

Heart Disease Prediction System Using Hybrid Technique of Opinion Mining Algorithms

Samiksha Arvind Kale¹, Prof. Dr A .B .Gadicha²

^{1,2} Dept of Computer Science & Engineering

^{1,2} P. R. Pote (Patil) College of Engineering & Management, Amravati-444605 2020-2021

Abstract- Most countries confront high and growing rates of heart illnesses or disorder. No matter the way that, best in school pharmaceutical is making the enormous measure of knowledge reliably, little has been done to use this open data to know the challenges that face a viable illustration of echocardiography examination comes about. To style a perceptive model for heart illnesses acknowledgment using data processing strategies that are fit enhancing the constancy of heart infections conclusion. Learning Discovery in Database strategy including nine iterative and instinctive advances was grasped to think basic cases from a dataset containing a few of echocardiography examination reports of heart patients over the world. Thereafter, we divide this data into Training and Testing Data Sets and use SVM technique to get relatively higher prediction accuracy. The first goal of this research paper is to plan out a model that provides a highly accurate prediction of heart condition. As we've done a mixture of Genetic and Naïve Bayes Technique, the Investigation developed a Hybrid model of both these techniques and called it Hybrid Genetic Naïve Bayes Model for predicting high accuracy in results.

Keywords- Heart Disease, Data Mining, Classification, Linear SVM, GA.

I. INTRODUCTION

Because of a good accessibility of superlative measure of data and a requirement to vary over this accessible huge measure of data to helpful data requires the use of data mining strategies. Information Mining and KDD (learning disclosure within the database) have clothed to be prominent as lately. the recognition of data mining and KDD (information revelation in database) should not be an amazement since the measure of the knowledge increases that are accessible are extremely extensive to be analyzed physically and even the techniques for programmed information investigation in sight of established insights and machine adapting frequently threaten issues when preparing large, dynamic information increases comprising of complex items [11]. Information Mining is that the centerpiece of data Discovery Database (KDD). Numerous individuals regard data processing as the same word for KDD since it is a key piece of

KDD process. There are sure stages of data mining that you simply will get to get comfortable with, and these are exploration, pattern identification, and deployment. Information mining is an iterative procedure that commonly includes the accompanying stage [17].

HEART DISEASE

A key challenge confronting healthcare organizations (hospitals, medical centers) is that the facility of quality services at reasonable prices. Quality amenities suggest diagnosing patients accurately and regulating medications that are effective. Poor clinical choices can prompt deplorable results, which are during this manner unsatisfactory. Hospitals should limit the value of clinical tests. They will accomplish these outcomes by utilizing fitting PC based data and additionally choice emotionally supportive networks [4], [6]. The guts is that the essential piece of our body. Life is itself reliant on effective working of the guts. Within the event that task of the guts isn't legitimate, it'll influence the opposite body parts of human, for instance, cerebrum, kidney then on. Coronary illness may be a sickness that effects on the activity of the guts. There's variety of elements which builds danger of Heart ailment [13],[17]. A number of them are listed below:

- The case history of heart condition
- Smoking
- Cholesterol
- High vital sign
- Obesity
- Lack of workout.

II. LITERATURE SURVEY

Heart disease might be a term that assigns to an oversized range of medical conditions associated with the guts. These medical conditions describe the abnormal health conditions that directly influence the guts and everyone its elements. A heart disease might be a serious unhealthiest in today's time. The foremost studies done by victimization neural networks with fifteen attributes has outperformed over all different processing techniques.

KiranJyoti(October - 2012) [24] used the synthetic neural network (ANN), usually mentioned as a "neural network" (NN), may be a mathematical model or machine model supported biological neural network. In alternative words, it's an emulation of the biological neural system. Disorder prediction system has been developed using fifteen attributes. Earlier thirteen attributes were used for prediction, however, this analysis work incorporated a pair of additional attributes, i.e. fatness and smoking for economical designation of a heart disease. The info mining tool weka 3.6.6 is used for the experiment. Initially, missing values were known within the dataset which they were replaced with acceptable values using Replace Missing Values filter from 3.6.6. Further, numerous processing techniques are analysed on heart condition information. Confusion matrix is obtained for each classifier.

Mohammad Taha Khan, Dr. ShamimulQamar and Laurent F. Massin(2012) [22] presented prototype model for the carcinoma additionally to disorder prediction using processing techniques. Two decision tree algorithms C4.5 and also the C5.0 is employed on these datasets for prediction and performance of every algorithm are compared. Pruning algorithm is employed to reduce a mistake and avoiding the over fitting. Pruning a tree is that the action to interchange a whole sub tree by a leaf. The replacement takes place if the expected error rate within the sub tree is bigger than within the single leaf. During this study, they started by generating the whole (generally over fitted) classification tree and alter it using pruning simply once.

MA.JABBAR, Dr. PRITI CHANDRA (October 2011) [14] implemented CBARBSN Cluster based Association Rule Mining supported Sequence number during which they projected a replacement rule which mixes the construct of sequence numbers and cluster. The entire information base is split into partitions of equal size, every partition is mentioned as a cluster. Every cluster is taken under consideration one at a time by loading the first cluster into memory and hard frequent item sets. Then the second cluster is taken under consideration equally and hard frequent item sets. This approach reduces main memory demand since it considers solely a touch cluster at a time and it's scalable and efficient.

Ms. Ishtake S.H (April 2013) [32] was implemented a model heart condition prediction system is developed using 3 processing classification modelling techniques specifically, Decision. Trees, Naïve Bayes and Neural Network The system extract hidden information from a historical heart condition information. DMX query language and functions are accustomed build and access the models. Five mining goals

are defined supported business intelligence and knowledge exploration. The goals are evaluated against the trained models. All three models might answer complicated queries, each with its own strength with regard to simple model interpretation, access to elaborate data and accuracy.

Dr. K. Usha Rani (September 2011) [18] used Neural Network approach for the analysis of disorder. Neural Networks have emerged as an important tool for classification. The advantages of Neural Networks helps for efficient classification of given information. To increase the potency of the classification method parallel approach is additionally adopted within the training part. The experiment is conducted with disorder dataset by considering the one and multilayer neural network modes. Back propagation algorithm with momentum and variable learning rate is used to coach the networks. The experimental results verified that neural networks technique provides satisfactory results for the classification task.

ShantakumarB.Patil (February 2009) [10] presented an efficient approach for extracting vital patterns from the guts sickness data warehouses for the efficient prediction of attack. The pre-processed disorder information warehouse was clustered to extract information most relevant to attack mistreatment K-means clump algorithm. The frequent things are mined effectively mistreatment MAFIA algorithm. Supported the calculated important weightage, the frequent patterns having worth larger than a predefined threshold were chosen for the valuable prediction of attack. In our future work, we've planned to style and develop a cheap attack prediction system with the assistance of these chosen vital patterns mistreatment AI techniques.

Dilip Roy Chowdhury (2011) [16] represents the use of artificial neural networks in predicting neonatal illness identification. The projected technique involves training a Multi-Layer Perceptron with a BP learning algorithm to acknowledge a pattern for the designation and prediction of neonatal diseases. The rear propagation algorithm was accustomed train the ANN design and also an equivalent has been tested for the varied classes of neonatal illness. Concerning ninety-four cases of varied sign and symptoms, parameter is tested during this model. This study exhibits ANN based prediction of neonatal illness and improves the identification accuracy of seventy-five with higher stability.

Milan Kumari (June 2011) [17] proposed study covers information science classification techniques like ripper classifier, decision Tree, Artificial neural networks (ANNs), and Support Vector Machine (SVM) are analysed on heart condition dataset. Performance of these techniques is

compared through sensitivity, specificity, accuracy, error rate, True Positive Rate, and False Positive Rate. 10-fold cross validation methodology was used to measure the unbiased estimate of these prediction models.

As per our results error rates for ripper, decision Tree, ANN, and SVM are a pair of .756, 0.2755, 0.2248 and 0.1588 respectively. The accuracy of the ripper, decision Tree, ANN, and SVM are eighty-one .08%, 79.05%, 80.06% and 84.12% respectively. The analysis shows that out of these four classification models SVM predicts heart condition with least error rate and highest accuracy.

Niti Guru (January-June 2007) [4] proposed a system that uses a neural network for prediction of disorder, vital sign, and sugar. A set of seventy-eight records with thirteen attributes are used for training and testing. He used a supervised network for diagnosing of disorder and trained it using back propagation formula. On the thought of unknown information is entered by a doctor the system can notice that unknown information from training data and generate an inventory of possible illness from that patient will suffer.

III. CONCLUSION

In this study, our aim was to style a heart condition prediction system using various data processing techniques and to perform the analysis of the results obtained for all implemented techniques. So for the completion of the guts disease prediction model survey, we've evaluated the favored and effective heart condition prediction methods from the literature survey and eventually select the foremost effective algorithms of Naïve Bayes and Genetic Algorithm for his or her performance analysis on the guts disease prediction. The performances of the models were evaluated using the genetic algorithm and naïve Bayes. During this paper, total 14 attributes are used to get more accurate results. From results, it has been seen that projected model provides correct results as compare to existing models. This system could also be additionally expanded. Alternative processing techniques can also be used for predication e.g. clustering, statistic, and Association rules. The text mining could also be used to mine an immense quantity of unstructured information out there in healthcare business info.

REFERENCES

- [1] Nada Lavrac, Selected techniques for data mining in medicine, Elsevier, 1999.
- [2] Carlos Ordóñez, Edward Omiecinski, Mining Constrained Association Rules to Predict Heart Disease, IEEE.
- [3] Marcel A.J. van Gerven, Predicting carcinoid heart disease with the noisy-threshold classifier, Elsevier, 2007.
- [4] Niti Guru, Anil Dahiya, NavinRajpal, Decision Support System for Heart Disease Diagnosis Using Neural Network, Delhi Business Review, Vol. 8, No. 1, January-June 2007
- [5] HumarKahramanli, NovruzAllahverdi, Design of a hybrid system for diabetes and heart diseases, Elsevier, 2008.
- [6] LathaParthiban and R.Subramanian, Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm, International Journal of Biological and Medical Sciences, 2008.
- [7] TanawutTantimongcolwat, ThanakornNaenna, Identification of ischemic heart disease via machine learning analysis on Magnetocardiograms, Elsevier, 2008.
- [8] Resul Das, Ibrahim Turkoglu, AbdulkadirSengur, Effective diagnosis of heart disease through neural networks ensembles, Elsevier, 2009.
- [9] Resul Das, Ibrahim Turkoglu, AbdulkadirSengur, Diagnosis of valvular heart disease through neural networks ensembles, Elsevier, 2009.
- [10] ShantakumarB.Patil, Y.S.Kumaraswamy, "Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network"; European Journal of Scientific Research, ISSN 1450-216X Vol.31 No.4, 2009.
- [11] Rupa G. Mehta, Dipti P. Rana, Mukesh A. Zaveri, A Novel Fuzzy Based Classification for Data Mining using Fuzzy Discretization; World Congress on Computer Science and Information Engineering, 2009.
- [12] M. Anbarasi, E. Anupriya, N.ch.s.n.Iyengar, Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm, International Journal of Engineering Science and Technology, 2010.
- [13] Asha Rajkumar, G. Sophia Reena, Diagnosis of Heart Disease Using Datamining Algorithm; Global Journal of Computer Science and Technology, Page 38 Vol. 10 Issue 10 Ver. 1.0 September, 2010.
- [14] Ma.jabbar, Dr.priiti Chandra, B.L. Deekshatulu, cluster based association rule mining for heart attack prediction, Journal of Theoretical and Applied Information Technology, 2011.
- [15] JyotiSoni, SunitaSoni et al. "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction"; International Journal of Computer Applications (0975 – 8887) Volume 17– No.8, March 2011
- [16] Dilip Roy Chowdhury, Mridula Chatterjee & R. K. Samanta, An Artificial Neural Network Model for Neonatal Disease Diagnosis, International Journal of

- Artificial Intelligence and Expert Systems (IJAE), Volume (2): Issue (3), 2011.
- [17] Milan Kumari, Sunila Godara, Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction, IJCST Vol. 2, Issue 2, June 2011
- [18] Dr. K. Usha Rani, analysis of heart diseases dataset using neural network approach, International Journal of Data Mining & Knowledge Management Process, 2011.
- [19] Venkatadri. M, Dr. Lokanatha C. Reddy a review on data mining from past to the future. International Journal of Computer Applications, 2011.
- [20] Nan-Chen Hsieh & Lun-Ping Hung & Chun-Che Shih, Intelligent Postoperative Morbidity Prediction of Heart Disease Using Artificial Intelligence Techniques, J Med Syst, 2012.
- [21] Chaitrali S. Dangare, Dr. Mrs. Sulabha S. Apte, A data mining approach for prediction of heart disease using neural networks, international journal of computer engineering and technology, 2012.
- [22] Mohammad Taha Khan, Dr. Shamimul Qamar and Laurent F. Massin, A Prototype of Cancer/Heart Disease Prediction Model Using Data Mining, International Journal of Applied Engineering Research, 2012.
- [23] N. Aditya Sundar, P. Pushpa Latha, M. Rama Chandra, performance analysis of classification data mining techniques over heart diseases data base, international journal of engineering science and advanced technology, 2012.
- [24] Nidhi Bhatla Kiran Jyoti, An Analysis of Heart Disease Prediction using Different Data Mining Techniques, International Journal of Engineering Research & Technology (IJERT), 2012.
- [25] T. John Peter, K. Somasundaram, study and development of novel feature selection framework for Heart disease prediction, International Journal of Scientific and Research Publications, 2012.
- [26] Shadab Adam Pattekari and Asma Parveen, prediction system for heart disease using naïve bayes, International Journal of Advanced Computer and Mathematical Sciences, 2012.
- [27] Chaitrali S. Dangare, Sulabha S. Apte, Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques, International Journal of Computer Applications, 2012.
- [28] Oleg Yu. Atkov, Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters, Elsevier, 2012.
- [29] M. Akhil Jabbar, Dr. Priti Chandra, Dr. B. L. Deekshatulu, Heart Disease Prediction System using Associative Classification and Genetic Algorithm, International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies, 2012.
- [30] Jesmin Nahar, Tasadduq Imama, Kevin S. Tickle, Yi-Ping Phoebe Chen, Association rule mining to detect factors which contribute to heart disease in males and females, Elsevier, 2013.
- [31] Jesmin Nahar, Tasadduq Imam, Computational intelligence for heart disease diagnosis: A medical knowledge driven approach, Elsevier, 2013.
- [32] Ishtake S.H, Prof. Sanap S.A., "Intelligent Heart Disease Prediction System Using Data Mining Techniques", International J. of Healthcare & Biomedical Research, 2013.
- [33] Indira S. Fal Dessai, Intelligent Heart Disease Prediction System Using Probabilistic Neural Network, International Journal on Advanced Computer Theory and Engineering, 2013.
- [34] Rashedur M. Rahman, Farhana Afroz, Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis, Journal of Software Engineering and Applications, 2013.
- [35] Adebayo Peter Idowu, Data Mining Techniques for Predicting Immunize-able Diseases: Nigeria as a Case Study, International Journal of Applied Information Systems, 2013.
- [36] Abhishektaneja, Heart Disease Prediction System Using Data Mining Techniques, Oriental Scientific Publishing Co., India, 2013.
- [37] Monali Dey, Siddharth Swarup Rautaray, Study and Analysis of Data mining Algorithms for Healthcare Decision Support System, International Journal of Computer Science and Information Technologies, 2014.