Expert system for diagnosis of coronary artery disease: A survey

Atomsa Y; Muhammad LJ*; Ishaq FS; Yusuf Abdullahi

1Mathematics & Computer Science Department, Federal University of Kashere, Gombe State, Nigeria. 2Biological Sciences Department, Federal University of Kashere, Gombe State, Nigeria.

Abstract
Coronary Artery Disease (CAD) is one of the most dangerous diseases which lead to sudden cardiac death. 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’ patients. This study presents a state-of-the-art of the methods and techniques used for the development of expert system to diagnose CAD by scholars. The study found fuzzy logic as the most frequently used and successful technique for the development of the expert system to diagnose CAD. While data mining techniques found to be the second and neural network as third most frequently used and successful technique for the development of the expert system to diagnose CAD. The study further found that, 40% of the studies reviewed to use hybrid approach where fuzzy logic and data mining techniques were used in the development of the expert system for diagnosis of CAD.

Keywords: expert system; artificial intelligence; fuzzy logic; data mining; coronary artery disease.

Introduction
From earliest moments in the history of the computer science, computer scientists have dreamed of creating an electronic brain of all the modern technological quests. The search led to create artificial intelligence systems which have one of the most ambitious and not surprisingly, controversial. It also seems that very early on, scientists and doctors alike were captivated by the potential such a technology might have in medicine. With intelligent computers able to store and process vast stores of knowledge, the hope was that they would become perfect ‘doctors in a box’, assisting or surpassing clinicians with tasks like diagnosis [8].

With such motivations, a small but talented community of computer scientists and healthcare professionals set about shaping a research program for a new discipline called Artificial Intelligence in Medicine (AIM) [10]. These researchers had a bold vision of the way AIM would revolutionize medicine, and push forward the frontiers of technology which led the emergence expert system [8].

In this study, research and conference papers published by reputable publishers that employed expert system for diagnosis of coronary artery disease were searched and reviewed in this work.

Coronary artery disease
Coronary Artery Disease (CAD) is also known as Ischemic Heart Disease (IHD) or Coronary Heart Disease (CHD) [34]. CAD is one of the deadliest diseases in the world. The disease occurs when plaque builds up inside the coronary arteries. The arteries supply heart muscle with oxygenated blood [13]. Plaque is made up of cholesterol, fat, calcium, and other substances
found in the blood [35,36]. Therefore, over time, plaque narrows and hardens arteries that resulted of reduction of blood flow to heart muscles. Eventually, an area of plaque can cause a blood clot to form on the surface of the plaque or rupture. If the clot becomes large enough, it can mostly or completely block the flow of oxygen-rich blood to the part of the heart muscle fed by the artery [21]. This can lead to the following

i. **Angina:** chest pain or discomfort that occurs when inadequate amount of oxygen-rich blood is flowing to an area of heart muscle. Angina may feel like pressure or squeezing in chest. The pain also may occur in shoulders, arms, neck, jaw, or back [21].

ii. **Heart attack:** occurs when blood flow to an area of the heart muscle is completely blocked. This prevents oxygen-rich blood from reaching that area of heart muscle, causing it to die. Without quick treatment, a heart attack can lead to serious problems or death [21].

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 [30]. In the developed countries CAD is one of the the number one killers with over 7.4 million deaths attributed to (World Health Organization [31]. It has been estimated that, CAD is one in every seven deaths in the United States is due to heart disease [37]. According to [6] 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 one of the most common cause of death in the United States. According [31], CAD is not just an American problem but however the problem of the world in general and the disease causes more deaths and disabilities and is responsible for more economic costs than any other single illness.

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 [38]. 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. Most of the victims of Coronary Artery Disease in Nigeria usually neglect or handle the initial symptoms casually and only consult medical attention when those symptoms become severe or critical. But at that time the treatment becomes complicated and sometimes due to the acuteness and severity of the disease, the patients die before getting proper medication.

Several research efforts have shown that CAD has overtaken infectious disease to become the foremost most fatal disease in Nigeria [7]. This epidemic invasion of the CAD into Nigerian society, combined with the challenges of an apparently weak health sector in Nigeria, which lacks the requisite capacity and needed human and financial resources prevent proper treatment of the disease and its risk factors, only aggravates the situation. Substantial proactive measures are therefore necessary to properly address this epidemic [39,40]. The use of information and computing technologies to improve on the processes of health care delivery and management is not a new endeavor [29,32].

**Expert system**

According to [15] expert system is a computer program that uses knowledge and inference procedures to solve problems that was difficult enough to acquire significant human expertise for their solutions. It can also be defined as a computer system that uses human knowledge to solve problems that normally would require human intelligence [4]. The expert systems represent the expertise knowledge as data or rules within the computer. These rules and data can be called upon when needed to solve problems [10]. Expert systems are only those systems where the knowledge for the definition of the system (regardless of the importance of its role in the performance, or the particular method of representation) comes from a human whose ability to solve problems is well beyond the common, up to the rank of unusual expertise [32]. According to [24], 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. 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 [16].

There are many application areas of expert system such as medicine, education, agriculture, oil and gas, environment, law, manufacturing, telecommunication and power systems etc. [10]. In contrast to conventional computer programs where the knowledge base is often embedded in the program codes, so that as the knowledge changes, the program has to be rebuilt, the knowledge-based expert systems collect the small fragments of human knowledge into a knowledge-base, which is used to reason through a problem, using knowledge that is appropriate [4]. An important advantage here is that within the domain of the knowledge-base, a different problem can be solved using the same program without programming efforts. Also, expert systems have the ability to explain the reasoning process and handle levels of confidence and uncertainty that conventional algorithms could not handle [12].

Expert systems are the commonest type of Artificial Intelligence systems in routine clinical use. They contain medical knowledge, usually about a very specifically defined task, and are able to reason with data from individual patients to come up with reasoned conclusions. There are many different types of clinical task to which expert systems can be applied.

i. **Diagnostic assistance:** When a patient’s case is complex, rare or the person making the diagnosis is simply inexperienced, an expert system can help come up with likely diagnoses based on patient data.

ii. **Therapy critiquing and planning:** Systems can either look for inconsistencies, errors and omissions in an existing treatment plan, or can be used to formulate a treatment based upon a patient’s specific condition and accepted treatment guidelines.

iii. **Image recognition and interpretation:** Many medical images can now be automatically interpreted, from plane X-rays through to more complex images like angiograms,
Methods

In this study research and conference papers published by reputable publishers that employed expert system for diagnosis of coronary artery disease were searched and reviewed in this work.

In Ref [2] an expert system for the diagnosis of the level of coronary heart disease by taking into account the problem of data imbalance has been developed. The first stage of the research was preprocessing, which included resampled non-stratified random sampling (R), the synthetic minority over-sampling technique (SMOTE), clean data out of range attribute (COR), and Remove Duplicate (RD). The second step was the sharing of data for training and testing using a k-fold cross-validation model and training multiclass classification by the K-star algorithm. The third step was performance evaluation.

In the study of [11] 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. The data used in the study was obtained from the Cardiac Unit, Department of Cardiology, Faculty of Medicine, Assiut University, Egypt of 200 patients containing 76 attributes. But however only six of the attributes were considered because the author of the study opinioned on these six attributes are relevant to coronary artery disease which include blood pressure, cholesterol, maximum heart rate, blood sugar, old peak, age.

In ref. [23] a web based Fuzzy Logic-based Expert System for the diagnosis of heart failure disease was developed. The system consists of a Knowledge Base (which is made up of a Database), a fuzzy logic component, a Fuzzy Inference Engine and a Decision Support Engine which comprises of cognitive and emotional filter as well as Tele medicine facilities. The system was implemented using Hypertext Preprocessor (PHP), JavaScript and Hypertext Mark-up Language (HTML) with My Structured Query Language (MySQL) as the Database Management System.

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 [32]. The decision tree was used to select the most significant attributes and the output is converted into crisp if-then rules. The crisp sets of rules are transformed into the fuzzy rules and these rules constitute the fuzzy rule base. Genetic Algorithm (GA) is used to tune the fuzzy membership functions and the optimized of membership functions using GA helps to achieve better accuracy.

In the study [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). Network parameters, introductory and lower parameters, were trained by back-propagation error (gradient descent) method. The proposed method was evaluated through data collected from medical files of 152 patients with coronary angiography in Kowsar Hospital, Shiraz, Iran during September, 2013. The performance indicators of this system were specificity and sensitivity. The indicators, as extracted from testing results, were found to be 0.88 and 1, respectively.

A fuzzy expert system for diagnosis of coronary artery disease by a non-invasive procedure was implemented in the study of [18]. The imperialist competitive algorithm is used to adjust the fuzzy membership functions. The proposed method has been evaluated with the Cleveland and Hungarian datasets. The advantage of the method adopted in the study compared with others, is its interpretability.

The adaptive neuro fuzzy inference system and Advanced fuzzy resolution mechanism has been proposed in the study of [2] to diagnose the heart disease. The Advanced fuzzy resolution mechanism was designed with predictive value and if then rules to diagnosis the heart disease.

In ref. [22] a computer intelligent based approach for the diagnosis of heart diseases have been developed. Apriori, Predictive Apriori and Tertius were the three different rule mining algorithms used to present rule extraction experiment on heart disease data and showed as efficiency algorithm for diagnosis task.

A fuzzy rule-based system which concentrated only on accuracy and interpretability of the system was proposed by [20]. Fuzzy decision tree method was used based on fuzzy Relational Database Management System (RDBMS) and rules were generated based on C4.5 algorithm known as Fuzzy Rule Generation System (FRGS) algorithm. A fuzzy decision tree was developed by converting a medical application of heart relational database to fuzzy heart relational database and then the decision tree using C4.5 algorithm. However, the study concentrates only on the interpretability of the output of the system rather than accuracy of the generated rules and knowledge acquisition of the system fuzzy decision tree method based on fuzzy Relational Database Management System (RDBMS).

In Ref. [27] an expert system that provided a heart disease patient with background for suitable diagnosis and treatment (Especially Angina Pectoris and Myocardial infarction) has been developed. The system was composed of four stages. The first stage was receiving the symptoms from the patient. The second stage was requesting from the patient to make some analysis and investigation to help the system to make a correct decision in the diagnosis. The third stage was doing diagnosis of patient according to information from patient (symptoms, analysis and investigation). The four stage was determining the name of appropriate medication or what should be done until the patient recovers (step therapy). The system was only able to give appropriate diagnosis and treatment for two heart diseases namely; angina pectoris and infarction. The programs used for diagnosis and system analysis include CLIPS and PROLOG.

In ref. [25] a weighted fuzzy rule-based Clinical Decision Support System (CDSS) for computer-aided diagnosis of the heart disease has been developed. The proposed CDSS for risk prediction of the heart patients contains two steps such as: generation of weighted fuzzy rules and developing of a fuzzy rule-based decision support system. The automatic procedure to generate the fuzzy rules was an advantage of the proposed system and the weighted procedure introduced in the proposed work was additional advantage for effective learning of the fuzzy system. These weighted fuzzy rules were used to build the CDSS using Mamdani fuzzy inference system.
The different data mining techniques such as neural networks, decision trees and naive bayes has been proposed by [28] for the study of heart disease prediction system. The multi-layer perceptron neural networks were used to map the input data onto the output data.

A coronary artery disease fuzzy expert system for microarray data classification using a novel Genetic Swarm Algorithm, has been proposed in the study of [33] for obtaining near rule set and membership function tuning. The convergence of genetic swarm algorithm and their classification accuracy were improved by using the advanced and problem specific genetic operators. The major disadvantages in the study is that, the probability distribution of genes is not here to compute the mutual information and the rule set is not tuned properly due to rounding off problems.

In the study of [26] a screening system has been developed for the early detection of Coronary Artery Disease. The system is able to identify the risk group of patients at an early stage for proper guidance and medication. A well-defined questionnaire was developed advised by domain experts and sample data are collected from the patients. The risk factors identified, are mainly based on Indian population and they have been identified from the past experience and the expertise of the medical experts. These factors are age, smoking habit, having obesity, hypertension, diabetes, and hyperlipidemia. Symptoms considered for first level information are angina pectoris or chest pain.

In the study of [17] a fuzzy rule-based system was designed to serve as a decision support system for diagnosis Coronary heart disease. A multi-objective genetic algorithm was employed optimize both the accuracy and transparency of the system. Furthermore, a new way for employing Ensemble Classifiers Strategy (ECS) method was also proposed to enhance the classification ability of the system. The result of the study showed the rules generated by the system are humanly understandable and the system is able to identify the uncertainty cases so that the physician can give a special consideration to deal with them.

In ref [5] a Fuzzy Expert System for heart disease diagnosis using V. A. Medical Center, Long Beach and Cleveland Clinic Foundation database has been developed. The system is design with in Matlab software and it is viewed as an alternative for existing methods to distinguish of heart disease presence.

In ref [3] a fuzzy decision support system for the diagnosis of coronary artery disease based on evidence has been developed. The coronary artery disease data sets were collected from University California Irvine (UCI). The knowledge base of the system is taken by using rules extraction method based on Rough Set Theory. The rules then are selected and fuzzified based on information from discretization of numerical attributes. The result of the study showed that the system is able to give the percentage of coronary artery blocking better than cardiologists and angiography.

The neural networks ensemble method has been proposed in the study of [9] for the effective diagnosis of the heart disease. This ensemble method was able to create new models by combining the posterior probabilities or the predicted values from multiple predecessor models.

A rule-based Decision Support System was developed and implemented in study of [19] for the diagnosis of Coronary Artery Disease. The generation of the decision support system is realized automatically using a three stage methodology which induction of a decision tree from a training set and extraction of a set of rules using id3 clarification algorithm, transformation of the set of rules into a fuzzy model and optimization of the parameters of the fuzzy model.

The fuzzy neural network and k-fold cross validation has been proposed in the study of [14] a hybrid system for the diabetes and heart diseases. The back propagation algorithm was used to train a fuzzy network. The classification accuracies of the datasets were obtained by the k-fold cross validation. The major disadvantage of study is that, the missing values are not handled properly.

Results and discussion

The results of this study have been drawn from the published articles reviewed in the literature review of the study which termed a methodology. Fuzzy logic has emerged the most frequently used and successful technique for the development of the expert system for diagnosis of CAD. While data mining techniques have emerged as the second while neural network emerged as third most frequently used and successful technique for the development of the expert system for diagnosis of CAD. It should be noted that, 40% of the studies reviewed in this study used hybrid approach where fuzzy logic and data mining techniques were used in the development of the expert system for diagnosis of CAD. Table 1 shown the summary of the reviewed papers with the method/tools/techniques used for development expert system to diagnose CAD.

<table>
<thead>
<tr>
<th>SN</th>
<th>Ref.</th>
<th>Method/Tools for development</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[30]</td>
<td>Data mining</td>
</tr>
<tr>
<td>2</td>
<td>[11]</td>
<td>fuzzy soft sets</td>
</tr>
<tr>
<td>3</td>
<td>[23]</td>
<td>Fuzzy Logic, web development tools PHP, JavaScript and Hypertext Mark-up Language, MySQL</td>
</tr>
<tr>
<td>4</td>
<td>[32]</td>
<td>fuzzy logic and Genetic Algorithm</td>
</tr>
<tr>
<td>5</td>
<td>[1]</td>
<td>fuzzy logic and MATLAB</td>
</tr>
<tr>
<td>6</td>
<td>[18]</td>
<td>fuzzy logic, competitive algorithm and MATLAB</td>
</tr>
<tr>
<td>7</td>
<td>[2]</td>
<td>adaptive neuro fuzzy</td>
</tr>
<tr>
<td>8</td>
<td>[22]</td>
<td>Apriori Predictive data mining technique</td>
</tr>
<tr>
<td>9</td>
<td>[20]</td>
<td>C4.5 Data mining algorithm/ DBMS and Fuzzy Logic</td>
</tr>
<tr>
<td>10</td>
<td>[27]</td>
<td>Logic, CLIPS and PROLOG.</td>
</tr>
<tr>
<td>12</td>
<td>[28]</td>
<td>as neural networks, decision trees and naive bayes</td>
</tr>
<tr>
<td>13</td>
<td>[32]</td>
<td>fuzzy logic and novel Genetic Swarm Algorithm</td>
</tr>
<tr>
<td>14</td>
<td>[26]</td>
<td>fuzzy logic and C4.5 data mining algorithm</td>
</tr>
<tr>
<td>15</td>
<td>[17]</td>
<td>fuzzy logic and Ensemble Classifiers Strategy (ECS) method</td>
</tr>
<tr>
<td>16</td>
<td>[5]</td>
<td>fuzzy logic, data mining algorithm and MATLAB</td>
</tr>
<tr>
<td>17</td>
<td>[3]</td>
<td>fuzzy logic, Rough Set Theory and data mining algorithm</td>
</tr>
<tr>
<td>18</td>
<td>[9]</td>
<td>neural networks</td>
</tr>
<tr>
<td>19</td>
<td>[19]</td>
<td>fuzzy logic and Id3 decision tree data mining algorithm</td>
</tr>
<tr>
<td>20</td>
<td>[14]</td>
<td>fuzzy logic and neural network</td>
</tr>
</tbody>
</table>
## References


