Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 50 (2015) 537 – 542

Procedia Computer Science

2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

Designing a Cloud based Framework for HealthCare System and applying Clustering techniques for Region Wise Diagnosis.

Maulik Parekh, Dr. Saleena B.,
[1]M. Tech CSE with Specialization in Cloud Computing, VIT University, Chennai 600048, India
[2] Associate Professor, SCSE, VIT University, Chennai 600048, India

Abstract

With the advancement of technology and the limitations of the conventional healthcare system, an improvised framework for healthcare system is needed. This paper presents another cloud based skeleton which relates key segments of any healthcare framework which are patient, doctor, symptom and disease. The paper fundamentally concentrates on how these parts are inter-related and how we can infer suitable data from them. As an implementation, it shows the basic healthcare analyser interface which takes data as input and mines the data by using some of the data mining techniques like clustering. It is convenient for government associations which point at investigating restorative issues and to enhance health conditions of India.

Peer-review under responsibility of scientific committee of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

Keywords: ***HealthCare system; Data mining; Cloud Computing; Open Stack; Cloud Foundry***

1. Introduction

The cloud is a large set of interconnected computers which provide services anytime. We have to pay only for the services that we are using. The softwares do not run from a personal computer, rather they are stored on the server and are accessed through the internet. The healthcare industry undergoes tremendous pressure to deliver quality service to patients and doctors across the globe. This is where the Information Technology comes as a rescue to the continuous and voluminous demands for the healthcare sector. A healthcare cloud is designed as an interconnection of extensive number of computers and servers specifically dedicated to meet the needs of the healthcare industry. The healthcare service is delivered via an internet connection to the user who can be a either a doctor or a patient. The cloud service enables registered users to access the hardware and software managed by the third parties at remote locations. It has brought substantial transformation in the way the information is stored and accessed. The service availability is guaranteed along with unlimited scalability, cost reduction (pay per use) and much better performance than the traditional computing techniques. Our paper is divided into following sections:

1) The key components of the system with a brief overview of cloud computing and its models are discussed.
2) Analysis of the existing systems is done and their drawbacks that limit their usages are discussed.
3) Our proposed model that has evolved by combining a number of technologies like Big Data, cloud are discussed thus, bringing together the advantages which are enjoyed in each of these technologies. The proposed model is

*Corresponding author. Tel.:+91-846-038-9229
E-mail address: maulikparekh2@gmail.com; parekh.maulik2013@vit.ac.in

Peer-review under responsibility of scientific committee of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)
doi:10.1016/j.procs.2015.04.029

compared with the existing system and enumerates way in which our model overcomes the limitations faced by the present system. We have shown our model as an end to end flow where healthcare customers – patients, doctors and government have been considered as the end users. Each of these end users are empowered with different sets of functionalities in our system.
4) The detailed implementation of our system is described and the limitations that our system fails to address are discussed.
5) Finally, the conclusions are drawn and the future works that can further enhance our proposed system is discussed.

2. Existing works

Chenguang He and Xiaomao Fan introduced a cloud based platform for Healthcare Services but there are no relationships between symptoms and diseases; diseases and their appropriate specialists/doctors. So by using this platform and adding other functionalities we have tried to propose a new architecture. [5] Ahn, Y.W., Cheng, A.M.K., Baek, J., Jo, M. and Chen, H.-H have invented a scaling mechanism for virtual resources to support mobile, pervasive, real-time healthcare applications based on the cloud but their focus is upon the virtualization on cloud. The questions that they seem to be focusing are- 'How can we add more virtual resources?' and 'How can we backup or migrate data?' .Thesel problems are solved in our cloud based proposed model [6]. 'Hybrid Clinical Decision Support System for rural Bangladesh' is also a good support system for healthcare. The authors have proposed an automated diagnostic system which could be used in the rural areas where there is a scarcity of doctors [7].

'A Hybrid Mobile-based Patient Location Tracking System for Personal Healthcare Applications' system uses cellular network infrastructure using GPS system for tracking the location of patients. So we have used the idea of GPS in our proposed model to locate patient and find out nearest specialist doctors in surrounding area [12]. NFC based patient appointment system is also useful in our model as our model is flexible and generic so we can add more features also [13].

2.1. Limitations:

There is no efficient correlation from symptoms to diseases and diseases to the specialist doctor in previous systems. Our proposed system predict appropriate diseases based on the symptoms and then display the list of specialist doctors in nearby areas. Currently all kind of data related to patients, doctors, symptoms and diseases are available, it can used to perform different analysis, to improve the Indian healthcare system. Our model is generic which we can expand easily by adding new functionalities and we can also make it efficient.

3. Proposed model

3.1. Overview:

An improvised healthcare model is designed over previous healthcare models. We have combined all the main components of the healthcare system together that are patients, doctors, symptoms and diseases. Fig 1 explains the complete working of our proposed model. It enables patients to get required health services through a mobile application with some clicks on the app's tabs. User/patient can login into the mobile application and enters the symptoms and his location. The symptoms are all the health issues that are being suffered by the user for which he expects a treatment.

Based on the symptoms entered by the user, it displays the nearest specialized doctors and hospitals with respect to the user's location. The system relates entered symptoms to diseases by analyzing historical data maintained by the system. The system analyzes by finding the closest match (disease) corresponding to the symptoms. Then it maps this matched disease to specialized doctors by acquiring details from the database. This mapping is displayed as a set of specialized doctors and their details at the user's end. User/patient can select a specialized doctor or a hospital and book an appointment. To book the appointment, the user must register into the mobile app. On the other hand, doctors can either accept or reject the appointment and an acknowledgement is displayed in the user account accordingly.

The patient visits the doctor with whom the appointment has been fixed and the doctor treats him. As per the treatment, the doctor enters the actual disease being suffered by the patient into the system through his account. The entry made by the doctor updates the overall database of our system that further may strengthen the validity of our database and enriches it. With more number of entries, the analysis of this database also gets enhanced and the user. Both the patient and the doctor have their own individual set of database. Using the 'Health History', the patient can access his own data while the doctor manages and maintains his own patient records that he can use for his own purpose.
The system updates its historical database on a daily basis.

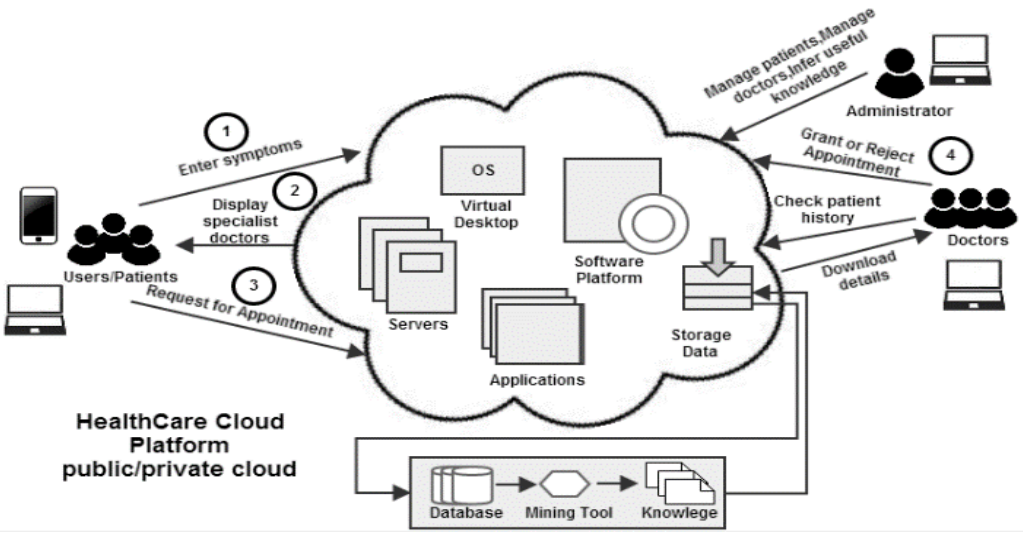


Fig.1. Architecture of the Proposed Model

3.2. Implementation:

The paper demonstrates the simple web analyzer interface which displays different data mining techniques' result. This interface is for demonstration purpose as to how we can integrate data mining techniques into the cloud based healthcare application and infer useful results. Here, we have used only clustering techniques for simplicity and for analysis purpose. For development side, we have used JAVA language and Weka API for integration of data mining techniques into eclipse IDE. Our web interface provides functionality to see different scenarios. When we click on a particular scenario it fetches data from the related database table by converting it into the CSV format. By using this CSV data file and Weka API, we use appropriate data mining techniques, i.e., clustering, to display result. These results can help different organizations and government to improve the healthcare system.

Healthcare System can use any deployment model based on service provider requirements. However, we have opted to deploy the cloud using open stack, such as, IaaS and Cloud foundry like PaaS. After configuring this, we can build SaaS healthcare application upon it. The reason being that after it works as a SaaS based application, we can use all advantages of the cloud. Healthcare cloud contains different components like servers, data storage, network, hardware, virtual desktop, applications and software platform. Each has its unique importance in cloud architecture.

To sum up everything, the following design specifications were proposed:

Cloud Computing:

Cloud Computing is not a technology. It's a model that empowers handy, on-interest system access to a pervasive pool of configurable processing assets that might be quickly provisioned and surrendered with insignificant administration exertion or convenience supplier communication. Here, the assets are networks, storage, applications, services and servers. [1]

Cloud Deployment model:

Hybrid Cloud: The cloud can be accessed both within organizations (hospitals/government authorities) as well as among public (users/patients who register to our system). For the doctors, hospitals and authorities; the health records maintained at their side can be analysed and studied thoroughly to understand the nature of their patients and the disease. For the patients, they can simply have access over their health history.
Thus, we need to deploy an hybrid cloud which is a consolidation of two or more clouds that remain distinct elements and yet are bound together by a portion of the technology that facilitate portability of application with related data like cloud blasting.[4]

Cloud Service models

- Infrastructure-as-a-Service (IaaS) provides physical environment on which the user can build their own platforms and

deploy applications upon it.

- Platform-as-a-Service (PaaS) provides a platform where consumers can create and deploy applications. It's a middleware between IaaS and SaaS.
- Software-as-a-Service (SaaS) accommodate complete application to a cloud customer. It can be access through web portals or through web services. It helps a consumer utilize outsourced applications to support the various operations of their businesses where their corporate data may or may not be stored in a cloud.
- Data-as-a-Service (DaaS) where corporate data can be overseen by a cloud computing environment.[4]

Data mining

Data mining means deriving knowledge from the data. There are different techniques which are useful to mine the data.

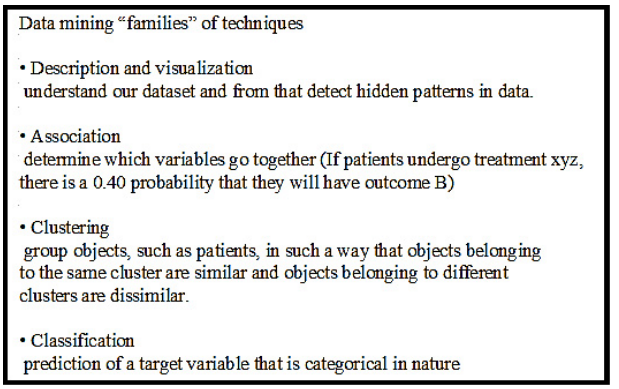


Fig. 2. Data mining Techniques

3.3 Analysis:

Once the system is enriched with data, it can perform different analysis by using data mining techniques. A sample data set's result is displayed using Weka API (Refer Fig 3 and 4). System uses "K-mean clustering techniques" for it and detail description of it is not in the scope of this paper. We can extend other techniques by integrating different algorithms in the web interface [8][14].In this dataset, age, sex, region, married, cancer, diabetes and blood pressure are the main factors of the datasets. We need to make all possible combinations for it and produce useful statistics.

From this result, we can say that cluster 4 has more instances of a married female aged around 44 , residing in the region-'Inner City', to may possibly have more chances of suffering from cancer and blood pressure.

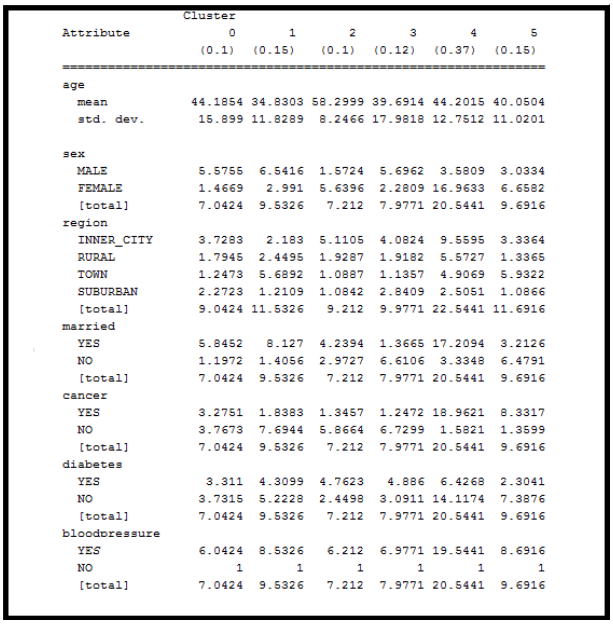
We can perform different analysis and produce as many statistics which can help to analyze the healthcare condition.

We can use various big data analyzing like machine learning, pattern recognition, Association rule learning etc. for analyzing huge amount of data. If we have data of different state, corresponding districts, different diseases, doctors and patients then we can say that in that particular region or district, the patients having a particular disease are more and that we need more medical treatment there. Some regions often have fewer doctors and more diseases so in such places, we require more medical services. Likewise, we can analyze and display useful results which is based on region-wise (state wise or district wise) studies[9].

3.4 Advantages:

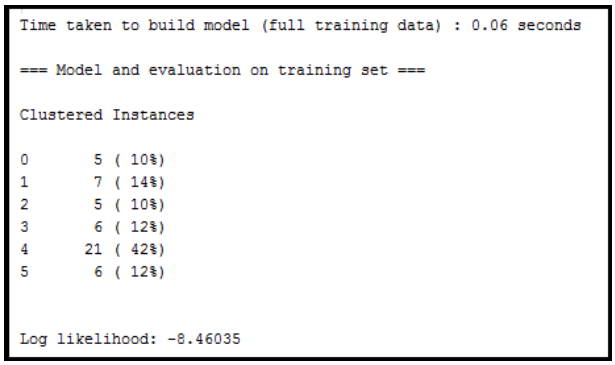
If a user is new to his location with no idea about the whereabouts of the nearby hospitals and specialized doctors, this application can prove to be a boon for them. Since, this model is based on cloud, therefore, it is more scalable and cost effective in comparison to conventional systems. This is an end to end system which can be helpful to doctors, patients and healthcare providers. It correlates appropriate symptoms to diseases and diseases to specialized doctors.

System administrator or provider can manage patient and doctors. They can access all the sharable data and apply data mining techniques and infer useful results which will help government and other NGOs to improve health conditions. It can be of a great help during the time of natural calamities. Our proposed system is covering every aspect of healthcare industry which no other existing system currently providing.



	Cluster					
Attribute	0	1	2	3	4	5
	(0.1)	(0.15)	(0.1)	(0.12)	(0.37)	(0.15)
age						
mean	44.1854	34.8303	58.2999	39.6914	44.2015	40.0504
std. dev.	15.899	11.8289	8.2466	17.9818	12.7512	11.0201
sex						
MALE	5.5755	6.5416	1.5724	5.6962	3.5809	3.0334
FEMALE	1.4669	2.991	5.6396	2.2809	16.9633	6.6582
[total]	7.0424	9.5326	7.212	7.9771	20.5441	9.6916
region						
INNER_CITY	3.7283	2.183	5.1105	4.0824	9.5595	3.3364
RURAL	1.7945	2.4495	1.9287	1.9182	5.5727	1.3365
TOWN	1.2473	5.6892	1.0887	1.1357	4.9069	5.9322
SUBURBAN	2.2723	1.2109	1.0842	2.8409	2.5051	1.0866
[total]	9.0424	11.5326	9.212	9.9771	22.5441	11.6916
married						
YES	5.8452	8.127	4.2394	1.3665	17.2094	3.2126
NO	1.1972	1.4056	2.9727	6.6106	3.3348	6.4791
[total]	7.0424	9.5326	7.212	7.9771	20.5441	9.6916
cancer						
YES	3.2751	1.8383	1.3457	1.2472	18.9621	8.3317
NO	3.7673	7.6944	5.8664	6.7299	1.5821	1.3599
[total]	7.0424	9.5326	7.212	7.9771	20.5441	9.6916
diabetes						
YES	3.311	4.3099	4.7623	4.886	6.4268	2.3041
NO	3.7315	5.2228	2.4498	3.0911	14.1174	7.3876
[total]	7.0424	9.5326	7.212	7.9771	20.5441	9.6916
bloodpressure						
YES	6.0424	8.5326	6.212	6.9771	19.5441	8.6916
NO	1	1	1	1	1	1
[total]	7.0424	9.5326	7.212	7.9771	20.5441	9.6916

Fig. 3. weka healthcare data set attributes



```
Time taken to build model (full training data) : 0.06 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       5 ( 10%)
1       7 ( 14%)
2       5 ( 10%)
3       6 ( 12%)
4      21 ( 42%)
5       6 ( 12%)


Log likelihood: -8.46035
```

Fig. 4. weka result no of clusters and instances

3.5 Challenges:

There are additionally noteworthy dangers that every health care association must face when transitioning to cloud-based facilitating. Turning over information, security, accessibility and control to an outsider implies that an organization has no true influence over where its information really exists. Confide in our cloud merchant tackles an entire diverse significance. Security and protection, center issues in the human services business, must be infallible. [3]

Other basic issues might be information accessibility, failure limits, catastrophe reinforcement and fast reaction times. Most merchants can have much a greater number of abilities than a singular client. Unapproved revelation of data brings about extreme outcomes to the association and critical expenses in recouping and restoring information and additionally advising influenced people.

At the point, recognizing a supplier's solidness, immovability of estimating structure and accessibility or uptime insurances are discriminating. Users must have sufficient data transfer capacity to oblige his needs. Nothing is more regrettable than being reliant on software that is so moderate it becomes unusable. So other issues are related to contracts which provides all the requirements of the customers. These are some of the challenges while dealing with a cloud based system [2][10].

4. Conclusion

The exhibited Health Cloud is a vigorous answer to Patients, specialists and a few health awareness suppliers. Patients can undoubtedly get on click service anywhere and anytime. Medicos and doctors both can manage patient history, making

it as a source to understand diseases, their causes and remedies that can help in enhancing the quality of treatment and medicines. Medicinal service suppliers do not need to stress over different concerns like operational cost of the system, data imparting and handling of large medical data. One of the essential viewpoints here is the means by which we can infer handy data from old therapeutic information to enhance health condition. The Cloud has lots of preferences and yet these focal points picked up can't be concealed by issues of trust, protection, and security.

5. Future Work

This paper only demonstrates the clustering techniques to infer utilizable information from the healthcare data. We can integrate some more techniques like association and classification to infer incipient patterns form the historical data which will be a subsidiary to patients, medicos, regime and different NGOs to amend the health condition.

Acknowledgements

I would like to convey my sincere gratefulness to my guide Dr. Saleena B. for the never-ending support of my research study. Furthermore, I extend my gratitude to Dr Nayeemulla Khan and Dr. Jeganathan L, Professor and Dean for their perpetual involution and subsidiary comments.

References

1. P., Mell, T., Grance. The NIST definition of cloud computing [Online], Available: http://csrc.nist.gov/groups/SNS/cloud computing/cloud-def. v15.doc, 2009. [Accessed: 15-July- 2011].
2. L. Savu (2011). Cloud Computing: Deployment Models, Delivery Models, Risks and Research Challenges, *Computer and Management (CAMAN), 2011 International Conference on*, pp. 1-4.
3. P., Metri, G., Sarote (2011).Privacy Issues and Challenges in Cloud Computing, *International Journal of Advanced Engineering Sciences and Technologies (IJAEST),* vol. 5, pp.1-6.
4. IBM (2010), Cloud Deployment and Delivery Models, Available from https://www.ibm.com/developerworks/mydeveloperworks/c2028fdc [Accessed13- July -2011].
5. Chenguang he, Xiaomao fan, Ye li (Jan 2013),Toward Ubiquitous Healthcare Services With a Novel Efficient Cloud Platform*, Biomedical Engineering, IEEE Transactions on*, Vol. 60, no. , pp. 230 – 234.
6. Ahn, Y.W.; Cheng, A.M.K.; Baek, J.; Jo, M.; Chen, H.-H (2013). An auto-scaling mechanism for virtual resources to support mobile, pervasive, Real-time healthcare applications in cloud, *Network, IEEE*, Volume: 27, pp.: 62 – 6
7. Al Iqbal, R (2012). Hybrid clinical decision support system: An automated diagnostic system for rural Bangladesh. Informatics, Electronics & Vision (ICIEV), 2012 *International Conference on*, Page(s): 76 – 81
8. Ilayaraja, M. ; Meyyappan, T. (2013).Mining medical data to identify frequent diseases using Apriori algorithm, Pattern Recognition, Informatics and Mobile Engineering (PRIME), *2013 International Conference on*, Page(s): 194 – 199.
9. Chauhan, R.; Kumar, A. (2013), Cloud computing for improved healthcare: Techniques, potential and challenges, *E-Health and Bioengineering Conference (EHB), 2013*, Page(s): 1 – 4
10. Huang, Feixiang ; Wang,Shengyong ; Chan,Chien- Chung(2012), Predicting disease by using data mining based on healthcare information system, Granular Computing (GrC), *2012 IEEE International Conference on*, Page(s): 191 – 19
11. Chew, S.H.;Biomed. Eng.Res.Centre,Nanyang Technol.Univ.;Chong, P.A.;Gunawan, E.;Goh, K.W.,A Hybrid Mobile-based Patient Location Tracking System for Personal Healthcare Applications, *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, Page(s):5188 – 5191, Aug. 30 2006-Sept. 3 2006
12. Yeo Sy Mey; Dept. of Comput. & Inf. Syst., Inst. Teknol .Brunei, Gadong, Brunei; Sankaranarayanan, S., Near Field Communication Based Patient Appointment, *Cloud & Ubiquitous Computing & Emerging Technologies (CUBE), 2013 International Conference on* ,Page(s):98 - 103
13. Xueying Wu; Inf. Sci. & Eng. Coll., Dalian Polytech. Univ.,Dalian,China ;Chunlong Yao,Application of improved K-means clustering algorithm in transit data collection*, Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*(Volume:7),Page(s):3028 - 3030
14. Poteras, C.M.; Fac. of Autom. Comput. & Electron., Univ. of Craiova, Craiova, Romania ; Mihaescu, M.C. ; Mocanu, M. ,An optimized version of the K-Means clustering algorithm*, Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on*, Page(s):695 - 699