Journal of Advances and Scholarly Researches in Allied Education
Vol. XV, Issue No. 2, (Special Issue) April-2018, ISSN 2230-7540

Health Prediction System Using Data Mining

Kalyani Shahaji Zodage[1]* Puja Sarage[2], Trupti Sudrik[3], Prof. Rashmi Sonawane[4]

Abstract – Here we propose a framework that enables clients to get moment direction on their medical problems through an astute social intelligent health care system online. The framework is bolstered with different symptoms and the disease or illness associated with those systems. Also the system allows user to share their symptoms and issues. Data Mining as a field of research has already well proven capabilities of identifying hidden patterns, analysis and knowledge applied on different research domains, now gaining popularity day by day among researchers and scientist towards generating novel and deep insights of these large biomedical datasets also. Uncovering new biomedical and healthcare related knowledge to support clinical decision making, is another dimension of data mining. Through massive literature survey, it is found that early disease prediction is the most demanded area of research in health care sector.

Keywords: Data Mining, Prediction, Health Care, Random Forest, Analysis, Classification.

----------X----------

1. INTRODUCTION

Information Mining is a non-unimportant extraction of verifiable, already obscure and potential valuable data about information. To put it plainly, it is a procedure of examining information from alternate point of view and assembling the learning from it. The found information can be utilized for various applications for instance social insurance industry. These days social insurance industry creates substantial measure of information about patients, sickness finding and so on. Information mining gives a set of strategies to find concealed examples from information. A significant test confronting Healthcare industry is nature of benefit .Nature of administration infers diagnosing sickness accurately and gives viable medications to patients. Through massive literature survey, it is found that early disease prediction is the most demanded area of research in health care sector. As healthcare domain is bit wider domain and having different disease characteristics, different techniques have their own prediction efficiencies, which can be enhanced and changed in order to get into most optimize way.

2. SENSOR INTRODUCTION

■ Heartbeat Sensor

A persons heartbeat is the sound of the valves in his/hers heart contracting or expanding as they force blood from one region to another. The number of times the heart beats per minute (BPM),is the heartbeat rate and the beat of the heart that can be felt in any artery that lies close to the skin is the pulse.

■ Temperature Sensor

The LM35 series are precision integrated-circuit temperature sensors, whose output voltage is linearly proportional to the Celsius(Centigrade)temperature. The LM35 thus has an advantage over linear temperature sensors calibrated in Kelvin, as the user is not required to subtract a large constant voltage from its output to obtain convenient Centigrade scaling.

3. RESEARCH METHODOLOGY

■ Random Forest Algorithm

Random forest or random decision forests Are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

Decision trees are a popular method for various machine learning tasks. Tree learning "come[s] closest to meeting the requirements for serving as an off-the-shelf procedure for data mining", say Hastie *et al.*, "because it is invariant under scaling and various other transformations of feature values, is robust to inclusion of irrelevant features, and produces inspect able models. However, they are seldom accurate"

4. LITERATURE SURVEY

[1] The prediction of survival of Coronary heart disease (CHD) has been a challenging research problem for medical society. The goal of this paper is to develop data mining algorithms for predicting survival of CHD patients based on 1000 cases .We carry out a clinical observation and a 6-month follow up to include 1000 CHD cases. The survival information of each case is obtained via follow up. Based on the data, we employed three popular data mining algorithms to develop the prediction models using the 502 cases. We also used 10-fold cross-validation methods to measure the unbiased estimate of the three prediction models for performance comparison purposes.

[2] Decision tree is one kind of inductive learning algorithms that offers an efficient and practical method for generalizing classification rules from previous concrete cases that already solved by domain experts. It is considered attractive for many real-life applications, mostly due to its interpretability. Recently, many researchers have been reported to endow decision trees with incremental learning ability, which is able to address the learning task with a stream of training instances. However, there are few literatures discussing the algorithms with incremental learning ability regarding the new attributes

[3] The data mining comprises of analysis of large data from various perspectives and obtaining summary of useful information. The information can be transferred into knowledge regarding future trends and history. Data mining has a very important role in the information technology domain. Huge amounts of complex data is generated by health care sector today. These data includes details about diseases, patients, diagnosis methods, electronic patient's details hospitals resources etc. The data mining methods are very helpful in making medicinal decisions in disease curing. The vast data collected by healthcare industry are not mined and hence information is hidden. And as a result the decision making is not effective.

5. PRAPOSED SYSTEM

Patient Registration: If Patient is a new user he will enter his personal details and he will user Id and password through which he can login to the system.

Patient Login: If patient have already an account then he/she can log into the system

View Details: Patient and Doctor both can view their entered details. Doctor can also view patient's details and Patients can view only Doctor's little information.

Diseases Prediction: Patient will specify the symptoms caused due to his illness. System will ask certain question regarding his illness and system predict the disease based on the symptoms specified by the patient and system will also suggest doctors based on the disease.

6. SYSTEM DESIGN

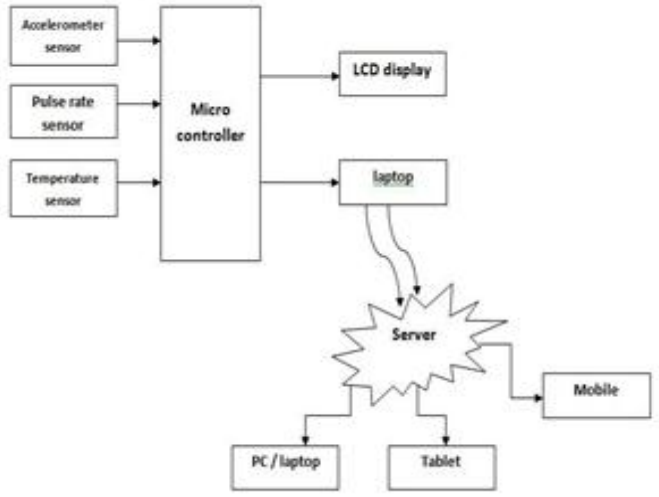

Fig. Architecture design

7. APPLICATIONS

- Can be used in hospitals.
- Health monitoring system.
- Patient self-assessment and monitoring.
- Remote patient monitoring in Healthcare
- Automatic Wireless Health Monitoring System

8. RESULTS

- **Home Page**

- **Symptoms Matching**

- **Login**

- **Result**

- **Sensor Data**

9. CONCLUSION

Tendency for data mining application in healthcare today is great, because healthcare sector is rich with information, and data mining is becoming necessity. Healthcare organizations produce and collect large volumes of information on daily basis. They need to be collected and stored in the organized forms, and their integration enables forming of hospital information system. Healthcare data mining provides countless possibilities for hidden pattern investigation from these data sets. These patterns can be used by physicians to determine diagnoses, prognoses and treatments for patients in healthcare organizations The propose system will be useful in urgent cases where patient is unable to reach doctor, for emergency cases that do not have doctors in an area, during late night emergencies and also for preliminary examination of patients.

10. REFFERENCES

Gitanjali J. et al. (2010). "APRIORI algorithm based medical data mining for frequent disease identification", IIJIT, Vol. 2, Issue 4 pp. 1-5, 2014, ISSN 2321-5976.

Michele Berlingeri (2007). "Mining Clinical Data with a Temporal Dimension: a Case Study", In proc. of IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2007), pp. 429 – 436, ISBN: 0- 7695-3031-1.

My Chau Tu AND Dongil Shin (2009). "A Comparative Study of Medical Data Classification Methods Based on Decision Tree and Bagging Algorithms" In IEEE proc of Eighth International Conference on Dependable, Autonomic and Secure Computing, pp. 183-187, ISBN: 978-1-4244-5421-1.

Ranjit Abraham, Jay B. Simha, Iyengar (2006). "A comparative analysis of discretization methods for Medical Data mining with Naïve Bayesian classifier", In proc of IEEE international conference on information Technology, pp. 235 - 236, 2006, ISBN: 0-7695-2635-7

SAM CHAO, FAI WONG (2009). "An Incremental Decision Tree Learninng Methodology Regarding Attributes in Medical Data Mining", In IEEE proceedings of the Eighth International Conference on Machine Learning and Cyb ernetics, Baoding, 12-15 July 2009, pp. 12-15, vol. 3, ISBN: 978-1-4244-3702-3

Yanwei Xing, Jie Wang and Zhihong Zhao (2007). "Combination data mining methods with new medical data to predicting outcome of Coronary Heart Disease" In IEEE proc of International Conference on Convergence Information Technology (ICCIT 2007), pp. 868 – 872, ISBN: 0-7695-3038-9.

Corresponding Author

Kalyani Shahaji Zodage*

E-Mail – zodagek@gmail.com