International Journal of Digital Information and Wireless Communications (IJDIWC) 6(4): 274-280
The Society of Digital Information and Wireless Communications, 2016 ISSN: 2225-658X (Online); ISSN 2412-6551 (Print)

Proposition of an Intelligent System for Predictive Analysis Using Medical Big Data

Basma Boukenze [1]*

[1]Computer, Networks, Mobility and Modeling laboratory
FST, Hassan 1st University, Settat, Morocco
Email: basma.boukenze@gmail.com

Abdelkrim Haqiq[2]

[2]Computer, Networks, Mobility and Modeling laboratory
FST, Hassan 1st University, Settat, Morocco
e-NGN Research Group, Africa and Middle East
ahaqiq@gmail.com

ABSTRACT

Emerging technologies such as mobile application, cloud computing, big data analytics, predictive analytics revolutionized all sectors. This is particularly true for the healthcare system as a sensitive sector. Nowadays, healthcare industry mainly depends on information technology to provide best services.
A big-data revolution in healthcare starts with the vastly increased supply of health data. In fact, these new technologies are applied to improve the medical sector.
This paper proposes an intelligent system that combines big-data analysis with data-mining and mobile healthcare techniques for self-monitoring. The system attempts to exploit the healthcare data through an intelligent process analysis and big data processing. This approach aims to extract useful knowledge to be used in decision making and to ensure a real-time medical monitoring.

KEYWORDS

Emerging Technologies, Healthcare Data Analysis, Mobile Application, Cloud Computing, Big Data Analytics, Data Mining, Learning Algorithm

1 INTRODUCTION

Nowadays, a number of technologies including cloud computing, mobile phone, big data and predictive analytics have been identified as Emerging Technologies (ET) [1]. (ET) are defined as innovation in a constantly and accelerating evolving and touch every sector resulting a big change in process and procedure, efforts are multiplied now to applying ET in sensitive sector like healthcare system.
Digitized information is omnipresent because Data is growing and moving faster than healthcare organizations can consume it. This mainly is due to the efforts of researchers in the medical field and their discoveries take as an example human DNA. Widespread use of the electronically medical records wishes totally transforms medical care [2]. the latest innovations concerning genetics and smart home or smart places enables patient self-monitoring and treatment by using simpler devices[3]. The appearance of sensing technology like M-health [4]; healthcare data appears like a digital flood creating puddles and lakes, creeks and torrents, of data , and this increases in parallel with the rapid growth in the use of mobile devices like smart phones, laptops, tablets, personal sensors .
Large data volumes at high velocities were originally an option that characterizes supercomputers, nuclear physics, military simulations and space travel. Late in the 20th

century, bigger and faster data appeared in airline and bank operations, particularly with the growth of credit cards. Starting in 1990, The Human Genome Project was the launch of Big Data in healthcare [5], and this was due to a statistic that showed that 80% of medical data is unstructured and is clinically relevant and much significant. This data resides in multiple places like individual EMRs, lab and imaging systems, physician notes, medical correspondence, claims, CRM systems, and finance.

The potential of Big Data analytics allows to slow the ever-increasing costs of care, help providers to practice more effective medicine, empower patients and healthcare providers, support fitness and preventive self-care, and dream about more personalized and predictive medicine. Yet, social media, cloud computing, and using the intelligent procedure for managing analyzing and extracting information from Data; this approach will transform healthcare system and gives the power to explore, predict and why not anticipate the cure. Big-data analysis promises and affirms that future is no longer mysterious.

We discuss the great role played by new technology in the field of health like healthcare analysis, and then we present our proposed system and its contribution.

The rest of this paper will present as follow: in section II, we present related works concerning technologies applied in the healthcare system and research's work in this field. Section III is reserved for description of our proposed platform. And the last section gives conclusions and perspectives.

2 RELATED WORKS

The medical industry has been swamped by new technologies because of many proofs, we take as an example the implementation of e -health systems. Such as Kagawa University that was designed and implemented for the academic health education. E-Healthcare is a form of private cloud service for university students who can get their health records from physical measurement devices with their authentication based on smart card, their health data is managed in the appropriate database, doctors / nurses can investigate the relevant data according to their requests and the necessary information to self-healthcare are provided by the Web service [6].

if we talk about Cloud computing as new technology applied in the healthcare system ,it brings many benefits ; by creating a network between doctors; patients and healthcare institutes and facilitates access to medical information anywhere and anytime [7],

Cloud computing provides healthcare a much appreciated services concerning data handling by ensuring [8, 9]:

- Resiliency: platforms offers by cloud service providers are characterized by a powerful infrastructure that provides redundancy and storage of any data quantity to ensuring high availability anytime and anywhere.

- Mobility: the cloud infrastructure is providing the backbone for medical personnel to access all sorts of information from any location and from a whole set of devices; the communication will be done in an easier way given that the facility of access will be the same to one patient or several in the same time.

- Privacy: cloud computing platforms are characterized by a very high level of security than local IT department in a hospital can ensure.

- External management: By cloud provider don't need doing updates or installing the certificates or repairing blocking systems .

In addition to all these benefits, cloud adapts to all situations to ensure ease of access at a high level.

Many researchers are currently focusing on the benefits of new technologies [10, 11], for their advantages and promises, including the great role

of cloud computing in the stage of managing healthcare data that are becoming increasingly large.

More than that, some of them give the design of a cloud computing-based Healthcare SaaS Platform (HSP) to deliver healthcare information services with low cost, high clinical value and high usability with the high level of security [12,13].

Big data analysis specially in healthcare area is considered as a revolutionary approach to improving the quality of healthcare service [14, 15], because analytics figures to play a pivotal role in the future of healthcare system and as a result of research to develop healthcare sector [16] systems found obliged to receive a new form of data such as: human DNA, data genetics; hence the necessity of leveraging all these resources and embitter human health. Analytics also are now applied in healthcare to compare the cost and effectiveness of interventions, treatments, public health policies, or medical devices to reduce failed investments.

In fact, this kind of analysis can give the best solution to prevent medical disasters. For example, infectious diseases can be predicted by data healthcare analysis and the health authority could manage this situation and save the humans.

Also we will soon be awash in genomic data [17], given the incredible size and dimensionality of these datasets, the field of analytics will need to borrow techniques to face it and to make it useful.

in addition to that , some predictive analytics platform for disease targets across varying patient cohorts using electronic health records (EHRs) are created to facilitate specific biomedical research workflows, such as refinement of hypotheses or data semantics [18].

A lot of tools are used now to create platforms for big data analytics, the most known is the open-source distributed data processing platform or Hadoop (Apache platform) [19]. It belongs to the class of technologies "NoSQL" that have evolved to managing data at high volume. Hadoop has the potential to process extremely large amounts of data mainly by allocating partitioned data sets to numerous servers (nodes), each of which solves different parts of the larger problem and then integrates them for the final result [20,21]. Hadoop can serve both roles of organizing and data analyzing tool .Hadoop can handle very large volumes of data with different structures or no structure at all. But Hadoop is a little difficult to install, configure and manage, and people with Hadoop skills are not easily found. In addition, for these reasons, it appears organizations are not quite ready to embrace Hadoop completely.

Knowing that the adoption of EHRs and electronics data, prepared a submitted base for applying analysis and become the norm in healthcare, it enables the building of predictive analytic solutions. These predictive models, as we know have the potential to lower cost and improve the overall health of the population. As predictive models become more pervasive, some standards appear to be used by all the parties involved in the modeling process: like The Predictive Model Markup Language (PMML) [22].It allows for predictive solutions to be easily shared between applications and systems. And it can be used to expedite the adoption and use of predictive solutions in the healthcare industry.

According to our research, we found that there are many efforts to creating platforms based on cloud computing for managing medical records and simplify access to data. The patient does not care about the way with his doctor manages his medical data. But the most important for him is what is the positive impact of this, on his health situation?

What we propose is a platform that combines the benefits of mobile healthcare and big data analysis. Making as the primary objective, exploration and extraction of useful knowledge and self-monitoring in real time for patients.

3 SYSTEM ARCHITECTURE

3.1 System Characteristics:

The proposed solution is an intelligent system that we gave the name of ‘’Intelligent Predictive Healthcare System (IPHCS)’’, it analyzes medical data which coming from different sources in a real-time, this process helps to decide about patient health condition by using the extracted

information from his own data . So the system will:

- ✓ be hosted in a cloud and can be accessed anytime, anywhere, and by any communication equipment,
- ✓ make a quick analysis in real-time to give accurate future information using intelligent and very specific tools,

The reaction of the two main actors (Patient / Doctor) in the system is as follows:

Doctor accedes to IPHCS for:

- Consult patients profiles
- Monitoring and controlling the health status of each patient
- Introducing new data (patient or subscription of treatment)

Patient has a dual interaction with the system:

Indirect interaction: the patient follows the guidelines of his doctor, who is based on turn on IPHCS to decide.

Direct interaction: the patient has a medical device such as (Smartphone, Smart watch, Bracelet) equipped with a sensor designed to detect for example (heart rhythm using infrared LEDs and photodiodes sensor; evaluate the intensity of effort by measuring your heart rate …etc).

Information received by the sensor will be managed by the IPHCS which then produces the following:

- ✓ A report sends to the doctor to alert him about the change of patient status.
- ✓ An emergency warning message sends to the patient in case of emergency to alert and warn him about his status, pending the intervention of the concerned doctor.

The captured information is sent and subsequently managed by the system, which monitors in real-time.
The doctor intervenes on the basis of the received report, and the patient will be contacted for the necessary. (Figure1).

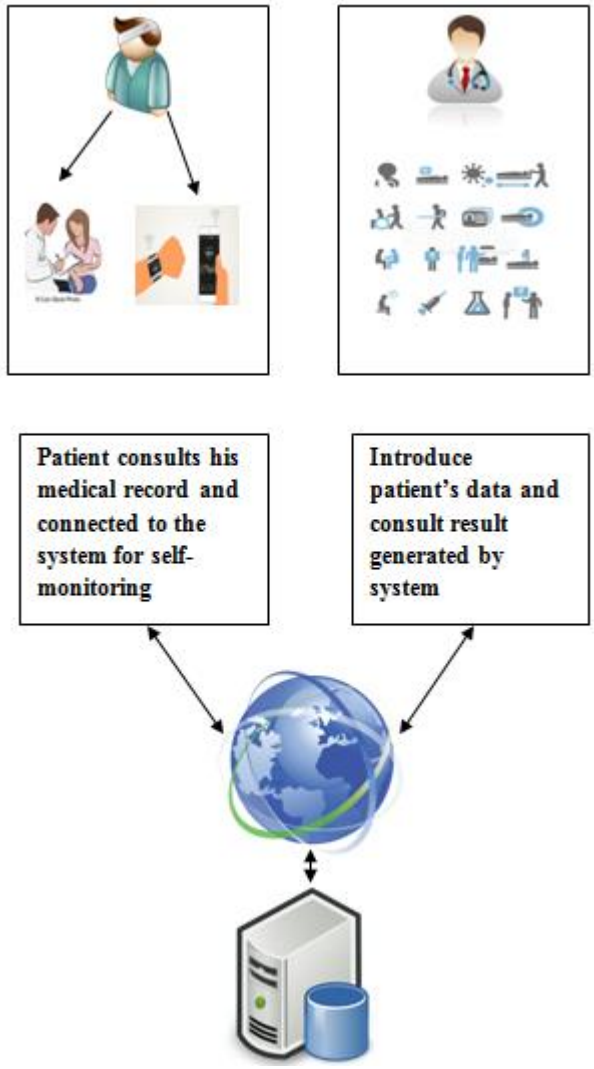

Figure1. Typical intelligent Healthcare system schema

3.2 System Architecture

IHPCS is capable to:

- Analyze a large amount of medical data;
- predict what the patient may have in the future as complexity and pathologies by data mining techniques;
- Anticipate the cure and treatment;
- Monitor patient in real time;

- Provide Patient opportunities to make a self- monitoring in real-time by the use of health mobile devices.

Our proposed system architecture is shown in Figure2 comprise of several steps.

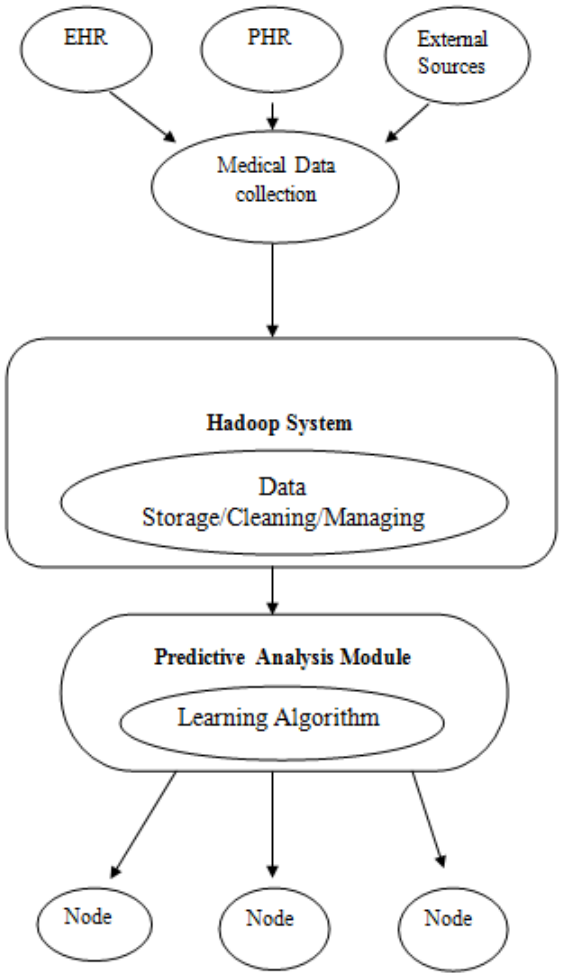

Figure 2.Architecture of the predictive analysis system-Health Care Application

Data collection: is the most important and sensitive phase because the data is the main element and the pivot of the system. We must mention that more data is accurate the predicted information is more accurate.

The voluminous medical data can come from various Electronic Health Record (EHR) / Patient Health Record (PHR), Clinical systems and external sources like government sources, laboratories, pharmacies, insurance companies etc, in various formats (flat files, .csv, tables, ASCII/text, etc.) .

Storage and processing: it is a very important phase seen it demands very powerful techniques and tools to manage and process the voluminous data.

Predictive analysis: is the master step in all this process, because it rests on the exploration of analyzed data to extract useful knowledge on the basis of data mining tools and algorithms to find links between the medical data.

Processing visualized reports: The results obtained after the predictive analysis process are Exploited by Doctors and Patients as follows:

• Doctor for help in decision making and giving a general view of the patient's status.

• Patient will have the results of this process by his doctor but he is always in interaction with the system by a mobile device that he owned.

3.3 Used Technologies

In the first layer, Hadoop is used as an open source framework designed to perform processing on massive medical data, the operating principle is as follows (please check whether you have clearly conveyed the principles.)

The infrastructure applies the well-known principle of grid computing, of dividing the execution of a process on multiple nodes or clusters of servers.

In Hadoop architecture logic, this list is divided into several parts, each part being stored on a different server cluster. Instead of lean processing in a single cluster, as is the case for traditional architecture, the distribution of information helps distribute the processing across all compute nodes on which the list is distributed [23].

To implement such a technical process, Hadoop is coupled to a file system called HDFS (Hadoop Distributed File System for). It manages the allocation of storage of user data in blocks of information on different nodes. HDFS was inspired by a technology used by Google to own these cloud services, and known as Google File System (GFS).

Map/Reduce: the distribution and management of the calculations are carried out by Map Reduce. This technology combines two types of function:
The Map function: which resides on the master node and then divides the input data or task into smaller subtasks, which it then distributes to worker nodes that process the smaller tasks and pass the answers back to the master node?
The subtasks are run in parallel on multiple computers.
The Reduce function: collects the results of all the subtasks and combines them to produce an aggregated Final result — which returns as the answer to the original big query.

The second layer is characterized by the great role of Map-Reduce module for the process of predictive analysis. And to reinforce more and more the system in matters of prediction, it must be equipped a powerful predictive algorithm or learning algorithm to ensure the important phases of the process and build a suitable model of prediction.
Data mining technology like a delicate process , executed by predictive algorithms, which have shown a strong effectiveness and efficiency in predicting , take as an example supporting vector machine (SVM) [24], decision tree(C4.5) [25] , and Naive Bayes (NB) [26], as They Are Currently classified Among the top 10 classification methods Identified by IEEE Python & Related Resources [27].
For that, our system should be equipped with a learning algorithm among the cited ones or a combination of several learning algorithms to benefit from its performances and build a powerful hybrid algorithm that will be applied to all types of medical prediction.

4 CONCLUSION

Certainly predict the future is no longer a difficult task, with emergent technologies , medical field will benefit from all the voluminous medical data to extract knowledge for helping to decision making, reduction of cost and go beyond the subscription processing up to the prediction of diseases before their apparition.
being given that medicine faces long certain diseases called silent diseases like some chronic diseases and that early diagnosis can eradicate them.
With predictive analysis of big data, we will have the power to solving the major problems that are a real challenge in front of the development of medicine just to save human life.

REFERENCES

1. Macharia, P., Singh,U.G., Zuva,T.: Improving, Quality and Access to Healthcare by Adopting Emerging Technologies, International Journal of Digital Information and Wireless Communications (IJDIWC) , 2016, vol,6 pp, 41,45
2. Richard, H., James, B., Anthony, B., Federico, G., Robin, M., Richard, S. and Roger, T.: Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs. Health Affairs, 24, no.5 (2005):1103-1117 .
3. Marianthi, T., Nikos, T.: Smart Home Solutions for Healthcare: Privacy in Ubiquitous Computing Infrastructures.
4. Santosh, K., Wendy, J.: Mobile Health Technology Evaluation ; ELSEVIER ; Volume 45, Issue 2, August 2013, Pages 228–236
5. Bonnie, F., Ellen M., Tobi, S.: Big Data in Healthcare Hype and Hope; Business Development for Digital Health; October 2012.
6. Miyazaki, M., Kamano, H., Imai, Y.: An e-Healthcare System for Ubiquitous and Life-long Health Management, International Journal of Digital Information and Wireless Communications (IJDIWC), 2016, vol. 6, pp 163-172.
7. Madhusudhana, R., Rambhupal , B.: Survey of Adapting Cloud Computing in Healthcare; International Journal of Advanced Research in Engineering and Applied Sciences ; 2278-6252
8. Barbara, C., Mario, C.: Cloud Computing In Healthcare And Biomedicine, SCPE; Volume 16, Number 1, pp. 1–18.
9. Sanjay, P., Sindhu, M., Jesus, Z.: A Survey of the State of Cloud Computing in Healthcare; Canadian Center of Science and Education. Vol. 1, No. 2, 2012 , September 19, 2012 .
10. Jonathan, N., Brian, M., Sharat, K.: Healthcare in the cloud: the opportunity and the challenge. MLD.
11. Lena, l. , Hans-Ulrich, P.: A scoping review of cloud computing in Healthcare ; BMC Medical Informatics and Decision Making (2015) ; DOI 10.1186/s12911-015-0145-7.

12. Sunkara, V., Akula, S.: Role of Cloud Computing in Health Monitoring System ; IJSEAT, Vol 2, Issue 10, October – 2014
13. Sungyoung, O., Jieun C.: Architecture Design of Healthcare Software-as-a- Service Platform for Cloud-Based Clinical Decision Support Service; Health Inform Res. 2015 April; 21(2):102-110.
pISSN 2093-3681 , eISSN 2093-369X.
14. Michael, J., Keith, A., Craig, M.: Applications of business analytics in healthcare. ELSEVIER. Business Horizons (2014) 57, 571—582.
15. Nura, E., Mohammad Reza, B.: Knowledge discovery in medicine: Current issue and future trend; ELSEVIER; Expert Systems with Applications 41 (2014) 4434–4463.
16. Srinath, S., Sameep, M.: Big Data Analytics. Springer ;December 20-23, 2014 ; Volume 8883 2014 .Available at http://link.springer.com/book/10.1007/978-3-319-13820-6
17. Peter, G., Basel, K.: The 'Big Data' Revolution in Healthcare; McKinney and Company April 2013. Center for US Health System Reform Business Technology Office. Accelerating-value-and-innovation.pdf
18. Kenney, N., Amol G.: PARAMO: A Parallel Predictive Modeling Platform for Healthcare Analytic Research Using Electronic Health Records; ELSEVIER. Journal of Biomedical Informatics 48 (2014) 160–170.
19. Wullianallur, R., Viju, R.: Big Data Analytics in Healthcare: Promise and Potential, Health Information Science and Systems 2014.
20. Borkar, V., Carey, M., Chen, L.: Big Data Platforms: What's Next? ACM Crossroads 2012, 19(1):44–49.
21. Zikopoulos, P., Eaton, C., DeRoos, D., Deutsch, T., Lapis, G.: Understanding Big Data – Analytics for Enterprise Class Hadoop and Streaming Data. McGraw-Hill: Aspen Institute; 2012.
22. Alex, G.: Predictive Analytics in Healthcare the Importance of Open Standards. ZEMENTIS IBM Developers Worker. 29 November 2011.
23. Alfredo, C., Yeol S.: Analytics Over Large-Scale Multidimensional Data: The Big Data Revolution; DOLAP'11, October 28, 2011, ACM 978-1-4503-0963-9/11/10.
24. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2013, 200 pp,
25. Rahman, R., Rabbi, F.: Using and comparing different decision tree classification techniques for mining ICDDR, B Hospital Surveillance data'', Elsevier, V, 38, pp 11421–11436
26. Ang, S., Ong, H., Chin Low, H. .: Classification Using the General Bayesian Network ,science and technology, 24 (1) 205 – 211 pp, (2016)]
27. Top Data Mining Algorithms Identified by IEEE & Related Python Resources: http://www.datasciencecentral.com/profiles/blogs/python-resources-for-top-data-mining-algorithms