AND METHOD

The methodological approach completed the entire advanced analytics flow approach, which begins collecting information as well as continues through cleaning the data, tagging, as well as separation. Afterward, when, the solution incorporates the physiological data to create the mathematical framework. Lastly, medical information must be interpreted as well as visualized. The technology offers to provide an innovative solution to the challenge of information sharing. Utilizing bridges among multiple sources as well as MapReduce Hadoop clusters, the suggested solution successfully tackles the information diversity challenge. It also makes use of cutting-edge classification algorithms to boost quality and reduce data service latency. Moreover, it ensures that client health records was known stably and safely. Cloud-based units offer the advantage of properly increasing efficiencies. Receiving IoT feeds through wearable technology was continuously processed by cloud servers. Those nodes rely on a collection of virtualization to provide well before information to the cloud technology tier for some further analysis. Cloud technology,

from the other side, covers the benefits like cloud infrastructure as a service (IaaS), platform as a service (PaaS), software as a service (SaaS), plus public utilities. IaaS, for instance, provides limitless internet data storage, but PaaS is in charge of running asset apps. SaaS controls access to a wide range of applications, including essential companies, as well as storing big quantities of information for wireless monitoring. Cloud computing was required to enable large-scale, actual data analysis through IoT devices. Moreover, cloud technology would be an essential step in our proposed methodology since it serves as a foundation for storage and computation in the cloud computing environment. As a result, the conceptual approach was effective in eliminating delay induced by the public cloud program's transportation network level. Ongoing inter gadget sensory streams quickly build a vast collection in healthcare

applications. As a result, pervasive computing machines were expected to manage this information properly in live time.

Design elements

The general design of their suggested number of layers is depicted in **Figure 1**. It uses mist processing units as well as a web computer to gather medical big data technologies through IoT portable tech. A diverse gadget which was before, analysis, connection, including statistics have all been guaranteed in the actual time of the foundation required to construct. Because cloud servers were nearer to the wearable technology phone's geographical place, certain services should be supplied during particular time windows, as depicted in **Figure 1**. The cloud services, transport massive quantities of data to the cloud environment, which processes the compute complex applications.

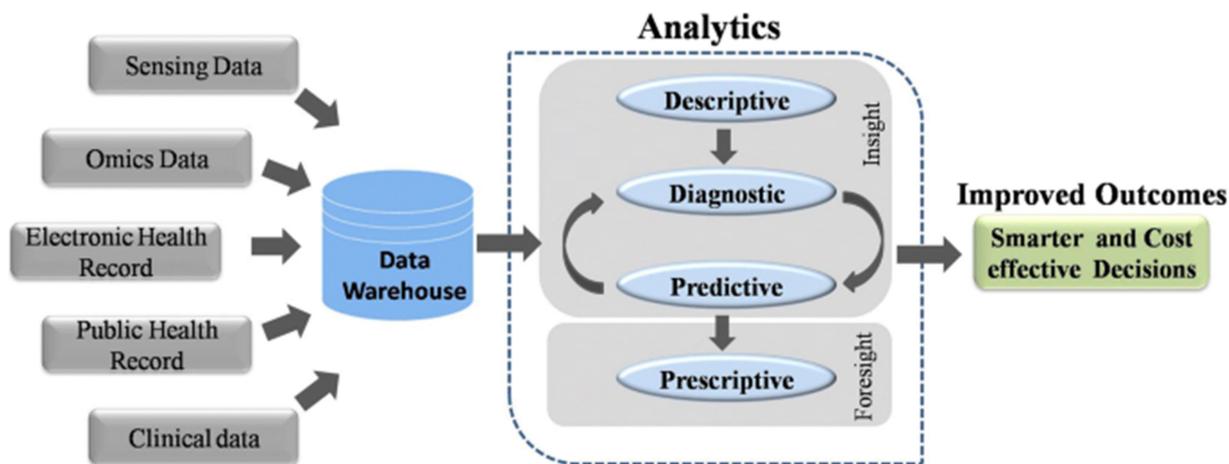


Figure 1: The architecture of the proposed framework

It was divided into three sections: (1) sensor information, including pretreatment, (2) characteristic queue, choice, plus separation, as well as (3) the prediction method. Data preprocessing, which essentially generalizes the influence of various information generated by sensors, can take several diverse features. The normalization procedure has been used to collect the information from various sensors. Ex. (1) As well as (2), however, offer normalization by base average squaring as well as normalization by upper and lower bounds.

$$E_s^* = \frac{E_{ti}}{\sqrt{\frac{1}{n} \sum_{j=1}^n E_{ti}}}, i \in [0, n] \quad (1)$$

$$E_s^* = (E_{ti} - Min_o) \frac{Max_n - Min_n}{Max_o - Min_o} + Min_n, i \in [0, n] \quad \dots(2)$$

Method 1 depicts a basic feature selection method. For selecting features, a variety of evaluation methods have been used. The Gini coefficient, in the example, evaluates the average disparity between the elements of a frequency and percentage that ranges from zero (total equal) through 1 (total disparity) (complete inequality). The Browns method in Eq. (3) Was commonly used to determine the Gini index, wherein E signifies Gini coefficient, represents accumulate a percentage with one character, and has been added up a percentage of the

target concept. Any other strategy for feature extraction was an Information Gain (IG), which would be founded on groundbreaking work in cognitive science.

$$H = |1 - \sum_{i=1}^n (y_i - y_{i-1})(y_i + y_{i-1})| \quad (3)$$

Healthcare smartwatches initiate information creation in the proposed model. That information is collected through multiple devices as well as transmitted to remote Healthcare portals, which act as a conduit here between sensors as well as the pervasive computing level. The information-gathering process was made possible by Communication devices that connect smartwatches to IoT controllers. M2M, as well as Messaging Queuing Telemetry Transport (MQTT), were two well-known instances of these systems. The IoT endpoints in their proposed model seem to be in charge of regional continuous data collection, storing, management, including screening. Moreover, known as clusters authenticity was assured utilizing strategy entry control with the help of the medical perimeter firewall. Since the quantity, speed, frequency diversity of features has always been increasing, active surveillance information produces a huge quantity of data. Considering the complexity as well as the scale of information, as well as the necessity for pretreatment, cleaning, screening, as well

as placing the original, health information analysis takes a very long time. Cloud-based routers act as a bridge among fitness trackers as well as the public cloud. This shortens the time consumption, which would be critical in time-critical situations.

The cloud receiver receives complex data and also doesn't utilize a theory. The information which was before, rule mining, categorization, forecasting, as well as visualization have all been services company offer. Mistakes, duplication, including anomaly management have all been part of the market phase. Moreover, the cloud servers must parse as well as filter every incoming IoT stream of data to convert them into a process commonly. The cloud technology layer processes the products and services in the proposed architecture (including historic data science, file systems, identification, authorization applications, including simple healthcare resources architecture). The public cloud effectively implements big data analysis that requires a lot of computing, including MapReduce as well as Apache Spark. To keep records of IoT healthcare devices as well as refresh cloud resources, the cloud system depends on bouncing back processing.

Healthcare Management System

The effectiveness of classification lies at the heart of every digital health platform. Because the provided information could be large, unlimited, even uneven, categorization becomes much more difficult. In addition, the standard deviation could be uneven, leading inside one class to get greater observations than the others. The set of features or characteristics seems to be the most important aspect of the classification modeling process. Supervised learning methods like Alex Net, LeNet, Multilayer Perceptron (MLP), as well as persistent extended shorter-term storage have now been shown to surpass several other traditional methods in terms of productivity, clarity, overall reliability when working with serious information.

The medical security measure was critical in ensuring that admittance credentials were correct and also that which was before the rules were followed. An IoT brokerage protects the information in a digital network by accepting or rejecting entry. Queries that have been recorded as well as verified then were transferred towards the next tier. Detection approach, as well as confidential information retrieval, assure information security. Various assaults could be made of IoT information obtained through smart technology. As a result, to manage these

assaults as well as to avoid data tampering or deletion, a strong privacy system must be in place. Those processes aren't taken into account throughout this work. Exercise capacity, safety, interaction safeguards, data - flow safeguards, data encryption safeguards, verification, health - care security mechanisms, healthcare handset tracking, as well as vulnerability assessment components, were included in the suggested scheme to block fraudulent threats, malware, as well as other factors that may influence trustworthiness.

DISCUSSION

Table 1 summarizes 4 main systems to their suggested structure in terms of launchpad, broadcasting kind, broadcasting fundamental, broadcasting resource, as well as packet delay in addition to assessing its applicability. Most seriously lack broadcasting (i.e., they merely batching), as well as the data transmission was significant (several minutes) in that and media, as shown in **Tables 1**. According to a reliance pause unloading approach, the suggested architecture outperforms previous platforms in terms of bandwidth. It makes use of a cloud environment to transport big amounts

of information to a computer company that manages computationally expensive throughput and reduces.

Furthermore, however, most quality healthcare monitoring systems were built on an application-level layer (like Apache Spark as well as Hadoop) plus internet connectivity (like Cassandra, spouts, or HDFS), our framework goes one step forward along with some additional advanced functionality. In comparison to certain other relevant models (**Table 2**), the conceptual approach outperforms them to ensure validity, durability, level isolation, as well as high availability. Those platforms were restricted to pictures only, were not extensible, as well as lacked privacy as well as data consistency. To protect the data, denial of service, including impersonating, key qualities include dependability, legality, security, as well as privacy. Some analogous systems, on the other hand, miss many of these characteristics due to the difficulty to memorize a PIN to obtain entry, access control difficulty, especially delayed confirmation schedule **Table 2**. The proposed framework comparison with features.

Table 1: Framework of proposed system

	Platform	Type	Primitive	Source	Transmission Delay
Mostafa RR[1]	Spark	Realtime	-	HDFS	High
SivaparthipanCB[2]	Opnet	Batch	Dstream	-	High
LvZ[4]	Opnet	Realtime	Keyvalue	HDFS	Medium
QaffasA[8]	Hadoop	Batch	Dstream	-	High

Proposed system	Spark	Batch/Realtime	Dstream	HDFS	Low
-----------------	-------	----------------	---------	------	-----

Table 2: Characteristics comparison of the proposed framework

	Sustainabi lity	Tolerance	Reliability	Compliance	confidentiality	validity	Integrity
Mostafa RR[1]	High	High	Low	Low	High	Low	High
SivaparthipanCB[2]	Low	Low	High	High	Low	Low	Low
LvZ[4]	High	Low	High	High	High	Low	Low
QaffasA[8]	Low	High	High	High	Low	Low	High
Proposed system	High	High	High	High	High	High	Low

CONCLUSION

Researchers developed a system for health insurance IoT advanced analytics that uses combined mist as well as internet technology throughout this study. It moreover makes use of Apache Spark elements to identify healthcare evidence and create fast judgments utilizing various computational methods. The suggested framework has been proposed with a complete assessment as well as an example. In principle, the current proposal might provide true alerts to clients whenever an issue comes, enabling individuals to intervene if required. Insights performed with the cloud layer aid in the management of extraordinarily high medical IoT streaming data coming from various origins. A thorough examination of the proposed methodology would be carried out in the coming, with just an emphasis on computation, extraction of features,

especially virtualized environment dumping issues.

REFERENCES

- [1] El-Hasnony IM, Mostafa RR, Elhoseny M, Barakat SI. Leveraging mist and fog for big data analytics in IoT environment. Transactions on Emerging Telecommunications Technologies. 2021 Jul;32(7):e4057.
- [2] Sivaparthipan CB, Muthu BA, Manogaran G, Maram B, Sundarasekar R, Krishnamoorthy S, Hsu CH, Chandran K. Innovative and efficient method of robotics for helping the Parkinson's disease patient using IoT in big data analytics. Transactions on Emerging Telecommunications Technologies. 2020 Dec;31(12):e3838.
- [3] Li W, Chai Y, Khan F, Jan SR, Verma S, Menon VG, Li X. A comprehensive survey on machine learning-based big data analytics for

- IoT-enabled smart healthcare system. Mobile Networks and Applications. 2021 Jan 6:1-9.
- [4] Lv Z, Qiao L. Analysis of healthcare big data. Future Generation Computer Systems. 2020 Aug 1;109:103-10.
- [5] Sánchez-Gallegos DD, Carrizales-Espinoza D, Reyes-Anastacio HG, Gonzalez-Compean JL, Carretero J, Morales-Sandoval M, Galaviz-Mosqueda A. From the edge to the cloud: A continuous delivery and preparation model for processing big IoT data. Simulation Modelling Practice and Theory. 2020 Dec 1;105:102136.
- [6] Khanra S, Dhir A, Islam AN, Mäntymäki M. Big data analytics in healthcare: a systematic literature review. Enterprise Information Systems. 2020 Aug 8;14(7):878-912.
- [7] Atitallah SB, Driss M, Boulila W, Ghézala HB. Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions. Computer Science Review. 2020 Nov 1;38:100303.
- [8] Manoja, I., Sk, N. S., & Rani, D. R. (2017, March). Prevention of DDoS attacks in cloud environment. In 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC) (pp. 235-239). IEEE.
- [9] Dr.P.Sivakumar an Mr.K Rajagobal, ‘Object Based Ring Routing Path Management Algorithm for Energy Efficient Nest Node of Sensor Network’, Journal of Computer and System Sciences, Volume 83, Issue 3, ISSN: 0022-0000, 2017, Pages 3-21. (Scopus Indexed))
- [10] Dr.P. Sivakumar, ‘Efficient Job Scheduling of Genetic Algorithm with Tabu Search and Round Robin’, International Journal of Printing, Packaging & Allied Sciences, (ISSN 2320-4387), vol. 4, no. 4, pp. 2864-2878,2016.
- [11] Latchoumi, T. P., & Parthiban, L. (2021). Quasi Oppositional Dragonfly Algorithm for Load Balancing in Cloud Computing Environment.
- [12] Chinnamahammad Bhasha, A., & Balamurugan, K. (2021). Studies on Mechanical properties of Al6061/RHC/TiC hybrid composite. International Journal of Lightweight Materials and Manufacture.

-
- [13] Pavan, M. V., Balamurugan, K., Srinivasadesikan, V., & Lee, S. L. (2021). Impact and Shear Behavior of PLA/12% Cu Reinforced Composite Filament Printed at Different FDM Conditions. *Arabian Journal for Science and Engineering*, 1-12.
- [14] Latchoumi, T. P., Vasanth, A. V., Bhavya, B., Viswanadapalli, A., & Jayanthiladevi, A. (2020, July). QoS parameters for Comparison and Performance Evaluation of Reactive protocols. In *2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE)* (pp. 1-4). IEEE.
- [15] Renugadevi N, Saravanan S, Sudha CN. Revolution of Smart Healthcare Materials in Big Data Analytics. *Materials Today: Proceedings*. 2021 May 9.
- [16] Adi E, Anwar A, Baig Z, Zeadally S. Machine learning and data analytics for the IoT. *Neural Computing and Applications*. 2020 Oct;32(20):16205-33.