ISSN 2347 - 3983
Volume 8. No. 5, May 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter50852020.pdf
https://doi.org/10.30534/ijeter/2020/50852020

HDP: Heart Disease Prediction Tool using Neural Network

Wan Aezwani Wan Abu Bakar[1], Mustafa Man[2], Wan Suryani Wan Awang[3], Nur Laila NajwaMohd Josdi[4], AimanNatasya Naquiyah[5], Nik Nurul Nadia Nik Pa[6]
[1,3,4,5,6]Faculty of Informatics and Computing, Universiti Sultan ZainalAbidin, Besut Campus, 22200 Besut, Terengganu.
[2]Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu.
[1]wanaezwani@unisza.edu.my; [2]mustafaman@umt.edu.my; [3]suryani@unisza.edu.my;[4,5,6]{sl3347, sl3347, sl3330}@unisza.edu.my

ABSTRACT

Heart disease may strike everyone and endangering individuals of all gender, age or weight. Malaysia Ministry of Health (MOH) in 2017 has revealed that one in two Malaysians suffers from high cholesterol, which contributes to heart disease. As the principal cause of death with 13.2% in 2016, heart diseases seem to be worst from time to time. There should be a proper initiative to help in preventing this disease. Healthcare industry has a big amount of data collection but unfortunately, they are not mined properly to extract hidden pattern and relationship. Plus, Malaysians sometimes are too busy to keep up their regular checkup to ensure their health. Some might be short in time and some might be short in finance as the cost of clinical test might be burdening. There might also happen so many times that you or someone yours need doctors help immediately, but they are not available due to some reason. Healthcare domain needs data mining technique to improve the efficiency of analytical methodology to detect and value the relationship within the health profile intelligently, in cheapest, fastest and safest way. In realizing this situation, we propose a system that uses a neural network techniqueto map a set of input data onto a set of appropriate output data. The technique helps to generate highly accurate prediction result based on the relationships within some medical factors using only a system that can be access anywhere without having to go the hospital and spending money.

Key words: Heart Disease, Prediction, health, neural network

1. INTRODUCTION

Heart disease is a cause of various conditions include blood vessel diseases like coronary artery disease, heart rhythm problems (arrhythmias) and heart defects since born with congenital heart defects. World Health Organization (WHO) reported the Coronary Heart Disease in Malaysia in 2017 has reached 30,598 or 22.13% in total deaths cases. The death rate is 137.02 over 100,000 ranks of population which made Malaysia to be the 63rd of Heart Disease Death in the world [1]. Compared to 2016, death cases are reported with 13.2%, pneumonia with 12.5%, cerebrovascular diseases with 6.9%, transport accidents 5.4% and malignant neoplasm of trachea, bronchus & lung with 2.2% [2]. The concern goes to the age of onset of heart disease in Malaysia, is much younger compared to neighboring countries and some western nations. Therefore, a serious awareness program on preventing atherosclerosis risk factors must be enforced. Common heart disease risk factors are age, gender, high blood pressure, high serum cholesterol levels, smoking and other factors could be used for early detection. A quick prediction tool for early heart disease detection is crucial to prevent this heart disease from widely spread and getting serious [2].

An efficient and accurate diagnosis or prediction tool is extremely needed and a crucial task for researchers. Normal practice of diagnosis depends upon a doctor's past experience and knowledge that may contribute to unreliable results such as errors, delayed treatments and incur in medical costs of treatments charged to patients. Since most people have limited time limitations, they have not been able to commit themselves to the health of their bodies. They would feel that periodic check-ups are not important as long as they are not attacked by any illness [3].

Thus, an instant response tool for medical diagnosis would give benefit for an early precaution. This research attempts to showcase how artificial neural network techniques can be deployed to help in generating highly accurate prediction result using only a convenient click on a smartphone that would give user the same result for a fraction of the price and a fraction of the time.

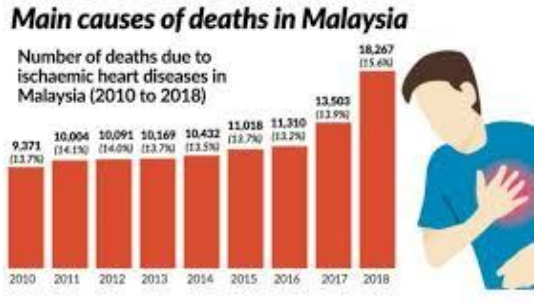

Figure 1: Heart Disease in Malaysia

2. ALGORITHM

Heart Disease Prediction (HDP) is conducted using Artificial Neural Network (ANN). This machine learning systems inspire by the biological neuron that constitute to an animal brains [4-6]. Such systems would learn from imposed training to perform tasks from examples, basically without being programmed with task-specific rules. Input data is given at the input layer for processing, prior to producing the results. The predicted result is extracted by subtracting from actual output and the error value is calculated.

Then, BackPropagation (BP) algorithm is used to adjust the weights [6-8, 15]. Weights adjustment begin between output layer nodes and last hidden layer nodes and propagates in reverse or backwards throughout the network. When it's done, then the upwards process starts again [9-10]. InBP algorithm, the gradient of the loss function at each respective weight in the rule of chain is calculated, the gradient of a layer at one particular time is also computed, and also backward iteration from the last layer is also undertaken to get rid of redundant calculations of intermediate terms in the whole chain rule [11-14]. The final result will be between 0-0.49 (No risk for heart disease) and 0.5-1 (Risk for heart disease). The formula for calculating the neuron's output as follows,

$$\sum weight_i \cdot input_i = weight1 \cdot input1 + weight2 \cdot input2 + weight3 \cdot input3$$

And a sigmoid function (SF) is used for normalizing output. SF is a mathematical function with "S"-shaped curve or sigmoid curve which ranges between 0 and 1

$$\frac{1}{1+e^{-x}}$$

From the normalizing output, the prediction is done based on the accuracy formula as follows,

$$PPV\ (Positive\ Predictive\ Value) = tp/tp + fp$$
$$NPV\ (Negative\ Predictive\ Value) = tn/tn + fn$$
$$Accuracy = tp + tn/tp + fp + fn + tn$$

Where tp is True Positive, tn is True Negative, fp is False Positive and fn is False Negative.

3. PROJECT DESCRIPTION

Heart Disease Prediction apps is developed to increase the level of awareness among youth as well as adults. Through mobile apps, users can use to check their cardiovascular fitness whenever they feel uneasy at any particular time. They can just select the attributes displayed and click submit to predict on their results. The normal attributes are age category, sex or gender, type ofchest pain while the medical history attributes such as blood pressure, sugar in fasting blood, electrocardiographic results and the maximum of heart rate (pulse) achieved will be evaluated using neural network algorithm in Weka prior to result indication.

3.1 Context Diagram of HDP

Figure 2 shows the overall process in Heart Disease Prediction Application using Neural Network. There are two external entities, user and Weka [9-10]. The user will fill in user details and log in to access the application. User can manage their health details and update their prediction details. The application gives prediction result to the user prior to filling up their health detail. Weka will provide the prediction weight to be used by the application to produce prediction result.

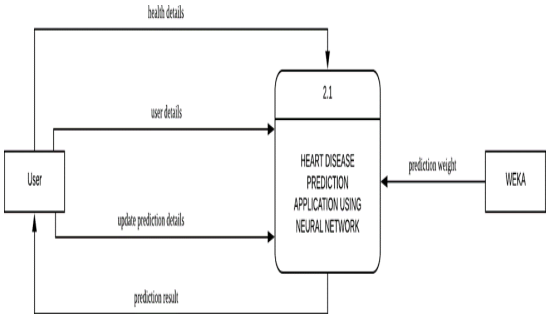

Figure 2: HDP context diagram

3.2 Data Flow Diagram of HDP

Data Flow Diagram (DFD) Level 0 consist of four main processes. The first process is to manage user where user will fill in user detail to be store in user profile data store. User can manage health information by providing their health details to be store in health profile data store. The user also can manage prediction by submitting their health details to prediction profile data store and the data store will later respond by giving out their prediction result requested by user. The prediction result then will be used for managing knowledge by Weka before the knowledge details then be sent to knowledge profile.

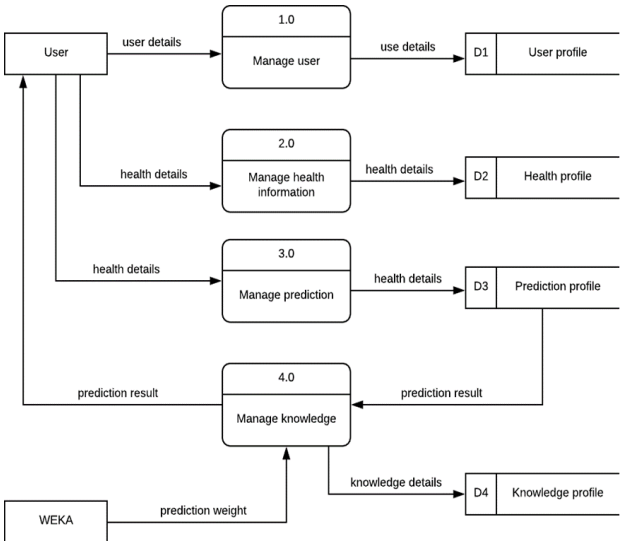

Figure 3: Health information in HDP

4. EXPERIMENTATION

All experimentations are done with i3 Processor, 4GB RAM, 2GB VRAM, 20 GB SSD and 1TB Hard disk storage. The HDP system is accessible through any browsers. Apache Tomcat is a web server engine to support the functionality of the modules. The programming languages are MySQL to manage for database, and JavaScript, Cascading Style Sheets (CSS) and PHP for page scripting.

4.1 Data collection

The data used for this project is collected from UCI Machine Learning Repository [9,11]. Statlog (Heart) Datasetconsist of 13 attributes with 270 observations and no missing values. However, only 7 attributes were used for this project.Then, the dataset is preprocessed to convert its format to.arff as Weka [12] only can process dataset with .arff and .csv file only. Weka is a data mining tool comprises of several machine learning algorithms to perform data mining tasks. Weka provides methods for data pre-processing, classification, regression, clustering, association rules, and visualization. After the preprocessing, the next step is to classify the training set using Multi-Layer Perceptron (MLP) with 1 hidden layer. The dataset is trained for several times until its accuracy achieved the maximum point. Then, the weights of all the attributes can be taken to calculate the prediction. Until then, the development of the application started with Waterfall Model.

There are 270 instances and 13 attributes where the attributes are depicted in Table 1.

Table 1: UCIHeart Disease Attributes

Attribute Names	Attributes Character
Age	Real
Gender	Binary
Type ofChest pain	Nominal
Blood Pressure	Real
Fasting Blood Sugar >120 Mg/Dl	Binary
Electrocardiographic Results	Nominal
Heart Rate (Pulse)	Real

4.2 Database Connection

HDP is connected through SQL Script for database connection.

```
$servername = "localhost";
$username = "root";
$password = "";
$dbname = "heartdisease";
```

Figure 4: Health information in HDP

5. RESULT AND DISCUSSION

The HDP interfaces are depicted in Figure 5 to Figure 7. To log in, users are required to insert their username and password. For user who does not have an account yet, they need to register by clicking the register link. For user who already has the account, clicking the "Login" button will redirect them to health detail page.

Figure 5: HDP Login

In Figure 6, user need to fill in the form provided with their health information. After answering all questions, they need to click the "Predict" to get the answer.

Figure 6: Prediction Attributes in HDP

User will get their prediction answer in this page. If the output is from 0-0.49, it will say 'NO' which means no risk of heart disease. If the output is from 0.5 to 1, it will say 'YES' which means there is a risk of heart disease. The result is depicted in Figure 7.

Figure 7: HDP results

6. CONCLUSION AND FUTURE WORKS

The contribution of this application to the user is they can have a prediction of getting heart disease with a fraction of price and time. This is because, user do not have to worry if they do not have enough money or time to get a regular check-up at the hospital. With the health detail that is filled by the user, neural network will calculate the result based on its theory and provide result to the user. So, user will likely have an accurate prediction of their heart's health status. This application can help user to alert on their health status and aware on heart disease factor,

whereas it can be controlled by them. The factors are supposedly can be fully controlled by humans if they have a healthy lifestyle. If he or she takes good care of these factors, they are able to reduce the risk for the heart disease.

ACKNOWLEDGEMENT

A sincere gratitude goes to Director of Nebula Technologies that has shown their letter of intent in using our apps. Also, to all faculty members and for supporting our work in reviewing for spelling errors and synchronization consistencies and also for the meaningful comments and suggestions.

REFERENCES

1. Malaysia Coronary Heart Disease, World Health Rankings from https://www.worldlifeexpectancy.com/malaysia-coronary-heart-disease
2. Aziz, Zariah A., Yvonne YL Lee, BahariAwangNgah, NorsimaNafizahSidek, Irene Looi, MdRafiaHanip, and Hamidon B. Basri. "Acute stroke registry Malaysia, 2010-2014: results from the National Neurology Registry." *Journal of Stroke and Cerebrovascular Diseases* 24, no. 12 (2015): 2701-2709.
3. Khoo, K. L., H. Tan, and T. H. Khoo. "Cardiovascular mortality in peninsular Malaysia: 1950-1989." *Med J Malaysia* 46, no. 1 (1991): 7-20.
4. Taneja, Abhishek. "Heart disease prediction system using data mining techniques." Oriental Journal of Computer science and technology 6, no. 4 (2013): 457-466. G. Benelli, C. Jeffries, T. Walker, “Biological control of mosquito vectors: past, present, and future”. Insects. 2016 Dec;7(4):52. https://doi.org/10.3390/insects7040052
5. Kim, Jae Kwon, and Sanggil Kang. "Neural network-based coronary heart disease risk prediction using feature correlation analysis." *Journal of healthcare engineering* 2017 (2017). A. Wilson, TN. Nguyen, “The zika virus epidemic: public health roles for nurses”. OJIN: The Online Journal of Issues in Nursing. 2017 Jan 1;22(1). https://doi.org/10.1155/2017/2780501
6. Albu, Adriana, and Loredana Ungureanu. "Artificial neural network in medicine." *Telemed. JE Health* 18, no. 6 (2012): 446-453.
7. AedesTech Mosquito Home. Available: http://www.onedream.com.my/index.php?ws=showproducts&products_id=1010579&cat=Mosquito-Killer-Repellent
8. Uyar, Kaan, and Ahmet İlhan. "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks." *Procedia computer science* 120 (2017): 588-593. Dengue Tech Challenge 2016. Available: http://www.platcomdiscovery.com/denguetech/winners.aspx https://doi.org/10.1016/j.procs.2017.11.283
9. Kim, Jae Kwon, and Sanggil Kang. "Neural network-based coronary heart disease risk prediction using feature correlation analysis." *Journal of healthcare engineering* 2017 (2017). DR. PatnaikPatnaikuni.”A Comparative Study of Arduino, Raspberry Pi and ESP8266 as IoT Development Board. International Journal of Advanced Research in Computer Science. 2017 May 15;8(5).
10. Amutha, A. Jerline, R. Padmajavalli, and D. Prabhakar. "A novel approach for the prediction of treadmill test in cardiology using data mining algorithms implemented as a mobile application." *Indian heart journal* 70, no. 4 (2018): 511-518. EC. https://doi.org/10.1016/j.ihj.2018.01.011
11. Prima, SS. Munifaha, R. Salam, . Aziz, AT. Suryani, “Automatic water tank filling system controlled using ArduinoTM based sensor for home application”. Procedia engineering. 2017 Jan 1;170:373-7. https://doi.org/10.1016/j.proeng.2017.03.060
12. Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. "The WEKA data mining software: an update." *ACM SIGKDD explorations newsletter* 11, no. 1 (2009): 10-18. https://doi.org/10.1145/1656274.1656278
13. Asuncion, Arthur, and David Newman. "UCI machine learning repository." (2007).
14. Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. "The WEKA data mining software: an update." *ACM SIGKDD explorations newsletter* 11, no. 1 (2009): 10-18. https://doi.org/10.1145/1656274.1656278
15. Back Propagation from https://en.wikipedia.org/wiki/Backpropagation