Hearty: A Mobile Expert System for Screening the Risk of Cardiovascular Diseases

Asia Pacific Journal of Multidisciplinary Research
Vol. 8 No.1, 42-46
February 2020
P-ISSN 2350-7756
E-ISSN 2350-8442
www.apjmr.com
ASEAN Citation Index

Iluminada Vivien R. Domingo (DBA), Princess L. Mansilungan, John Vincent C. Carig
Polytechnic University of the Philippines
dvrdomingo@yahoo.com,cessmansilungan@gmail.com, vnccrg3@gmail.com

Date Received: September 27, 2019; Date Revised: February 3, 2020

Abstract *– Cardiovascular diseases become number one cause of death among people. People must be guided based on proper diet. This paper presented current issues on CVDs and proposed HearTy. HearTy is an expert system embedded on Android Operating System for the screening of cardiovascular diseases or CVDs. It is a rule-based expert system on Android phone for screening the risk as well as the likelihood of having CVDs based on the risk factors present and symptoms felt by the patient. The system covers five types of CVDs which includes Stroke, Coronary Heart Disease, Heart Attack, Heart Failure, and Peripheral Vascular Disease. The system applies Backward Chaining and RETE Algorithm. Backward Chaining is a type of reasoning for the inference engine which starts with a list of goals and RETE Algorithm is a pattern matching algorithm which identifies a rule that satisfy the pattern in order to generate the outputs. Result of the study showed accuracy of 68.75%, 66.59% sensitivity, error rate of 31.5% and 71.565% specificity which led to require further development of the developed software.*

Keywords *– Android, Backward Chaining, Cardiovascular Diseases (CVD), Expert System, RETE Algorithm*

INTRODUCTION

Cardiovascular diseases (CVDs), known as heart ailments by the general public had been the top cause of mortality in the Philippines since 1991. [1] Furthermore, the paper presented symptoms for early detection of the illness such as easily getting short of breath during exercise or activity, easily tiring during exercise or activity, swelling in the hands, ankles or feet[1]. As observed by many Filipinos, eating excess fatty and salty food is another basic cause for the disease [5].

Objective of the study specifically is to provide android application which screens patients with similar symptoms and provide healthy solutions to prevent the symptoms from recurring. Through reading related researches, the proponents sought to find the probable ways to screen a person's risk for cardiovascular diseases as well as if there is a likelihood of having one or not. This paper evaluated Hearty's accuracy, sensitivity, error rate and specificity of the computed values.

Through the span of time given to the research, the proponents have recognized that there are several factors that could affect a person and could trigger the existence of cardiovascular diseases in them. These factors are simply called risk factors which has two types: the non-modifiable risk factors or the factors that a person cannot controlled and the modifiable risk factors or those factors that can be controlled. Non-modifiable factors include age, gender, family history of diseases and others. Modifiable factors include physical inactivity, unhealthy diet, high blood pressure (hypertension), smoking (active and second-hand smokers), high blood cholesterol, social isolation and the lack of quality support. [5]

However, diagnosis of heart diseases can only be done when doctors or cardiologists do their checkups. Through the use of the created android application software, people might be guided and warned for possible heart diseases. A person should be aware about all of these factors to prevent having greater risk for the cardiovascular diseases.

One of the factors to keep in mind is age. On the onset of this research conducted by the proponents, it was said that as age increases, there is a greater possibility that a person's risk could increase. At the age of 65 or older, about four out of five people die because of heart attacks. Another factor is gender. The risk of heart disease to women rises dramatically after menopause, when they stop producing estrogen. [2] Men generally develop CVDs earlier than women. [3] Then next is family history of heart disease. According to research, if a first-degree blood relative had heart diseases at the age of 55 or lower for a male relative or 65 years old and

lower for a female relative, the risk increases. [4] To conclude, risk factors do affect the probable existence and/or further development of a cardiovascular disease.

On a research in California of the Filipino-Americans, it showed that almost 80% of the community residents get their health information from family and friends, television and health-care providers and television. [5] Moreover, we Filipinos, most especially those who are old, tend to practice 'self-diagnosis' and only seek the help of the doctors when their illnesses are already pretty serious. [6] If they have sought for help from an expert and knew that they have this type of diseases, they might have been more knowledgeable about the risk factors. A person should be aware about all of these factors to prevent having greater risk for the cardiovascular diseases [7].

As a result, the proponents devised an expert system which would inform users about their current condition, if there is a possibility or likelihood of having cardiovascular disease/s or not. Since the use of smart phones in the Philippines is immensely growing, and almost every people own it, it led the proponents to make a mobile-based expert system, particularly for Android phones [8].

SYSTEM ARCHITECTURE

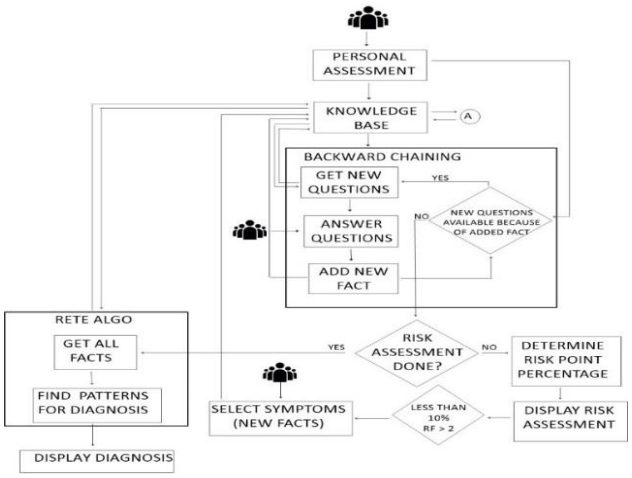

Figure 1: Hearty System Architecture

The information that was used in the system architecture, the systems' knowledge about the subject, would only include those which were gathered from the internet and books as well as through the interviews with doctors done by the proponents to make sure that the system's output would be able to replicate accurately the knowledge of the experts [9]. The system did not include anything related to medicines such as what to take, when to take it and other factors concerning medicines as well as other laboratory tests to further support the diagnosis [10].

The process of the system starts with the user providing all the needed information by the system [11]. The first information to be fed to the system first is the user's name, age, and gender. This will then be used during the inference process [12]. These data would be used by the system to further ask the user more about each of the risk factors both modifiable and non-modifiable [13]. These facts are then used to throw new questions to gather more data it needs to assess the user's risk of having the disease as applied using decision support system[14]. The backward chaining process involves throwing questions to user every time the data that the system needs is not present inside the knowledge base yet. This is the inference process [15].

As soon as the user finishes answering all the questions about risk factors, the results for the risk assessment are presented. The information that were fed to the system previously would now be used to assess the patient's risk of having cardiovascular disease using the Framingham Risk Score which provides the risk score and percentage of a person in developing heart disease within a 10-year basis [16]. If a person's risk percentage/equivalent is lower that 10%, he/she would be subject to the assessment of the other risk factors that were not considered on the pointing system [17]. If two or more risk factors are present, he would be asked to continue and assess his/her symptoms. This situation, along with if a patient has 10% and above risk percentage would always be subject to assess symptoms [18].

During the assessment of symptoms, symptoms would first be displayed, and then follow up questions would be thrown to the user. With the information from the previous answers of the users to questions, the Backward Chaining algorithm again searches for the next possible questions or information it should look at or ask from the user. After completing that process, the user would now be prompted of the results. The results are made by the RETE Algorithm. The RETE Algorithm has patterns stored in the knowledge base that would be used to identify what heart disease a patient might be experiencing. To identify what patterns are satisfied, the RETE algorithm uses the facts that were fed to the system while doing the Backward Chaining process. All the information/questions answered by the user are facts used in the process of the RETE Algorithm in making the diagnosis. Upon knowing what it is, the user is presented information about his/her probable heart disease.

DATA GATHERING PROCEDURE

The researchers used interview as one of the methods of collecting the needed data. It is a data-gathering device which, in a sense, is an oral questionnaire. [21] Interview is a purposeful conversation that usually involves a face to face relationship between two persons, one of whom called the interviewer and the other called the interviewee [22]. In this study, the researchers are the interviewer and the experts are the interviewee specifically the heart doctors from University of the East Ramon Magsaysay Medical Center. The researchers used this method simply because it can yield a more complete and valid information [23]. The total number of heart doctors in UERM at the time of writing is twelve, among whom eight doctors were interviewed. Twenty heart patients were also interviewed. Likewise, two university clinic doctors and two staff of Polytechnic University of the Philippines Sta. Mesa campus were also interviewed.

The researchers also used survey questionnaires as another method of collecting data [24]. A questionnaire is a research instrument consisting of a series of questions and other prompts for the purpose of gathering information from respondents [25].

EVALUATION PROCESS

Accuracy is the condition or quality of being true, correct, or exact. It is a proportion of the results, in a population, both true positive and true negative [20]. It is the one which measure the rate of veracity or the truthfulness of a diagnostic test on a condition, which is in this study, a disease [26].

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \qquad (1)$$

Where:

True positive (TP) = **Correct** diagnosis of the **existence** of a cardiovascular disease compared to the expert's diagnosis.

False positive (FP) = **Incorrect** diagnosis of the **existence** of a cardiovascular disease compared to the expert's diagnosis.

True negative (TN) = **Correct** diagnosis of the **inexistence** of a cardiovascular disease compared to the expert's diagnosis.

False negative (FN) = **Incorrect** diagnosis of the **inexistence** of a cardiovascular disease compared to the expert's diagnosis.

Sensitivity is the portion of true positives which were identified as existing by a diagnostic test. It shows how good the test, which is the system, is at detecting a condition or disease. The numerical values represent the probability of a test to identify patients who actually have the disease. The higher the sensitivity, the lesser it would return a false-positive result.

$$Sensitivity = \frac{TP}{TP+FN}\ x\ 100 \qquad (2)$$

Specificity is the portion of true negatives which were correctly identified by the system. It also suggests how good the diagnostic test (the system) is at identifying a normal or negative condition. The numerical value of specificity represents the likelihood of a test to diagnose diseases without giving false-positive results.

$$Specificity = \frac{TN}{TN+FP}\ x\ 100 \qquad (3)$$

Error is used measuring of the proportion of items that got wrong.

$$Error = \frac{FP+FN}{TP+TN+FP+FN} \qquad (4)$$

EVALUATION RESULTS

The following table summarizes the results of computed values as answered by the respondents where Expert 1 is composed of ten heart doctors and Expert 2 is composed heart patients and clinic staff as follows:

Table 1 Summary of Results

Expert	Accuracy (%)	Error Rate (%)	Sensitivity (%)	Specificity (%)
# 1	72.5	27.5	69.56	76.47
# 2	65	35	63.63	66.66
Overall	68.75	31.25	66.59	71.565

Expert 1 has a percentage of 72.5% accuracy while Expert 2 is represented by heart patients and clinic staff has a percentage of 65%. The obtained overall average accuracy of the system is equivalent to 68.75%. This only shows that 31.25% of the generated output of the system must be enhanced in order to achieve 100% accuracy.

Expert 1 has a higher percentage of sensitivity than Expert 2. Expert 1 has a percentage of 69.56% while Expert 2, on the other hand, has a percentage of 63.63%. Based on the obtained overall average sensitivity of 66.59%, as computed by the proponents, this numerical value represents the probability of a test to identify patients who actually have the disease.

The highest specificity appears to be in evaluation made by Expert 1. It has a percentage of 76.47%. Expert 2 on the other hand has a percentage 66.66%. Based on

the obtained overall average specificity of 71.565%, as computed by the proponents, this represents the probability of a test to identify normal or negative conditions or inexistence of the disease.

CONCLUSIONS

Based from the findings of the study entitled "HearTy: A Mobile Expert System for Diagnosing the Risk of Cardiovascular Diseases", the researchers conclude that HearTy can determine if a person has the likelihood of having CVDs based on risk factors and symptoms with the accuracy of 68.75%. This also means that 31.25% of the generated output of the system must be enhanced in order to achieve 100% accuracy.

HearTy can detect the likelihood of cardiovascular diseases with a success percentage of 66.59% sensitivity. Sensitivity is the proportion of positives that are correctly identified by the system. In short, this is the proportion of the values of True Positives (TP) or correct diagnosis of the existence of a cardiovascular disease compared to the expert's diagnosis and the False Negatives (FN) or incorrect diagnosis of the inexistence of a cardiovascular disease compared to the expert's diagnosis

The system can also diagnose users without heart disease with a success percentage of 71.565% specificity. Specificity is the ability of the system to correctly identify those without the disease. Therefore, this is the proportion of the values that are False Positives (FP) or incorrect diagnosis of the existence of a cardiovascular disease compared to the expert's diagnosis and True Negatives (TN) or correct diagnosis of the inexistence of a cardiovascular disease compared to the expert's diagnosis.

However, in terms of adequacy of contents, as commented and suggested by the experts in the field, HearTy needs to provide more information and explanation of the disease.

In the conclusion of Expert System for Screening of Neuromuscular Disorders which the proponents have used as one of the related studies, it was said that this expert system can be help medical professionals and patients of the said disease to be given an overview and treatment option for the disease they possess [27]. It was the same statement that our experts from the University of the East Ramon Magsaysay Memorial Medical Center had told us, our system, Hearty's output is leading. Which means that, it can help people as well as doctors in finding out the probable things that might affect the existence of a heart disease on the patient as well to understand well his/her symptoms. Both studies have attained the same level of relevance for experts, however in different fields [28].

RECOMMENDATIONS

Broaden the scope of the diseases that would be considered by the system. Or better yet, focus on just one type of disease to have a higher accuracy of outputs or diagnosis. Add more follow up questions in order to further validate the symptoms felt by the user. This would also be resulting to a higher accuracy. Implement the system in other mobile operating systems such as iOs and Windows. Instead of using SQLite, try other database with Xamarin because there are times that the Android application fails because of the database. Improve the graphical user interface (GUI) of the application to make it more appropriate to the medical field and also for the application to be more user-friendly. Have a Filipino translation to make the system in order for the users to easily understand and since this system is developed for Filipino citizens.

REFERENCES

[1] DOH. (2013). Leading Causes of Mortality. Retrieved October 11, 2015 from *Department of Health*: http://www.doh.gov.ph/node/2573

[2] Black, H. R. (2012). Cardiovascular Risk Factors. In Sharpe, G. J. et al., eds., *Heart Book*..

[3] WHO. (2014). Cardiovascular Diseases (CVD). Retrieved October 11, 2015 from *World Health Organization*: http://www.who.int/mediacentre/factsheets/fs317/en/

[4] WHF. (2015). Cardiovascular disease risk factors. Retrieved October 11, 2015 from *World Heart Federation*: http://www.world-heart-federation.org/cardiovascular-health/cardiovascular-disease-risk-factors/,

[5] NHLBI. (2016). *Cardiovascular Risk in the Filipino Community*. National Heart, Lung, and Blood Institute; Asian & Pacific Islander American Health Forum; West Bay Pilipino Multi-Services, Inc.

[6] Stanford School of Medicine. (n.d.). Health Beliefs and Behaviors. Retrieved October 8, 2015 from *Stanford School of Medicine*: https://geriatrics.stanford.edu/ethnomed/filipino/fund/health_beliefs.html

[7] Aquino, G. V. (2015). In *Fundamentals of Research*. National Book Store, Inc, Mandaluyong City.

[8] Aguire, J.V., et al. (2014). Cadio Vascular Application: A mobile application for Food and Drug Administration.

[9] Ajoudanian, S. & Jazi, M.D. (2009). Deep Web Content Mining, World Academy of Science, Engineering and Technology.

[10] Basole, R. & Cho, R. (2004). Mobile Decision Support System and Their Effect on User Performance.

[11] Bharamagoudar, G.R., Totad, S.G.,Reddy, P. (2012), "Literature Survey on Web Mining," *IOSR Journal of Computer Engineering (IOSRJCE)*, 5 (4), pp. 31-36.
[12] Cho, J. (2010). Crawling the Web: Discovery and Maintenance of Large-Scale Web Data.
[13] Druzdzel, M. J. & R. R. (2009). Flynn Decision Support Systems", Encyclopedia of Library and Information Science..
[14] Druzdzel, M.J. & Flynn, R.R., (2012). Decision Support Systems, Encyclopedia of Library and Information Science.
[15] Edwards, J., McCurley, K., & Tomlin, J. (2001, April). An adaptive model for optimizing performance of an incremental web crawler. In *Proceedings of the 10th international conference on World Wide Web* (pp. 106-113).
[16] Erl, T., "SOA Principles of Service Design," The Prentice Hall Service-Oriented Computing Series. Prentice Hall, (July, 2017)
[17] Fielding, R.T. (2013). Architectural Styles and the Design of Network-based Software Architectures, University of California, Irvine.
[18] Kausar, A., Daka, V.S., Singh, S.K., "Implementation of Parallel Web Crawler through .NET Technology," IJMECS International Journal of Modern Education and Computer Science, (August, 2014)
[19] Liu, B., "In Web DATA Mining", Springer-Verlag Berlin Heidelberg, New York, (2017), pp16-22
[20] Mawale, M.V. and Chavan, V., "Issues and Challenges in Complex Analysis of Variable," (IJCSIT) International Journal of Computer Science and Information Technologies, 5, 4, (2014)
[21] Pedrinaci, C., Domingue, J., "Toward the Next Wave of Services: Linked Services for the Web of Data", Journal of Universal Computer Science, 16, 13, (2015), pp. 694-1719
[22] "RadioShack Converts 40% to 60% of Mobile Store-Locator Clicks into Store Visits", download only from https://www.thinkwithgoogle.com/case-studies/radioshack-converts-mobile-clicks.html, (November, 2012).
[23] Sarawagi, S. (2017). Information Extraction", Foundations and Trends in Databases, 1(3), 261-377.
[24] Tan, J. S., Lu, E. H. & Tseng, V. S. (2012). Preference-oriented mining techniques for location-based store search, Department of Computer Science and Information Engineering, Springer Verlag London Limited.
[25] Turban, E. (2016) "Decision support and expert systems: management support systems,".
[26] Englewood Cliffs, N.J., Prentice Hall. (2015).
[27] Yang, Q. & Zhang, H.H., (2017). "Web-Log Mining for Predictive Web Caching", IEEE
[28] Transactions on Knowledge and Data Engineering, 15(4), (July/August) 2013.