Journal Of Clinical & Medical Images Case Reports

Case Report

Open Access, Volume - 2

Feature Selection Based Fuzzy Expert System for Efficient Diagnosis of Coronary Artery Disease

Y Atomsa[1]; LJ Muhammad[2*]; FS Ishaq[3] and Yusuf Abdullahi[4]

[1]Computer Science Department, Federal University of Kashere, Gombe State, Nigeria.

[2]Computer Science Department, Federal University of Kashere, Gombe State, Nigeria.

[3]Computer Science Department, Federal University of Kashere, Gombe State, Nigeria.

[4]Biological Sciences Department, Federal University of Kashere, Gombe State, Nigeria.

Received Date : Feb 11, 2022
Accepted Date : Mar 04, 2022
Published Date : Mar 18, 2022
Archived : www.jcmimagescasereports.org

***Corresponding Author:** LJ Muhammad, Computer Science Department, Federal University of Kashere, Gombe State, Nigeria.
Email: lawan.jibril@fukashere.edu.ng

Abstract

Coronary artery disease (CAD) is one of the most dangerous diseases which lead to sudden cardiac death. According to Word Health Organization, CAD is the number one killer in the developed world, with over 7.4 million deaths attributed it. Before, CAD is not common disease in Nigeria, however, is at this moment gaining much popularity in the country following the rising number of health issues related to CAD diseases, including higher death rate, which is mostly due to lack of proper awareness among the common people. The diagnosis of CAD is very expensive and time consuming which made computer scientists to use artificial techniques such as expert system to diagnose CAD's patients. Therefore, in this work, fuzzy based expert system for efficient diagnosis of coronary artery disease has been developed, implemented and evaluated. Hence, the system has archived 90.08% overall accuracy which is very excellent, thus the accuracy determines the proportion of the total number of predictions that were correct. At the same time, the system has 91.30% accuracy to classify of normal patients correctly by the system (specificity) and 90.24% accuracy to classify abnormal patients correctly by the system (sensitivity). This showed that, the system performed efficiently and excellently to diagnose CAD.

Keywords: CAD; Artificial Intelligence; Expert System; Fuzzy Logic; dataset; diagnosis.

Introduction

Coronary Artery Disease (CAD) is one of the deadliest diseases in the world. It has been estimated that nearly one half of all middle-aged men and one third of middle-aged women in the United States have been affected with the CAD disease [46, 50]. In the developed countries CAD is one of the number one killers with over 7.4 million deaths attributed to [29, 48]. It has been estimated that, CAD is one in every seven deaths in the United States is due to heart disease. CAD is the primary cause of death in women, taking more lives than all cancers combined The proportion of deaths in the United States that are due to CAD has been decreasing slowly but continuously over the past half century. Nonetheless, CAD remains the single most common cause of death in the United States, according [16, 19, 46]. In Nigeria, CAD is at the moment gaining much popularity following the rising number of health issues related to the disease, including higher death rate, which is mostly due to lack of proper awareness among the common people. According to the latest WHO data published in May 2014 Coronary Artery Disease Deaths in Nigeria reached 53,836 or 2.82% of total deaths. The rising figures of health issues (cases) and mortality rate related to CAD in Nigeria recently are alarming [2, 7, 26-27]. Research outputs have suggested a rather increasing occurrence of instances of the CAD in the Nigerian community, which however, the populace seems not to be well informed and/or alarmed about it, and also, given the health emergencies response level of Nigeria, the Nigerian health systems do not seem very readily capable to deal with the menace of CAD in both an immediate and a continuous strategy to lower and/or overcome its adverse effects on the population [26-27]. However, the major problem in traditional method of medical diagnosis is inadequate guarantee of precision and accuracy. There are huge data management tools available within health care systems, but analysis tools are not sufficient to discover hidden relationships amongst the data. Most of medical information is vague, imprecise and uncertain [8]. Extracting correct information from this data is considered an art [16, 12-13]. It can be said to be an art because it is complicated by many factors and its solution involves literally all of a human's abilities including technical expertise and

Citation: Y Atomsa, LJ Muhammad, FS Ishaq, Yusuf Abdullahi. Feature Selection Based Fuzzy Expert System for Efficient Diagnosis of Coronary Artery Disease. J Clin Med Img Case Rep. 2022; 2(2): 1103.

intuition. Hence, it becomes necessarily essential to leverage computing technologies that support analytical processing of this data to extract hidden correlations and intelligences within the data in order to improve accuracy and precision of medical diagnosis [15].

Expert systems have been specifically applied in a variety of life sciences support systems development, ranging from storage and retrieval of medical records, diagnostics, up to expert knowledge/decision support systems [3-5, 9]. Expert system defined as an intelligence system that extracted its knowledge using appropriate technique with the perception of human expert, to solve the problems or make decision as human being does [10, 44]. Expert System is defined as an intelligence system which uses extracted knowledge from past domain expert decision making reasoning in form of rules to solve problems that ordinarily require human expertise for their solution, and has the capability to update its rule-base as new knowledge is discovered. There are many application areas of expert system such as medicine, education, agriculture, oil and gas, environment, law, manufacturing, telecommunication and power systems etc. [11]. This research work aims to develop a fuzzy based expert system for supporting and completing human expertise in the diagnosis of Coronary Artery Disease. The focus is leveraging computing technology systems for the efficient, precise and re-accessible diagnosis procedure for the Coronary Artery Disease.

Related Work

In this study research articles and conference papers published by reputable pub¬lishers that employed expert system for diagnosis of coronary artery disease were searched and reviewed in this section. In work [47] an expert system for the diagnosis of the level of coronary heart disease by taking into account the problem of data imbalance developed. In the study of [14], a fuzzy soft sets expert system to predict patients suffer coronary artery disease was developed. The research was a pioneering approach in applying fuzzy soft sets to a medical diagnosis problem in the form of predicting patients who may be suffering from coronary artery disease. In study of [39], a web based Fuzzy Logic-based Expert System for the diagnosis of heart failure disease was developed. An evolutionary fuzzy expert system is proposed for the diagnosis of the Coronary Artery Disease (CAD) based on Cleveland clinic foundation datasets for heart diseases in the study of [49]. In the study of [1], patients with coronary artery disease were identified and classified through the neuro-fuzzy network with the capacity of automatically extracting fuzzy rules. Fuzzy expert system was implemented using facilities and functions of MATLAB software (7.12.0 version). A fuzzy expert system for diagnosis of coronary artery disease by a non-invasive procedure was implemented in the study of [33]. The adaptive neuro fuzzy inference system and Advanced fuzzy resolution mechanism has been proposed in the study of [17] to diagnose the heart disease. The work of [38] has developed a computer intelligent based approach for the diagnosis of heart diseases. A fuzzy rule-based system which concentrated only on accuracy and interpretability of the system was proposed by Ref [36] and system that provided a heart disease patient with background for suitable diagnosis and treatment. Ref. [42] developed a weighted fuzzy rule-based Clinical Decision Support System (CDSS) for computer-aided diagnosis of the heart disease. The different data mining techniques such as neural networks, decision trees and naive bayes has been proposed in the work of [45] for the study of heart disease prediction system. A coronary artery disease fuzzy expert system for microarray data classification using a novel Genetic Swarm Algorithm, has been proposed in the study of [18] for obtaining near rule set and membership function tuning. In the study of [43] screening system has been developed for the early detection of Coronary Artery Disease. In the study of [33], a fuzzy rule-based system was designed to serve as a decision support system for diagnosis Coronary heart disease. In the work [6] a Fuzzy Expert System for heart disease diagnosis using V.A. Medical Center, Long Beach and Cleveland Clinic Foundation database was designed and system is being designed with in Matlab software and it is viewed as an alternative for existing methods to distinguish of heart disease presence.

Table 1: *Description of the Dataset Features.*

SN	Feature	Units	Range
1	Age	Years	1 – 150
2	Sex	Male (1), Female (0)	0,1
3	Family History	Yes (1), No (0)	0,1
4	Smoking	Yes (1), No (0)	0,1
5	Diabetes	Yes (1), No (0)	0,1
6	Hypertension	Yes (1), No (0)	0,1
7	Hyperlipimedia	Yes (1), No (0)	0,1
8	Blood Pressure	mmHg	90 – 190
9	Glucose	mg/dL	37 – 295
10	Cholesterol	mg/dL	128 – 575
11	Triglyceride	mg/dL	40 – 690
12	HDL	mg/dL	10.6 – 73
13	LDL	mg/dL	10 – 220
14	Creatinine	mg/dL	0.6 – 3.3
15	Body mass index	kg/m^2	20.28 – 40.25
16	Heart rate	Bpm	42 – 124
17	Chest pain	Typical Angina (4), Atypical Angina(3), Non- Anginal pain(2), Asymptomatic (1)	1 – 4
18	Diagnosis of CAD	Positive (1), Negative (2)	0,1

NB: **mmHg** stands for millimeters of mercury, **mg/d**L stands for milligrams per deciliter, **kg/m²** stands for Kilogram-Meter Squared and **Bpm** stands for beats per minute

Methods and Materials

Dataset

The medical expert diagnostic dataset for coronary artery disease obtained at the Federal Teaching Hospital, Gombe State, Nigeria were prepared in the appropriate format with the helped of medical experts in the hospitals and only data instances of the dataset without missing values were considered and collected. Therefore, there are only one thousand two hundred and one data instances of the dataset without missing value. The dataset is labeled one with eighteen features including demographic, history and clinical features of the patient's CAD. The feature of the dataset are age, sex, CAD family history, smoking, type of the chest pain, diabetes, glucose, hypertension, blood pressure, cholesterol, Hyperlipidemia, high density lipoprotein (HDL), Triglyceride, low density lipoprotein (LDL), Creatinine, BodyMass, HeartRate and Diagnosis (**Table 1**) shows the description of the dataset features.

Feature Selection with Correlation Analysis

Correlation coefficient analysis was carried out on the dependent and independent features of the CAD Dataset [20]. Correlation coefficient is used to determine the strength relationship that exists between the dependent and independent features which can either be positive or negative. The r value is a set of infinite number between -1 to +1 which show the existing relationship either positive or negative between the dependent and independent features [22-24]. The feature can be evaluated by the equation (1) below:-

$$Importance = \frac{\overline{kavg(corr_{fc})}}{\sqrt{k + k\,(\,k-1)\,\overline{avg(corr_{ff})}}} \qquad (1)$$

Where the Importance is the correlation coefficient between dependent feature set and independent feature and is the ranking criteria for evaluating the set of feature, (avg(corrfc)) is the average of the correlation between the dependent feature and the independent feature, avg(corrff) is the average of the correlation between feature set, and k is the number of features. Correlation coefficient analysis was carried out on the dependent features of the CAD dataset which include Age, Sex, Family History, Smoking, Chest Pain, Diabetes, Glucose, Hypertension, Blood Pressure, Cholesterol, Hyperlipidemia HDL, Triglyceride, LDL, Creatinine, Body Mass and Heart Rate and Diagnosis feature which is an independent features of the CAD Dataset. (**Table 2**) and shows the r value of dependent feature against the independent feature of the dataset while (**Figure 1**) shows entire the correlation coefficient analysis matrix of the dataset features.

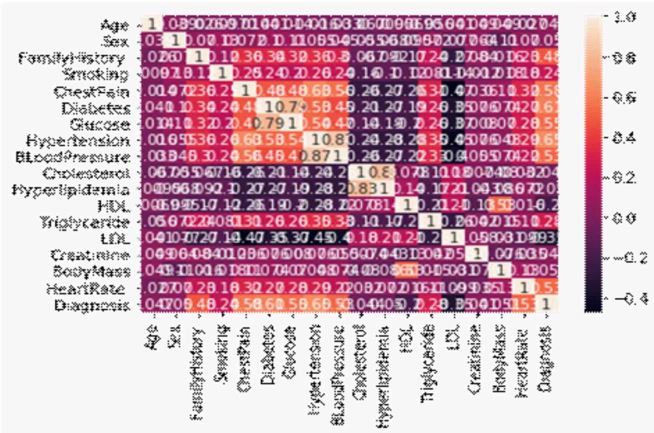

Figure 1: *The correlation coefficient analysis matrix of the dataset features.*

Table 2: *r value of the correlation coefficient analysis.*

SN	Dependent Feature	Independent feature	r value	correlation coefficient relationship
1	Age	Medical Diagnostic Result	0.42	Moderate uphill positive correlation coefficient relationship
2	Sex	Medical Diagnostic Result	0.50	Moderate uphill positive correlation coefficient relationship
3	FamilyHistory	Medical Diagnostic Result	0.48	Moderate uphill positive correlation coefficient relationship
4	Smoking	Medical Diagnostic Result	0.24	Weak uphill positive correlation coefficient relationship
5	ChestPain	Medical Diagnostic Result	0.58	Moderate uphill positive correlation coefficient relationship
6	Diabetes	Medical Diagnostic Result	0.61	Strong uphill positive correlation coefficient relationship
7	Glucose	Medical Diagnostic Result	0.55	Moderate uphill positive correlation coefficient relationship
8	Hypertension	Medical Diagnostic Result	0.65	Strong uphill positive correlation coefficient relationship
9	BLoodPressure	Medical Diagnostic Result	0.53	Moderate uphill positive correlation coefficient relationship
10	Cholesterol	Medical Diagnostic Result	0.44	Moderate uphill positive correlation coefficient relationship
11	Hyperlipidemia	Medical Diagnostic Result	-0.50	Moderate uphill negative correlation coefficient relationship
12	HDL	Medical Diagnostic Result	-0.20	weak uphill negative correlation coefficient relationship
13	Triglyceride	Medical Diagnostic Result	0.28	Weak uphill positive correlation coefficient relationship
14	LDL	Medical Diagnostic Result	0.35	Moderate uphill positive correlation coefficient relationship
15	Creatinine	Medical Diagnostic Result	0.40	Moderate uphill positive correlation coefficient relationship
16	BodyMass	Medical Diagnostic Result	0.50	Moderate uphill positive correlation coefficient relationship
17	HeartRate	Medical Diagnostic Result	0.53	Moderate uphill positive correlation coefficient relationship

We remove the all independent attributes that have less than 0.50 positive correlation coefficient relationships with dependent attributes of the dataset as shown in (**Table 3**).

Mining Dataset with selected features

Interpreting mined patterns: in this phase, the visualization the extracted hidden patterns or knowledge using C4.5 algorithm was generated and it is called decision tree. However, only independent attributes with more with 0.50 positive correlations and above were used to build the decision tree with C4.5 algorithm. Figure 2 has shown the visualization decision tree of the algorithm. C4.5 is one of the data mining classification algorithms and it is an extension of Quinlan's earlier ID3 algorithm. C4.5 algorithm was proposed by Ross Quinlan in 1993 to overcome some of the limitations of ID3 algorithm [25, 28, 51]. One the limitation of ID3 algorithm overcomes by C4.5 is ID3 sensitivity to features with number of values [30, 41].

Table 3: *Independent Attributes with more with 0.50 positive correlations and above.*

SN	Dependent Feature	Independent feature	r value	Correlation coefficient relationship
1	Sex	Medical Diagnostic Result	0.50	Moderate uphill positive correlation coefficient relationship
2	ChestPain	Medical Diagnostic Result	0.58	Moderate uphill positive correlation coefficient relationship
3	Diabetes	Medical Diagnostic Result	0.61	Strong uphill positive correlation coefficient relationship
4	Glucose	Medical Diagnostic Result	0.55	Moderate uphill positive correlation coefficient relationship
5	Hypertension	Medical Diagnostic Result	0.65	Strong uphill positive correlation coefficient relationship
6	BLoodPressure	Medical Diagnostic Result	0.53	Moderate uphill positive correlation coefficient relationship
7	BodyMass	Medical Diagnostic Result	0.50	Moderate uphill positive correlation coefficient relationship
8	HeartRate	Medical Diagnostic Result	0.53	Moderate uphill positive correlation coefficient relationship

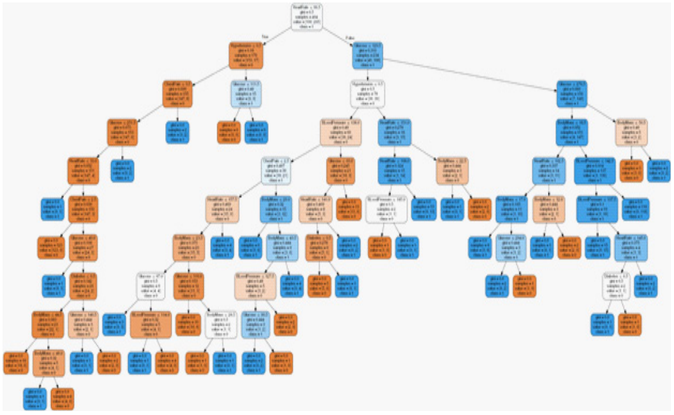

Figure 2: *Decision tree of C4.5 using CAD Dataset.*

The decision tree generated with C4.5 algorithm was converted or transformed into or crisp rules. Below are the corresponding crisp rules of generated from the decision tree in (**Figure 2**).

IF (HeartRate <99.5 mg/dl and BP <152.5 mg/dl and Hypertension =No and Diabates < 0.15 and Sex < 47) THEN Negative

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension= Yes and Diabates < 0.15 and Sex > 47) THEN Positive

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension = Yesand Diabates <2.85) THEN Negative

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension = No and Diabates =>1.52 and Chest pain= non_anginal) THEN Negative

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension =No and Diabates =>2.2 and Chest pain= asymt and BMI >=19) THEN Negative

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension =Yes and Diabates = Yes and Chest pain= asymt and BMI >=19 and Sex < 65) THEN Positive

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension >=152.2 and Diabates = Yes and Chest pain= atyp_angina and Glucose < 69.5) THEN Positive

• IF (HeartRate <99.5 mg/dl and BP < 152.5 mg/dl and Hypertension >=152.2 and Diabates = Yes1 and Chest pain= atyp_angina and Glucose >= 69.5) THEN Positive

• IF (HeartRate <=99.5 mg/dl and BP < 152.5 mg/dl and Hypertension >=152.2 and Diabates = No and Chest pain= atyp_angina and Glucose >= 69.5) THEN Negative

a) Sex: This input has two instances; either the patient is male or female. Male = 1 and Female = 0 Hence, there is no fuzziness or overlap for this input. The membership functions of the linguistic variables of the sex input is shown is (Figure 3).

Discussion

Autoimmune abnormalities are a known complication of immunotherapy drugs, which can be used to treat a wide range of malignancies including lung and breast cancer and melanoma, to name a few [2]. Specifically, autoimmune endocrine adverse effects resulting from cancer immunotherapy drugs include hypophysitis, hyperthyroidism, hypothyroidism and adrenal insufficiency. There are several immune checkpoint inhibitor drugs already on the market for cancer immunotherapy, such as the anti-CTLA-4 (ipilimumab and tremelimumab) and anti-PD1 antibodies (pembrolizumab, nivolumab and pidilizumab), all of which have been shown to have autoimmune side effects. Ipilimumab, pembrolizumab and nivolumab, however, are the most commonly used and studied for their

Table 4: *Input variables, their ranges and Linguistic Terms.*

Input (Variable)	Range	Linguistic Term
Sex	1 2	Male Female
Chest Pain	1 2 3 4	Typical angina Atypical angina Non-angina Asymptomatic
Diabetes	1 0	Yes No
Glucose	<108 mg/dL 100–126 mg/dL >120 mg/dL	Low (Normal) Normal (Prediabetes) High (Diabetes)
Hypertension	1 0	Yes No
Blood Pressure (BP)	< 134 mmHg 128 - 154 mmHg > 147 mmHg	Low (Hypotension) Normal (Nomotension) High (Hypertension)
Body Mass Index (BMI)	< 10 kg/m^2 8- 25 kg/m^2 > 22 kg/m^2	Underweight Normal Obese
Heart Rate (HR)	< 50 bpm 45 - 75 bpm > 70 bpm	Low Normal Fast
Diagnosis Result	< 4 2 - 6 4- 8 > 6	Healthy Mild Moderate Severe

a) Sex: This input has two instances; either the patient is male or female. Male = 1 and Female = 0 Hence, there is no fuzzi-

ness or overlap for this input. The membership functions of the linguistic variables of the sex input is shown is (**Figure 3**).

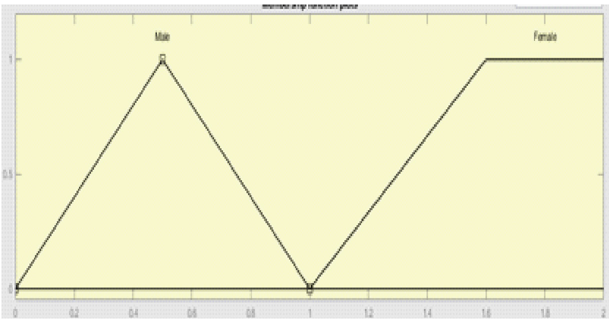

Figure 3: *Membership functions of the linguistic variables of sex.*

a. Chest pain: This input has four Chest Pain types: Typical Angina, Atypical Angina, NonAngina, and Asymptomatic. One Patient can have only one type of Chest Pain at a time. To represent Chest Pain, 1 = Typical Angina, 2 = Atypical Angina, 3 = Non-Angina and 4 = Asymptomatic. Hence, there is no fuzziness or overlap for this input, thus there are crisp set of values because the patient patient has just one chest pain at a time and the membership functions of the linguistic variables of the chest pain input is shown is (**Figure 4**).

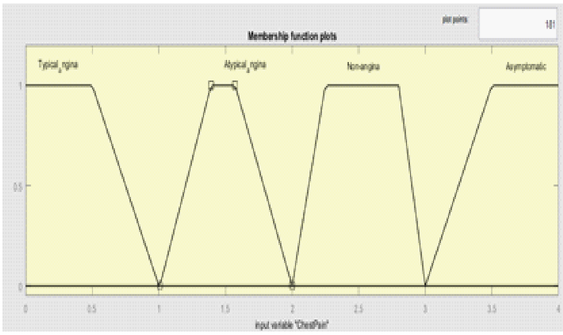

Figure 4: *Membership functions of the linguistic variables of Chest Pain.*

a. Diabetes: This input has two instances; either the patient is positive or negative. Positive = 1 and Negative = 0 Hence, there is fuzziness or overlap for this input. The membership functions of the linguistic variables of the diabetes input is shown is (**Figure 5**).

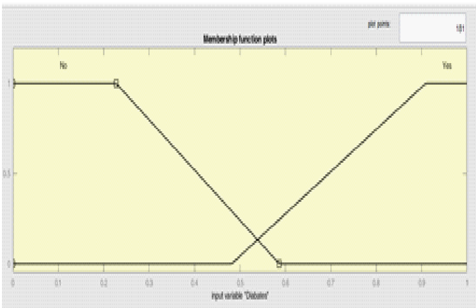

Figure 5: *Membership functions of the linguistic variables of Diabetes.*

b. Result of Diagnosis (Output): The output consists of four fuzzy sets and their linguistic variables are Healthy, Mild, Moderate and Severe. Each Linguistic variable has membership function associated with it. The membership functions of the linguistic variables of Diagnosis output is shown is (**Figure 6**).

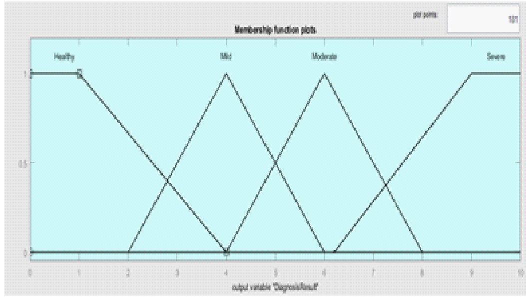

Figure 6: *Membership functions of the linguistic variables of Diagnosis.*

Fuzzification

Moreover, after determining the linguistic variable of each attribute, converting crisp value into fuzzy values, the crisp set rules generated earlier were converted into fuzzy set of rules. Fuzzification is the process of transforming crisp values into grades of membership for linguistic terms of fuzzy sets. The membership function is used to associate a grade to each linguistic term. Below are some of the fuzzy rules.

- IF (HR is Normal and BP is Normal and Glucose is Low and Diabetes is Yes and Sex is Female) THEN Healthy
- IF (HR is Normal and BP is Low and Glucose = Normal is High and Diabetes = Yes and Sex is Male) THEN Mild
- IF (HR is Low and BP is Normal and Glucose is Norma and Diabetes is Yes and Sex is Male) THEN Moderate
- IF (HR is Low and BP is Low and Glucose is Low and Diabetes is No and Sex is Female) THEN Severe
- IF (HR is Normal and BP is Normal and Glucose is High and Diabete) THEN Healthy
- IF (HR is Low and BP is Low and Glucose is High and Diabetes is Yes) THEN Mild
- IF (HR is Normal and BP is Normal and Glucose is High and Diabetes is No and Chest pain is non_anginal) THEN Healthy
- IF (HR is Low and BP is Low and Glucose is High and Diabetes is Yes and Chest pain is non_anginal) THEN Mild
- IF (HR is Low and BP is Normal and Glucose is High and Diabetes is No and Chest pain is asymt and and BMI is Normal) THEN Healthy
- IF (HR is Normal and BP is Low and Glucose is High and Diabetes is No and Chest pain is asymt and and BMI is High)THEN Mild
- IF (HR is Low and BP is Normal and Glucose is High and Diabetes is No and Chest pain is asymt and and BMI is Normal and Sex is Female) THEN Moderate

Knowledge inference or Knowledge reasoning

Knowledge inference or Knowledge reasoning: this involves application of logical rules to the knowledge to deduce new information. The inference engine draws conclusions from the replicated human expertise in the knowledge base of the expert system, which is the hallmark of the expert system [37, 40]. Mamdani inference technique is used to stimulate reasoning of expert physicians in field of diagnosis of coronary artery disease in this work.. The usage of Mamdani technique is shown in (**Figure 7**).

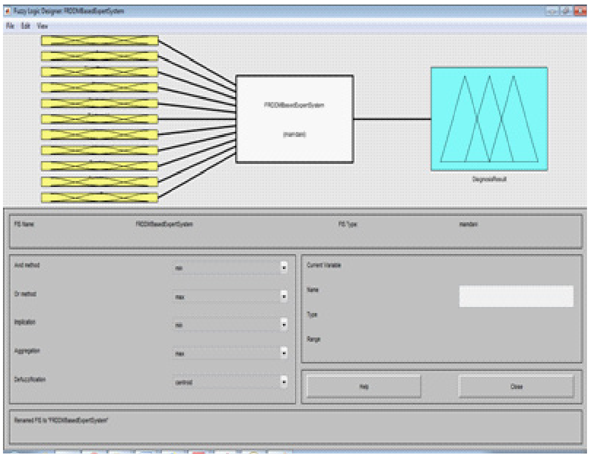

Figure 7: *Usage of Mamdani technique.*

Defuzzification

Defuzzification is the process of converting the output of inference engine (fuzzy values) into crisp values. As the name implies, defuzzification is the opposite operation of fuzzification. Since in the first procedure the crisp values of input variables are fuzzified into degree of membership with respect to fuzzy sets, the last procedure extracts a precise quantity out of the range of fuzzy set to the output variable [34-35]. The Defuzzifier technique employed in this work is Centroid. Centroid defuzzification returns the center of area under the curve [31]. Centroid Method (also called center of area or center of gravity) which is the most prevalent and physically appealing of all the defuzzification methods. It is adopted in this study. It is given by the algebraic expression below:-

$$Z_{COA} = \frac{\int_z \mu A(z).zdz}{\int_z \mu A(z)} \qquad (2)$$

Where z is the output variable, and $\mu A(z)$ is the membership function of the aggregated fuzzy set A with respect to z. the .

Fuzzy Based Expert System for Efficient Diagnosis of Coronary Artery Disease

Like any other Fuzzy Inference System, 4.6. Fuzzy Based Expert System for Efficient Diagnosis of Coronary Artery Diseasehas been implemented in MATLAB and it has five primary graphical user interfaces (GUIs) can all interact and exchange information to each other as shown in (**Figure 14**). Any of the interfaces can read and write both to the workspace and to the disk, (the read-only viewers can still exchange plots with the workspace and/or the disk). Like any fuzzy inference system, any or all of these five GUIs can be opened. (**Figure 8**) shown GUI of the expert system and other GUIs of the system include Membership Editor Viewer, Rule Editor Viewer, Rule Viewer and Surface Viewer. GUI of Membership Editor Viewer is shown in (**Figure 9**), GUI of Rule Editor Viewer is shown in (**Figure 10)**, GUI of Rule Viewer is shown in (**Figure 11** and **Figure 12**) has shown the GUI of Surface Viewer.

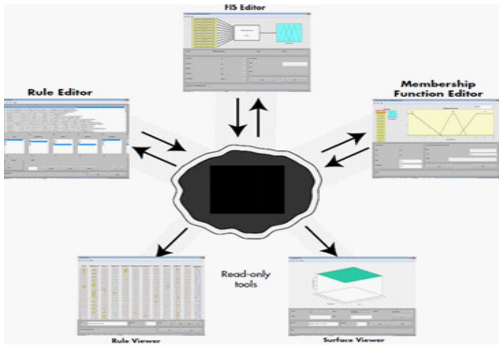

Figure 8: *GUIs of Feature Selection Based Fuzzy Expert System for Efficient Diagnosis of Coronary Artery Disease.*

Figure 9: *GUI of Membership Editor Viewer.*

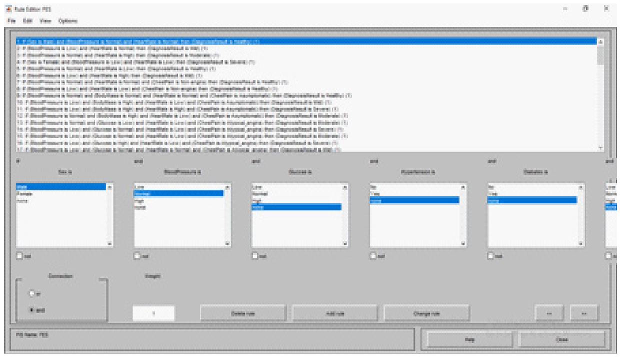

Figure 10: *GUI of Rule Editor Viewer.*

Figure 11: *GUI of Rule Viewer.*

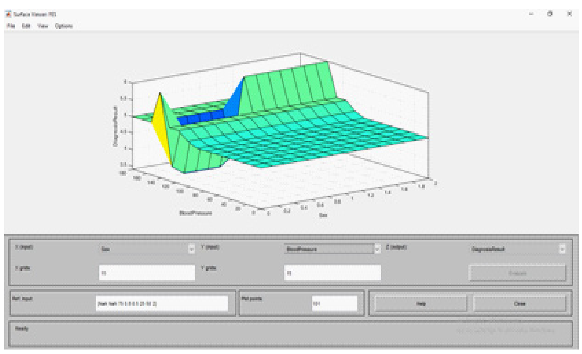

Figure 12: *Surface Viewer.*

Performance Evaluation of the Fuzzy Based Expert System

There are four classes of linguistics variable of the system output which include Healthy, Mild, Moderate, and Severe. Both Healthy, Mild patients are considered as normal patients and both Moderate and Severe are considered as abnormal (CAD) patients while evaluating the performance of the fuzzy expert system model. For the performance evaluation of the system were considered based on the following techniques:-

i. True positive (TP): It denotes the number of abnormal patients correctly classified by the framework.

ii. True negative (TN): It denotes the number of normal patients correctly classified by the framework.

iii. False positive (FP): It denotes the number of normal patients wrongly classified as abnormal patients by the framework.

iv. False negative (FN): It denotes the nu+mber of abnormal patients wrongly classified as normal patients by the framework.

v. Specificity: It is defines as percentage of normal patients classified correctly by the framework. It is determined as

$$\text{Specificity} = \frac{\text{tn}}{\text{tn} + \text{fp}} \quad (3)$$

i. Sensitivity: It is defined as the percentage of abnormal patients classified correctly by the model. It is determined as

$$\text{Sensitivity} = \frac{\text{tp}}{\text{tp} + \text{fn}} \quad (4)$$

ii. Accuracy: It denotes the percentage of correctly classified patients. In the present work (4-class problem) it is determined as:-

$$\text{Accuracy} = \frac{\text{h} + \text{i} + \text{j} + \text{k}}{\text{N}} \quad (5)$$

Where h is the number of correctly classified rules as healthy, i the number of correctly classified rules as mild, j the number of correctly classified rules as moderate, j the number of correctly classified rules severe and N is the total number of rules in the system knowledge base.

Table 5: *General Classes of patients.*

Class	Total number of rules	No. of rules Correctly classified	No. of rules wrongly classified
Healthy	44	38	4
Mild	48	46	2
Moderate	42	36	6
Severe	40	38	2
Grand total	**174**	**158**	**14**

Table 6: *Classes of normal patients.*

Class (Normal Patient)	Total number of rules	No. of rules Correctly classified	No. of rules wrongly classified
Healthy	44	38	4
Mild	48	46	2
Grand total	**92**	**84**	**6**

Table 7: *Classes of Abnormal Patients.*

Class (Abnormal Patient)	Total number of rules	No. of rules Correctly classified	No. of rules wrongly classified
Moderate	42	36	6
Severe	40	38	2
Grand total	**82**	**74**	**8**

The knowledge base of the Fuzzy System has 174 rules with four classes as shown in (**Table 8**). These rules were validated against the human expert driven data collected Federal Teaching Hospital, Gombe State. (**Table 5**) has shown the general classes of patients. There were forty four rules classified as healthy and forty eight rules classified as mild which were considered for normal or healthy patients. So total number of rules classified for healthy patient is ninety two as shown in (Table 6). while for abnormal patients, forty two rules were classified as moderate and forty rules as severe. So total number of rules classified for abnormal patients are forty-one as shown in (**Table 7)**.

Out of forty four rules classified as healthy, thirty eight rules were correctly in conformity with the expert system, so only four rules were not in conformity with the data. While Out of forty eight rules classified as mild, forty six rules were correctly in conformity with the expert system, so only two rules were not in conformity with the data. Therefore, the system correctly classified ninety two normal rules and however only six rules were wrongly classified as normal. While Out of the forty two rules classified as moderate, thirty six rules were correctly in conformity with data of the expert system, so only six rules were not in conformity with the data. While Out of forty rules classified as severe, thirty eight rules were correctly in conformity with data of the expert system, so only two rules were not in conformity with the data. Therefore, the system correctly classified seventy four abnormal rules and however eight rules were wrongly classified as abnormal. The validation result of the expert system is as follows:-

True positive (TP) = 37

True negative (TN) = 84

False positive (FP) = 4

False negative (FN) = 6

Specificity =84/92 = 91.30%

Sensitivity =37/41 = 90.24%

Accuracy =158/174 = 90.08%

Therefore, the system has achieved 90.08% overall accuracy which is very excellent, thus the accuracy determines the proportion of the total number of predictions that were correct. At the same time, the system has 91.30%accuracy to classify of normal patients correctly by adopting the proposed framework (specificity) and 90.24% accuracy to classify abnormal patients correctly by adopting the proposed framework (sensitivity). This showed that, the system performed efficient and excellently to diagnose CAD.

Conclusion

Coronary Artery Disease (CAD) is one of the deadliest diseases in the world and in Nigeria, the CAD is at the moment gaining much popularity following the rising number of health issues related to the disease, including higher death rate, which is mostly due to lack of proper awareness among the common people. Expert systems have been specifically applied in a variety of life sciences support systems development, ranging from storage and retrieval of medical records, diagnostics, up to expert knowledge/decision support systems. In this work, a fuzzy based expert system for supporting and completing human expertise in the diagnosis of CAD has been developed and evaluated. The system archived 90.08% accuracy, 91.30% specificity and 90.24% sensitivity respectively, which showed that, the system performed efficiently and excellently to diagnose CAD and it can deployed and used in the hospitals in Nigeria.

Acknowledgement

This work was supported by the Tertiary Education Trust Fund, Nigeria (TETFUND), as an Institution Based Research Fund (IBR) for Federal University of Kashere, Gombe State, Nigeria.

References

1. A Saeed and K Asieh. “Identification and Classification of Coronary Artery Disease Patients using Neuro-Fuzzy Inference Systems”, Journal of mathematics and computer Science. 2014; 13:136-141.

2. AA Haruna, LJ Muhammad, BZ Yahaya et al. “An Improved C4.5 Data Mining Driven Algorithm for the Diagnosis of Coronary Artery Disease”. International Conference on Digitization (ICD), Sharjah, United Arab Emirates. 2019; 48-52.

3. AV Senthil Kumar. “Diagnosis of heart disease using advanced fuzzy resolution mechanism”, Journal of Artificial Intelligence. 2013; 1-9.

4. AS Noor, PA Venkatachalam and FH Ahmad. “Diagnosis of Coronary Artery Disease Using Artificial Intelligence Based Decision Support System”, Proceedings of the International Conference on Man-Machine Systems (ICoMMS). 2019; 11-13.

5. A Abraham, “Rule-based Expert Systems” Handbook of Measuring System Design, Wiley & Sons, Ltd, 2005.

6. A Ali and N Mehdi “A Fuzzy Expert System for Heart Disease Diagnosis” Proceedings of the International Multi Conference of Engineers and Computer Scientists. 2010; 1:17-19.

7. American Heart Association (AHA). Heart disease and stroke statistics -at a glance Retrieved from https://www.heart.org/idc/groups/ahamah-public/@wcm/@sop/@smd/documents/downloadable/ucm_480086.pdf Date: 16th January, (2019).

8. CU Nwaneli, “Changing Trend in Coronary Heart Disease in Nigeria”, Africa medical Journal. 2010; 1(1):1-4.

9. C Enrico. “The Guide to Health Informatics”, 3rd Edition. London, Arnold, 2003.

10. D Resul, T Ibrahim and S Abdulkadir. “Effective diagnosis of heart disease through neural networks ensembles”, Expert Systems with Applications. 2009; 36:7675-7680.

11. E Turban and E Jay. “Decision Support Systems and Expert Systems”. 6th Edition, Prentice Hall, Upper Saddle River, NJ, 2011.

12. FS Ishaq, LJ Muhammad, YZ Yahaya, et al. Fuzzy-Based Expert System for Diagnosis of Diabetes Mellitus. International Journal of Advanced Science and Technology. 2020; 136:39-50

13. FS Ishaq, LJ Muhammad, BZ Yahaya, Y Atomsa. “Data mining driven models for diagnosis of diabetes mellitus: a survey”, Indian Journal of Science and Technology. 2018; 11:41.

14. H Nasruddin, RS Osama, MK Ahmed and AG Mohammed. “Fuzzy Soft Expert System in Prediction of Coronary Artery Disease”, International Journal of Fuzzy Systems, 2016.

15. J Giarratano and G Riley. “Expert Systems: Principles and Programming”, PWS-Kent Publishing Co, Boston, M.A, 1989.

16. KP Debabrata, AV Aruldoss, P Renukadevi, D Devaraj. “Design of fuzzy expert system for microarray data classification using a novel Genetic Swarm Algorithm”, Expert Systems with Applications. 2012; 39:1811-1821.

17. K Humar and A Novruz. “Design of a Hybrid System for the Diabetes and Heart Diseases”, Expert Systems with Applications. 2008; 35:82-89.

18. K P Ganesh, A V Aruldoss, P Renukadevi and D Devaraj. “Design of fuzzy expert system for microarray data classification using a novel Genetic Swarm Algorithm”, Expert Systems with Applications. 2012; 39:1811-1821.

19. LN Rani, S Defit, LJ Muhammad. “Determination of Student Subjects in Higher Education Using Hybrid Data Mining Method with the K-Means Algorithm and FP Growth”, International Journal Of Artificial Intelligence Research. 2021; 5(1).

20. LJ Muhammad, EA Algehyne. “Fuzzy based expert system for diagnosis of coronary artery disease in Nigeria”, Health Technol. 2001; 11:319-329.

21. LJ Muhammad, EJ Garba, ND Oye, GM Wajiga, AB Garko. “Fuzzy rule-driven data mining framework for knowledge acquisition for expert system”. In: Translational Bioinformatics in Healthcare and Medicine. Elsevier, Academic Press. 2021; 201-214.

22. LJ Muhammad, AH Ahmad, AM Ibrahim, A Mansir, B Bature, MA Jamila. “Performance Evaluation of Classification Data Mining Algorithms On Coronary Artery Disease Dataset”, IEEE 9th International Conference on Computer and Knowledge Engineering (ICCKE 2019), Ferdowsi University of Mashhad, October, 2019.

23. LJ Muhammad et al. “Deep Learning Models for Predicting COVID-19 Using Chest X-Ray Images”. In Trends and Advancements of Image Processing and Its Applications. EAI/Springer Innovations in Communication and Computing. Springer, Cham, 2022.

24. LJ Muhammad, EA Algehyne, SS Usman. “Predictive Supervised Machine Learning Models for Diabetes Mellitus”. Springer Nature Computer Science, 2020.

25. LJ Muhammad, M Besiru Jibrin, BZ Yahaya, IA Mohammed Besiru Jibrin, A Ahmad and JM Amshi, “An Improved C4.5 Algorithm using Principle of Equivalent of Infinitesimal and Arithmetic Mean Best Selection Attribute for Large Dataset,” 2020 10th International Conference on Computer and Knowl-

edge Engineering (ICCKE), Mashhad, Iran, 2020; 006-010.

26. LJ Muhammad, EJ Garba, ND Oye, and GM Wajiga. “On the Problems of Knowledge Acquisition and Representation of Expert System for Diagnosis of Coronary Artery Disease (CAD)”, International Journal of u- and e- Service, Science and Technology. 2018; 11(30):49-58.

27. LJ Muhammad, I Al-Shourbaji, AA Haruna et al. “Machine Learning Predictive Models for Coronary Artery Disease”. SN COMPUT. SCI. 2021; 2:350.

28. LJ Muhammad et al. Using Decision Tree Data Mining Algorithm to Predict Causes of Road Traffic Accidents, its Prone Locations and Time along Kano –Wudil Highway. International Journal of Database Theory and Application, 2017; 10:197-208.

29. LJ Muhammad, SS Usman. Power of Artificial Intelligence to Diagnose and Prevent Further COVID-19 Outbreak: A Short Communication 2020.

30. LJ Muhammad, MM Islam, Usman SS. et al. Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients’ Recovery. Springer Nature Computer Science, 2020.

31. LJ Muhammad, BZ Yahaya, A Garba. et al. Multi Query Optimization Algorithm Using Semantic and Heuristic Approaches, International Journal of Database Theory and Application, 2016; 6(9).

32. L Adel, N Raja, Z Roziati and B Awang. “Design of a Fuzzy-based Decision Support System for Coronary Heart Disease Diagnosis”, Journal of Medical System, 2012.

33. M Zahra and SA Mohammad. “CADICA: Diagnosis of Coronary Artery Disease Using the Imperialist Competitive Algorithm”, Journal of Computing Science and Engineering, 2014; 8(2):87-93.

34. MG, Tsipouras, TP Exarchos, DI Fotiadis, A Kotsia, A Naka and LK Michalis, “A Decision Support System for the Diagnosis of Coronary Artery Disease”, Proceedings of the IEEE Symposium on Computer-Based Medical Systems, 2008.

35. M Islam, S Mahmud, LJ Muhammad. et al. Wearable Technology to Assist the Patients Infected with Novel Coronavirus (COVID-19). SN COMPUT. SCI. 2020; 1:320.

36. M Idris, A Pervez, JA Tariq and SZ Syed. “Fuzzy Rule Based Classification for Heart Dataset using Fuzzy Decision Tree Algorithm based on Fuzzy RDBMS” World Applied Sciences Journal. 2013; 28(9):1331-1335.

37. ND Wong. “Epidemiological studies of CHD and the evolution of preventive cardiology.”. Nature reviews. Cardiology. 2014; 11(5):276-89.

38. N Jesmin and I Tasadduq. “Association rule mining to detect factors which contribute to heart disease in males and females”, Journal of Expert Systems with Applications. 2013; 40:1086–1093.

39. OC Akinyokun, GB Iwasokun, SA Arekete and RW. Samuel Fuzzy logic-drive expert system for the diagnosis of heart failure disease. Artificial Intelligence Research, 2015; 4(1).

40. OA Sarumi, O Aouedi, LJ Muhammad. “Potential of Deep Learning Algorithms in Mitigating the Spread of COVID-19”. In Understanding COVID-19: The Role of Computational Intelligence. Studies in Computational Intelligence, 2021; 963.

41. O Alsayed, MSM Rahim, IA Bidewi, et al. Selection of the Right Undergraduate Major by Students Using Supervised Learning Techniques. Appl, 2021.

42. PK Anooj. “Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules, Journal of Computer and Information Sciences”, 2012; 24:27-40.

43. P Debabrata, KM Mandana, P Sarbajit, S Debranjan and C Chandan. “Fuzzy expert system approach for coronary artery disease screening using clinical parameters”, Journal of Knowledge-Based Systems, 2012; 36:162-174.

44. RA Soltan, MZ Rashad and B El-Desouky. “Diagnosis of Some Diseases in Medicine via computerized Experts”, International Journal of Computer Science & Information Technology (IJCSIT), 2013; 5(5).

45. SD Chaitrali and SA Sulabha,“Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques, International Journal of Computer Applications, 2012; 47(7):0975 -888 .

46. TA Ibrahim, AS Adejoke, A Moses, CA Tony, BZ Ibrahim, and HW Muhammad. “ICT Knowledge, Utilization and Perception among Healthcare Providers at National Hospital Abuja, Nigeria”, American Journal of Health Research. Special Issue: Health Information Technology in Developing Nations: Challenges and Prospects Health Information Technology. 2015; 3(1):47-53.

47. WM Wiharto, K Hari, and H Herianto. “Intelligence System for Diagnosis Level of Coronary Heart Disease with K-Star Algorithm”. Healthcare Informatics Research, 2016; 22(1):30-38.

48. World Health Organization (WHO). (2015). Cardiovascular diseases, factsheet#317.Retrieved from http://www.who.int/mediacentre/factsheets/fs317/en/ Date: 16th January, 2019.

49. YD Niranjana and S Anto. “An Evolutionary-Fuzzy Expert System for the Diagnosis of Coronary Artery Disease”, International Journal of Advanced Research in Computer Engineering & Technology, 2014; 3(4).

50. Y Atomsa, LJ Muhammad, FS Ishaq, et al. “Expert system for diagnosis of coronary artery disease: A survey” Journal of Clinical Images and Medical Case Reports, 2021; 2.

51. Z Yahaya, LJ Muhammad, N Abdulganiyyu, FS Ishaq. “An improved C4. 5 algorithm using L’Hospital rule for large dataset”, Indian Journal of Science and Technology, 201.