Scientific Research Publishing

Int. J. Communications, Network and System Sciences, 2023, 16, 77-96
https://www.scirp.org/journal/ijcns
ISSN Online: 1913-3723
ISSN Print: 1913-3715

Development of a Data Mining Model to Detect Cardiovascular Disease

Emmanuel Nwabueze Ekwonwune[1], Chinyere Ihekachi Ubochi[1], Augustine Ekekwe Duruoha[2], Ngozi Amarachi Duru[3]

[1]Department of Computer Science, Imo State University, Owerri, Nigeria
[2]Department of Computer Science, Gregory University, Abia, Nigeria
[3]Department of Computer Science, Federal University of Technology, Owerri, Nigeria
Email: ekwonwuneemmanuel@yahoo.com

How to cite this paper: Ekwonwune, E.N., Ubochi, C.I., Duruoha, A.E. and Duru, N.A. (2023) Development of a Data Mining Model to Detect Cardiovascular Disease. *Int. J. Communications, Network and System Sciences*, **16**, 77-96.
https://doi.org/10.4236/ijcns.2023.165006

Received: November 11, 2022
Accepted: May 16, 2023
Published: May 19, 2023

Abstract

In the previous publication on Volume 15 No 9, September 30, 2022 of IJCN, we analyzed "Data Mining as a Technique for Healthcare Approach". In this edition, emphasis has been made on the "Development of Data Mining Model to Detect Cardiovascular Diseases (CVD)". A Software was developed using the internationally accepted Software Engineering Methodology (SSADM), coding by OOP and packing by prototyping methodologies. Among others, this paper discusses; Cardiovascular diseases, Data Mining Algorithm, Analysis and Information flow of the Present System, Data flow and High level flow of the Proposed System, Modulating, System Design and Development, Hardware and Software Specifications, System Testing, Evaluation and Documentation.

Keywords

Data Mining, Algorithm, System Design, Healthcare, CVD, Methodology

1. Introduction

Data Mining is one of the most vital and motivating areas of research with the objective of finding meaningful information from huge data sets. In present era, Data Mining is becoming popular in healthcare field because there is a need for efficient analytical methodology for detecting unknown and valuable information in health data. In health industry, Data Mining provides several benefits such as detection of the fraud in health insurance, availability of medical solution to the patients at lower cost, detection of causes of diseases and identification of medical treatment methods. It also helps the healthcare researchers for making efficient healthcare policies, constructing drug recommendation sys-

tems, developing health profiles of individuals etc. [1]. The data generated by the health organizations is very vast and complex due to which it is difficult to analyze the data in order to make important decision regarding patient health. This data contains details regarding hospitals, patients, medical claims, treatment cost etc. So, there is a need to generate a powerful tool for analyzing and extracting important information from this complex data. The analysis of health data improves the healthcare by enhancing the performance of patient management tasks. The outcome of Data Mining technologies is to provide benefits to healthcare organization for grouping the patients having similar type of diseases or health issues so that healthcare organization provides them with effective treatments. It can also be useful for predicting the length of stay of patients in hospital, for medical diagnosis and making plan for effective information system management. Recent technologies are used in medical field to enhance the medical services in cost effective manner. Data Mining techniques are also used to analyze the various factors that are responsible for diseases for example type of food, different working environments, education level, living conditions, availability of pure water, health care services, cultural, environmental and agricultural factors as shown in **Figure 1**.

2. Conceptual Framework

The healthcare industry battles with millions of digitally recorded data and patterns being collected at enormous speed due to the wide spread usage of powerful computer devices nowadays [2]. These data collected are mostly unorganized

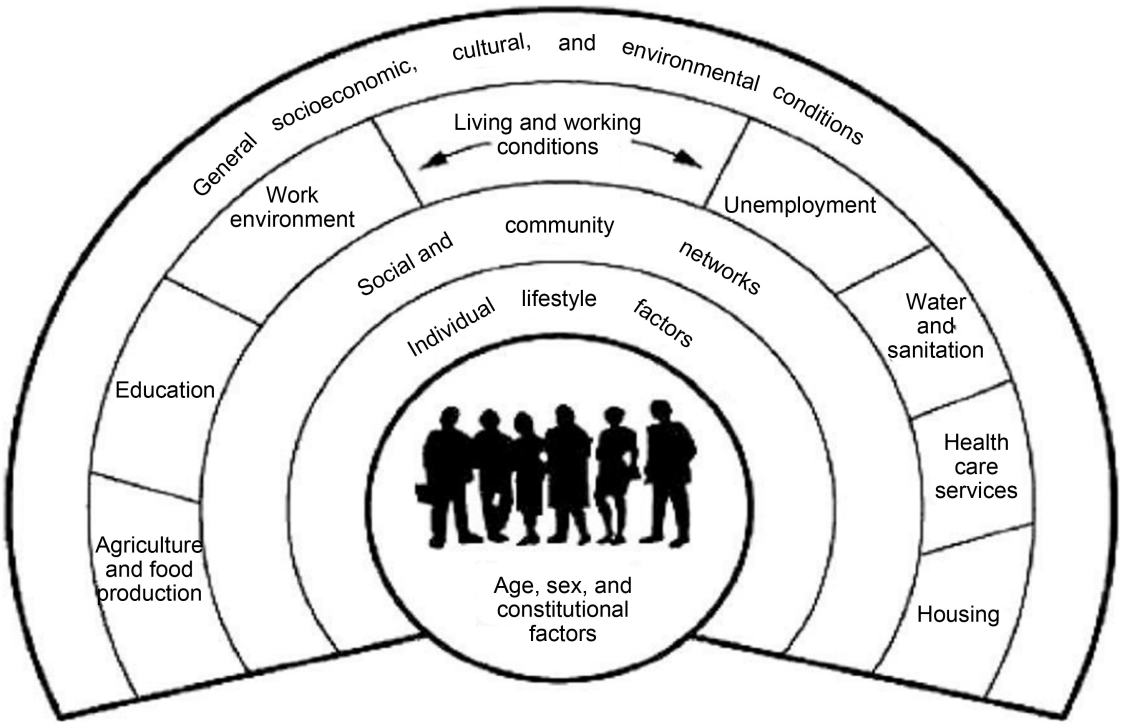

Figure 1. Factors responsible for diseases [2].

and have not been used properly for appropriate applications, thus, imposing new challenges regarding their management including their modeling, storage, and retrieval capabilities. There is often interesting knowledge in the data that is not readily evident. The spread of electronic patient records, with their computer-readable entries e.g. Magnetic Resonance Imaging (MRI), signals like ECG (Electrocardiography), clinical information like blood sugar, blood pressure, cholesterol levels etc. as well as the physician's interpretation is opening new possibilities for medical data mining and a world of virtual research [3].

Knowledge Discovery in Databases (KDD) and Data Mining (DM) provide a solution to the information flood problem by extracting valid, novel, potentially useful, and ultimately understandable patterns from data [4]. Patterns constitute compact and rich in semantics representations of raw data [5]; compact by means that they summarize, to some degree, the amount of information contained in the original raw data and rich in semantics by means that they reveal new knowledge hidden in the abundance of raw data.

Several pattern types exist in the literature, mainly due to the wide heterogeneity of data and data mining applications, as well as due to the large variety of pattern extraction techniques as a result of the different goals that a mining process tries to accomplish (*i.e.*, what are the data characteristics that the mining task tries to highlight). Different data mining tasks achieve different insights over the data: classification captures the class of data or a new item, clusters reveal natural groups in data, decision trees detect characteristics that predict (with respect to a given class attribute) the behavior of future records, and so on [6]. This unorganized data requires processing to be done to generate meaningful and useful information from the large databases. In order to organize large amount of data, you implement the concept of Database Management Systems (DBMS) such as Oracle, and SQL Server. These Database Management Systems require you to use SQT, a specialized query language to retrieve data from a database. However, the use of SQT is not always adequate to meet the end user requirements of specialized and sophisticated information from an unorganized large data bank. Database researchers pay more attention to the issues related to the volume of data and also concerned with the effective use of the available database techniques such as efficient data retrieval mechanisms. This therefore necessitates you to look for certain alternative techniques to retrieve information from large and mostly unorganized sources of data.

2.1. Theoretical Framework

Nowadays, data stored in medical databases are growing in an increasingly rapid way. Analyzing that data is crucial for medical decision making and management. It has been widely recognized that medical data analysis can lead to an enhancement of health care by improving the performance of patient management tasks. There are two main aspects that define the need for medical data analysis [7]:

1) Support of specific knowledge-based problem solving activities through the analysis of patients' raw data collected in monitoring.

2) Discovery of new knowledge that can be extracted through the analysis of representative collections of example cases, described by symbolic or numeric descriptors. For these purposes, the increase in database size makes traditional manual data analysis to be insufficient. To fill this gap, new research fields such as knowledge discovery in databases (KDD) have rapidly grown in recent years. KDD is concerned with the efficient computer-aided acquisition of useful knowledge from large sets of data. The main step in the knowledge discovery process, called data mining, deals with the problem of finding interesting regularities and patterns in data. The terms Knowledge Discovery in Databases (KDD) and Data Mining are distinct. KDD refers to the overall process of discovering useful knowledge from data. It involves the evaluation and possibly, interpretation of the patterns to make the decision of what qualifies as knowledge. It also includes the choice of encoding schemes, preprocessing, sampling, and projections of the data prior to the data mining step [7].

Data Mining refers to the application of algorithms for extracting patterns from data without the additional steps of the KDD process [8]. The KDD process is often to be nontrivial; however, we take the larger view that KDD is an all-encompassing concept. KDD is a process that involves many different steps. The input to this process is the data, and the output is the useful information desired by the users. However, the objective may be unclear or inexact. The process itself is interactive and may require much elapsed time. To ensure the usefulness and accuracy of the results of the process, interaction throughout the process with both domain experts and technical experts might be needed.

Data mining is the step in the process of knowledge discovery in databases, that inputs predominantly cleaned, transformed data, searches the data using algorithms, and outputs patterns and relationships of interest in a particular representational form or a set of such representations as classification rules or trees, regression and clustering, to the interpretation/evaluation step of the KDD process. The definition clearly implies that what data mining (in this view) discovers are hypotheses about patterns and relationships. Those patterns and relationships are then subject to interpretation and evaluation before they can be called knowledge.

A simple data mining process model includes the following steps [9]:

1) Select a target data set. 2) Data preprocessing. 3) Data transformation. 4) Data mining. 5) Interpretation/evaluation. 6) Presentation. 7) Documentation.

2.2. Cardiovascular Diseases

These are conditions that involve narrowed or blocked blood vessels that can lead to a heart attack, chest pain (angina) or stroke. Other heart conditions, such as those that affect your heart's muscle, valves or rhythm, also are considered forms of heart disease. It can also be referred to more than one disease of the

circulatory system including the heart and blood vessels, whether the blood vessels are affecting the lungs, the brain, kidneys or other parts of the body. Cardiovascular diseases are the leading cause of death in adult Canadian men and women [10].

The following six types of cardiovascular disease are highlighted below as shown in Figure 2:

Ischemic heart disease is the most common type of cardiovascular disease in Canada and other industrialized countries around the world. It refers to problems with the circulation of blood to the heart muscle. A partial blockage of one or more of the coronary arteries can result in a lack of enough oxygenated blood (ischemia) thus causing symptoms such as angina (chest pain) and dyspnea (shortness of breath). A complete blockage of an artery causes necrosis (damage to the tissues) or a myocardial infarction, commonly known as a heart attack.

Cerebrovascular disease (Stroke) refers to a problem with the circulation of blood in the blood vessels of the brain. A blockage with effects lasting less than 24 hours is referred to as a transient ischemic attack. A complete blockage with long-term effects is referred to as a cerebrovascular thrombosis (clot) or accident or a stroke. Sometimes, a blood vessel in the brain can burst resulting in long term effects.

Peripheral vascular disease affects the circulation primarily in the legs. Patients with this disease typically complain of pain in their calves especially when walking.

Heart failure occurs when the pumping action of the heart cannot provide enough blood to the rest of the body as it is needed. This can happen as a result of damage to the heart muscle, for example from a heart attack, or from excessive consumption of alcohol, or because of a heart muscle disease also called a cardiomyopathy. Patients with heart failure usually suffer from shortness of breath and swelling of the legs.

Rheumatic heart disease once common in Canada is a major problem in many poor countries. This disease begins with a bacterial infection in childhood, affecting joints and heart valves. The heart problems appear many years later. Often the valves have to be replaced by an operation.

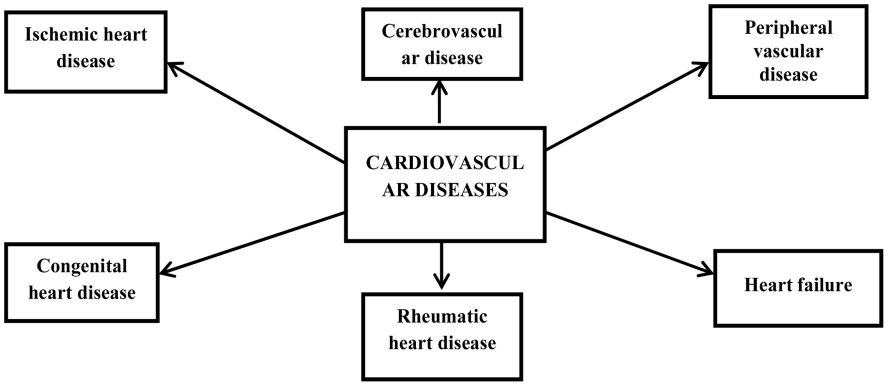

Figure 2. Cardiovascular disease: field work.

Other infections can occur attacking the inner tissues of the heart including the valves (endocarditis) and the outer tissue overlying the heart (pericarditis).

Congenital heart disease is a problem with the structure of the heart arising because of a birth defect. These anatomical defects can be as simple as a small hole in one of the inside walls of the heart or they can be very complex, affecting the way blood flows through the heart and lungs. Some congenital heart problems result in death unless immediately corrected by surgical intervention. Others cause disability to varying degrees and are treated by surgery later in life with correction of the problem sometimes requiring more than a single operation.

3. Methodology and System Analysis

Research methodology is an investigation in order to discover new factors through planning and systematic collection, analysis and interpretation of data. It is a structured approach to arriving at dependable solution to problems through the planned, systematic and objective search for new knowledge of the object of study.

Research methodology is therefore a detailed description of what the researcher planned and procedure adopted in gathering new facts relevant to the project work.

It is therefore an established fact that without data, there can be no analysis. This is crux of social science research; data can be defined simply as basic facts or figure resulting from business, economic and social activities of an individual. Trying to develop new system that will manage cardiovascular diseases, there exist the unreliable need to properly study the existing system so as to define:

1) Basic processes; 2) Modus operandi; 3) Strength; 4) Weaknesses; 5) Preferences; 6) Exceptions.

Data and information relevant to the above mentioned will be collected and analyzed in the subsequent sections of this project work, this research methodology describes the process involved in attracting data about a system problem through any of the available data collection methods such as interview, observation, experiment, questionnaire, etc.

Methodology can be seen as the framework used to structure, plan, and control the overall process of developing an information system. According to Camillus Alberto (2014), it is used to refer to a specific series of steps or procedure which governs the analysis and design of a particular project. It includes the methods, techniques and procedure which are used to collect and analyze information. Some of the various methodology adopted by software developers include the SSADM (Structure System Analysis and Design Methodology), Prototyping, Water Sluice, UML and lots more. In designing and developing the proposed software, the SSADM methodology will be adopted. This methodology was chosen owing to its flexibility, reusability, entity behavior modeling among others. The use of this methodology also enables the researcher to thoroughly investigate and analyze the present system in order to have in depth under-

standing of the system.

3.1. Analysis of the Present System

Umezuruike Hospital formally opened its door to the public on the 15th of October, 1971 soon after the end of the Nigerian civil war. Its first home was in rented premises at 78/79 Wetheral Road, Owerri, a small fifteen bedded clinic. That was an epoch-making occasion and it made news in the national dailies at the time, because it was the second fully equipped privately operated clinic in the metropolis of Owerri, offering both out-patient and in-patient care of high quality. Before then, the only option was the ever crowded General Hospital, Owerri. It can be said that Umezuruike Hospital right from its inception was a trail blazer. The purpose-built site of the Hospital was put into use in 1975 but the building was not completed until 1978. Every journey in life has a beginning and for an individual, an ultimate end. However, institutions may continue to survive and Umezuruike Hospital has always been structured in the belief that it will survive its founders. The founding of Umezuruike was indeed a mixture of planning and chance from the late father of the proprietor, Chief Oparaeburuo Umez-Eronini of blessed memory. He nurtured the dream while the Nigerian civil war was the accidental catalyst. Umezuruike hospital comprises of various medical practitioners ranging from cardiologists, gynecologists, pediatricians and lots more as shown in Figure 3. Their services include: general out-patient clinic, surgical out-patient clinic, antenatal and postnatal clinic, pediatric clinic, ultrasonography services and lots more.

Information flow of the present system: The information flow is shown in Figure 3 below.

3.2. Research Methodology

Research methodology is an investigation in order to discover new factors through planning and systematic collection, analysis and interpretation of data. It is a structured approach to arriving at dependable solution to problems through the planned, systematic and objective search for new knowledge of the object of study. Research methodology is therefore a detailed description of what the researcher planned and procedure adopted in gathering new facts relevant to the project work.

It is therefore an established fact that without data, there can be no analysis. This is crux of social science research; data can be defined simply as basic facts or figure resulting from business, economic and social activities of an individual. Trying to develop new system that will manage cardiovascular diseases, their exists the unreliable need to properly study the existing system so as to define: 1) Basic processes; 2) Modus operandi; 3) Strength; 4) Weaknesses; 5) Preferences; 6) Exceptions.

Data and information relevant to the above mentioned will be collected and analyzed in the subsequent sections of this project work, this research methodology

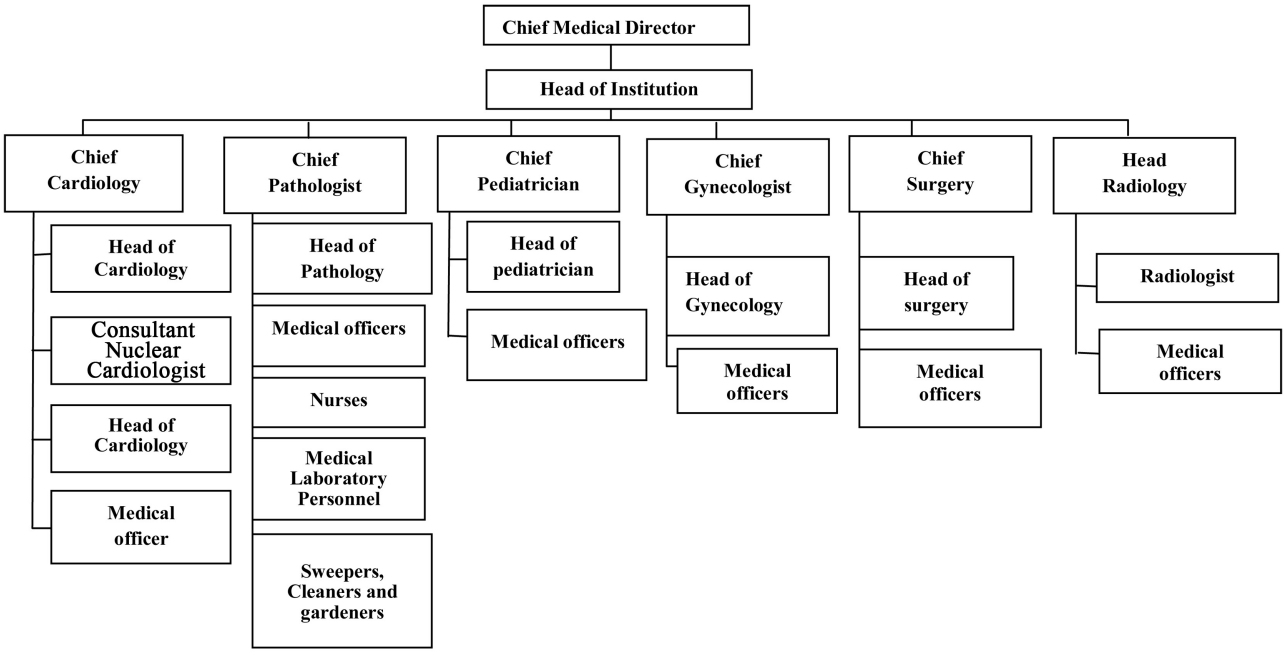

Figure 3. Organogram of the present system. Source: [10].

describes the process involved in attracting data about a system problem through any of the available data collection methods such as interview, observation, experiment, questionnaire etc. [10].

3.2.1. Modulus Operandi

Information obtained from a visit to Umezuruike hospital Owerri revealed that it is usually a herculean task obtaining data regarding possible cardio vascular diseases. An interview with Doctor Parray Obinna a cardiologist with the hospital, revealed that most times owing to delay in test results, poor data storage system and lack of adequate prompt knowledge on various cardio vascular diseases, most patients end up in very bad states and in some cases they die.

He further pointed out that if there was a remedy to be made, he would strongly suggest a system which will provide prompt cardio vascular answers, prompt test result and remedies as well as symptoms to these cardio diseases. Dr Parray further pointed out that if patients could regularly undergo cholesterol test to ascertain their various cholesterol levels, the high rate of mortality would be drastically reduced as ones cholesterol level is a great determining factor.

3.2.2. Weaknesses of the Existing System

Below are some of the major problems associated with the current method of mining data associated with cardio vascular diseases in Umezuruike hospital Owerri.

- Difficulty in accessing patients' health details promptly in order to ascertain if they have had similar health issues.
- Delay in obtaining relevant information regarding a cardio vascular disease from junks of cardio stored information.

- Unavailability of cardio vascular information to all health personnel within the hospital.
- Delay in obtaining patients cholesterol level which happens to be a major attribute of cardio vascular diseases.
- Poor record up keep.

3.2.3. Proposed System

Here the proposed system is data mining software for detecting cardio vascular diseases. This system was developed after careful observation of the weaknesses of the existing system and it was also proposed to curb the weaknesses of the existing system.

- This new system will have the ability to provide real time answers regarding cardio vascular diseases.
- A database containing the numerous types of cardio vascular diseases, their symptoms, causes and remedies.
- A system with the ability to test and analyze patients' cholesterol level, which has been proven to be one of the greatest determining factors of cardio vascular diseases.
- A system with the ability to obtain information regarding a specific cardio vascular disease through its search medium.

Process Flow of the Proposed SysteM: The process flow of the Proposed System is shown in Figure 4 below. It discusses the various examination centres in the System of Umezuruike Hospital which is the chosen as our case.

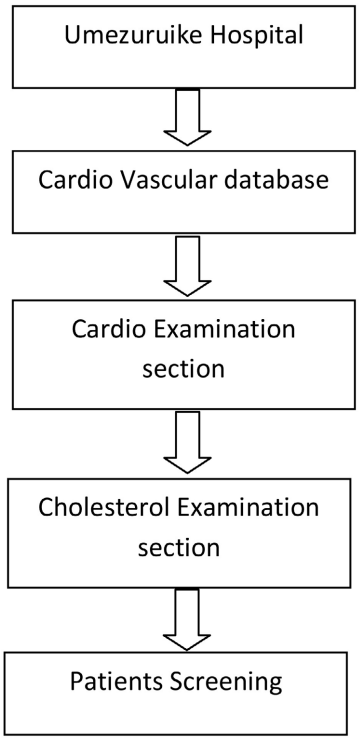

Figure 4. Process flow of the proposed system.

3.2.4. Data Flow Diagram of the Propsed System

The Data Flow Diagram of the Proposed System is shown in Figure 5 below. It displays the flow of Data from Keyboard to processing of Data, Screen. It also shows Data flow from Database File to processing of Data and Display.

3.2.5. High Level Model of Proposed System

The High level model of the Proposed System is shown in Figure 6 below. It shows the Login Screen, Main menu, Patient, Cardio Administration and Logout menus, etc.

3.2.6. Expectations of the New System

The expectations of users when interacting with the proposed system have been identified below as:

- The system should perform in real time.
- It should always show the last data on the system except it has been updated.
- The system should be able to show the data of users.
- Should be able to calculate and output patients' cholesterol level and detect possible cardio vascular disease.

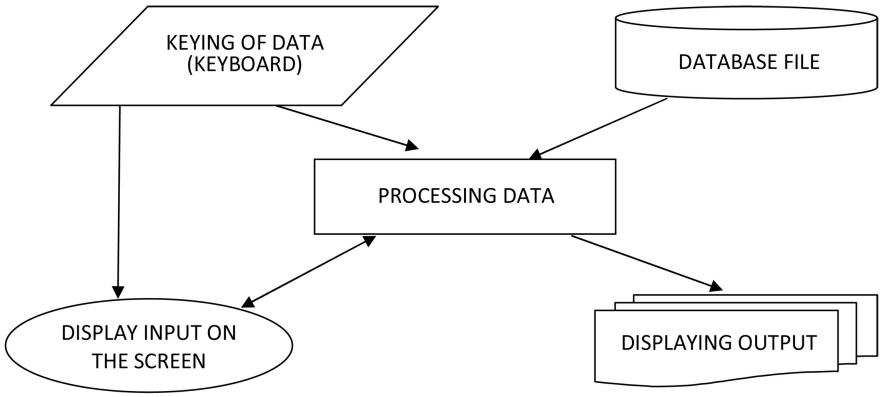

Figure 5. Data flow diagram of the proposed system.

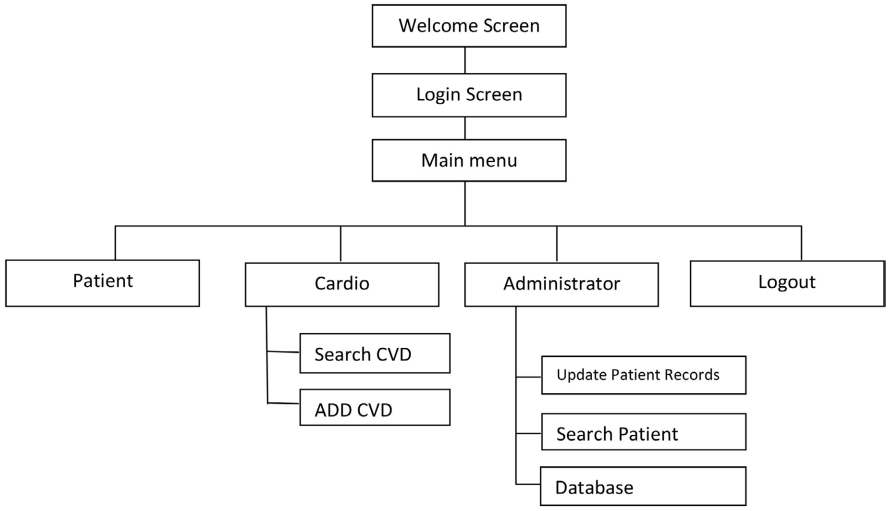

Figure 6. High level model of the proposed system.

- It should allow clinic administrator to input data.
- It should permit the administrator to view the list of all patients and all cardio vascular diseases in the system.
- Enable the user search for specific cardiovascular diseases, its symptoms and possible remedies.
- The system should be able to keep in its database all work done on the system.

3.2.7. Modularity

The program of this research work was designed in modules, which could be complied into a single application. Each module is written separately and linked up together and complied into a single unit "data mining system".

Splash Screen Module: This shows the name of the program, the developers name and details of the program.

Login Screen Module: Here the user is required to input some certain details in order to gain access to the system.

Main Menu Module: This module is made up of four sub systems namely the Patient Registration System, the Cardio Vascular System, the Administrative System, and the Cholesterol System. These four subsystems are user friendly and their various functions are self-explanatory. They also have some user defined functions and sub functions.

Patient Registration Module: This module is responsible for registering patients with various cardio vascular diseases.

Administration Module: This module explains where all registered patients are stored also known as the database.

Cardiovascular Module: This module is responsible for comparing, accepting and searching various cardio vascular diseases.

Logout Module: Where the user exists the system.

3.3. Feasibility Study

A feasibility study was carried out for the new system. Most doctors accepted the invention of a data mining software as it will aid eliminate the various inconsistencies associated with the current system. Quite a good number of medical personnel welcomed the need of a data mining software and opined that after its development, it should be available to all medical personnel irrespective of their level on the food chain as this would also help reduce the work load.

3.3.1. Operational Feasibility Study

Question asked for this study were:

- Will the system produce poorer result in any respect or area?
- Will the system be used if it is developed and implemented?
- Would there be support for the users in the development of the system?

3.3.2. Technical Feasibility Study

Here, I verified the availability and easy usage of my chosen programming lan-

guage. The visual basic is a GUI (graphical user interface) environment that allows visual tools and supports computer user interactions. It also supports systems with 4 GB RAM, 250 MB hard disk space. Its operating system ranges from windows vista up to windows 8.

3.3.3. Financial Feasibility Study

The system developed and installed will be a good benefit to the users of the system. The system will be developed and operated in the existing hardware and software infrastructure, since most laptops come with already installed Microsoft software there is no need of additional hardware and software for the system.

3.4. Data Mining Algorithm

An algorithm in data mining (or machine learning) is a set of heuristics and calculations that creates a model from data. To create a model, the algorithm first analyzes the data you provide, looking for specific types of patterns or trends. The algorithm uses the results of this analysis over many iterations to find the optimal parameters for creating the mining model. These parameters are then applied across the entire data set to extract actionable patterns and detailed statistics. There are quite a number of data mining algorithm some of which are: C4.5, K-means, Support Vector Machines (SVM), Apriori etc.

3.5. Data Mining Algorithm Adopted

The Clustering algorithm is the data mining algorithm adopted for this research work this is owing to the fact that it is an algorithm that learns which items in a data are commonly associated with each other. This can be very useful for grouping similar items together in tables.

For example, a hospital might have a database of patients and diseases. By running the clustering algorithm on this data, you might find that cardiovascular diseases are commonly bought together with cholesterol. You could use this information to display cardiovascular diseases at the right time to a particular patient.

4. System Design and Implementation

System design is the architectural drawing and outlining of the specification document. It is the blue print of the system that is to be developed therefore this section discusses the application, the various phases adopted in the design and development and implementation of the new system. It also discusses the various menus and screen displays, how the system can be installed, the hardware and software specifications/requirements for the installation of the system and how to start up the system. The proposed system has been designed in such a way that user access must be granted to users before they can gain access to the software [10].

The program design was considered in the following areas:

1) Modular design: creation of modules was necessary since we realized that the system would be made of different units which would be made up of differ-

ent units which would be difficult to design as one unit. We therefore created representative modules for the complex whole.

2) Actual design: at this point, the individual modules so far created were transformed into actual working design; this design stage involves the creation of forms and planning of necessary objects on these forms.

4.1. System Design

4.1.1. System Control Center

The system control center which depicts the new system's basic functional unit is shown in Figure 4.

4.1.2. Description of the New System

The Structured System Analysis and Design methodology which is mostly used in software development was adopted in the design and development of the new system. This methodology was adopted owing to its simplicity and because an overview of the new system has to be formulated first, which specifies and details each subsystem. Each sub system is then refined in yet greater detail, sometimes in many additional subsystem levels, until the entire specification is reduced to base elements.

4.2. System Flowchart

The new system will be an integration of the graphical user interface of the modules of the system to be designed. Its functionalities are as follows, see Figure 7 below:

- It will be able to assign user names and passwords to both new and existing users.
- It will be able to store patients' data.
- Provide real time answers to cardio vascular diseases.
- Secure patients data.
- Provides prompt answers to cholesterol test.

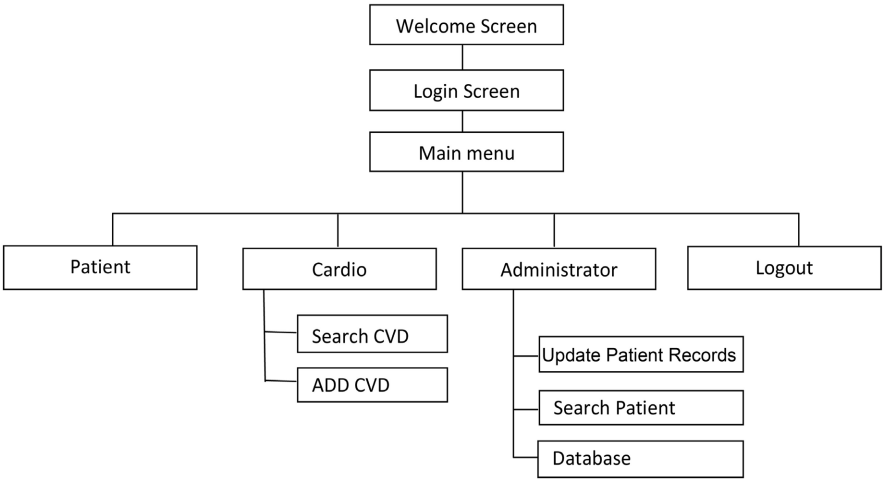

Figure 7. System design (system control centre).

- Able to search specific cardio vascular disease.

The System flowcharts are shown in Figures 8-10.

4.3. Specifications

4.3.1. Database Tables

The following tables specify the various database used in the development of the new system. The User Login Table, Patient Registration Table, Admin Login Table and the Cardiovascular Table are shown in Tables 1-4 respectively.

4.3.2. Overall Flowchart of the New System

The overall flowchart of the new system is shown in figure below. Similarly, the Patient Registration flowchart and Cardio Registration flowchart are shown in Figure 9 and Figure 10 respectively.

1) Patient Registration Flowchart

2) CVD Registration Flow Chart

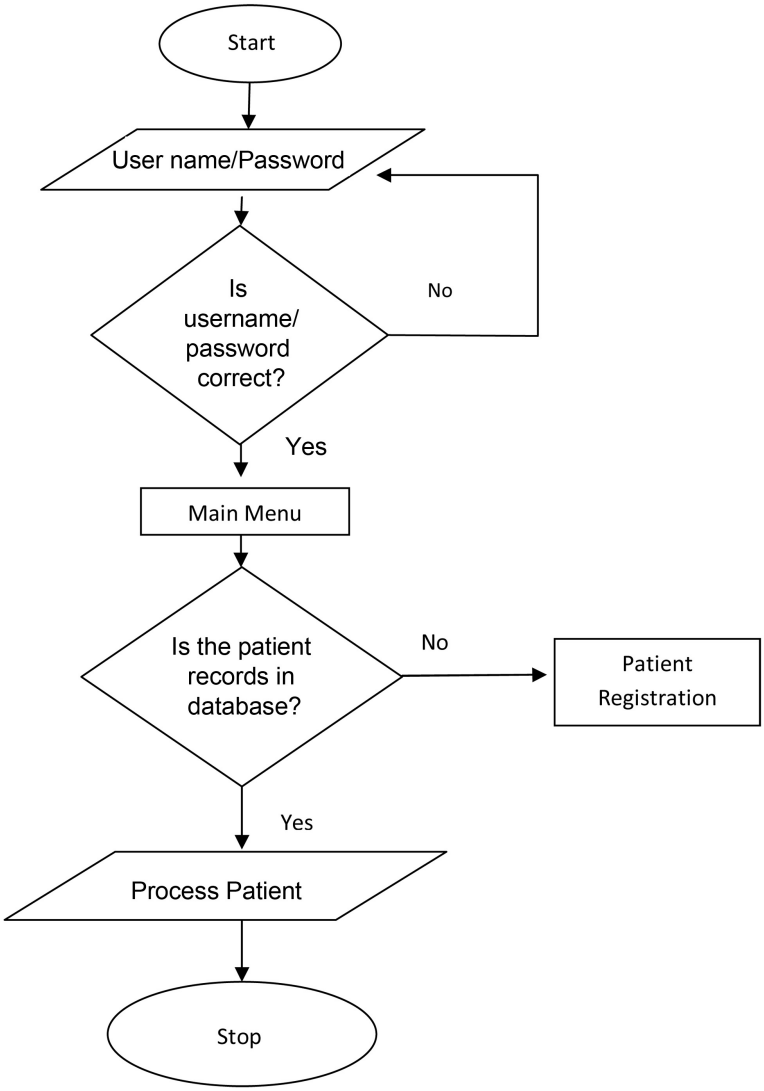

Figure 8. Overall data flow diagram of the system.

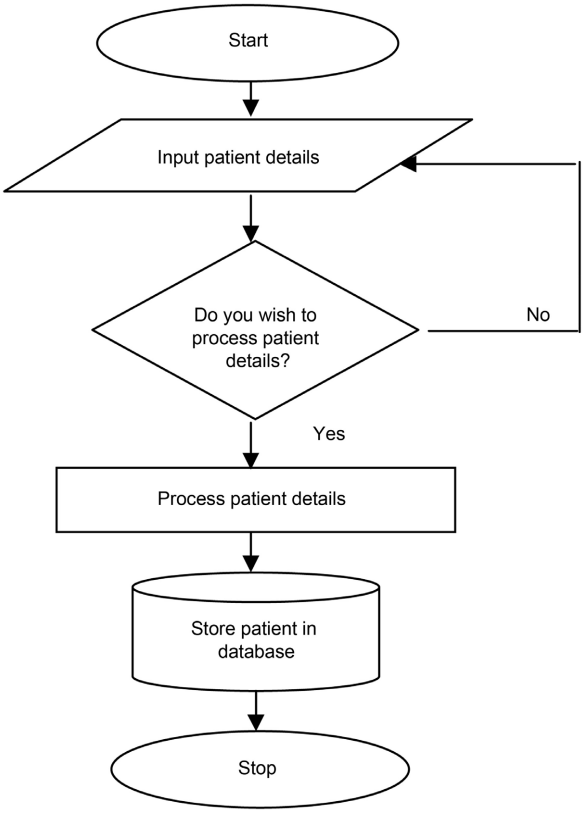

Figure 9. Patient registration flow-chart.

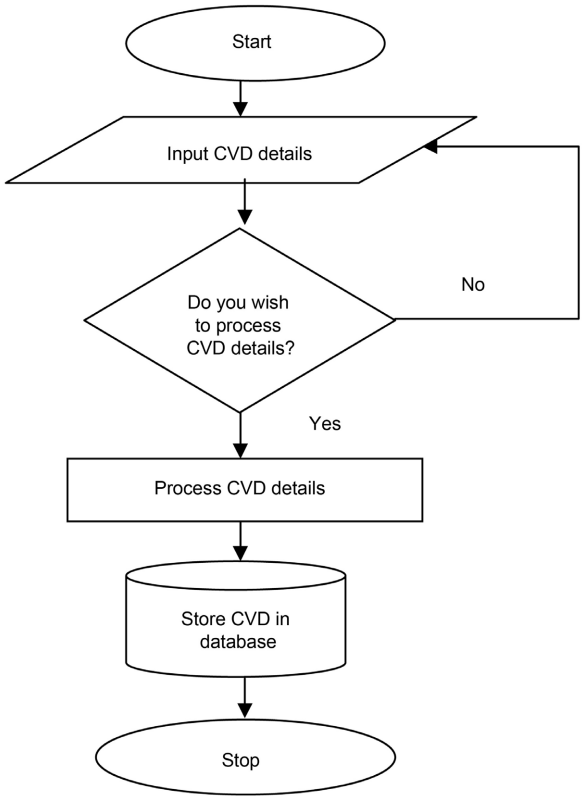

Figure 10. CVD registration flow-chart.

Table 1. User login table.

S/N	File Name	Data Type	Width
1.	Username	Var/Character	5
2.	Password	Var/Character	5

Table 2. Patient registration table.

S/N	File Name	Database	Width
1.	Surname	Character	20
2.	Other Names	Character	30
3.	Picture	Object Ole	12
4.	Age	int	5
5.	Phone	Int	10
6.	Gender	Character	20
7.	Residential Address	VAR Character	20
8.	Date of birth	Var Character	20
9.	Date of registration	Date/time	10
10.	Cholesterol level	Character	20
11.	Any cardio vascular disease?	Boolean	3

Table 3. Admin login table.

S/N	File Name	Data Type	Width
1.	Username	Var/Character	5
2.	Password	Var/Character	5

Table 4. Cardio vascular table.

S/N	File Name	Data Type	Width
1.	Cardio Vascular Type	Var/Character	50
2.	Symptom	Var/Character	50
3.	Remedy	Var/Character	50
4.	Cholesterol Level	Var/Character	50

4.4. Algorithm

1) Overall Program

Step 1: Start

Step 2: Accept Username and Password

Step 3: Check if Username == "admin" && Password == "admin" if yes go to step 4 else go to step 5

Step 4: Main menu

Step 5: Display Message "Invalid username or password"

Step 6: If Main menu on click = = "Home" go to step 7

else if Main menu == "Patient" go to step 10

else if Main menu == "Cardio_vascular" go to step 11

else if Main menu =="Logout" go to step 12

Step 7: Accept Username and Password

Step 8: Check if Username == "admin" && Password == "admin" if yes go to step 9 else go to step 5

Step 9: View database

Step 10: View Patient registration, Patient search and Patient update module

Step 11: View Cardio_vascular registration and search module

Step 12: Stop

2) Patient Registration

Step 1: Start

Step 2: Accept patient details

Step 3: Check if Patient_name = " " Or Address = " " Or Cardio_years= " " Or Gender = " " Or Address = " " if yes go to step 4 else go to step 5

Step 4: Display "Please fill all fields"

Step 5: Display "Registration was successful"

Step 6: Print

Step 7: Stop

3) Cardiovascular Registration

Step 1: Start

Step 2: Accept CVD details

Step 3: Check if CVD name = " " Or Symptoms = " " Or Possible_Remedy= " " Or Definition = " " Or Causes = " " if yes go to step 4 else go to step 5

Step 4: Display "Please fill all fields"

Step 5: Display "Registration was successful"

Step 6: Print

Step 7: Stop

4.5. Hardware/Software Specifications

Installing software and hardware in a computer comprises the most basic of computer related knowledge today. In today's day and age, computers have become part and parcel of our lives. We can find them everywhere around us. They have a variety of significant and important applications in several different fields, including industries, agriculture, research, medicine etc for most of us however; the term computer signifies the PCs that we use at our homes and offices. We use them to watch movies, listen to songs, play games, make related document etc. it is necessary that we have some understanding about at least the installation of its basic components, namely software and hardware. This knowledge can prove to be vital, especially when we wish to upgrade our PCs. In the very basic sense, computer software constitutes the monitor, keyboard, and all the

other peripheral devices that are connected to it. Both these components are essential for the proper functioning of any computer system [10].

The following are the various hardware and software system specification needed for the proper installation and usage of the new system.

4.5.1. Hardware Specifications

- Processor—Pentium 4 and above
- Motherboard—ATX
- Monitor—SVGA colour monitor
- Cache memory—256 kb
- CD/DVD ROM
- RAM—512 MB and above
- Keyboard standard—101/102 keys
- Hard disk drive—Least 40 GB free space

4.5.2. Software Specifications

In addition to the above listed hardware specifications, the following software specifications are also essential for the proper operation of the system.

- Operating system—Windows 7 or 8
- Platform—Visual Basic 6.0
- Back end—Access Database management

4.6. Testing and Evaluation

This refers to the testing of the new system. Real transactions are simulated and are used to test the system. Three types of testing are carried out; Unit testing, System testing and Integrated testing.

- **Unit Testing:** in this a program module that has been written and successfully complied, is tested instruction by instruction with a well prepared test data. Each of these modules can be put together as a single program and tested.
- **Integrated Testing:** modules or programs are combined into groups and tested; the purpose is to find how these programs would interact with one another.
- **System Testing:** this utilizes special testing data meant for that purpose. The expected result of the system should be predetermined, to see if the system would give us exactly what we desire from it.

4.7. System Documentation

Documentation is the formalization record of detailed analysis and design of the new system. It outlines the techniques and methods used to correct the problem area of the existing system. All stages in a system network require documentation which constitutes of link between one activity of the new system and the other, the existence of a documentation standard would make communication between analyst and system user in particular easier, it will also make develop-

ment of a new system easier and quicker since documentation standard would give guide lines to be followed by the developer. The content of the documentation manual also include hardware specification and program flow chart.

5. Summary and Conclusions

In this research, we have looked at specific applications of data mining techniques in the detection of cardiovascular diseases. We can conclude that data mining is of great importance in the easy detection of cardiovascular diseases. Data mining, however, is not a "silver bullet" capable of solving all cardiovascular diseases, but rather it aims to provide possible prevention methods, remedies and symptoms. After the implementation of this project, there would be fundamental changes in the way patients are being registered for cardio vascular sections, as well as in the way information regarding the various cardio vascular diseases are obtained.

Cardiovascular diseases pose a serious life-threatening risk globally across all races and age groups. But most of them could be controlled and risk reduced through a proper medical checkup, proper dieting, nutrition, healthy eating habits and exercising all of which are incorporated into our data mining software for easy access. With the cardiovascular data mining software, we can achieve the following:

- Obtain real time information regarding various cardiovascular diseases.
- Able to always show the last data on the system except it has been updated.
- Store patients' past and present cardio vascular records.
- A system with the ability to show the data of users.
- Enable the user to search for specific cardiovascular diseases, their symptoms and possible remedies.
- With this new system, we can reduce the risk of a number of most serious life-threatening cardiovascular diseases and enjoy life to its full.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Kohn, H.C. and Tan, G. (2005) Data Mining Application in Healthcare. *Journal of Healthcare Information Management*, **19**.

[2] Savage, N. (2011) Mining Data for Better Medicine. *MIT Technology Review*, **38**, 235-237.

[3] Joshi, S., Shenoy, P.D., Venugopal, K.R. and Patnaik, L.M. (2010) Data Analysis and Classification of Various Stages of Dementia Adopting Rough Set Theory. *International Journal on Information Processing*, **4**, 86-89.

[4] Fayyad, U. and Uthurusamy, R. (1996) Data Mining to Knowledge Discovery in Databases. *Communications of the ACM*, **39**, 24-26. https://doi.org/10.1145/240455.240463

[5] Rizzi, S., Bertino, E., Catania, B., Golfarelli, M., Halkidi, M., Terrovitis, M., Vassiliadis, P., Vazirigiannis, M. and Vrachnos, E. (2003) Towards a Logical Model for Patterns. In: Song, I.Y., Liddle, S.W., Ling, T.W. and Scheuermann, P., eds., *Conceptual Modeling—ER* 2003, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-39648-2_9

[6] Ntoutsi, I. (2008) Similarity Issues in Data Mining—Methodologies and Techniques. University of Piraeus, Greece, 31-32.

[7] Ekwonwune, E, Ubochi, C. and Duroha, A. (2022) Data Mining as a Technique for Healthcare Approach. *International Journal of Communication, Network and System Sciences*, **15**, 149-163. https://doi.org/10.4236/ijcns.2022.159011

[8] Li, J.-S., Yu, H.-Y. and Zhang, X.-G. (2011) Data Mining in Hospital Information System. *New Fundamental Technologies in Data Mining*, **1**, 143-156.

[9] Balasunda, V., Devi, T. and Saravanan, N. (2012) Development of a Data Clustering Algorithm for Predicting Heart. *International Journals of Computer Applications*, **48**, 8-13.

[10] Chinyere, U. (2015) Data Mining Techniques for Detecting Cardiovascular Diseases.