

# An Overview of Data Mining in Health Prediction System

B. Srinivas

Associate Professor

Department of Computer Science

Madda Walabu University, Ethiopia

*Abstract*— This system supports an end user and online consultation. 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. As health care 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.

**Key words:** Data Mining, Embedded System, Hill Climbing, SVM, JAVA, MySQL, Eclipse

## I. 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. Poor determination can prompt unfortunate outcomes which are unsatisfactory. The analysis of infections is a key and perplexing employment in prescription. The acknowledgment of coronary illness from various highlights or signs is a multi-layered issue that isn't free from false suspicions and is every now and again joined by imprudent impacts. The medicinal services industry gathers immense measure of social insurance information which shockingly are "not mined" to find concealed data for successful basic leadership. Number test ought to be performed to determination the coronary illness and it frames different examinations with the information accessible and it is changed from patient to persistent for medicinal services associations, information mining methods are connected to foresee the Heart Ailment and helpful in planning learning base for encourage treatment. It was discovered that more than one out of three grown-ups are discovered ailing due to

heart issues according to the reports of the World Health Organization (W.H.O.).

## II. GOALS AND OBJECTIVES

### A. Goals

To predict various diseases using Data mining.

### B. Objectives

To predict diseases using Data mining applications is challenging task but it drastically reduces the human efforts and increases the diagnostic accuracy.

Go Green and reduce paper work.

## III. MOTIVATION

Different data mining tool are used to predict the accuracy level in different health care problem, so we decided to implement system which will helpful to identify and predict the multiple diseases.

## IV. PROBLEM STATEMENT

To design a Health Prediction System for medical data classification and early disease prediction by using SVM and Hill-climbing algorithm. It might have happened so many times that you or someone need doctor's help immediately, but they are not available due to some reason. People cannot identify his symptoms and take medicines without consulting doctors. Some medicines are very much harmful to health. So user needs online consultation.

## V. 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 additionally utilized 10-overlap cross-approval techniques to quantify the fair-minded gauge of the three forecast models for execution examination purposes. The outcomes showed that the SVM is the best indicator with 92.1 % exactness on the holdout test simulated neural systems turned out to be the second with 91.0% precision and the choice trees models turned out to be the most exceedingly terrible of the three with 89.6% exactness. The similar investigation of various expectation models for survival of CHD patients alongside a 10-overlay cross-approval gave us a knowledge into the relative forecast capacity of various information.

- 2) Decision tree is viewed as appealing for some genuine applications, for the most part because of its interpretability. Anyway, there are a few literary works about the calculations with incremental learning capacity in regards to the new properties. In this paper, i+Learning (Intelligent, Incremental and Interactive Learning) hypothesis is proposed to supplement the customary incremental choice tree learning calculations by concerning new accessible ascribes notwithstanding the new approaching occurrences. The trial comes about uncover that i+Learning strategy offers the guarantee of settling on choice trees an all the more capable, adaptable, exact and profitable worldview, particularly in medicinal information mining group.
- 3) Data mining has an imperative part in the data innovation area. Immense measures of complex information is produced by social insurance part today. These information incorporates insights about ailments, patients, determination strategies, electronic patients points of interest clinics assets and so forth, the information mining strategies are extremely useful in settling on restorative choices in malady curing. The immense information gathered by human services industry are not mined and consequently data is covered up. What's more, subsequently the basic leadership isn't compelling. The information found can be utilized by the social insurance directors for upgrading the administration quality. In this paper, a strategy for distinguishing recurrence of maladies specifically land area for a given timeframe utilizing Apriori information mining method in light of affiliation rules is proposed.
- 4) Medical information mining has been a prominent information mining theme recently. Particularly, diagnosing of the coronary illness is one of the essential issue and numerous specialists examined to create wise restorative choice emotionally supportive networks to help the doctors. In this article, the utilization of choice tree C4.5 calculation is proposed, packing with choice tree C4.5 calculation and stowing with Naïve Bayes calculation to recognize the coronary illness of a patient and look at the adequacy, remedy rate among them. The information we think about is gathered from patients with coronary vein infection.
- 5) Clinical databases store a lot of data about patients and their medicinal conditions. Information mining methods can separate connections and examples holding in this abundance of information, and along these lines be useful in understanding the movement of ailments and the viability of the related treatments. A normal structure of medicinal information is a grouping of perceptions of clinical parameters taken at various time minutes. In this sort of settings, the fleeting measurement of information is an essential variable that ought to be considered in the mining procedure and returned as a feature of the extricated learning. In this way, the traditional and entrenched structure of consecutive example mining isn't sufficient, in light of the fact that it just spotlights on the sequentiality of occasions, without extricating the average time slipping by between two specific occasions. Time-clarified arrangements (IAS), is a novel mining worldview that takes care of this issue. As of late

characterized in our research center together with an effective calculation for removing them, IAS are successive examples where each progress between two occasions is clarified with an ordinary change time that is discovered regular in the information. In this paper we report a true medicinal contextual investigation, in which the IAS mining worldview is connected to clinical information with respect to an arrangement of patients in the follow-up of a liver transplantation. The point of the information examination is that of surveying the adequacy of the extracorporeal photograph pheresis (ECP) as a treatment to anticipate dismissal in strong organ transplantation. For every patient, an arrangement of biochemical factors is recorded at various time minutes after the transplantation. The IAS designs separated demonstrate the estimations of interleukins and other clinical parameters at particular dates, from which it is workable for the doctor to survey the viability of the ECP treatment. We trust that this contextual analysis does not just demonstrate the intriguing quality of separating IAS designs in this specific setting in any case, more yearningly, it proposes a general technique for clinical information mining, at whatever point the time measurement is an essential variable of the issue in investigation.

## VI. SYSTEM ARCHITECTURE

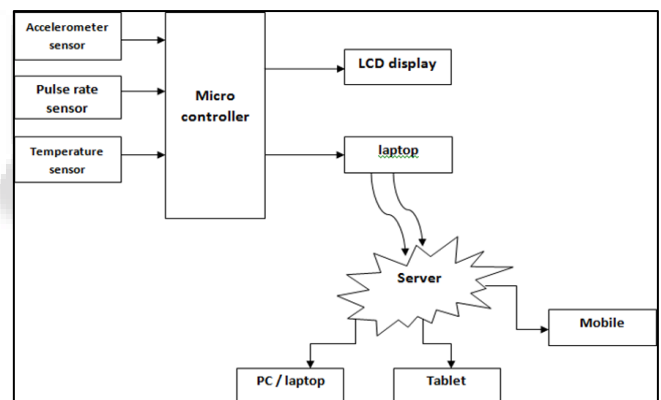


Fig. 1: System Architecture

## VII. CONCLUSION

We are going to design biomedical sensor based health prediction system which will help to older people, children, etc. Biomedical sensors will sense body parameters real time and server system will give predictions and recommendations to tackle with the conditions.

## REFERENCES

- [1] SAM CHAO, FAI WONG, "An Incremental Decision Tree Learning Methodology Regarding Attributes in Medical Data Mining", In IEEE proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009, pp. 12-15, vol. 3, ISBN:978-1-4244-3702-3
- [2] Gitanjali J. et al., "APRIORI algorithm based medical data mining for frequent disease identification", IJIT, Vol. 2, Issue 4 pp. 1-5, 2014, ISSN 2321-5976 2010.

- [3] My Chau Tu AND Dongil Shin, "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, 2009
- [4] Michele Berlinger, "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, 2007
- [5] Shoban Babu Sriramoju, "Analysis and Comparison of Anonymous Techniques for Privacy Preserving in Big Data" in "International Journal of Advanced Research in Computer and Communication Engineering", Vol 6, Issue 12, December 2017, DOI 10.17148/IJARCCCE.2017.61212 [ ISSN(online) : 2278-1021, ISSN(print) : 2319-5940 ]
- [6] Shoban Babu Sriramoju, "OPPORTUNITIES AND SECURITY IMPLICATIONS OF BIG DATA MINING" in "International Journal of Research in Science and Engineering", Vol 3, Issue 6, Nov-Dec 2017 [ ISSN : 2394-8299 ].
- [7] Shoban Babu Sriramoju, "Heat Diffusion Based Search for Experts on World Wide Web" in "International Journal of Science and Research", <https://www.ijsr.net/archive/v6i11/v6i11.php>, Volume 6, Issue 11, November 2017, 632 - 635, #ijsrnet
- [8] Dr. Shoban Babu Sriramoju, Prof. Mangesh Ingle, Prof. Ashish Mahalle "Trust and Iterative Filtering Approaches for Secure Data Collection in Wireless Sensor Networks" in "International Journal of Research in Science and Engineering" Vol 3, Issue 4, July-August 2017 [ ISSN : 2394-8299 ].
- [9] Dr. Shoban Babu, Prof. Mangesh Ingle, Prof. Ashish Mahalle, "HLA Based solution for Packet Loss Detection in Mobile Ad Hoc Networks" in "International Journal of Research in Science and Engineering" Vol 3, Issue 4, July-August 2017 [ ISSN : 2394-8299 ].
- [10] Shoban Babu Sriramoju, "A Framework for Keyword Based Query and Response System for Web Based Expert Search" in "International Journal of Science and Research" Index Copernicus Value(2015):78.96 [ ISSN : 2319-7064 ].
- [11] Sriramoju Ajay Babu, Dr. S. Shoban Babu, "Improving Quality of Content Based Image Retrieval with Graph Based Ranking" in "International Journal of Research and Applications" Vol 1, Issue 1, Jan-Mar 2014 [ ISSN : 2349-0020 ].
- [12] Dr. Shoban Babu Sriramoju, Ramesh Gadde, "A Ranking Model Framework for Multiple Vertical Search Domains" in "International Journal of Research and Applications" Vol 1, Issue 1, Jan-Mar 2014 [ ISSN : 2349-0020 ].
- [13] Mounika Reddy, Avula Deepak, Ekkati Kalyani Dharavath, Kranthi Gande, Shoban Sriramoju, "Risk-Aware Response Answer for Mitigating Painter Routing Attacks" in "International Journal of Information Technology and Management" Vol VI, Issue I, Feb 2014 [ ISSN : 2249-4510]
- [14] Mounika Doosetty, Keerthi Kodakandla, Ashok R, Shoban Babu Sriramoju, "Extensive Secure Cloud Storage System Supporting Privacy-Preserving Public Auditing" in "International Journal of Information Technology and Management" Vol VI, Issue I, Feb 2012 [ ISSN : 2249-4510]
- [15] Shoban Babu Sriramoju, "Multi View Point Measure for Achieving Highest Intra-Cluster Similarity" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 2, Issue 3, March 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798]
- [16] Shoban Babu Sriramoju, Madan Kumar Chandran, "UP-Growth Algorithms for Knowledge Discovery from Transactional Databases" in "International Journal of Advanced Research in Computer Science and Software Engineering", Vol 4, Issue 2, February 2014 [ ISSN : 2277 128X ]
- [17] Ajay Babu Sriramoju, Dr. S. Shoban Babu, "Analysis on Image Compression Using Bit-Plane Separation Method" in "International Journal of Information Technology and Management", Vol VII, Issue X, November 2014 [ ISSN : 2249-4510 ]
- [18] Shoban Babu Sriramoju, "Mining Big Sources Using Efficient Data Mining Algorithms" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 2, Issue 1, January 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]
- [19] Ajay Babu Sriramoju, Dr. S. Shoban Babu, "Study of Multiplexing Space and Focal Surfaces and Automultiscopic Displays for Image Processing" in "International Journal of Information Technology and Management" Vol V, Issue I, August 2013 [ ISSN : 2249-4510 ]
- [20] Dr. Shoban Babu Sriramoju, "A Review on Processing Big Data" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol-2, Issue-1, January 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]
- [21] Shoban Babu Sriramoju, Dr. Atul Kumar, "An Analysis around the study of Distributed Data Mining Method in the Grid Environment : Technique, Algorithms and Services" in "Journal of Advances in Science and Technology" Vol-IV, Issue No-VII, November 2012 [ ISSN : 2230-9659]
- [22] Shoban Babu Sriramoju, Dr. Atul Kumar, "An Analysis on Effective, Precise and Privacy Preserving Data Mining Association Rules with Partitioning on Distributed Databases" in "International Journal of Information Technology and management" Vol-III, Issue-I, August 2012 [ ISSN : 2249-4510 ]
- [23] Shoban Babu Sriramoju, Dr. Atul Kumar, "A Competent Strategy Regarding Relationship of Rule Mining on Distributed Database Algorithm" in "Journal of Advances in Science and Technology" Vol-II, Issue No-II, November 2011 [ ISSN : 2230-9659 ]
- [24] Shoban Babu Sriramoju, Dr. Atul Kumar, "Allocated Greater Order Organization of Rule Mining utilizing Information Produced Through Textual facts" in "International Journal of Information Technology and

- management” Vol-I, Issue-I, August 2011 [ ISSN : 2249-4510 ]
- [25] Ramesh Gadde, Namavaram Vijay, “A SURVEY ON EVOLUTION OF BIG DATA WITH HADOOP” in “International Journal of Research in Science and Engineering”, Vol-3, Issue-6, Nov-Dec 2017, 92-99 [ ISSN : 2394-8299 ].
- [26] Namavaram Vijay, S Ajay Babu, "Heat Exposure of Big Data Analytics in a Workflow Framework" in “International Journal of Science and Research”, Volume 6, Issue 11, November 2017, 1578 - 1585, #ijsrnet
- [27] Ajay Babu Sriramoju, Namavaram Vijay, Ramesh Gadde, “SKETCHING-BASED HIGH-PERFORMANCE BIG DATA PROCESSING ACCELERATOR” in “International Journal of Research in Science and Engineering”, Vol-3, Issue-6, Nov-Dec 2017, 92-99 [ ISSN : 2394-8299 ].
- [28] Guguloth Vijaya, A. Devaki, Dr. Shoban Babu Sriramoju, “A Framework for Solving Identity Disclosure Problem in Collaborative Data Publishing” in “International Journal of Research and Applications” (Apr-Jun © 2015 Transactions), Vol 2, Issue 6, 292-295
- [29] Siripuri Kiran, 'Decision Tree Analysis Tool with the Design Approach of Probability Density Function towards Uncertain Data Classification', International Journal of Scientific Research in Science and Technology(IJSRST), Print ISSN : 2395-6011, Online ISSN : 2395-602X, Volume 4 Issue 2, pp.829-831, January-February 2018. URL : <http://ijsrst.com/IJSRST1841198>
- [30] Ajmera Rajesh, Siripuri Kiran, " Anomaly Detection Using Data Mining Techniques in Social Networking" in “International Journal for Research in Applied Science and Engineering Technology”, Volume-6, Issue-II, February 2018, 1268-1272 [ ISSN : 2321-9653], [www.ijraset.com](http://www.ijraset.com)