IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sep 2018 E-ISSN: 2347-2693

Diagnosis of Heart Disease using Cultural Algorithm with Neural Network

S. Shrivastava[1*], S. Shrivastava [2]

[1] Dept. of Computer Science & Engineering, IES-IPS Academy, RGPV, Indore, India
[2] Dept. of Computer Science & Engineering, IES-IPS Academy, RGPV, Indore, India

[*]*Corresponding Author: shiva.shri16@gmail.com, Tel.: +91-9406669949*

Available online at: www.ijcseonline.org

Accepted: 27/Sept/2018, Published: 30/Sept/2018

***Abstract*—**Heart disease detection is considered as the most complicated task in the world of medical sciences. There arises a necessity to progress the work and to develop a decision support system to find out a heart disease of a patient. To achieve a correct and cost effective treatment computer-based and support systems can be developed to make good decision. These information which exists contains the huge amounts of data which are organized in the form of images, text, charts and numbers. Hence, there is necessity which is motivating to create an excellent and useful project which will help physicians and cardiologists to predict the heart disease before it damage the health. It can be able to solve complicated enquiries for detecting heart disease in a patient and as a result it will assist medical practitioners to make more accurate and precise clinical decisions which traditional decision support systems were not able to decide. By providing effective and respective solution, it will surely help to reduce costs of treatment.

***Keywords*—** Back Propagation, Cardiovascular disease Confusion Matrix, Genetic Algorithm, Neural Network.

I. INTRODUCTION

Cardiovascular diseases are a major public health problem. Such a claim corresponds to a current reality; however, it is anticipated that cardiovascular pathologies will continue to be an important health problem in the future, mainly due to the increase in their prevalence in developing countries and population aging [1].

More specifically, in high-income countries, heart disease and cerebrovascular disease represent the first and second leading cause of death among men and women [2]. In Europe alone, in 2008, more than 4.3 million deaths were attributed to cardiovascular diseases, which represented 48% of the total number of deaths on the continent [3] However, despite the serious problem that cardiovascular diseases currently pose, knowledge of their main modifiable risk factors makes prevention possible. The three most important modifiable cardiovascular risk factors are smoking, hypertension and hypercholesterolemia, followed by diabetes, overweight / obesity, sedentary lifestyle [4]. The objective of this study is to analyze the effect that the preventive programs of the cardiovascular diseases applied in primary care have on the avoidable hospitalization specific of these diseases.

This section which is the I section one gives a detailed explanation about heart disease, classification, diagnosis and treatment of heart disease, datamining and finally data mining in health care is explained in this overview, motivation and objective of the research work. Section II reviews about the previous work on data mining approaches in prediction of heart disease and the problem which is been found is explained in this research section. Section III shows the shows the proposed methods which are used in the work. Section IV mentions the details how the algorithm is been implemented and the worked with the data and produced the results in the Confusion Matrix for the cultural Algorithm. Section V shows the result of the work which is been proposed. This chapter summarizes the contributions of the dissertation.

II. CARDIOVASCULAR PREVENTION

Cardiovascular disease (CVD) is defined as a group of conditions that affect the heart and blood vessels. Despite the decreasing trend in developed countries in the last three decades, CVD as a whole is the main cause of mortality and hospitalization in the Spanish population [5]. Primary cardiovascular prevention is based on the identification and control of cardiovascular risk factors in the healthy population, thus attempting to prevent the onset of the disease [6].

In the coming years, there is an increase in the number of hospitalizations for these diseases, as a result of the technological development that will allow patients to offer new diagnostic and therapeutic instruments, the greater survival of patients with these health problems and the aging

of the Spanish population [7].In 2009, hospitalizations for CVD recorded the highest number of hospital discharges (12.8 per 100 high). CVD were the second cause of hospitalization among women (10.6%), only surpassed by pregnancy, delivery and puerperium. In men, it was the first cause of hospitalization, causing 15.3% of the total number of care incidences in that year [8].

III. PROPOSED METHODOLOGY

Heart disease is often detected when a patient reaches advanced stage of the disease. Hence the risk factors are analyzed from various sources [9]-[10]. The dataset was composed of 12 important risk factors which were sex, age, family history blood pressure, Smoking Habit, alcohol consumption, physical inactivity, diabetes, blood cholesterol, poor diet, obesity. The system indicated whether the patient had risk of heart disease or not. The data for 50 people was collected from surveys done by the American Heart Association [10]. Most of the heart disease patients had many similarities in the risk factors [11]. Table1 shows the identified important risk factors and the corresponding values and their encoded values in brackets, which were used as input to the system.

Table 1. Risk factors values and their encodings [12]

S. No.	Risk Factors	Values
1	Sex	Male (1), Female (0)
2	Age (years)	20-34 (-2), 35-50 (-1), 51-60 (0), 61-79 (1) , >79 (2)
3	Blood Cholesterol	Below 200 mg/dL - Low (-1) 200-239 mg/dL - Normal (0) 240 mg/dL and above - High (1)
4	Blood Pressure	Below 120 mm Hg- Low (-1) 120 to 139 mm Hg- Normal (0) Above 139 mm Hg- High (-1)
5	Hereditary	Family Member diagnosed with HD - Yes (1) Otherwise –No (0)
6	Smoking	Yes (1) or No (0)
7	Alcohol Intake	Yes (1) or No (0)
8	Physical Activity	Low (-1), Normal (0) or High (-1)
9	Diabetes	Yes (1) or No (0)
10	Diet	Poor (-1), Normal (0) or Good (1)
11	Obesity	Yes (1) or No (0)
12	Stress	Yes (1) or No (0)
Output	Heart Disease	Yes (1) or No (0)

Data analysis has been carried out in order to transform data into useful form, for this the values were encoded mostly between a range [-1, 1]. Data analysis also removed the inconsistency and anomalies in the data. In this paper, cultural genetically optimized Neural Network approach is used to determine optimum number of clusters in analyzed data. These methods are described below.

A. Diagnosis using Cultural algorithm

Cultural algorithm corresponds to modeling inspired by the evolution of human culture [13]. Thus, just as we speak of biological evolution as the result of a selection based on genetic variability, we can speak of a cultural evolution resulting from a selection exercising on the variability Cultural development. The following figure presents the basic CA framework.

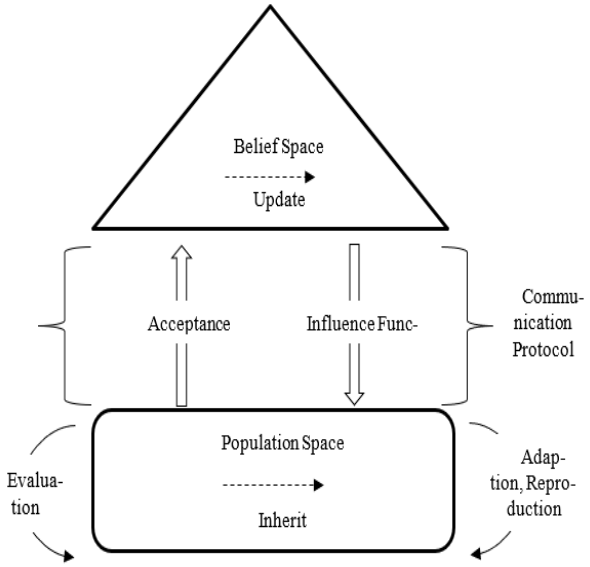

Figure 1.1 CA framework [14]

As Figure 1 shows, the population space and the belief space can evolve respectively. In the process of the CA evolution, the population space is initialized with candidate solution agents at random, meanwhile, the initial knowledge sources in the belief space are built. At first the two spaces evolve independently. Then the selected agents from the population space are used to update the belief space. After the knowledge sources being updated, the belief space will reversely guide the evolution of the population space. repeat till a termination condition has been reached.

Table 2. Patient's case study data in encoded form

No	Sex	Age	Blood Cholesterol	Blood Pressure	Hereditary	Smoking	Alcohol Intake	Physical Activity	Diabetes	Diet	Obesity	Stress	Heart Disease
1	F	35	High	Normal	No	No	Yes	Low	Yes	Poor	Yes	Yes	Yes
2	M	70	Low	Low	No	No	Yes	High	Yes	Normal	No	No	No
3	F	60	High	High	No	No	No	Normal	Yes	Poor	Yes	Yes	Yes
4	F	36	Low	Normal	No	No	No	Normal	No	Good	No	No	No
5	M	30	Low	Normal	No	No	Yes	High	No	Normal	No	No	No
6	F	39	Low	Normal	Yes	No	Yes	High	Yes	Normal	No	Yes	No
7	F	41	High	Normal	No	No	No	Low	No	Poor	Yes	No	No
8	M	70	High	Normal	No	No	Yes	Low	No	Poor	Yes	No	Yes
9	M	65	Normal	High	Yes	Yes	Yes	Normal	Yes	Poor	Yes	No	Yes
10	M	30	Normal	High	No	Yes	No	Normal	No	Good	No	Yes	No
11	F	31	Low	Normal	No	No	No	High	No	Normal	No	No	No
12	F	29	Low	Normal	No	No	Yes	High	No	Good	No	No	No
13	M	30	Low	Normal	No	No	Yes	Normal	No	Normal	No	No	No
14	F	45	Normal	High	Yes	Yes	No	Normal	Yes	Normal	Yes	Yes	No
15	M	25	High	Normal	Yes	Yes	Yes	Low	Yes	Normal	No	No	Yes
16	F	37	Normal	Normal	No	No	No	Normal	Yes	Poor	No	Yes	No
17	F	37	Normal	High	No	Yes	Yes	High	No	Poor	No	Yes	No
18	M	53	High	Low	No	Yes	No	Normal	Yes	Normal	No	Yes	No
19	M	57	High	Normal	No	Yes	No	Low	No	Poor	Yes	Yes	Yes
20	M	52	High	Low	No	No	No	Normal	Yes	Poor	Yes	No	No
21	M	48	Normal	Normal	Yes	Yes	Yes	Normal	No	Normal	No	No	Yes
22	M	62	High	High	No	Yes	Yes	Normal	Yes	Normal	No	No	Yes
23	M	56	Normal	High	No	Yes	Yes	Low	No	Poor	Yes	Yes	Yes
24	F	27	Low	Normal	No	No	No	High	No	Good	No	No	No

.

The data for risk factors related to heart diseases collected from 50 people is provided in Table 2

No	Sex	Age	Blood Cholesterol	Blood Pressure	Hereditary	Smoking	Alcohol Intake	Physical Activity	Diabetes	Diet	Obesity	Stress	Heart Disease
25	M	33	Normal	Normal	No	No	No	Normal	Yes	Good	No	No	No
26	F	33	Normal	Normal	No	No	Yes	Low	Yes	Poor	No	Yes	No
27	M	37	High	Normal	No	No	Yes	Normal	No	Normal	No	Yes	No
28	M	43	Normal	High	No	No	No	Normal	Yes	Poor	Yes	Yes	Yes
29	M	46	Low	Normal	No	No	No	Normal	Yes	Poor	Yes	Yes	No
30	F	36	Low	Normal	No	No	No	Normal	No	Normal	No	No	No
31	F	29	Low	Normal	No	No	No	Normal	No	Good	No	No	No
32	F	47	Normal	Normal	No	No	Yes	High	Yes	Normal	No	Yes	No
33	M	58	High	High	No	Yes	Yes	Normal	Yes	Normal	No	Yes	Yes
34	M	44	High	Normal	Yes	Yes	Yes	Normal	No	Normal	Yes	Yes	Yes
35	F	36	Normal	High	No	No	No	Normal	No	Good	Yes	No	Yes
36	M	42	Low	Normal	Yes	No	Yes	Low	No	Poor	No	Yes	No
37	F	25	Low	Normal	No	No	No	High	No	Poor	No	No	No
38	F	28	Low	Normal	No	No	Yes	High	No	Normal	No	No	No
39	F	26	Low	Normal	Yes	No	No	Normal	No	Normal	Yes	No	Yes
40	M	28	Low	Normal	No	No	No	Normal	No	Poor	No	No	No
41	F	45	High	Normal	No	No	Yes	Low	Yes	Poor	Yes	Yes	Yes
42	M	63	Low	Low	No	No	Yes	High	Yes	Good	No	No	No
43	F	55	High	High	No	No	No	Normal	Yes	Normal	Yes	Yes	Yes
44	F	44	Low	Normal	No	No	No	Normal	No	Normal	No	No	No
45	M	35	Low	Normal	No	No	Yes	High	No	Normal	No	No	No
46	F	42	Normal	Normal	No	No	Yes	High	Yes	Good	No	No	No
47	F	43	Normal	Normal	No	No	No	Low	No	Poor	Yes	No	No
48	M	65	Normal	Normal	No	No	Yes	Low	No	Normal	Yes	Yes	Yes
49	M	74	Normal	High	No	Yes	Yes	Normal	Yes	Normal	Yes	Yes	Yes
50	M	36	Normal	High	No	Yes	No	Normal	No	Poor	No	No	No

IV. SIMULATION AND RESULTS

The performance of proposed technique has been studied by means of MATLAB simulation.

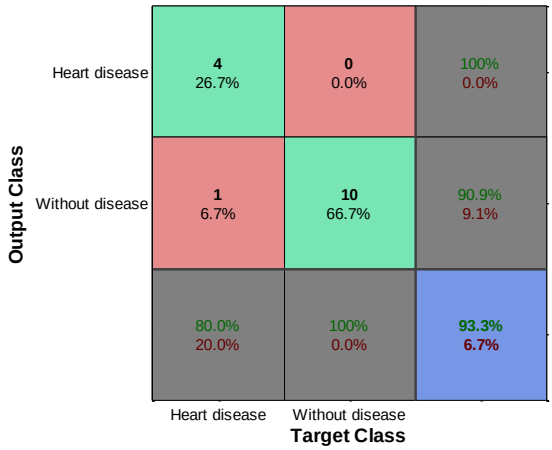

Figure 2. Confusion matrix for Cultural Algorithm

The row and column are the labels for detection and no detection of heart disease from database. There are 2 sets of classes and each class having different set of detection. Total 15 samples of patients are taken out of which 4 patients are classified correctly and none of the samples were misclassified but in normal category 10 samples are classified correctly out of 15 samples and 1 sample is misclassified. The confusion plot indicates the accuracy i.e. 93.3% for this approach.

- Population Size=30
- Acceptance ratio=0.40
- Alpha=0.3
- Maximum iteration to solution=50

According to the above Confusion Matrix, we have,

M=4, N=0, O=1, P=10

And we know that, Total=M+N+O+P,

So, substituting the values on the formulas we get,

(4/4+0)*100=100%

(0/4+0)*100=0%

(1/1+10)*100=9.1%

(10/1+10)*100=90.9%

(4/4+1)*100=80.0%

(1/4+1)*100=20.0%

(0/0+10)*100=0%

(10/0+10)*100=100%

Now, on substituting the values to obtain accuracy,

Accuracy= P+M / Total

Accuracy= [(4+10)/4+0+1+10]*100

Accuracy=93.33%

Hence, it's provided the same results as obtained manually.

V. CONCLUSION AND FUTURE WORK

In this paper, Cultural Algorithms are a class of models derived from the cultural evolution process. The weights which are used for back-propagation can be optimized first and then given as input to our network to give much better results and it was found that it is effective to predict the risk of heart disease when the person provide the required attributes value. On observing the confusion matrix, it was found that the CA based approach with the accuracy of 93.3%.

REFERENCES

[1] Jones, D.S. and Greene, J.A., 2013. The decline and rise of coronary heart disease: understanding public health catastrophism. American journal of public health, 103(7), pp.1207-1218.

[2] Mathers, C.D. and Loncar, D., 2006. Projections of global mortality and burden of disease from 2002 to 2030. Plos med, 3(11), p.e442.

[3] World Health Organization, 2011. Global status report on noncommunicable diseases 2010. Geneva: World Health Organization.

[4] Buttar, H.S., Li, T. and Ravi, N., 2005. Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation. Experimental & Clinical Cardiology, 10(4), p.229.

[5] Jenkins, C.D., 1988. Epidemiology of cardiovascular diseases. Journal of consulting and clinical Psychology, 56(3), p.324.

[6] World Health Organization, 2007. Prevention of cardiovascular disease. World Health Organization.

[7] Cowie, M.R., Anker, S.D., Cleland, J.G., Felker, G.M., Filippatos, G., Jaarsma, T., Jourdain, P., Knight, E., Massie, B., Ponikowski, P. and López- Sendón, J., 2014. Improving care for patients with acute heart failure: before, during and after hospitalization. ESC Heart Failure, 1(2), pp.110-145.

[8] Ajay, V.S. and Prabhakaran, D., 2010. Coronary heart disease in Indians: Implications of the INTERHEART study. The Indian journal of medical research, 132(5), p.561.

[9] Centre for Disease Control and Prevention, Online available at: http://www.cdc.gov/heartdisease/risk_factors.htm

[10] American Heart Association, Online available at: http://www.heart.org/HEARTORG/Conditions

[11] D. Isern, D. Sanchez, and A. Moreno, "Agents Applied in Health Care: A Review", International Journal of Medical Informatics, Vol. 79, Issue 3, pp. 146-166, 2010.

[12] Syed Umar Amin, Kavita Agarwal, Dr. Rizwan Beg, "Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors", Proceedings of IEEE Conference on Information and Communication Technologies (ICT), pp. 1227–1231, April 2013.

[13] Reynolds, R.G., 1994, February. An introduction to cultural algorithms. In Proceedings of the third annual conference on evolutionary programming (Vol. 131139). Singapore.

[14] Reynolds, R.G. and Peng, B., 2004, November. Cultural algorithms: modeling of how cultures learn to solve problems. In Tools with Artificial Intelligence, 2004. ICTAI 2004. 16th IEEE International Conference on (pp. 166-172). IEEE.

[15] S. Sabeena, V. Sujitha "Prediction of Heart Disease Using AI and NLP Techniques", International Journal of Computer Sciences and Engineering, Vol-6, Issue-2, Jan 2018

Authors Profile

Shiva Shrivastava has completed her Bachelor of Engineering in Computer Science and Master of Engineering in Computer Science (Rajiv Gandhi Technical University of Bhopal in year 2014 and 2017 respectively. She has published her 6 papers which belongs to communication and technical skills in VIPSAR , IJESRT, IJDACR Journals and presented papers in National Conferences. She is working as a Technical content writer and a Business Architect in an IT company.

Shreya Shrivastava is pursuing Bachelor of Computer Science Engineering from IES, IPS Academy, Indore(Rajiv Gandhi Technical University of Bhopal,). She has actively participated in various social causes like NGO. She is silver medalist in MBD, awarded was Ex. Chief Minister Mr.Babu Lal Gaur. She is the winner of various olyampiads. She was also a student of UCMAS- Mental Artimatic System. She is the CBSE Ajmer district topper in the senior and senior seconday examinations.

