

A system based on data mining techniques for predicting heart diseases

Prof. Rohini Hanchate¹, Jayesh Ramani², Mayur Lalchandani³, Ramendra Singh⁴

¹Department of Computer Engineering, DY Patil Institute of Engineering and Technology, Ambi, Pune, India

²Department of Computer Engineering, DY Patil Institute of Engineering and Technology, Ambi, Pune, India

³Department of Computer Engineering, DY Patil Institute of Engineering and Technology, Ambi, Pune, India

⁴Department of Computer Engineering, DY Patil Institute of Engineering and Technology, Ambi, Pune, India

Abstract- The prosperous application of information minings in extremely visible fields like e-business, promoting and retail has crystal rectifier to its application in different industries and sectors. Among This sectors simply discovering is health care. The Health care trade is usually "information rich", however sadly not all the info are well-mined that is needed for locating hidden patterns & effective; deciding .Discovery of hidden patterns and relationships typically goes unexploited . Advanced information minings modelling techniques will facilitate remedy this case. This analysis paper intends to use information minings Classification Modelling Techniques, namely, call Trees, Naïve mathematician and Neural Network, together with weighted association Apriori algorithmic rule, Random Forest and MAFIA algorithmic rule in Heart Diseases Prediction. Exploitation medical profiles like age, sex, pressure and glucose it will predict the chance of patients obtaining heart diseases.

Keywords- Random Forest, SVM, Diseases, etc.

I. INTRODUCTION

In today's present cardiovascular diseases is that the most fatal one. This diseases attacks an individual thus instantly that it hardly gets any time to induce treated with. Thus diagnosis patients properly on timely basis is that the most difficult task for the medical fraternity. A wrong diagnosing by the hospitals ends up in earn a foul name and loosing name. At a similar time treatments of the aforementioned diseases is sort of high and not reasonable by most of the patients notably in Republic of India. The aim of this paper is to develop a price effective treatments exploitation information minings technologies for facilitating data base call web. Most the hospitals use some hospitals management system to manage health care in patients. Sadly most of the systems seldom use the large clinical

information wherever important info is hidden. As This systems produce immense quantity of information in varied forms however this data is rarely visited and stay untapped. So, during this direction scores of efforts are needed to create intelligent selections. The diagnosing of this diseases exploitation completely different options or symptoms may be a complicated activity. During this paper exploitation varied Information minings technologies and effort is formed to help within the diagnosing of the diseases in question.

II. LITERATURE SURVEY

In the existing system, practical use of various collected data is time consuming. There are only few decision support systems available in medical industry whose functionalities are very limited. As mentioned earlier, medical decisions are made with doctor's intuition and not from the rich data from the medical database. Wrong treatments due to misdiagnosis causes serious threat in medical field. In order to solve these issues data minings solution was with help of medical databases was introduced.

Research on health care is literally the most vital part of science for humans, as none of us are immune to physical ailments. The existing literatures are diverse and roughly follow four lines of research: information extraction.

Information extraction from medical text is the basis for other higher order analytics, such as representation, classification, and clustering. The work in utilized SVM to recognize the medication related entities in hospitals discharge summaries, and classified this atomic elements into Preened categories, such as treatments and conditions. The greying of society, escalating costs of health care and burgeoning computer technologies are together driving more consumers to spend longer time online to explore health information. One survey

in shows that 59 percent of U.S. adults have explored the internet as a diagnostic tool in 2012.

Another survey in reports that the average U.S. consumer spends close to 52 hours annually online to wellness knowledge, while only visits the doctors three times per year in 2013. This have heightened the importance of online health resources as springboards to facilitate patient- doctor communication.

- 1) Read already published work in the same field.
- 2) Goggling on the topic of your research work.

III. Architecture

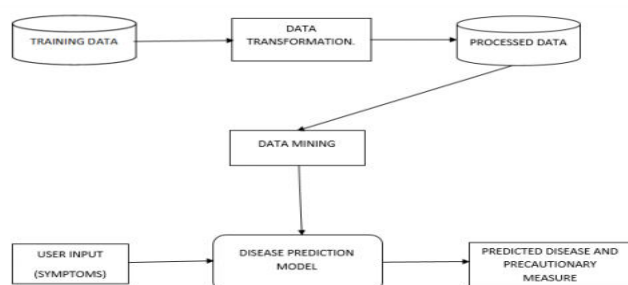


Fig: 3.1

The general diseases prediction system predicts likelihood of presence of a diseases gift in a very patient on the premise of their symptoms. It will conjointly suggest necessary precaution measures needed to treat the anticipated diseases. The system can at first be fed information from completely different sources i.e. patients, the info can then be pre-processed before any method is administered, this can be done thus on get clean information from the raw initial data, because the data would be clattery, or flawed. This information are going to be processed exploitation information minings algorithms, the system, are going to be trained thus on predict the diseases supported the input file given by the user.

IV. GET PEERREVIEWED

In today's modern world cardiovascular disease is the most lethal one. This disease attacks a person so instantly that it hardly gets any time to get treated with. So diagnosing patients correctly on timely basis is the most challenging task for the medical fraternity. A wrong diagnosis by the hospital leads to earn a bad name and loosing reputation. At the same time treatment of the said disease is quite high and not affordable by most of the patients particularly in India.

V. IMPROVEMENT AS PER REVIEWERCOMMENTS

Today, many hospitals manage healthcare data using healthcare information system; as the system contains huge amount of data, used to extract hidden information for making intelligent medical diagnosis. The main objective of this research is to build Intelligent Heart Disease Prediction System that gives diagnosis of heart disease using historical heart database. To develop this system, medical terms such as sex, blood pressure, and cholesterol like 13 input attributes are used. To get more appropriate results, two more attributes i.e. obesity and smoking are used, as these attributes are considered as important attributes for heart disease. The data mining classification techniques viz. Neural Networks, Decision Trees, and Random Tree are used

VI. CONCLUSION

The overall objective of our work is to predict more accurately the presence of heart diseases. In this topic two more input attributes obesity and smoking are used to get more accurate results. Data minings classification techniques were applied namely Random Forest
This technique will help us to achieve 97% accuracy as per algorithm.

APPENDIX

- NP-Complete and NP-Hard are: A decision problem is in P if there is a known polynomial-time algorithm to get that answer. The collection of all problems that can be solved in polynomial time is called P. That is, a decision question is in P if there exist an exponent k and an algorithm for the question that runs in time $O(n^k)$ where n is the length of the input.

REFERENCES

- [1] Frawley and G. Piatetsky -Shapiro, Knowledge Discovery in Databases: An Overview. Published by the AAAI Press/ The MIT Press, Menlo Park, C.A 1996.
- [2] Yanwei, X.; Wang, J.; Zhao, Z.; GAO, Y., "Combination data mining models with new medical data to predict outcome of coronary heart disease".

Proceedings International Conference on Convergence Information Technology 2007, pp. 868 – 872

[3] Sellappan Palaniappan, Rafiah Awang, “Intelligent Heart Disease Prediction System Using Data Mining Techniques” IEEE Conference, 2008, pp 108-115.

[4] A. S. Abdullah and R. R. Rajalaxmi, “A data mining model for predicting the coronary heart disease using random forest classifier,” in *Proc. Int. Conf. Recent Trends Comput. Methods, Common. Controls*

[5] J. P. Kelwade and S. S. Salankar, “Radial basis function neural network for prediction of cardiac arrhythmias based on heart rate time series,” in *Proc. IEEE 1st Int. Conf. Control, Meas. Instrum. (CMI)*, Jan. 2016, pp. 454–458.

[6] V. Krishnaiah, G. Narsimha, and N. Subhash, “Heart disease prediction system using data mining techniques and intelligent fuzzy approach: A review,” *Int. J. Comput. Appl.*, vol. 136, no. 2, pp. 43–51, 2016.

[7] P. S. Kumar, D. Anand, V. U. Kumar, D. Bhattacharyya, and T.-H. Kim, “A computational intelligence method for effective diagnosis of heart disease using genetic algorithm,” *Int. J. Bio-Sci. Bio-Technol.*, vol. 8, no. 2, pp. 363–372, 2016.

[8] M. J. Liberatore and R. L. Nydick, “the analytic hierarchy process in medical and healthcare decision making: A literature review,” *Eur. J. Oper. Res.*, vol. 189, no. 1, pp. 194–207, 2008.

[9] T. Mahboob, R. Irfan, and B. Ghaffar, “Evaluating ensemble prediction of coronary heart disease using receiver operating characteristics,” in *Proc. Internet Technol. Appl. (ITA)*, Sep. 2017, pp. 110–115.

[10] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, “Computational intelligence for heart disease diagnosis: A medical knowledge driven approach,” *Expert Syst. Appl.*, vol. 40, no. 1, pp. 96–104, 2013. doi: 10.1016/j.eswa.2012.07.032.