Expert Systems with Applications 38 (2011) 14037–14043

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Web based medical decision support system application of Coronary Heart Disease diagnosis with Boolean functions minimization method

Fatih Başçiftçi [a,*], Hayri İncekara [b]

[a] Department of Electronics and Computer Education, Technical Education Faculty, Selcuk University, 42003 Selçuklu, Konya, Turkey
[b] Graduate School of Natural and Applied Sciences, Selcuk University, 42003 Selçuklu, Konya, Turkey

ARTICLE INFO

Keywords:
Expert System
Boolean function minimization
Coronary Heart Disease
Heart attack
Reduced rule base

ABSTRACT

Heart and vein diseases are one of the most important health problems. The number of the people who died of heart and vein diseases is more than the number of the people who died of all other health problems and natural disasters. In order to decrease this number, the intervention must be started very earlier and people must be informed about this subject. In this study, Medical Expert System has been used all the probabilities of CHD has been determined (9 symptoms, $2^9 = 512$ different occasions) and from these probabilities, an accuracy table has been formed. This accuracy table has been simplified with the Boolean functions simplifying methods. 94 rules have been achieved by the Boolean functions minimization method. The rules achieved have formed the Expert System's rule base. 303 patients' values have been compared with the realized Expert System. The Medical Expert System which was formed as a result of the evaluation has evaluated men with 86.5%, women with 84.5% and in general with 86.1% accuracy rate.

1. Introduction

In developed countries Coroner Heart Diseases (CHD), leads the reasons of death and lameness (Shapiron, Sadock, & Sadock, 2000). According to the latest formal numbers of World Health Organization (WHO) and World Heart Federation's (WHF), the number of people who died of heart and vein diseases is 17.5 millions and there are 2 millions in the European Union (EU). The people who died of heart and vein diseases are more then the people who died of natural disasters and all other health problems (Tüzüner, 2005). According to the explanations of WHF and European Heart United Association's, it has become more updated and important that this increasing rate will grow by 2030. WHO explains that when we come to 2020, there will have been 25 million people who have died of heart and vein diseases in the world (www.tkv.org.tr) (www.tkd-online.org).

The hearth diseases which cause hundreds of people to die or make them live an unhappy, unhealthy and inefficient life is one of the most important socio-economical problems which lead billion dollars of economical loss. Though they are not the exact numbers, it is estimated that heart and vein diseases are a big burden to the EU with its cost about 170 billion Euros. It means that it costs 372 Euros per one person. The amount of the loss of production because of not working and death as a result of heart and vein diseases is calculated as 35 billion Euros for EU (http://www.tkd-online.org/UKSP/TKD_UlusalKalpSagligiPolitikasi_Taslak.pdf, 2010). To come up with these problems, it is very important that all the people be informed about the disease and precautions be taken against in the treatment and the diagnosis of these diseases.

In our country more then 12 million people (1 of 3 people) are under the risk of having heart-vein diseases (Sarıoğlu, 2009). For this reason, by looking at the 9 (9 symptoms, $2^9 = 512$ different occasions) significant symptoms of 14 different symptoms, it can be evaluated if one has CHD or not. In order to evaluate the disease, a Medical Expert System (MES) has been developed by using the data from the scientific articles and expert doctors in their fields. Instead of evaluating the disease in 512 different occasions, the Expert System (ES) has been able to evaluate the disease in less occasions by using the minimization of Boolean functions method. The 94 results we have got in the result of the minimization of Boolean functions method constitute the rule base of MES. The realized MES has been put in to the web portal and has been served for everybody's use. Through the agency of this system, we can benefit from time, place, money etc.

2. Material and methods

2.1. Coronary Heart Disease

The diseases which come out due to the lack of oxygen because the coroner arteries can't carry the necessary blood to the heart are called as CHD (Topsakal, 2006). CHD are the leading reasons of all

* Corresponding author. Tel.: +90 (332) 223 33 53; fax: +90 (332) 241 21 79.
E-mail addresses: basciftci@selcuk.edu.tr (F. Başçiftçi), incekara86@hotmail.com (H. İncekara).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.04.211

death in our country as it is in the whole world. One of the most important reasons of CHD is the atherosclerosis with the rate of 99% (Selek, 2006) (Libby, 2001). Atherosclerosis consists of a process in which fat increases which narrows the veins and which thickens to the sides of the veins which we call as coroner veins that feed our health. As a result of vein hardening or atherosclerosis which is known as decreased fat in the veins, the flow of blood decreases and the vein gets narrower (Taşkıran, 2005). Instead of CHD term, atherosclerosis hearth diseases or lack of coroner terms are also used (Selek, 2006; Topsakal, 2006).

2.2. The symptoms of Coronary Heart Disease (data set)

There are four data sets on the heart diseases. These data sets have been collected from these centers: (Detrano et al., 1989; Gennari, Langley, & Fisher, 1989; Türker, Tokan, & Yıldırım, 2005) (www.cs.northwestern.edu/~ddowney/courses/349/psets/heart-disease.names):

1. Cleveland Clinic Foundation (cleveland.data);
2. Hungarian Institute of Cardiology, Budapest hungarian.data);
3. V.A. Medical Center, Long Beach, CA (long-beachva. data);
4. University Hospital, Zurich, Switzerland switzerland.data).

In this study, in order to predict a heart disease, the Cleveland data set which has been formed by Robert D. of the VA. M.C. Long Beach, CA, has been used. In this data set, the aim is to determine if one has a heart disease or nor by looking at the results of many medical tests done to the patients. This data set consists of 303 samples which the most important 14 of 76 test results that can be used to identify a heart disease (Türker, Tokan, & Yıldırım T., 2005). In this study, the aim is to diagnose if there is a problem about the heart not by decreasing the number of five classes (healthy, 1st level diseased, 2nd level diseased, 3rd level diseased, 4th level diseased) to two classes (healthy, diseased). 164 (54%) of the samples belong to the healthy and 139 (46%) of the samples belong to the unhealthy class. 9 of the 14 qualities in the data set are distinguishing symptoms and 5 of them are the other symptoms of the disease and after classifying these this way, it can be shown as follows.

2.2.1. Distinguishing symptoms of the disease

1. Thal: 3 = normal; 6 = fixed defect; 7 = reversible defect,
2. Chest pain type (typical angina, atypical angina, non-angina pain, asymptomatic),
3. Thalach: maximum heart rate achieved,
4. Oldpeak: ST depression induced by exercise relative to rest,
5. ca: number of major vessels (0–3) colored by fluoroscopy,
6. Exang: exercise induced angina (1 = yes; 0 = no),
7. Slope: the slope of the peak exercise ST segment (up sloping, flat, down sloping),
8. Chol: mg/dl cholesterol serum,
9. Restecg: resting electrocardiographic results (normal, having ST–T wave abnormality, showing probable or definite left ventricular hypertrophy by Estes' criteria).

2.2.2. Other symptoms of the disease

10. Age: age in years,
11. Sex: sex (1 = male; 0 = female),
12. Trestbps: resting blood pressure (in mm Hg on admission to the hospital),
13. fbs: (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false),
14. 0 = Healthy; 1 = Sick.

In Fig. 1 below, the separation of importance of the symptoms of CHD values are given comparatively:

2.3. Expert System (ES)

The first studied about ES were done in the Stanford University. Edward Feigenbaum, who is one of the Professor of Stanford University, defined ES as smart computer programs which solves difficult problems requiring expert knowledge by using the extraction procedures (Tatlı, 2000). ES are the computer systems which can model the deciding and comparing jobs that an expert man or men can do (Nabiyev, 2003). A well improved ES has qualities like imitating some processes such as design, planning, defining, commenting, summarizing, generalizing, controlling and suggesting which experts can do (Kılağız, 1996). An ES basically consists of knowledge base, inference engine, rule adjuster and user interface. In addition, to run the ES, a working memory is needed. The basic form of ES can be seen in Fig. 2 (Allahverdi, 2002; Babalık & Güler, 2007; Chang & Tseng, 2008).

2.3.1. The working principle of Expert System

Basic ES working principles are like this: The person who uses the program gives ES the realities and he gets expert advice or expertise. ES basically consists of two main elements. The first one is the Knowledge Base and it includes the truths that are known earlier. The second element, Inference Engine concludes

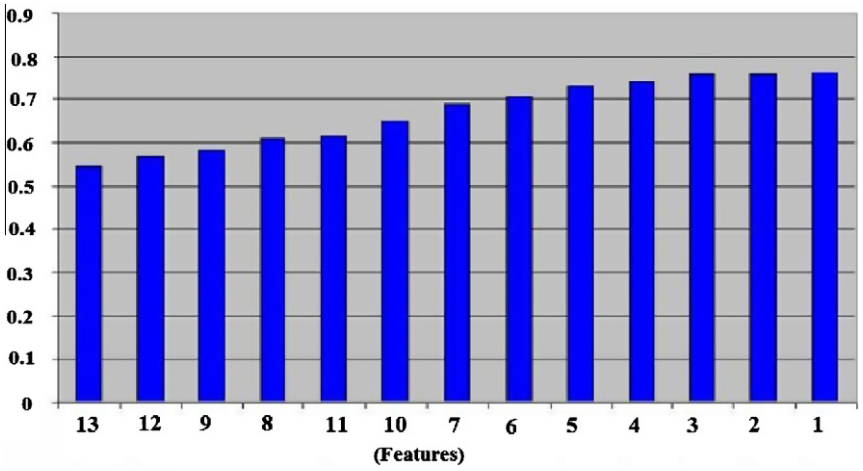

Fig. 1. Importance fragments of the symptoms.