ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

Rational Heart Disease Prediction System Using Dual Sentimental Analysis

[1] D.Roselin, [2] M.Priya, [3] Mr.L.Jerart Julus.
[1]UG Student, Dept .of IT, National Engineering College, Kovilpatti.
[2]UG Student, Dept. of IT, National Engineering College, Kovilpatti.
[3]Asst.Professor (Senior Grade), Dept. of IT, National Engineering College, Kovilpatti.

Abstract: Heart disease is a single foremost source of death in developed countries and one of the main contributors to disease burden in developing countries. The prediction of heart disease survivability has been a challenging research problem for many researchers. Hence there is a need to develop a decision support system for predicting heart disease of a user. Rational Heart Disease Prediction System using Dual Sentimental Analysis (RHDPS-DSA) is an end user support and online consultation project. The main motivation for using Dual Sentimental Analysis is to address the polarity shift problem in sentiment classification. Here we propose a system that allows users to get instant guidance on their health issues through online. It enables momentous knowledge, e.g. patterns, relationships among medical factors related to heart disease, to be established. Data Mining techniques, namely Duple Prediction and Duple Training Algorithm that analyse on Heart disease database can help the users in detecting the heart disease status based on their clinical data. The system is fed with various heart disease related symptoms which allows user to share their symptoms and issues. It then processes users symptoms to predict either the user is at risk or normal. When the user at risk, then he/she will be directed to those specialists.

Keywords: Dual sentimental analysis, Duple prediction, Duple training

I. INTRODUCTION

Rational Heart Disease Prediction System Using Data mining. Data mining is a new powerful technology which is of high interest in computer world. It is a sub field of computer science that uses already existing data in different databases to transform it into new researches and results. It makes use of Artificial Intelligence, machine learning and database management to extract new patterns from large data sets and the knowledge associated with these patterns. The actual task is to extract data by automatic or semi-automatic means. The different parameters included in data mining includes clustering, forecasting, path analysis and predictive analysis.

Data mining has many applications in the fields of telecommunication industry, financial data analysis biological data analysis and much more. With the growing research in the field of health informatics a lot of data is being produced. The analysis of such a large amount of data is very hard and requires excessive knowledge. E-healthcare applies data mining and telecommunication techniques for health diagnosis. E-health was primarily used for patient data analysis and disease diagnosis at various levels. Since the knowledge of its vast use more and more attention has been paid to this field from clinical data analysis to record management of users.

The system is first taught with various symptoms related to herat disease associated with each system. User gives the knowledge of symptoms he/she is dealing with. The machine processes these symptoms and provides the results. With advancement of technology more and more smart systems are being designed with better data mining technologies to give the most accurate results that could be associated with the disease. It then processes users symptoms to predict either the user is at risk or normal. The system have information about the doctors phone number, address along with feedback and administrator control panel for system processes. The patients different parameters such as voice, images, movements and daily activities are considered as parameter and systems are generated using quantitative analysis and pattern recognition.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

II. RELATED WORKS

Shusaku Tsumoto [1] projected that data mining methods will find interesting patterns from databases as use again of stored data and be important for clinical research and perform, because human beings cannot arrangement with such a huge amount of data. In this paper alert on the characteristics of clinical data and talk about how data miners deal with clinical data.

Carlos Ordonez [2] did study on prediction of heart disease with the help of Association rules. They used a simple mapping algorithm. This algorithm constantly treats attributes as numerical or categorical. This is used to convert medical records to a transaction format. An improved algorithm is used to mine the constrained association rules. A mapping table is prepared and attribute values are mapped to items. The decision tree is used for mining data because they automatically Split numerical values [2]. The split point chosen by the Decision tree are of little use only. Clustering is used to get a global understanding of data.

Sellappan Palaniappan, et.al. [3] developed Intelligent Heart Disease Prediction System (IHDPS) using data mining techniques, i.e. Decision Trees, Naïve Bayes and Neural Network. Each method possesses its own power to gain suitable results. The hidden patterns and relationships among them have been used to construct this system. The IHDPS is user-friendly, web-based, scalable, reliable and expandable.

C.Ordonez [4] uses association rules a technique in data mining to get better heart disease prediction result. The author have worked on the limitation of association rule which is nothing but mining the entire data set without validation on an independent sample. The modified algorithm with search constraints was introduced to trim down the number of association rules and validated using train and test approach. They have studied two complementary tasks: predicting the absence and predicting the existence of heart disease.

P.Chandra, M.Jabbar [5] produced class association rules using feature subset selection to detect a heart disease. Association rule determines relations between attributes values and classification to predict the class in the patient dataset [5]. Feature selection measures Like genetic search determines attributes which helps in predicting heart diseases.

Usha Rani [6] have proposed a system for predicting heart disease with the help of artificial neural network which is a combination of feed forward and back propagation algorithm. The experiment is carried out by considering single and multilayered neural network models. Parallelism is implemented to speed up the learning process at each neuron in all hidden and output layer.

R.Setthukkarasi,[7] have developed a novel neuro fuzzy technique to diagnose the fact of the disease from the set of the patient report .A generalized database is constructed for decision making from the reduced attributes set which is output of genetic algorithm. A four layered fuzzy neural network for efficient modeling and reasoning with temporal dependencies under uncertainty is used.

Chaitrali Dangare [8] has implemented system to predict heart disease three data mining classification techniques were applied that is Decision trees, Naive Bayes & Neural Networks. From results it has been seen that Neural Networks better than Decision trees & Naive Bayes.

M.Akhil jabbar, B.LDeekshatulu, Priti Chandra [9] propose a new algorithm which combines KNN with Genetic Algorithm for effective classification. To provide optimal solution genetic algorithms perform global search on complex large and multimodal Dataset . From the results it is also observed that hybridizing GA with KNN Performs well and give great accuracy.

Shadab Adam Pattekari and Asma Parveen [10] developed a Decision Support in Heart Disease Prediction System using Naive Bayesian Classification technique. The system discovers the hidden knowledge from a past heart disease database. This is the most effective model to predict patients with heart disease. This model could respond to complex queries, each with its own strength with respect to ease of model interpretation, access to detailed information and accuracy.

III. PROBLEM DEFINITION

Many hospital information systems are designed to support decision support systems which is utilized only by doctors and not by patients. Users themselves are not able to analyze about the severity of heart disease with the symptoms they experience.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

IV. EXISTING SYSTEM

In existing approach they have used Naïve Bayes Algorithm which is simple and quite efficient but the time taken to derive the expected result is large. Moreover, the accuracy is not upto the level. It can be visualized from the graph given below Fig 1.

V. PROPOSED METHODOLOGY

To overcome this problem we propose a novel approach of Dual Sentimental Analysis. Here we use two main algorithms i.e, Duple Prediction and Duple Training Algorithm along with decision tree which predicts the results with high accuracy and low running time. It can be visualized from the graph given below Fig 2.

DUAL SENTIMENTAL ANALLYSIS

In Dual Sentimenal Analysis (DSA) approach, we propose a duple training (DT) algorithm and a duple prediction (DP) algorithm respectively, to make use of positive and negative samples in pairs for training a statistical classifier and make predictions. In DT, the classifier is learnt by maximizing a combination of likelihoods of the positive and negative training data set. In DP, predictions are made by considering positive and negative samples.

DUPLE TRAINING

In Duple Training stage, the collected data sets will be trained to construct test sets which is going to be displayed in the survey window. The purpose of constructing test set is to train the data sets which may either be positive or negative.

DUPLE PREDICTION

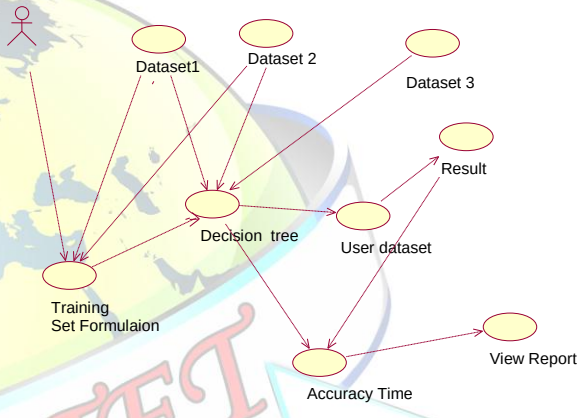
In Duple Prediction, the trained sets obtained from Duple Training is brought to decision tree in which the data are processed to predict the accurate results. Duple Prediction algorithm plays the major role in the whole process.

VI. SYSTEM ARCHITECTURE

System design is the process or art of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. One could see it as the application of systems theory to product development. There is some overlap and synergy with the disciplines of systems analysis, systems architecture and engineering. Object methods are becoming the most widely used methods for computer system design. The UML has become the standard language used in Object-oriented analysis and design. It is widely used for modeling software systems and is increasingly used for hi designing non-software systems and organizations.

USE CASE DIAGRAM

A use case diagram is a graphic depiction of the interactions among the elements of a system. A use case is a methodology used in system analysis to identify, clarify, and organize system requirements. The actors, usually individuals involved with the system defined according to their roles

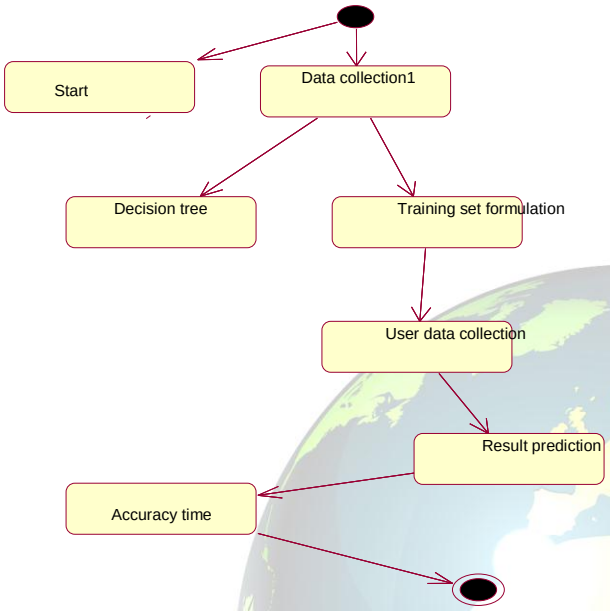


ACTIVITY DIAGRAM

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice iteration and concurrency. In the Unified Modeling Language, activity diagrams are intended to model both computational and organizational processes (i.e.workflows)

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017



VII. EXPERIMENT RESULT

The following tabular column represents the results that we have obtained from the alogithms,

		Accuracy	Time
1	Result1	6.6	2.4
2	Result2	7	4.4
3	Result3	6.2	3.8
4	Result4	6.8	2.8

Tab 7.1 Values obtained from Naïve Byes Algorithm

		Accuracy	Time
1	Result1	9.5	0.4
2	Result2	8.8	0.433
3	Result3	9	0.8
4	Result4	9.2	2.8

Tab 7.1 Values obtained from Dual Sentimental Analysis

From the values of table, the graphical representation is as follows

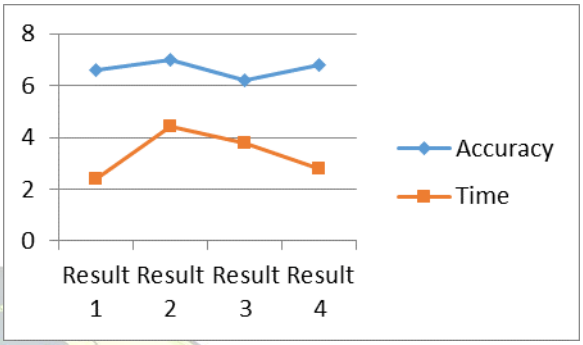


Fig 1.Comparison of data processing values using Naïve Byes Algorithm

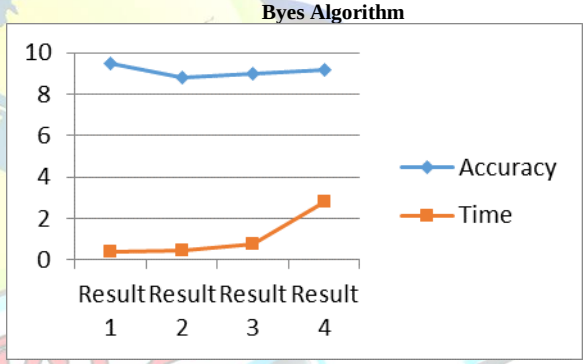


Fig 2.Comparison of data processing values using Dual Sentimental Analysis

VIII. CONCLUSION

Decision Support in Heart Disease Prediction System is developed using Dual Sentimental Analysis technique. The system extracts hidden knowledge from a historical heart disease database.Dual Sentiment Analysis is more effective to predict patients with heart disease. This model could answer simple queries, each with its own strength with ease of model interpretation and an easy access to detailed information and accuracy. The system is expandable in the sense that more number of records or attributes can be incorporated and new significant rules can be generated using underlying Data Mining technique. Presently the system has been using 15 attributes of medical diagnosis. It can also incorporate other data mining techniques and additional attributes for prediction.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

REFERENCES

[1]. D. Ratnam, P. HimaBindu, V. MallikSai, S. P. Rama Devi and P. RaghavendraRao, "Computer-Based Clinical Decision Support System for Prediction of Heart Diseases Using Naïve BayesAlgorithm", International Journal of Computer Science and Information Technologies, vol. 5, no. 2 (2014) pp.2384- 2388.

[2]. T. Santhanam and E. P. Ephzibah, "Heart Disease Prediction Using Hybrid Genetic Fuzzy Model", Indian Journal of Science and Technology, vol.8, no. 9, (2015), pp.797–803.

[3]. G. Purusothaman and P. Krishnakumari, "A Survey of Data Mining Techniques on Risk Prediction: Heart Disease",Indian Journal of Science and Technology, vol. 8, no. 12, (2015).

[4]. K. Srinivas, G. RaghavendraRao and A. Govardhan, "Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal Mining Regions Using Data Mining Techniques", Proceedings of 5th International Conference on Computer Science & Education, China, (2010), pp. 24–27.

[5]. K. Thenmozhi and P. Deepika, "Heart Disease Prediction Using Classification with Different Decision TreeTechniques", International Journal of Engineering Research and General Science, vol. 2, no. 6, (2014), pp.6-11.

[6]. S. Chaitrali, Dangare and S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", International Journal of Computer Applications, vol. 47, no. 10, (2012), pp. 0975 –888

[7]. J Peter and K. Somasundaram, "An Empirical Study on Prediction of Heart Disease Using Classification DataMining Techniques", Proceedings of IEEE International Conference on Advances In Engineering, Science And Management (ICAESM), (2012), pp. 514-518.

[8]. H. D. Masethe and M. A. Masethe, "Prediction of Heart Disease using Classification Algorithms", Proceedings of the World Congress on Engineering and Computer Science(WCECS), San Francisco, USA. (2014).

[9]. S. J. Gnanasoundhari, G. Visalatchi and M. Balamurugan, "A Survey on Heart Disease PredictionSystem Using Data Mining Techniques", International Journal of Computer Science and Mobile Application, vol. 2, no. 2 (2014), pp. 72-77

[10]. K. Lokanayaki and A. Malathi, "Exploring on Various Prediction Model in Data Mining Techniques for Disease Diagnosis", International Journal of Computer Applications, vol. 77, no. 5, (2013), pp. 0975 – 8887.