MAKERERE UNIVERSITY

COLLEGE OF COMPUTING AND INFORMATION SCIENCES

DEPARTMENT OF NETWORKS

Research Paper for Coronary Heart Disease Diagnosis System.

13th January 2023

Prepared by BSE23-27

NAME	REGISTRATION NUMBER
NDUHURA ANDREW	19/U/12060/PS
SSEKYANZI JONATHAN	19/U/13748/PS
KANNABI SSERWADDA MUHAMMED	19/U/13513/PS
RUTUNGU HENRY	19/U/8542/EVE

ABSTRACT

Coronary heart disease is one of leading killer diseases in Uganda. In this study, the project members of the Coronary Heart Disease Diagnosis System investigate how Coronary Heart Disease is diagnosed in Uganda. The research team used questionnaires and interviews to obtain information about diagnosis of coronary heart disease from medical specialists at Uganda Heart Institute and other medical facilities around the country. We managed to attain 50 responses from medical specialists who answered the questionnaire as a Google form. At the end of the study the research and basing on the data collected, the team found out that conducting a successful diagnosis of the disease takes a long time and quite difficult to conduct the diagnosis. Basing on the results of this study, the team proposed to address the issues through developing a Coronary Heart Disease Diagnosis System to effectively diagnose coronary heart disease accurately in a short period of time.

ACKNOWLEDGMENT:

We are grateful to the Almighty God for without His graces and blessings, this research would not have been possible.

Immeasurable appreciation and deepest gratitude for help and support are extended to the following people who in one way or another have made this research possible.

Our dear parents who have extended the necessary facilitation required not only in this study but throughout our education life.

Dr. Bernard Muwonge Sajjabi, our project supervisor who advised and attended to all the efforts of the team during this study.

Dr. Mary Nsabagwa, who is the overall supervisor of the project for her timely engagements with the team that enabled the team to polish through the work of this study.

All the respondents who have enabled this study to be a success.

Table of Acronyms

Abbreviation	Full meaning
AI	Artificial Intelligence
CHD	Coronary Heart Disease
CHDD	Coronary Heart Disease Diagnosis
IHDPS	The Intelligent Heart Disease Prediction System
UHI	Uganda Heart Institute

List of Figures

Table of Contents

1 Introduction

1.1 Background to the study

Globally, heart disease is the leading cause of death for both men and women with more than half of the deaths occurring in men [2]. The most common type of heart condition is coronary heart disease. Coronary heart disease (CHD) can be caused due to risk factors like high blood pressure, high blood cholesterol, tobacco use, obesity, unhealthy diet, physical inactivity, diabetes, advancing age, inherited disposition [6]. [2] suggests that coronary heart disease that can cause heart attacks kills more than 370,000 people per year. This can also be attributed to medical errors which cause tens of thousands of deaths in United States' hospitals [1]. Due to these highly costly and harmful medical errors, there is need for reliable and powerful clinical decision support systems in order to reduce the time of diagnosis and increase diagnosis accuracy especially for heart disease diagnosis [1]. Hospitals must also minimize the costs of clinical tests through automation in the heart disease diagnosis. Clinical databases that contain patient information are tapped into by data mining techniques [4]. In the same sense, artificial intelligence (AI) techniques such as machine learning can play a critical role in facilitating precision for diagnosis [5].

This research was conducted to obtain information about coronary heart disease that will help us in designing and developing the **CHDD system** which will predict the presence of the disease basing on Artificial Intelligence.

The goal of this project is to develop the **CHDD** system that will predict the presence of CHD, provide graphical analysis of the rate of spread of the disease, keep track of patients' results and notify patients diagnosed about the next check-up.

The project will be completed by June **2023.**

1.2 Problem Statement

This project addresses coronary heart disease diagnosis problems such as long time spent by experts trying to analyze and confirm the presence of the disease, limited number of heart disease doctors capable of accurately detecting the disease and the inaccessibility of diagnosis machines throughout the country.

1.3 Research questions

This research aimed at answering the following questions:

1. What are the major causes of coronary heart disease in Uganda?
2. What are the signs and symptoms of coronary heart disease?
3. How do you diagnose coronary heart disease?
4. How long does it take an expert to diagnose a single patient of CHD?
5. Would you say the diagnosis of this disease is hard or easy?
6. How much does the diagnosis cost the patient?
7. What are the main challenges faced during the diagnosis of coronary heart disease?
8. How many specialists are present at the hospital where you are currently working?
9. Do patients require next checkups? If yes, which means of communication do you use to remind them about the check ups?
10. How do you store patients' records?

1.4 Objectives of the study

1.4.1 Main Objective of the study

The main objective of this research study is to develop the "Coronary Heart Disease Diagnosing System" that will predict the presence of CHD, provide graphical analysis of the rate of spread of CHD, keep track of patients' results and notify patients about the next checkup.

1.4.2 Specific objectives

1. To gather and analyze the requirements of Coronary Heart Disease Diagnosis system.
2. To design and implement Coronary Heart Disease Diagnosing system basing on its requirements.
3. To test Coronary Heart Disease Diagnosing system to ensure that it meets its requirements.
4. To deploy and maintain Coronary Heart Disease Diagnosis system so that it can be accessed by its intended users.

1.5 Scope of the Study

This study focuses on obtaining first-hand information about coronary heart disease in Uganda. Doctors at Uganda Heart Institute, medical personnel in other hospitals and medical students constitute the scope of this research study.

2 Literature Review

This literature review examines the use of data mining in the automated diagnosis of heart diseases. Additionally, it illustrates how machine learning is used as a core to aid physicians detect the heart disease early in order determine the most appropriate treatment for the patients.

2.1 The construct of data mining in diagnosis alongside machine learning

Data mining is the turning of a collection of data into knowledge. Data mining is the process of finding previously unknown patterns and trends in databases and using that information to build predictive models [5]. Data mining tools have been used widely and they help reduce errors and investigation time for medical practitioners. There are various data mining techniques such as decision trees, artificial neural networks, Bayesian networks, and support vector machines. The following information indicates some of the reviewed systems.

[7] developed a system to predict more accurately the presence of heart disease. In this paper, three data mining classification techniques were applied namely Decision trees, Naive Bayes & Neural Networks. From results it has been seen that Neural Networks provides accurate results as compare to Decision trees & Naive Bayes.

[8] developed the computer aided heart disease prediction system that helps the physician as a tool for heart disease diagnosis. Some Heart Disease classification system is reviewed in this paper. From the analysis it is concluded that, data mining plays a major role in heart disease classification. Neural Network with offline training is a good for disease prediction in early stage and the good performance of the system can be obtained by pre-processed and normalized dataset. The classification accuracy can be improved by reduction in features.

The Intelligent Heart Disease Prediction System (IHDPS) [9] is developed using Decision Trees, Naïve Bayes and Neural Network techniques. Its models are trained and validated against a test dataset using Lift Chart and Classification Matrix methods to find out the model providing the highest correct prediction. Tabular and graphical visualization methods are used in IHDPS for better interpretation of results. The records of 909 elements with 15 medical attributes were split equally into two datasets: training dataset (455 records) and testing dataset (454 records). The most effective model to predict patients with heart disease appears to be Naïve Bayes followed by Neural Network and Decision Trees.

[10] designed a system to identify the chances of a coronary heart disease. They have divided all parameters in two levels according to criticality and each level is assigned separate weight age. Finally, both levels are considered to derive a final decision. They have implemented neuro-fuzzy integrated approach at two levels. So, error rate is very low and work efficiency is high. In this paper, they have performed the analysis for coronary heart disease.

[3] implemented an Intelligent System consisting of a web-based application in which the user answers the predefined questions. It retrieves hidden data from stored database and compares the user values with trained data set. It can answer complex queries for diagnosing heart disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot. By providing effective treatments, it also helps to reduce treatment costs.

[11] details that the electronic health records possess critical predictive information for machine-learning- based diagnostic aids. However, many traditional machine learning methods fail to simultaneously integrate textual data into prediction process because of its high dimensionality. In this paper, they present supervised method using Laplacian Eigenmaps to enable existing dimensional representations of textual data and accurate predictors based on these low-dimensional representations at the same time. This preserves the local similarities among textual data in high- dimensional space. The proposed implementation performs alternating optimization using gradient descent. For the evaluation, they applied Laplacian Eigenmaps method to over 2000 patient records from a large single- center pediatric cardiology practice to predict if patients were diagnosed with cardiac disease.

[12] describes that it is important to monitor heart rate during cycling. By monitoring heart rate during cycling, Cyclists can control the cycling session such as cycling cadence to determine the intensity of exercise. By controlling the intensity of cycling, cyclists can avoid risks measured by heart rate of cyclist. The heart rate can be measured by wearable sensor. But there are data that are

not recorded by the sensor at a regular time for example, one second, two second, etc. So we need a prediction model of heart rate to complete the missing data. The purpose of this study is to create a predictive model for heart rate based on cycling cadence using Feed forward Neural Network. The inputs are heart rate and cadence on the second. The output is the predictive value if heart rate on the next second. Feed forward Neural Network is used as a mathematical model of the relationship between heart rate and cycling Cadence.

This clinical decision support system (CDSS) [13] for diagnosis of cardiovascular heart disease (CVD) is based on classifier ensemble method. It combines a set of four different classifiers with ensembles. Support vector machines (SVM), neural networks (ANN), Decision trees (DT) and Bayesian networks (BN) are used as classifiers. The different classifiers in the ensemble are used to analyze patient's serum microarray chip data. This system, Apta-CDSS-E (Aptamer biochip-based CDSS – ensemble), provides diagnosis information to the physicians and it also helps clinicians by providing a set of biomarker candidates which can be used for CVD diagnosis. In AptaCDSS-E, the patient's blood sample is collected and an Aptamer biochip is created with the serum separated from the patient blood and protein expression levels are scanned. Then, a new work list is created by the scanner interface and analyzed by the decision engine of AptaCDSS-E trained with prior sample sets. At last the system provides integrated analysis results and clinical analysis facts. The physician's, decision results can then be saved into the system database for future model updates. AptaCDSS-E adopted the ensemble approach to generate enhanced results by grouping a set of classifiers of each SVM, ANN, DT, and BN. Each ensemble is constituted with several classifier models of each classification method. The training data are augmented by bagging and fed to each classifier

2.2 Conclusion

The above study reveals some of the weaknesses in the reviewed systems that we intend to improve in our proposed system.

Reviewed system	Weakness	Solution from proposed system

[13] AptaCDSS-E	low reliability, high cost of operation	Proposed system seeks to be available on the web accessible to medical practitioners throughout the country. It is aimed to achieve a low cost limited to internet data costs only to any device that can access to the internet
[3] Intelligent System, The Intelligent Heart Disease Prediction System (IHDPS) [9],	Generalization These systems are general and do not specify which diseases could or could not be detected.	Our proposed system focuses on the diagnosis of coronary heart disease as the Intelligent systems cannot cover all heart diseases.

3 Research methods

The table below describes the research methods used in achieving the research objectives and how the research methods were used during the **CHDD** project research.

Research Objective	**Research Method Used**	**How Research Method was Used**
To find out the diagnosis process of coronary heart disease.	1 Questionnaires 2 Interviews	1 Online questionnaires were issued out to specialists who filled them. 2 An expert at UHI was interviewed by the project members using an interview guide.
To investigate the difficulty of diagnosing the disease.	Questionnaires	Questionnaires were issued out to experts who provided feedback.
To find out how long it takes a doctor to diagnose a single patient for CHD	Questionnaires	Questionnaires were issued out to medical specialists who latter answered them

To investigate the number of specialists that are present in a particular hospital.	Questionnaires	Online questionnaires were filled by medical practitioners.
To find out whether patients are reminded about their next check-up.	Questionnaires	Questionnaires were issued out to medical specialists who latter filled them.

4 Results of the study

4.1

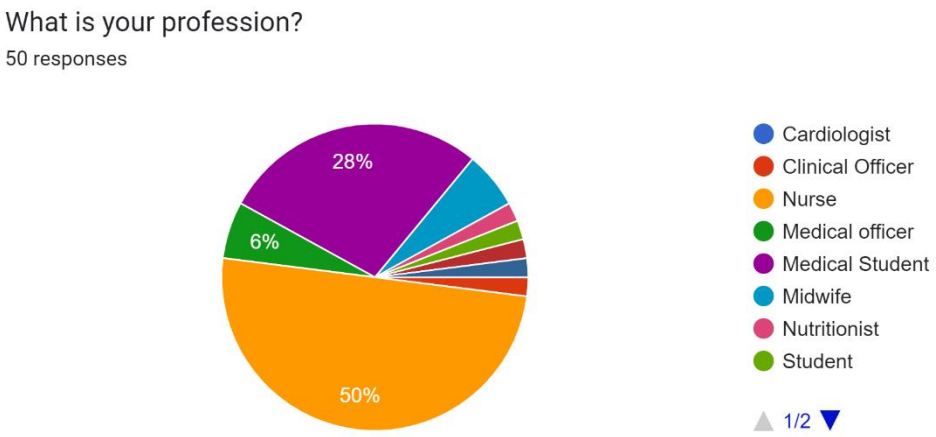

Figure 1. The professions of respondents

4.2

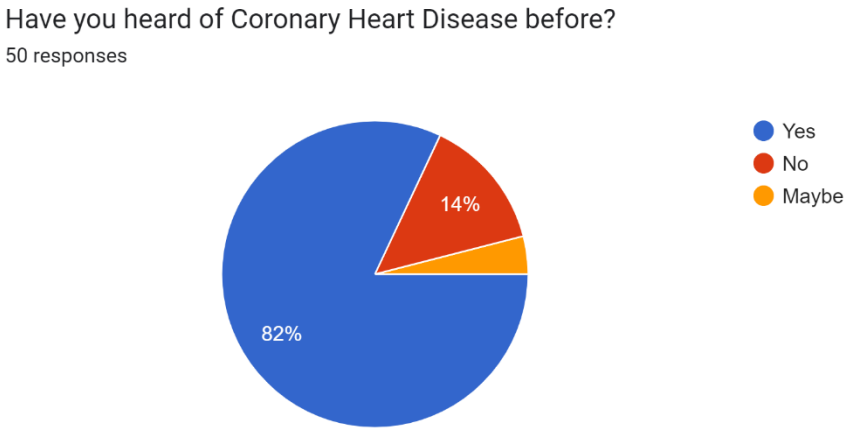

Figure 2. Respondents knowledge about coronary heart disease.

4.3

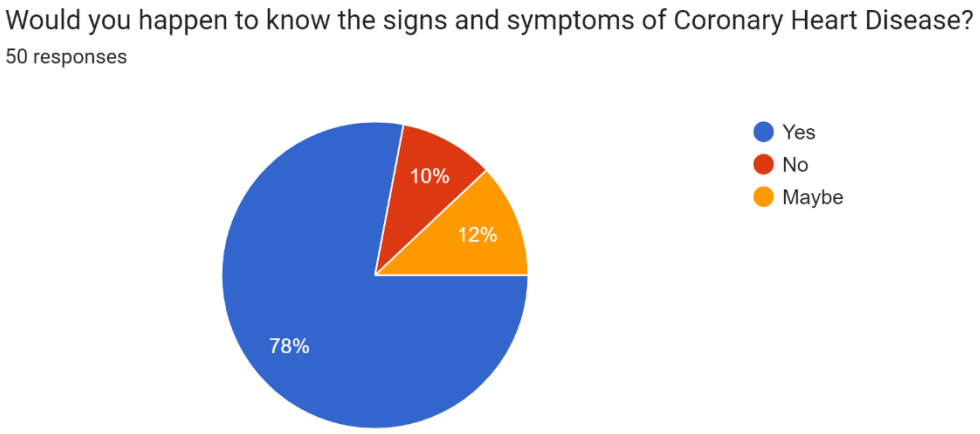

Figure 3. Knowledge about the signs and symptoms of Coronary Heart disease

4.4

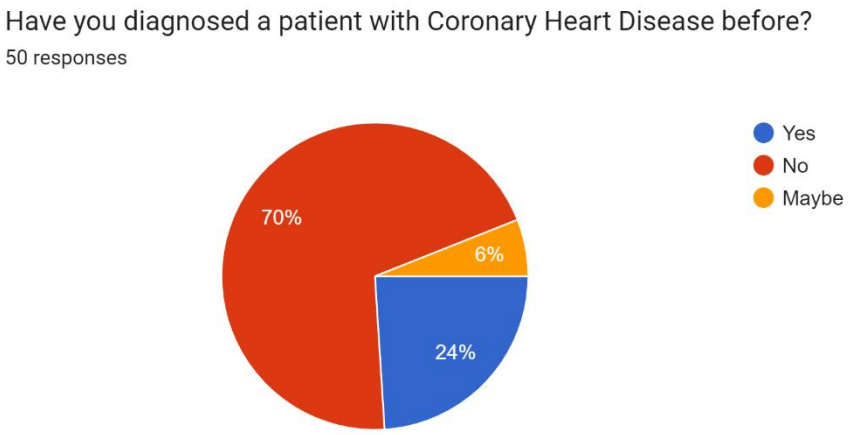

Figure 4 Possibility of having diagnosed a patient with Coronary Heart Disease

4.5

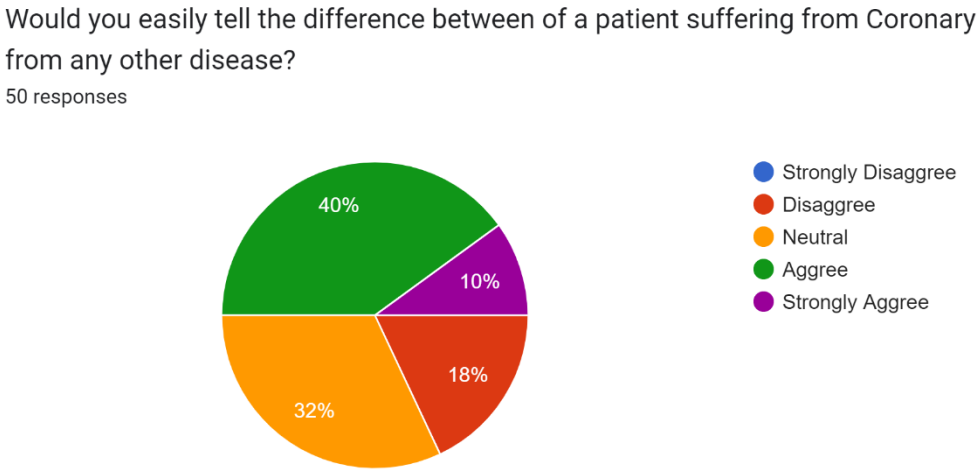

Figure 5 Differentiating a patient suffering from Coronary Heart Disease and any other disease

4.6

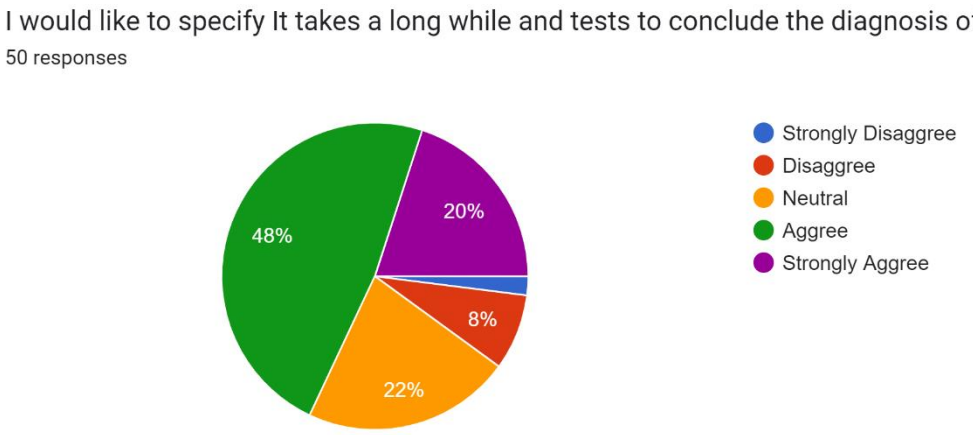

Figure 6 How long it takes to diagnose Coronary Heart Disease.

4.7

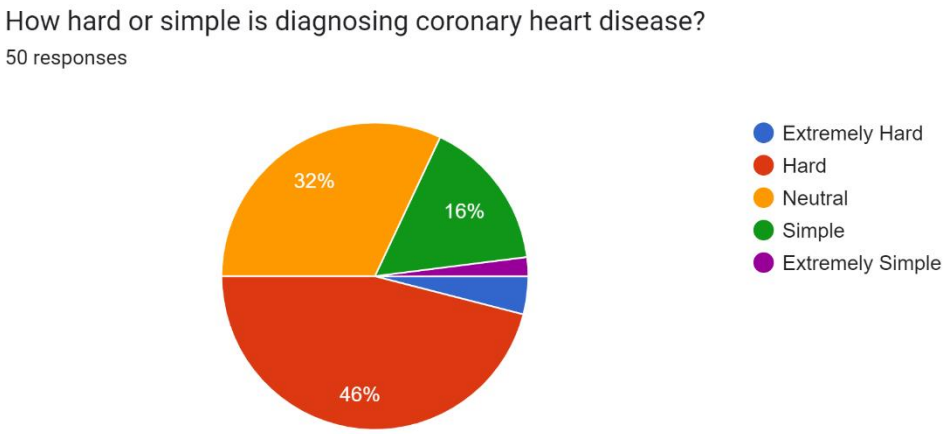

Figure 7 How hard or simple it is diagnosing Coronary Heart Disease

4.8

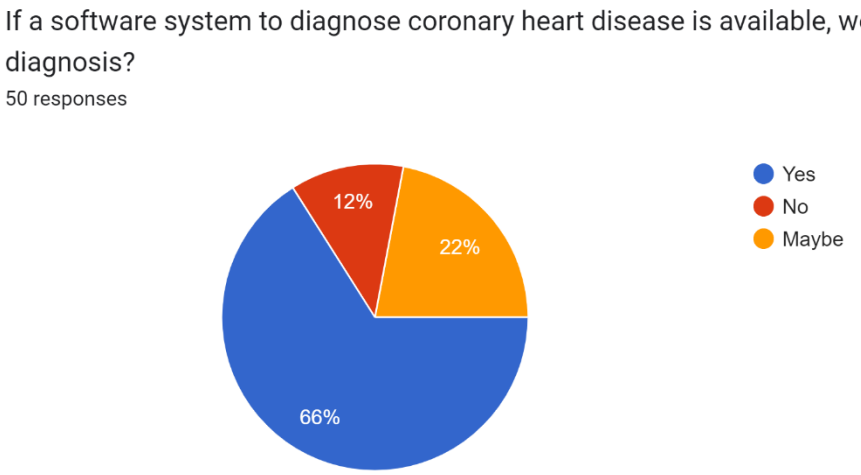

Figure 8 Willingness to use the proposed software system

4.9

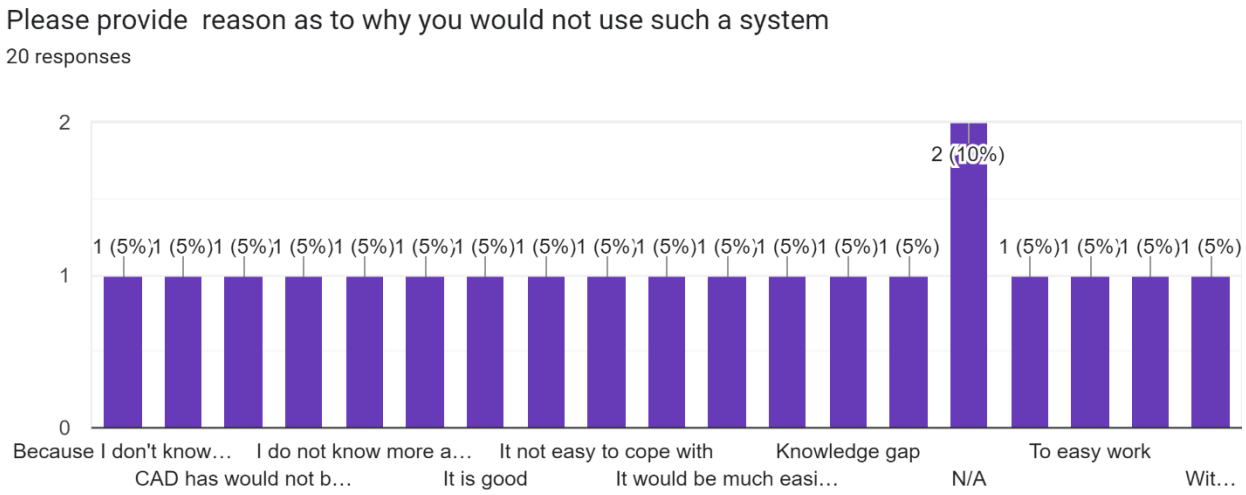

Figure 9 Reason for unwillingness in using the proposed software system

4.10

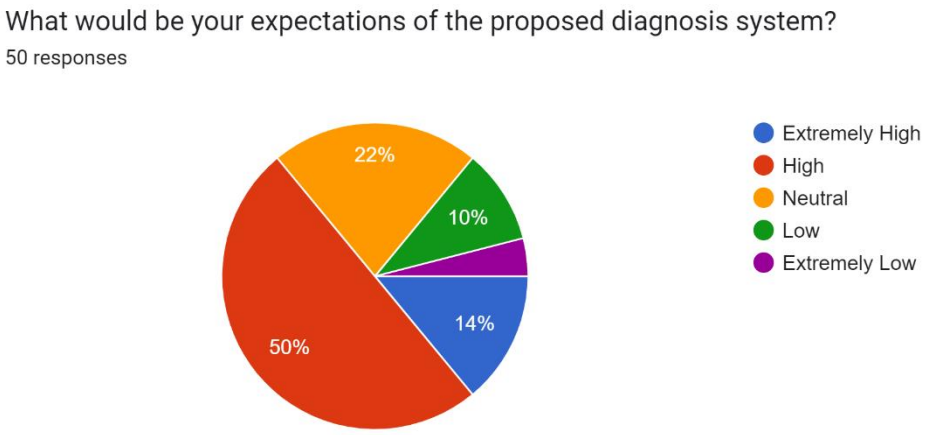

Figure 10 Expectations of the proposed software system

5 Discussion of the Results

Basing on the results obtained, we found that majority of the medical specialists know the signs and symptoms of the Coronary Heart Disease.

The results obtained indicate that majority of the medical specialists have never diagnosed the Coronary Heart Disease before.

The results indicate that majority of the specialists can easily differentiate a patient suffering from Coronary Heart Disease and any other disease.

The results show that the majority of specialists agree that diagnosing Coronary Heart Disease takes a long time.

The results show that there is a high level of difficulty in diagnosing Coronary Heart Disease.

The obtained results show that the medical specialists are willing and also have high expectations from the use of the Coronary Heart Disease Diagnosis system.

6 Conclusions and Recommendations

Since majority of the respondents know the signs and symptoms of Coronary Heart Disease, the Coronary Heart Disease Diagnosis system will have to use the signs and symptoms as an input in diagnosing the disease.

Since majority of the medical specialists have never diagnosed the Coronary Heart Disease before, this indicates the need for the Coronary Heart Disease Diagnosis system to ensure accuracy in the diagnosis results.

Since majority of specialists agree that diagnosing Coronary Heart Disease takes a long time, Coronary Heart Disease Diagnosis system is expected to provide diagnosis results in a short period of time less than a minute long.

Since the results show that a high level of difficulty in diagnosing Coronary Heart Disease, Coronary Heart Disease Diagnosis system is expected to make diagnosis simpler for the medical specialists.

Since most of the medical specialists are willing and also have high expectations from the use of the Coronary Heart Disease Diagnosis system shows the need for developing and maintaining the proposed system.

REFERENCES

[1] S.U Amin., K Agarwal., R Beg., (2013*) "*Data Mining in Clinical Decision Support Systems for Diagnosis, Prediction and Treatment of Heart Disease" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol 2, no. 1.

[2] K.H Miao, J.H Miao., G.J Miao., (2016) "Diagnosing Coronary Heart Disease using Ensemble Machine Learning" International Journal of Advanced Computer Science and Applications (IJACSA), vol 7, no. 10.

[3] S.A Pattekari., A Parveen., (2012) "Prediction system for heart disease using Naïve Bayes" International Journal of Advanced Computer and Mathematical Sciences, Vol. 3, no. 3

[4] ISF Dessai., (2013) "Intelligent heart disease prediction system using probabilistic neural network" International Journal on Advanced Computer Science and Engineering (IJACTE), vol. 2, no. 5

[5]G.C Anusha ., M.S Apoorva., N Deepthi ., V Dhanushree ., (2019). "Heart Disease Diagnosis Using Machine Learning." International Journal of Engineering Research & Technology (IJERT), vol 7, no. 10.

[6] K Cinetha., P.U Maheswari., (2014) "Decision support system for Precluding Coronary Heart Disease(CHD) using fuzzy logic" International Journal of Computer Science Trends and Technology (IJCST), vol. 2, no. 2,

[7] Sulabha S. Apte.,(2012) "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", International Journal of Computer Applications Volume 47, No.10

[8] R Chitra., V Seenivasagam., (2013) "Review of heart disease prediction system using data mining and hybrid intelligent techniques", ICTACT Journal on soft computing, volume 03, No. 04

[9]R Awang., S Palaniappan., (2008) “Intelligent heart disease predication system using data mining technique”. IJCSNS International Journal of Computer Science and Network Security. Vol. 8, No. 8

[10] S.B Patel., DPAK Sen., S Shukla., (2013) “A Data Mining Technique for Prediction of Coronary Heart Disease Using Neuro-FuzzyIntegrated Approach Two Level”, International Journal Of Engineering And Computer Science , Volume 2 no. 9

[11] P.T Ernest., HZK Zhou., P Frias., D Zeng., M Braunstein.(2013) "Supervised embedding of textual predictors with applications in clinical diagnostics for pediatric cardiology." Journal of the American Medical Informatics Association 21, no. e1

[12] B Gnaneswar ., MRE Jebarani.,(2017) "A review on prediction and diagnosis of heart failure." In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

[13] H.J Kim, S.,B Zhang., (2008)“AptaCDSS-E: A classifier ensemble based clinical decision support system for cardiovascular disease level prediction,” Expert Syst. Appl. Vol 34, No. 1

APPENDIX

Appendix A: Interview Guide

BSE23-27
INTERVIEW GUIDE

Introduction (Greetings, introducing ourselves and what we do)

Questions

1. What is coronary heart disease?
2. What are the major causes of coronary heart disease in Uganda?
3. What are the signs and symptoms of coronary heart disease?
4. Which other diseases present similar symptoms that may be confused with coronary heart disease during diagnosis?
5. Does the disease bring about other diseases? If so, which ones are they?
6. What is the possibility of suffering from the disease among the different age groups?
7. Which regions in Uganda are mostly affected by coronary heart disease?
8. How much does the diagnosis cost the patient?
9. What is the rate of increase or decrease of coronary heart disease in Uganda?
10. What are the tools are used during diagnosis?
11. What is the cost of such tools?
12. What major steps are taken during diagnosis?
13. Would you say the diagnosis of this disease is hard or easy?
14. What are the main challenges faced during the diagnosis of coronary heart disease?

Appendix B: Questionnaire

BSE23-27
Survey about Diagnosis and willingness to adopt our system about Coronary Heart Diagnosis.

This questionnaire has been prepared by a team of 4 Software engineering students doing their final year project at Makerere university. We would like to develop a system that can diagnose Coronary Heart Disease in a patient basing on the signs and symptoms. The signs and other variables will be entered into the system by the medical personnel and the system detects whether the patient is normal or requires more attention. This system will be made accessible anywhere in the country. We would like to assess how the disease is diagnosed as well as the willingness to adapt our system and the information shared in this document is confidential and will be used only for research purposes.

Dear Respondent,
We would greatly appreciate it if you could take a few minutes to complete this evaluation form.
Thank you
(Tick where the answer applies)

What is your profession?

Cardiologist		Clinical Officer	
Nurse.		Medical Officer	

Have you heard of Coronary Heart Disease before?

YES		NO	

Would you happen to know the signs and symptoms of Coronary Heart Disease?

YES		NO	

Have you diagnosed a patient with Coronary Heart Disease before?

YES		NO	

Would you easily tell the difference between of a patient suffering from Coronary Heart Disease from any other disease?

Strongly Disagree		Disagree		Strongly Agree	
Neutral		Agree			

How hard or simple is diagnosing coronary heart disease?

Extremely hard		Hard		Neutral	
Extremely simple		Simple			

I would like to specify
It takes a long while and tests to conclude the diagnosis of the patient.

Strongly Disagree		Disagree		Strongly Agree	
Neutral		Agree			

If a software system to diagnose coronary heart disease is available, would you use it during diagnosis?

YES		NO	

What would be your expectations of the proposed diagnosis system?

YES		NO	

Do you have any particular suggestion on how the system should operate in your own perspective?

……
.

Thank you for accepting to take part in this survey.